Александр Васильевич Якунин

Санкт-Петербургский государственный университет

a.yakunin@spbu.ru

**Генеративная визуализация в контексте дизайн-мышления**

Предлагаемое исследование посвящено сопоставительному анализу возможностей двух распространенных нейросетей для генерации изображений. Результаты исследования показывают различия в алгоритмах искусственного интеллекта обоих инструментов визуализации и степень их соответствия замыслу дизайнера.

Ключевые слова: искусственный интеллект, генеративное изображение, нейросети, техногенное проектирование.

Развитие технологий искусственного интеллекта обусловило появление на медиарынке доступных инструментов для создания изображений, созданных на основе алгоритмов нейросетей.

Генерация изображений на основе ИИ позволяет разрабатывать визуальные концепции быстрее и с меньшими затратами ресурсов, предлагая дизайнерам целый ряд дополнительных возможностей для проектирования. Искусственный интеллект способен значительно повысить эффективность дизайн-мышления [3]. Благодаря технологиям искусственного интеллекта дизайнеры получили возможность определять оптимальный путь проектирования на основе генерации множества концепций стиля, формы и детализации, что существенно ускоряет работу. Кроме того, подобные технологии усиливают способность дизайнера мыслить нестандартно [2].

 Однако в контексте подобных тенденций и возможностей остро встает вопрос об эффективности коммуникации между ИИ и его пользователем. В частности, способен ли алгоритм в основе ИИ-инструмента адекватно понимать описания, составленные дизайнером? С другой стороны, насколько близки сгенерированные изображения к первоначальным ожиданиям пользователя? Эти вопросы будут еще долго сопровождать внедрение систем ИИ в медиапрактику по мере их технологической эволюции.

В предлагаемом исследовании генерация изображений рассматривается как вспомогательное средство проектирования, эффективность которого зависит от степени взаимопонимания пользователя и нейросети. В качестве критерия оценки такой эффективности выступает оценка близости результата генерации к первоначальной идее дизайнера.

 Цель данного исследования – проверка точности изображений, полученных на основе текстового описания-эталона с помощью двух сервисов искусственного интеллекта. В качестве технологических инструментов были использованы нейросети Stable Diffusion и Leonardo AI. Изображения существующих дизайнерских продуктов-брендов были описаны с помощью атрибутивных прилагательных, описания преобразованы в промт-запросы, которые были введены в обе нейросети.

В исследовании приняли участие шесть студентов-дизайнеров с опытом работы в сфере дизайна. Их задача состояла в том, чтобы описать три стула и три чайника с использованием списка прилагательных, распределенных по четырем категориям: «стиль», «форма», «материал» и «цвет». Затем выбранные прилагательные были переведены в структуру промт-запроса, на основе которых была запущена генерация в нейросетях. Затем полученные результаты оценивались на предмет сходства с оригиналом – для каждого продукта участникам предлагалось выбрать из сгенерированных изображений наиболее близкое к оригинальному продукту и оценить данное сходство по шкале от 1 до 10. Данная методика оценки сходства между генерацией нейросетей и реальными образцами основана на предыдущих исследованиях [1].

 На завершающем этапе исследования оценивалось воспринимаемое юзабилити интерфейса двух нейросетевых приложений. Степень субъективной удовлетворенности оценивалась с помощью опросника SUS (шкалы удобства использования системы). Также было проведено полуструктурированное интервью для выяснения различий между двумя программными продуктами в пользовательском опыте, в частности, оценки их потенциала в расширении возможностей мышления.

 Как свидетельствуют результаты эксперимента, единые принципы алгоритмов существующих на медиарынке нейросетей существенно различаются в художественной достоверности генерируемых образов. В частности, Leonardo AI лучше, чем Stable Diffusion, выражает визуальные характеристики оригинального продукта – цвет, форму и технические детали. С другой стороны, Stable Diffusion обладает бóльшими возможностями в репрезентации генерируемых объектов с разных ракурсов. С другой стороны, исследование показало положительную корреляцию между способностью нейросетей генерировать валидные визуальные образы и степенью субъективной удовлетворенности их интерфейсом.

Литература

1. Chuang M. C., Chen C. Exploring the relationship between the product form features and feature composition and user’s Kansei evaluation // Journal of Design. 2004. No 9. P. 43–58.
2. Hung Y. C., Lin Y. H., Lin C. C., Liu C. T. The application of kansei engineering and morphological analysis in product form design // International Journal of Systematic Innovations. 2014. Vol. 3.
3. Karras T., Laine S., Aittala M., Hellsten J., Lehtinen J., Aila T. Analyzing and improving the image quality of StyleGAN // Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition. 2020. P. 8110–8119.