

УДК 004.032.26

РАЗРАБОТКА МЕТОДОВ И АЛГОРИТМА ФОРМАЛИЗАЦИИ ТЕКСТОВОГО ЗАПРОСА К ОНЛАЙН-СЕРВИСАМ, ГЕНЕРИРУЮЩИМ ИЗОБРАЖЕНИЯ ПОСРЕДСТВОМ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ

Ю. Л. Леохин, д.т.н., профессор, проректор по научной работе МТУСИ, Москва, Россия;
orcid.org/0000-0003-3321-4497, e-mail: y.l.leokhin@mtuci.ru

Т. Д. Фатхулин, к.т.н., доцент кафедры МК и ИТ МТУСИ, Москва, Россия;
orcid.org/0000-0003-0998-1055, e-mail: t.d.fatkhulin@mtuci.ru

Рассматривается задача исследования влияния составленного пользователем текстового запроса на качество изображения, сгенерированного нейросетевыми онлайн-сервисами. Целью работы является разработка методов составления формализованного запроса к нейросетевым сервисам и алгоритма, которые позволяют получить корректные результаты генерации изображения. Актуальность работы обусловлена возрастающим интересом к использованию нейросетевых технологий в самых различных областях жизни человека. Рассматриваются отечественные онлайн-сервисы «Шедеврум» и «Kandinsky 2.1», которые реализуют диффузионную и генеративно-сопоставительную технологии генерации изображения соответственно. Приводится методика, позволяющая формализовать запрос для генерации изображения нейронными сетями, и реализующий ее алгоритм. Приведена апробация разработанных методов и алгоритма путем ввода естественно-языковых и формализованных запросов и анализа сгенерированных нейронными сетями изображений. В заключении представлены результаты проведенного анализа, сделаны выводы.

Ключевые слова: нейронные сети, сервис, изображение, генерация, рекомендации, методика, алгоритм, запрос, методы.

DOI: 10.21667/1995-4565-2023-85-82-95

Введение

В настоящее время почти во всех сферах жизни человека все большее внедрение находят методы машинного обучения (МО, Machine Learning – ML) [1-4]. При этом одними из самых эффективных и часто используемых на практике являются методы, базирующиеся на искусственных нейронных сетях (НС, Neural Network – NN). Они позволяют обрабатывать изображения, видео- и аудио данные. Примером такого использования НС может служить обработка снимков дорожного полотна для поиска его дефектов [5-7], а также анализ видеопотока данных, который получают с видеокамер транспортных средств в режиме реального времени для ориентирования беспилотных автомобилей и общественного транспорта на дороге [1, 4, 5, 7]. Кроме того, нейросети могут использоваться для оптимизации маршрутизации в современных сетях связи [8-12]. Также НС дают возможность осуществлять прогнозирование, предсказывание значений численных величин на основе анализа определенных признаков объектов реального мира, что часто применяется, например, в экономических прогнозах на определенный временной промежуток в будущем [13].

Еще одним направлением развития НС является генерация текста и изображений. Нейросети научились создавать изображения, которые трудно отличить от реальных, созданных человеком. С учетом этого нейросети становятся мощным рабочим инструментом маркетинга в социальных сетях, редакторов, авторов, маркетологов, которые используют НС для облегчения своей работы при составлении рекламных слоганов, создании необычных рекламных баннеров [14], что обуславливает актуальность рассматриваемой предметной области.

Постановка задачи

Задача разработки методики и алгоритма формирования запроса, позволяющего сгенерировать качественное изображение в разумные временные интервалы, становится особо актуальной в свете широкого внедрения нейросетевых технологий в самых разнообразных сферах человеческой жизни, даже в тех, в которых раньше казалось их применение невозможным. Так в статье «10 примеров использования ChatGPT и Midjourney в бизнесе и не только», которая была опубликована в авторитетном российском медиахолдинге РБК [14], рассматривается, как в творческих сферах (архитектура, учебные онлайн-курсы, анимация и театр) и в бизнесе (производство, SMM, маркетинг, ретейл, общепит) используются современные достижения в области искусственного интеллекта (ИИ, Artificial Intellect – AI) и НС. При работе с текстом НС позволяют генерировать идеи для новых текстов, не зависеть от большого числа людей, их проблем, создавать тестовые материалы к учебным курсам по различным предметам, а также писать деловые письма партнерам по бизнесу. При работе с изображениями НС дают возможность создавать референсы, которые передают идеи авторов, рендеры, графические образы героев мультипликации и их костюмов, визуальные образы декораций в театре, дизайн ювелирных изделий и реклам в общепите. Однако большинство профессионалов отмечают следующие сложности при работе с ИИ и НС:

- во-первых, сгенерированный текстовый или графический материал требует доработки человеком, разбирающимся в проблеме;
- во-вторых, полученные материалы могут служить только базой, на основе которой будет создан готовый для конечного потребителя продукт;
- в-третьих, при составлении запроса к НС может уходить очень много времени на его формулировку и уточнение, если полученный материал не удовлетворяет поставленной человеком цели;
- в-четвертых, для генерации качественных изображений требуется составлять очень детализированные запросы.

Таким образом, важно учитывать, что во многом успех при создании изображения нейросетью зависит от правильно составленного запроса (текстового запроса). Соответственно возникает необходимость разработки алгоритма оптимального описания объекта, явления, процесса для максимально верной и быстрой генерации изображения нейронной сетью.

В настоящей работе будет разработан алгоритм, решающий эту задачу. Возможно, что эмпирически введенный запрос позволит получить более качественное изображение, чем по разрабатываемой методике, для одной архитектуры НС, в то время как для другой архитектуры НС результат генерации изображения может оказаться хуже. В исследовании осуществляется усреднение, унификация алгоритма формирования запроса для разных архитектур. Настоящее исследование является одной из первых попыток разработать методику формализации запроса с целью получения качественного сгенерированного изображения НС в короткий срок. Разрабатываемая методика и алгоритм создаются универсальными под разные архитектуры НС.

Теоретические исследования

Рассматриваемые технологии и инструменты генерации изображений

Для генерации изображений выделяют 2 типа (метода) нейронных сетей: генеративно-состязательные сети (GAN) [15] и диффузионные сети. Последние нашли в настоящее время большее применение [14, 16, 17]. Более подробно сравнение этих способов генерации изображений приведено в работе [15]. Также при выборе НС в настоящем исследовании предполагается использование только бесплатных, открытых решений, доступных всем пользователям сети Интернет. Следует рассматривать сервисы, которые не требуют от конечного пользователя специальной подготовки и знаний по настройке НС. Работа с сервисами должна проводиться только в вопросно-ответной форме («ответ» – сгенерированное изображение), и

должны использоваться только словесные описания запросов, без использования дополнительных опций, ключей настройки и т.п.

С учетом этого были выбраны следующие сервисы, основанные на нейросетевых методах генерации изображений: «Шедеврум» – диффузионная сеть, разработчик «Яндекс» [18] и «Kandinsky 2.1» – генеративно-сопоставительная сеть, разработчик «Сбер» [19]. Выбор этих НС определяется и тем, что «Kandinsky 2.1» и «Шедеврум» являются отечественными разработками, которые ориентированы в первую очередь на русскоязычную аудиторию. Они изначально предполагают, что запрос вводится на русском языке, они адаптированы под это. На момент написания статьи нейросетевой сервис «Шедеврум» в виде приложения для мобильных устройств может генерировать по запросу сразу 4 изображения, также он позволяет создавать новый текст по введенному краткому описанию ситуации. НС сервис «Kandinsky 2.1» дает возможность пользователю в Интернет-браузере по запросу сгенерировать изображение и даже немного редактировать его, после чего пользователь может его сохранить на своем устройстве.

Необходимо также отметить и популярные зарубежные сервисы для генерации изображений, такие как «Stable Diffusion», «Midjourney» и «DALL-E» [16, 17]. Среди профессионалов большую популярность получила НС «Stable Diffusion». «Stable Diffusion» в отличие от НС «Midjourney» генерирует самые качественные изображения из всех НС, но является платной [16, 17]) и «DALL-E» (открытое решение, имеются проблемы с генерацией лиц и человеческих фигур [16, 17]). «Stable Diffusion» генерирует довольно качественные изображения, в базовой версии требует от пользователя достаточно глубоких знаний архитектуры НС, что отражается в сложности формулируемых запросов. Учитывая указанные факторы, эти НС не будут рассмотрены в работе.

Методика описания объекта, явления, процесса и реализующий ее алгоритм

Разработаем методику описания объекта, явления, процесса с целью формирования запроса для генерации изображения нейронными сетями и реализующий ее алгоритм. Для начала в целях создания алгоритма описания объекта, явления или процесса рассмотрим лингвистический аспект данной проблемы. При формировании запроса для нейросети необходимо учитывать 2 фактора: семантику и синтаксис. Под *семантикой* в настоящем исследовании будем рассматривать значение слов или идиоматических выражений, входящих в запрос. Под *синтаксисом* – порядок и взаимосвязь слов в запросе.

Процесс создания изображения нейросетью непредсказуем. В то время как настоящий художник поэтапно планирует и осуществляет процесс создания изображения, в нейросети имеется только введенный пользователем запрос и выдаваемый нейросетью результат (ответ). Все что происходит от запроса до результата – «terra incognita». Результат порой оказывается совершенно отличным от ожидаемого. Для получения требуемого изображения иногда затрачивается слишком много времени. Пользователь, составляя запрос, может либо использовать готовые клише, подставляя в них свои лексические единицы, либо идти путем проб и ошибок, на практике пробуя разные варианты, меняя порядок слов, что требует массу времени. Следовательно, необходимо научиться таким образом составлять запрос, чтобы нейросеть максимально быстро правильно его понимала.

Итак, для создания изображения в нейросети необходимо ввести текстовый запрос. Нужно сразу определиться, что это не простые описания объекта, явления или процесса, как в естественных языках, которые предназначены для общения людей. Формулировки типа «серебряный век», «костер рябины красной», «бальзаковский возраст» возможны в естественных языках, так как предполагают либо определенный образовательный и культурный задел («educational background»), либо свойственную только человеку метафоричность мышления.

Чтобы избежать некорректного результата, в запросе для нейросети должны быть четко, однозначно отображены основные свойства реального объекта, изображение которого требуется получить. Для этого необходимо провести замену обычного описания объекта, свой-

ственного для естественных языков, его формальным описанием, когда в словесной форме необходимо четко выразить связи и существенные признаки объекта. Этот процесс можно назвать *формализацией*. На рисунке 1 представлены этапы формализации. На первом месте находится естественное описание объекта, свойственное для человеческого общения. Затем в процессе анализа на первое место помещается объект, так как именно он будет в центре изображения, а также выделяются его существенные признаки. Далее происходит непосредственно формализация, формулируется запрос для НС.



Рисунок 1 – Этапы преобразования естественно-языковой формулировки в запрос для НС
Figure 1 – Stages of converting natural language formulation into a request for NN

Для понимания, как НС генерирует изображение, необходимо разнообразить формулировки запросов, описаний объектов, явлений и процессов, а затем изучать результат [14, 16, 17]. К настоящему времени накоплен определенный практический опыт формализации описания или корректного составления запросов для нейросети [14, 16, 17], на основе которого можно сформулировать следующие методы, входящие в разрабатываемую методику:

1) назовите объект, процесс или явление (например: «театр», «школа», «космос», «человек», «тигр», «погоня», «полет», «оттепель» и т.д.);

2) добавьте признаки объекта: «красный шар», «молодой мужчина», «большой дом» (цвет, размер, настроение, возраст и т.п.), при этом, если признак выражен не качественным, а относительным прилагательным, то его лучше давать **после** самого объекта, используя так называемое пост-позитивное определение (для НС лучше дать формулировку не «медный самовар», «царственный орел», а «самовар из меди», «орел, похожий на царя»);

3) свойство или характеристику объекта возможно также задать через предлоги («ребенок на самокате», «кот внутри коробки», «машина в космосе», «гостиница, как средневековый замок»);

4) можно объединить два объекта или свойства («пес-самурай», «рюкзак-самокат», «пчела-фея», «орел-царь» и т.п.);

5) чтобы помочь нейросети сделать максимально верный выбор, можно добавить к прилагательному синонимы («большой», «огромный», «гигантский» и т.п.);

6) чтобы итоговое изображение было динамичным, укажите действие, которое выполняет объект: «летит», «прыгает», «смеется» и т.п.;

7) если объектов несколько, необходимо указать их точное количество (корректно на изображении будут сгенерированы **не более трех-четырех** объектов, при запросе объекта во множественном числе или добавлении слов «несколько», «много», «толпа» – НС будет рисовать случайное произвольное число объектов, чем больше будет этих объектов, тем **хуже будет детализация** изображения), числительное можно ввести как цифрой, так и словом;

8) если нейросети предстоит сгенерировать абстрактное понятие, лучше конкретизировать запрос, по возможности заменяя абстракции объектами, используя свои личные ассоциации, чтобы нейросеть поняла, что вам нужно (например, вместо «мир сквозь розовые очки» лучше запросить «мир счастливый и беспечный» или «мир радостный и беззаботный»);

9) не рекомендуется в запросах к НС употреблять слова и словосочетания «без», «кроме», «за исключением», так как нейросеть не обращает на них внимания, и на запрос «моло-

дой человек без усов» вероятнее всего сгенерирует мужчину с усами, лучше ввести запрос «тщательно выбритый молодой человек».

Также стоит отметить, что в современных нейросетевых сервисах возможно указать в запросе, в каком разрешении, с каким уровнем детализации, каких размеров, на каком фоне и в каком стиле следует сгенерировать изображение [16, 17]. Так, например, для этого можно использовать слова и словосочетания «детализированный», «очень детализированный», «в 4К» или «в 16К», а также «анимационный» или «мультипликационный». Имеется возможность представить сгенерированное изображение так, как будто оно сделано на «виртуальную камеру» с ее соответствующими настройками [вид камеры («Kodak», «Nikon», «Polaroid»), особенность используемой линзы, эффекты камеры («блюр»), и расстояние до объекта] [16, 17]. В рамках настоящего исследования не рассматриваются эти способы составления запросов, влияющие на качество и вид генерируемого изображения. В перспективе они будут рассмотрены при продолжении исследований.

На рисунке 2 представлен алгоритм формирования запроса к НС для генерации изображения, разработанный в соответствии с приведенными выше рекомендациями.

Далее приводится апробация представленного алгоритма формирования запроса к НС для генерации изображения посредством нейросетевых онлайн-сервисов «Kandinsky 2.1» и «Шедеврум».

Экспериментальные исследования

С учетом указанных ранее рекомендаций по составлению запросов апробируем разработанный алгоритм и определим визуально его эффективность по сравнению с обычными неформализованными естественно-языковыми запросами, которые может вводить пользователь. Для этого введем в указанные нейросетевые онлайн-сервисы приведенные ниже запросы.

Запрос 1 (естественно-языковой): «Отдыхающий на диване кот».

Запрос 2 (формализованный): «Кот отдыхает на диване».

Результаты генерации изображений по Запросам 1 и 2 приведены в таблице 1. Запросы являются довольно простыми, сгенерированные изображения по смыслу являются идентичными, имеются только различия в визуализации.

Запрос 3 (естественно-языковой): «Молодой человек без усов».

Запрос 4 (формализованный): «Тщательно побритый молодой человек».

Результаты генерации изображений по Запросам 3 и 4 приведены в таблице 2. При реализации Запроса 3 имеется критическая ошибка в полученном изображении (нейросеть рисует усы). Результат формализованного Запроса 4 соответствует ожиданиям пользователя.

Запрос 5 (естественно-языковой): «Совы на ветке».

Запрос 6 (формализованный): «10 сов на ветке».

Запрос 7 (формализованный): «3 совы на ветке».

Результаты генерации изображений по Запросам 5-7 приведены в таблице 3. Из полученных результатов видно, что неопределенное число сов в запросе приводит к их случайному числу на сгенерированных изображениях. При запросе изображения с 10 совами ни одна НС не сгенерировала результат корректно, к тому же качество полученных изображений невысокое, имеются графические искажения. При запросе изображения с 3 совами получается более корректный результат.

Запрос 8 (естественно-языковой): «Мир сквозь розовые очки».

Запрос 9 (формализованный): «Мир счастливый и беспечный».

Запрос 10 (формализованный): «Мир радостный и беззаботный».

Результаты генерации изображений по Запросам 8-10 приведены в таблице 4. Абстрактное понятие «Мир сквозь розовые очки» воспринимается НС буквально, генерируются розовые очки, пейзаж, наблюдаемый сквозь очки розового цвета. Сгенерированные по Запросам 9 и 10 изображения чуть ближе к семантике данного идиоматического выражения.

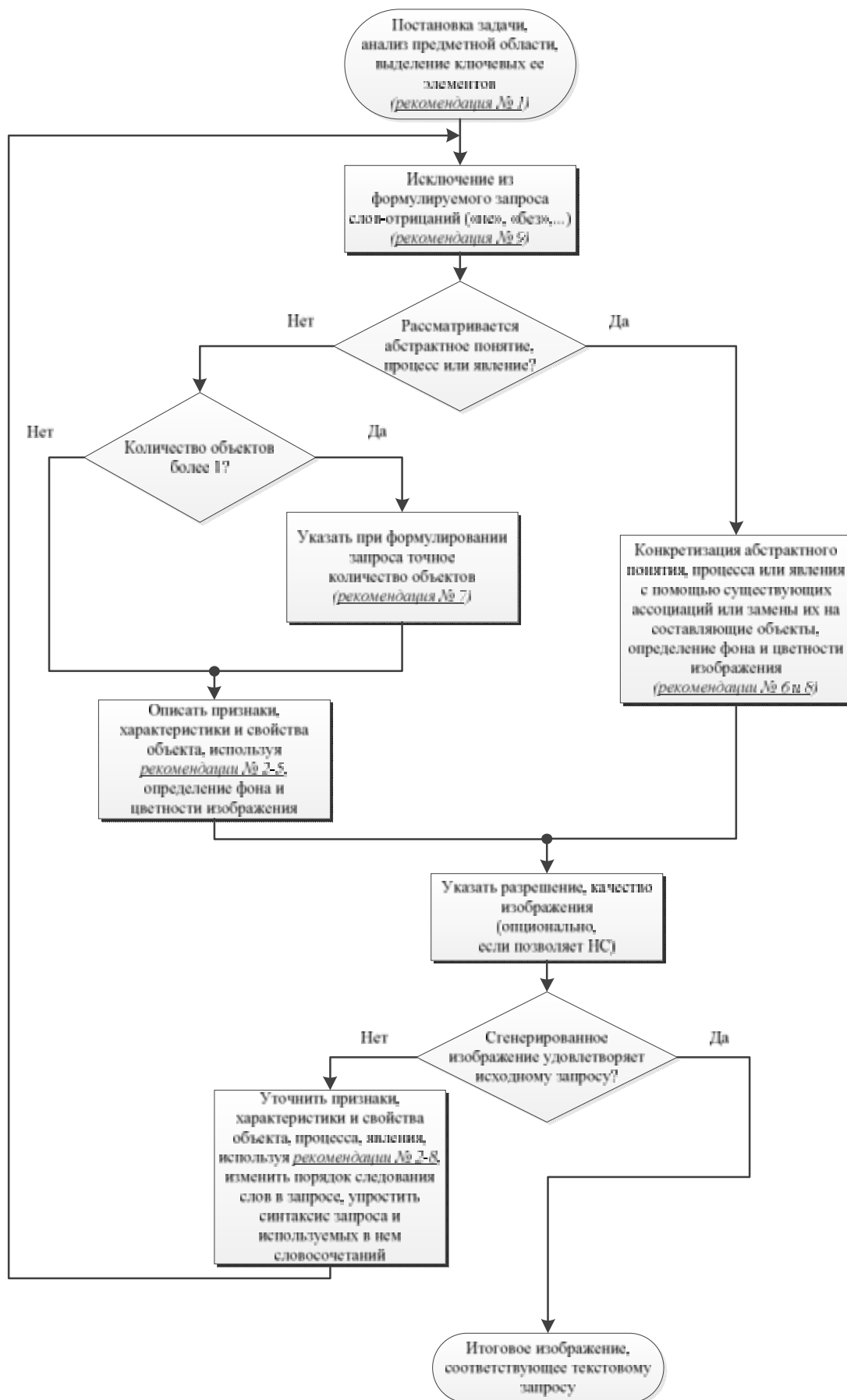


Рисунок 2 – Разработанный алгоритм формирования запроса к НС для генерации изображения
 Figure 2 – Developed algorithm for generating a request to NN for image generation

Таблица 1 – Результаты генерации изображений НС в соответствии с Запросами 1 и 2
Table 1 – Results of NN Images Generation According to Requests 1 and 2

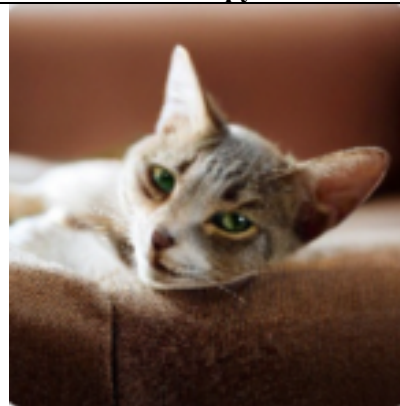
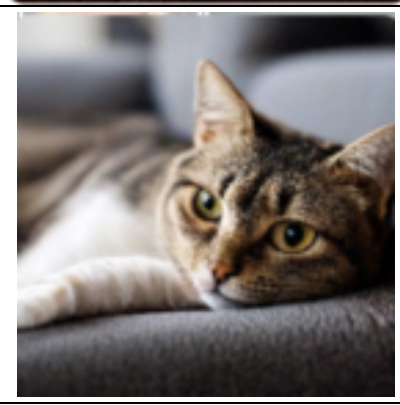
	«Кандинский 2.1»	«Шедевр» (2×2)	«Шедевр»
Запрос 1			
Запрос 2			

Таблица 2 – Результаты генерации изображений НС в соответствии с Запросами 3 и 4
Table 2 – Results of NN Images Generation According to Requests 1 and 2

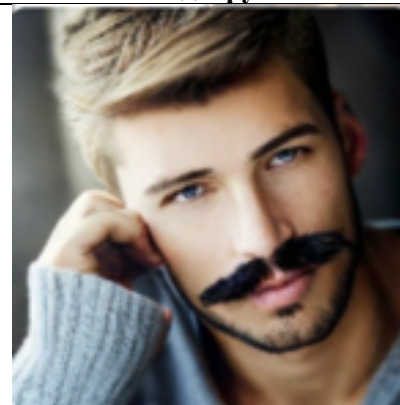
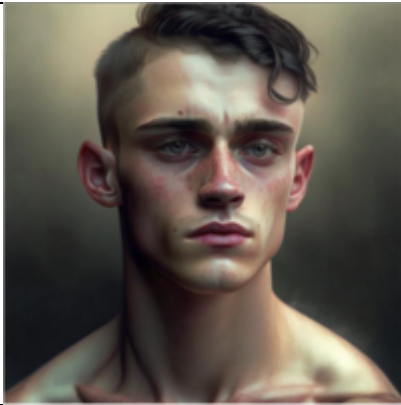
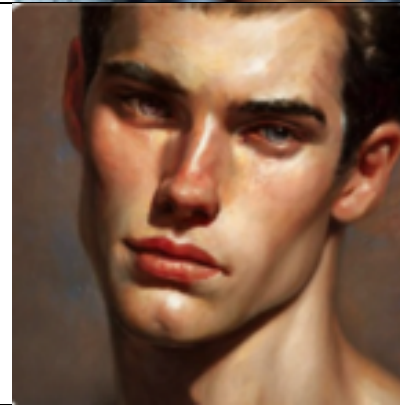
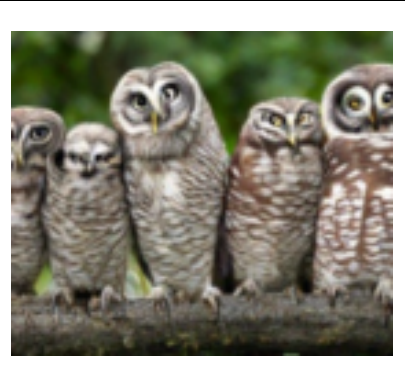
	«Кандинский 2.1»	«Шедевр» (2×2)	«Шедевр»
Запрос 3			
Запрос 4			

Таблица 3 – Результаты генерации изображений НС в соответствии с Запросами 5-7
Table 3 – Results of NN Images Generation According to Requests 5-7

	«Кандинский 2.1»	«Шедевр» (2×2)	«Шедевр»
Запрос 5			
Запрос 6			
Запрос 7			

Запрос 11 (естественно-языковой): «Царственный орел».

Запрос 12 (формализованный): «Орел-царь».

Результаты генерации изображений по Запросам 1 и 2 приведены в таблице 5. В Запросе 11 НС игнорирует определение «царственный», так как не может верно интерпретировать глубокий смысл качественного прилагательного. В Запросе 12 прилагательное заменяется именем существительным, т.е. конкретным объектом, НС легче справляется с таким запросом.

Как видно из таблицы 3 даже при вводе запроса «3 совы на ветке» имеется изображение, сгенерированное нейросетевым сервисом «Шедевр», на котором видно только 2 совы. В ходе проведения именно этого эксперимента (ввода Запросов 5-7) было проблематично определить сколько объектов (в данном случае сов) может корректно сгенерировать каждая из рассматриваемых нейронных сетей. Эмпирически было определено, что более-менее корректный результат получается при запросе 3 или 4 объектов (сов). Наилучшее соответствие результата вводимому запросу наблюдается именно для 3 объектов, что обуславливает приведение именно этого результата в настоящей работе. Со временем нейросетевые сервисы позволят получать корректные результаты для большего числа запрашиваемых объектов.

Таблица 4 – Результаты генерации изображений НС в соответствии с Запросами 8-10
Table 4 – The Results of NN Images Generation According to Requests 8-10

	«Кандинский 2.1»	«Шедеврум»(2×2)	«Шедеврум»
Запрос 8			
Запрос 9			
Запрос 10			

В результате проведенного исследования было проведено сравнение изображений, сгенерированных согласно введенным естественно-языковым и формализованным по разработанному алгоритму запросам. При генерации изображений посредством отечественных нейросетевых онлайн-сервисов «Кандинский 2.1» и «Шедеврум» было выявлено следующее.

Нейросетевой сервис «Шедеврум» генерирует изображения по введенному запросу дольше и чуть менее детализировано, чем сервис «Кандинский 2.1», однако на выбор пользователю предлагается сразу 4 изображения, причем в разных стилях без предварительного задания их в запросе. Это обстоятельство дает возможность выбора и может подтолкнуть пользователя к более удачному выбору изображения для своих целей (например, при поиске идей для рекламного баннера, визуала, референса и т.п.).

В свою очередь нейросетевой сервис «Кандинский 2.1» генерирует 1 изображение в более высоком разрешении и с большей детализацией в случайном стиле, наиболее подходящем к запросу по оценке самой НС, если стиль не указан в запросе пользователем самостоятельно.

Таблица 5 – Результаты генерации изображений НС в соответствии с Запросами 11 и 12
Table 5 – The Results of NN Images Generation According to Requests 11 and 12

	«Кандинский 2.1»	«Шедевр» (2×2)	«Шедевр»
Запрос 11			
Запрос 12			

Заключение

В ходе проведенных экспериментов по составлению запросов к нейросетевым сервисам «Кандинский 2.1» и «Шедевр» полученные результаты генерации изображений показали следующее. Разработанный алгоритм формирования запроса к НС для генерации изображения позволяет получить более корректный результат по сравнению с обычным естественно-языковым запросом. Особенно это заметно на запросах, содержащих отрицательные частицы и предлоги («без», «не»), относительные прилагательные («царственный орел»), предполагающих представление абстрактного понятия («мир сквозь розовые очки»), а также при использовании числительных в запросе. При вводе простых запросов и запросов, содержащих качественные прилагательные («отдыхающий на диване кот»), нейросетевые сервисы выдают примерно одинаковые смысловые результаты на изображениях как при использовании разработанного алгоритма, так и без его использования.

Таким образом, разработанный алгоритм формирования запроса к НС для генерации изображения показал свою эффективность. В перспективе планируется его модификация и модернизация с учетом увеличивающегося функционала современных нейросетевых сервисов как отечественных, так и зарубежных разработчиков.

Библиографический список

1. M. G. Gorodnichev, M. D. Gromov, K. A. Polyantseva, and M. S. Moseva. «Research and Development of a System for Determining Abnormal Human Behavior by Video Image Based on Deepstream Technology», 2022 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), St. Petersburg, Russian Federation, 2022, pp. 1-9. DOI: 10.1109/WECONF55058.2022.9803767.
2. M. G. Gorodnichev, K. A. Dzhabrailov, K. A. Polyantseva, and R. A. Gematudinov. «On Automated Safety Distance Monitoring Methods by Stereo Cameras», 2020 Systems of Signals Generating and

Processing in the Field of on Board Communications, Moscow, Russia, 2020, pp. 1-8, DOI: 10.1109/IEEECONF48371.2020.9078616.

3. **M. G. Gorodnichev, R. A. Gematudinov, K. A. Dzhabrailov, and T. D. Potapchenko.** «The Concept of an Automated Weight and Size Control System for Measuring the Mass of Freight Vehicles in a Traffic Flow (Weight-In-Motion)», 2019 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), Russia, 2019, pp. 1-5, DOI: 10.1109/SYNCHROINFO.2019.8814120.

4. **O. I. Sheluhin, D. V. Kostin, and M. G. Gorodnichev.** «Monitoring Anomalous States of Computer Systems by Intellectual Analysis of Data of System Journals», 2020 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russia, 2020, pp. 1-6, DOI: 10.1109/IEEECONF48371.2020.9078632.

5. **K. A. Polyantseva, M. G. Gorodnichev, M. S. Moseva, and T. D. Potapchenko.** «On the Applicability of Neural Networks in the Tasks of Detecting Dangerous Movement», 2019 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), 2019, pp. 1-4, doi: 10.1109/WECONF.2019.8840654.

6. **M. G. Gorodnichev, K. A. Polyantseva, M. S. Moseva, and M. V. Yashina.** «On Some Methods for Correcting Accelerometer and Gyroscope Data Using Machine Learning Algorithms», 2019 International Conference «Quality Management, Transport and Information Security, Information Technologies» (IT&QM&IS), 2019, pp. 186-191, doi: 10.1109/ITQMIS.2019.8928397.

7. **M. S. Moseva, M. G. Gorodnichev, K. A. Polyantseva, G. G. Vlasov, and M. D. Gromov.** «Development of a System for Fixing Road Markings in Real Time», 2022 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), Arkhangelsk, Russian Federation, 2022, pp. 1-7, DOI: 10.1109/SYNCHROINFO55067.2022.9840971.

8. **Леохин, Ю. Л., Фатхулин, Т. Д.** Оценка возможности предоставления гарантированной скорости передачи данных в программно-конфигурируемой оптической сети // Вестник Рязанского государственного радиотехнического университета. 2020. № 71. С. 45-59. DOI 10.21667/1995-4565-2020-71-45-59. EDN GQTGUP.

9. **D. Perepelkin, and T. Nguyen,** «Research of Multipath Routing Processes in Software Defined Networks Based on Ant Colony Optimization», 2022 11th Mediterranean Conference on Embedded Computing (MECO), Budva, Montenegro, 2022, pp. 1-6, DOI: 10.1109/MECO55406.2022.9797090.

10. **Перепелкин, Д. А., Нгуен, В. Т.** Исследование и анализ процессов многопутевой маршрутизации и балансировки потоков данных в программно-конфигурируемых сетях на основе генетического алгоритма // Вестник Рязанского государственного радиотехнического университета. 2022. № 79. С. 31-48. DOI 10.21667/1995-4565-2022-79-31-48. EDN XGSYWX.

11. **Перепелкин, Д. А., Фам, А. М.** Математические модели планирования упорядоченного набора операций для распределения разнородных ресурсов в промышленных телекоммуникационных сетях // Вестник Рязанского государственного радиотехнического университета. 2022. № 79. С. 56-67. DOI 10.21667/1995-4565-2022-79-56-67. EDN BMTQGM.

12. **Перепелкин, Д. А., Анисимов, К. В.** Разработка архитектуры и системы нечетких метрик каналов связи программно-конфигурируемой сети устройств Интернета вещей // Вестник Рязанского государственного радиотехнического университета. 2022. № 80. С. 53-66. DOI 10.21667/1995-4565-2022-80-53-66. EDN UPDKKO.

13. **E. G. Kukharensko, I. A. Korkunov, M. G. Gorodnichev, and T. U. Salutina,** «On the Introduction of Digital Economics in the Transport Industry», 2019 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russia, 2019, pp. 1-5, DOI: 10.1109/SOSG.2019.8706797.

14. 10 примеров использования ChatGPT и Midjourney в бизнесе и не только [Электронный ресурс]: // Обзор функционала и проблем популярных нейросетевых сервисов. – Режим доступа: https://www.rbc.ru/spb_sz/11/06/2023/6483048b9a794751683574bf, свободный. – Загл. с экрана. (дата обращения 11.06.2023).

15. **Митрофанов, А. О., Степанов, М. Н., Фатхулин, Т. Д.** Анализ нейросетевых методов генерации изображения по текстовому запросу // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2022. № 1. С. 19-23. EDN CWRLQA.

16. Как правильно составлять запросы для нейросетей: формулы, шаблоны и примеры [Электронный ресурс]: // Описание способов составления запросов к нейронным сетям. – Режим доступа: <https://smmplanner.com/blog/kak-pravilno-sostavliat-zaprosy-dlia-nieirosietiei-formuly-shablony-i-primiery/>, свободный. Загл. с экрана. (дата обращения 11.06.2023).

17. Как формулировать запросы для нейросети [Электронный ресурс]: // Описание способов составления запросов к нейронным сетям для генерации изображений. Режим доступа: https://pikabu.ru/story/kak_formulirovat_zaprosyi_dlya_neyroseti_9980895, свободный. Загл. с экрана. (дата обращения 11.06.2023).

18. «Яндекс» представил нейросеть «Шедеврум» для превращения текста в изображения, и выпустил для неё приложения [Электронный ресурс]: // Описание функциональных возможностей нейросетевого сервиса «Шедеврум». Режим доступа: <https://3dnews.ru/1084572/yandeks-predstavil-neyroset-shedevrum-dlya-generatsii-izobrazheniy-i-eyo-prilogeniya>, свободный. Загл. с экрана. (дата обращения 11.06.2023).

19. Новая нейросеть Kandinsky 2.1 от «Сбер» создает уникальные изображения на нескольких языках включая русский [Электронный ресурс]: // Описание функциональных возможностей нейросетевого сервиса «Kandinsky 2.1». Режим доступа: <https://dzen.ru/a/ZCхXuqhnwGKCamvF>, свободный. Загл. с экрана. (дата обращения 11.06.2023).

UDC 004.032.26

DEVELOPMENT OF METHODS AND ALGORITHM FOR FORMALIZING A TEXT REQUEST TO ONLINE SERVICES GENERATING IMAGES THROUGH NEURAL NETWORK TECHNOLOGIE

Yu. L. Leokhin, Dr. Sc. (Tech.), full professor, Deputy Rector for Scientific Work, MTUCI, Moscow, Russia; orcid.org/0000-0003-3321-4497, e-mail: y.l.leokhin@mtuci.ru

T. D. Fatkhulin, Ph.D. (Tech.), associate professor, department of MC and IT, MTUCI, Moscow, Russia; orcid.org/0000-0003-0998-1055, e-mail: t.d.fatkhulin@mtuci.ru

*The problem of studying the influence of a text request compiled by a user on the image quality generated by neural network online services is considered. **The aim** of the work is to develop methods for compiling a formalized request to neural network services and an algorithm, that allow obtaining correct results of image generation. The relevance of the work is due to the growing interest in the use of neural network technologies in various areas of human life. The domestic online services «Shedevrum» and «Kandinsky 2.1» are considered, which implement diffusion and generative-adversarial generation technologies of image generation, respectively. A technique is given that allows to formalize a request for image generation by neural networks and an algorithm implementing it. The approbation of the developed methods and algorithm is given by entering human language requests and formalized requests and analyzing images generated by neural networks. In conclusion, the results of the analysis are presented and findings are shown.*

Key words: neural networks, service, image, generation, recommendations, technique, algorithm, request, methods.

DOI: 10.21667/1995-4565-2023-85-82-95

References

1. **M. G. Gorodnichev, M. D. Gromov, K. A. Polyantseva, and M. S. Moseva.** «Research and Development of a System for Determining Abnormal Human Behavior by Video Image Based on Deepstream Technology», 2022 *Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, St. Petersburg, Russian Federation, 2022, pp. 1-9, DOI: 10.1109/WECONF55058.2022.9803767.
2. **M. G. Gorodnichev, K. A. Dzhabrailov, K. A. Polyantseva, and R. A. Gematudinov.** «On Automated Safety Distance Monitoring Methods by Stereo Cameras», 2020 *Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, Russia, 2020, pp. 1-8, DOI: 10.1109/IEEECONF48371.2020.9078616.
3. **M. G. Gorodnichev, R. A. Gematudinov, K. A. Dzhabrailov, and T. D. Potapchenko.** «The Concept of an Automated Weight and Size Control System for Measuring the Mass of Freight Vehicles in a Traf-

fic Flow (Weight-In-Motion)», 2019 *Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)*, Russia, 2019, pp. 1-5, DOI: 10.1109/SYNCHROINFO.2019.8814120.

4. **O. I. Sheluhin, D. V. Kostin, and M. G. Gorodnichev.** «Monitoring Anomalous States of Computer Systems by Intellectual Analysis of Data of System Journals», 2020 *Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, Russia, 2020, pp. 1-6, DOI: 10.1109/IEEECONF48371.2020.9078632.

5. **K. A. Polyantseva, M. G. Gorodnichev, M. S. Moseva, and T. D. Potapchenko.** «On the Applicability of Neural Networks in the Tasks of Detecting Dangerous Movement», 2019 *Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, 2019, pp. 1-4, doi: 10.1109/WECONF.2019.8840654.

6. **M. G. Gorodnichev, K. A. Polyantseva, M. S. Moseva, and M. V. Yashina.** «On Some Methods for Correcting Accelerometer and Gyroscope Data Using Machine Learning Algorithms», 2019 *International Conference «Quality Management, Transport and Information Security, Information Technologies» (IT&QM&IS)*, 2019, pp. 186-191, doi: 10.1109/ITQMIS.2019.8928397.

7. **M. S. Moseva, M. G. Gorodnichev, K. A. Polyantseva, G. G. Vlasov, and M. D. Gromov.** «Development of a System for Fixing Road Markings in Real Time», 2022 *Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)*, Arkhangelsk, Russian Federation, 2022, pp. 1-7, DOI: 10.1109/SYNCHROINFO55067.2022.9840971.

8. **Leokhin, Yu. L., Fatkhulin, T. D.** Otcenka vozmozhnosti predstavleniya garantirovannoi skorosti peredachi dannyh v programmno-konfiguriruemoi opticheskoi seti. *Vestnik Ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta*. 2020, no. 71, pp. 45-59. DOI 10.21667/1995-4565-2020-71-45-59. EDN GQTGUP.

9. **D. Perepelkin, and T. Nguyen,** «Research of Multipath Routing Processes in Software Defined Networks Based on Ant Colony Optimization», 2022 *11th Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro, 2022, pp. 1-6, DOI: 10.1109/MECO55406.2022.9797090.

10. **Perepelkin, D. A., Nguyen, V. T.** Issledovanie i analiz processov mnogoputevoj marshrutizatsii i balansirovki potokov dannyh v programmno-konfiguriruemym setyah na osnove geneticheskogo algoritma // *Vestnik Ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta*. – 2022. – no. 79. – pp. 31-48. – DOI 10.21667/1995-4565-2022-79-31-48. – EDN XGSYWX.

11. **Perepelkin, D. A., Fam, A. M.** Matematicheskie modeli planirovaniya uporyadochennogo nabora operatsii dlya raspredeleniya raznorodnyh resursov v promyshlennykh telekommunikatsionnykh setyah. *Vestnik Ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta*. 2022, no. 79, pp. 56-67. DOI 10.21667/1995-4565-2022-79-56-67. EDN BMTQGM.

12. **Perepelkin, D. A., Anisimov, K. V.** Razrabotka arhitektury i sistemy nechetkikh metrik kanalov svyazi programmno-konfiguriruemoi seti ustroystv Interneta veshhej. *Vestnik Ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta*. 2022, no. 80, pp. 53-66. DOI 10.21667/1995-4565-2022-80-53-66. EDN UPDKKO.

13. **E. G. Kukharensky, I. A. Korkunov, M. G. Gorodnichev, and T. U. Salutina,** «On the Introduction of Digital Economics in the Transport Industry», 2019 *Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, Russia, 2019, pp. 1-5, DOI: 10.1109/SOSG.2019.8706797.

14. 10 primerov ispol'zovaniya ChatGPT i Midjourney v biznese i ne tol'ko [Elektronnyy resurs]: Obzor funktsionala i problem populyarnykh nejrosetevykh servisov. Rezhim dostupa: https://www.rbc.ru/spb_sz/11/06/2023/6483048b9a794751683574bf, svobodnyj. Zagl. s jekrana. (data obrashheniya 11.06.2023).

15. **Mitrofanov, A. O., Stepanov, M. N., Fatkhulin, T. D.** Analiz nejrosetevykh metodov generatsii izobrazheniya po tekstovomu zaprosu. *Trudy Severo-Kavkazskogo filiala Moskovskogo tekhnicheskogo universiteta svyazi i informatiki*. 2022, no. 1, pp. 19-23. EDN CWRLQA.

16. Kak pravil'no sostavljat' zaprosy dlya nejrosetej: formuly, shablony i primery [Elektronnyy resurs]: Opisanie sposobov sostavleniya zaprosov k nejronnym setjam. Rezhim dostupa: <https://smmplanner.com/blog/kak-pravilno-sostavliat-zaprosy-dlya-nejrosietiei-formuly-shablony-i-primery/>, svobodnyj. Zagl. s jekrana. (data obrashheniya 11.06.2023).

17. Kak formulirovat' zaprosy dlja nejroseti [Jelektronnyj resurs]: Opisanie sposobov sostavleniya zaprosov k nejronnym setjam dlja generacii izobrazhenij. Rezhim dostupa: https://pikabu.ru/story/kak_formulirovat_zaprosyi_dlya_neyroseti_9980895, svobodnyj. Zagl. s jekrana. (data obrashhenija 11.06.2023).

18. «Yandex» predstavil nejroset' «Shedevrum» dlja prevrashhenija teksta v izobrazhenija, i vypustil dlja nejo prilozhenija [Jelektronnyj resurs]: Opisanie funkcional'nyh vozmozhnostej nejrosetevogo servisa «Shedevrum». Rezhim dostupa: <https://3dnews.ru/1084572/yandeks-predstavil-neyroset-shedevrum-dlya-generatsii-izobrazheniy-i-eyo-prilogeniya>, svobodnyj. Zagl. s jekrana. (data obrashhenija 11.06.2023).

19. Novaja nejroset' Kandinsky 2.1 ot «Sber» sozdaet unikal'nye izobrazhenija na neskol'kih jazykah vključaja russkij [Jelektronnyj resurs]: Opisanie funkcional'nyh vozmozhnostej nejrosetevogo servisa «Kandinsky 2.1». Rezhim dostupa: <https://dzen.ru/a/ZCxXuqhnwGKCamvF>, svobodnyj. Zagl. s jekrana. (data obrashhenija 11.06.2023).