

DOI: <https://doi.org/10.17816/DD108982>

Существование машинного интеллекта, цифрового искусства и диагностики: возможно ли оно?

А.В. Власов^{1, 2}¹ Научно-практический клинический центр диагностики и телемедицинских технологий, Москва, Российская Федерация² Научно-исследовательский институт медицины труда имени академика Н. Ф. Измерова, Москва, Российская Федерация

АННОТАЦИЯ

Развитие машинного интеллекта и применение генеративных изображений, созданных с его помощью, является перспективным направлением коммуникационного дизайна и человеко-машинного взаимодействия. Письмо в редакцию представляет собой авторское видение применения генеративных изображений в области диагностики состояний человека.

Использование машинного интеллекта как интерактивного и интеллектуального инструмента диагностики позволит психологу и врачу эффективно дополнить терапевтические процессы контролируемого взаимодействия их участников.

Сейчас уже существуют библиотеки моделей и наборы приложений с text-to-image алгоритмами, которые могут быть задействованы инженерами и дизайнерами в процессе создания объектов современного цифрового искусства, и также могут быть использованы в исследованиях новых парадигм с помощью визуальных коммуникаций, их прикладного применения в экспериментальной диагностике.

Ключевые слова: генеративные изображения; text2image; визуальное восприятие; диагностика.

Как цитировать

Власов А.В. Существование машинного интеллекта, цифрового искусства и диагностики: возможно ли оно? // *Digital Diagnostics*. 2022. Т. 3, № 3. С. 324–330. DOI: <https://doi.org/10.17816/DD108982>

DOI: <https://doi.org/10.17816/DD108982>

Coexistence of machine intelligence, cyber art, and diagnostics: is it possible?

Andrey V. Vlasov^{1, 2}

¹ Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russian Federation

² Izmerov Research Institute of Occupational Health, Moscow, Russian Federation

ABSTRACT

The development of machine intelligence and the application of generative images created using it is a promising area of communication design and human–machine interaction. This letter to the editor represents the author’s vision of the use of generative images for diagnosing human conditions.

The use of machine intelligence as an interactive and intelligent diagnostic tool will allow a psychologist and a physician to effectively complement the therapeutic processes of controlled interactions of their users.

Libraries of models and sets of applications with text-to-image algorithms are already available that can be used by engineers and designers in the process of creating objects of modern digital art. They can also be applied in the investigation of new paradigms using visual communications and their application in experimental diagnostics.

Keywords: generative images; text2image; visual perception; diagnostics.

To cite this article

Vlasov AV. Coexistence of machine intelligence, cyber art, and diagnostics: is it possible? *Digital Diagnostics*. 2022;3(3):324–330.

DOI: <https://doi.org/10.17816/DD108982>

Received: 28.06.2022

Accepted: 25.07.2022

Published: 22.08.2022

DOI: <https://doi.org/10.17816/DD108982>

机器智能、数字艺术和诊断的共存：有可能吗？

Andrey V. Vlasov^{1,2}

¹ Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russian Federation

² Izmerov Research Institute of Occupational Health, Moscow, Russian Federation

简评

机器智能的发展和在其帮助下创建的生成图像的使用是通信设计和人机交互的一个有前途的方向。致编辑的信提出了作者对生成图像（图1）应用于人类状况诊断的设想。

使用机器智能作为交互式 and 智能诊断工具将允许心理学家和医生有效地补充其参与者受控交互的治疗过程。

现在已经有了带有文本生成图像算法的模型库和应用程序集，可供工程师和设计师在创建当代数字艺术对象的过程中使用，也可以用于研究使用视觉的新范式通信，其在实验诊断中的应用。

关键词： 生成图像； 文本生成图像； 视觉感知； 诊断。

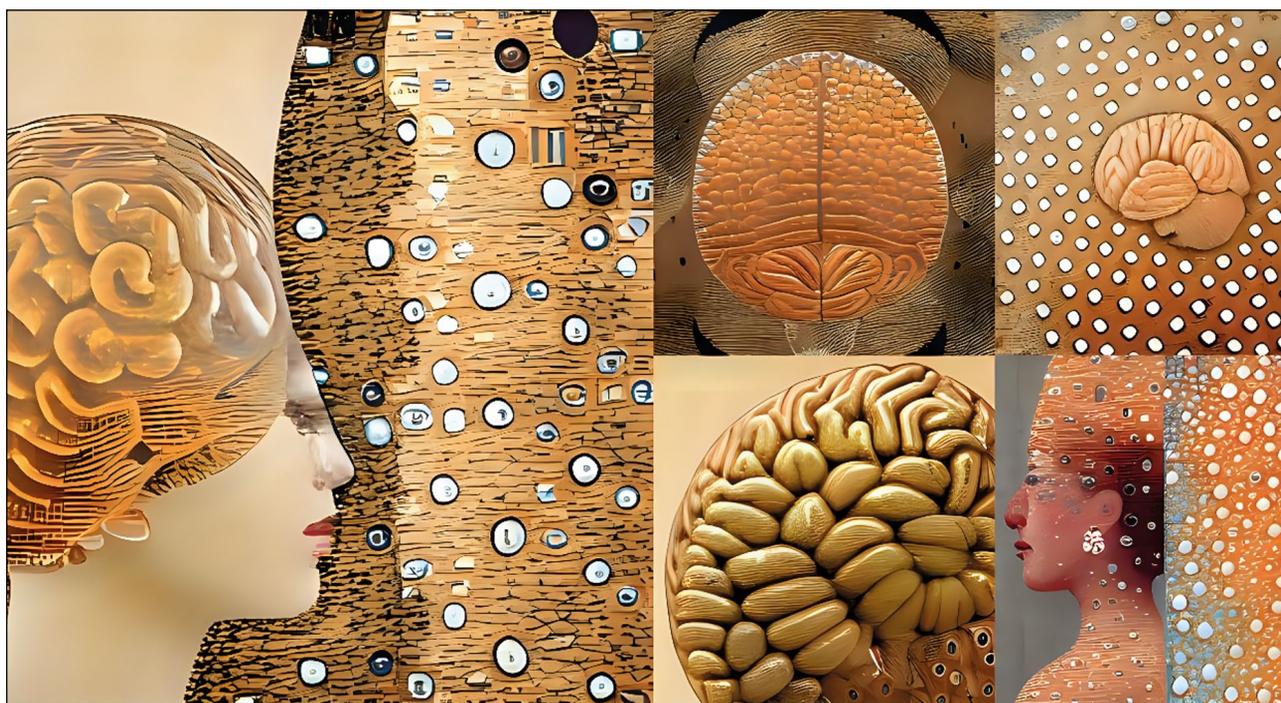
To cite this article

Vlasov AV. 机器智能、数字艺术和诊断的共存：有可能吗？ *Digital Diagnostics*. 2022;3(3):324–330. DOI: <https://doi.org/10.17816/DD108982>

收到：28.06.2022

接受：25.07.2022

发布日期：22.08.2022



ВИЗУАЛЬНОЕ ВОСПРИЯТИЕ ОБРАЗОВ

Машинное обучение (machine learning, ML) широко используется для диагностических целей, решений задач классификации, поиска и визуализации патологий, в том числе одной из самых изучаемых тематик, с позиций публикационной активности, является диагностика болезни Альцгеймера [1, 2]. Наряду с применением алгоритмов ML (в частности, метода опорных векторов) и расширением инструментария врача-диагноста активное развитие получают искусственные нейронные сети и генеративные модели по созданию визуального контента — text2image. Под моделью text2image понимается алгоритм, который позволяет генерировать изображение по текстовому запросу.

Восприятие визуальных образов, таких как художественные изображения, в современной культуре напрямую связано с эмоциональными и когнитивными процессами, личностными особенностями их восприятия и интерпретации индивидуально каждым человеком. В действительности то, как мы воспринимаем, например, абстракцию (рис. 1), может многое рассказать нам о нас самих. M.F. Koich и F. Pessotto в своей работе [3] показали, что искажение эмоционального восприятия изображений связаны с индивидуальными особенностями личности. В исследовании авторов чувство радости при предъявлении определённых изображений коррелировало с общительностью, а чувство страха — со способностью противостоять агрессии и отстаивать личные границы.

Перспективной технологией для транслирования художественного контента является технология виртуальной реальности, где пользователь (пациент) создаёт собственную реальность, «переходную» между внутренним миром и внешней реальностью, которая может быть исследована совместно с психологом или врачом [4]. Благодаря технологии виртуальной реальности в руках исследователей появляются новые инструменты с уникальными возможностями. Так, например, F. Paladines-Jaramillo с колл. [5] адаптировали тест Розенцвейга для диагностики фрустрации, для чего стимульный материал с рисунками различных ситуаций перенесли в виртуальную среду.

Постепенно технологии, в том числе искусственный интеллект общего назначения, смогут стать натуралистической частью терапевтических процессов, в которые будут интегрированы. Однозначно для адаптации и массового внедрения в практику необходимы исследования и разработка специальных терапевтических приложений и систем.

МАШИННЫЙ ИНТЕЛЛЕКТ

Возможности машинного интеллекта расширяются стремительно, не отставая от технологий виртуальной реальности. В течение последнего года мы наблюдаем ошеломительные результаты в создании генеративных объектов цифрового искусства¹, объектов дизайна, фотореалистических картин, живописных изображений

¹ Например, DALL-E 2 OpenAI (режим доступа: <https://openai.com/dall-e-2>); ruDALL-E (Далли) Sber, SberDevices (режим доступа: <https://rudalle.ru>).

с помощью генеративных состязательных нейросетей (generative adversarial network, GAN) и диффузионных моделей (diffusion models, DM), таких как DALL-E-2, Imagen, ruDALL-E, VQGAN, Stable Diffusion, Latent Diffusion, Disco Diffusion и др., работающих по принципу преобразования вводимого текста в изображение.

Следствием совместного взаимодействия человека, разрабатывающего алгоритм и вводящего текстовый запрос, и GAN (или DM) уже является дополнительный творческий эффект [6]. В данном случае вычислительным результатом работы модели text2image становится цифровой объект — 2D-изображение.

Интересен такой факт, что GAN-подобные модели используются для анализа данных нейровизуализации (компьютерная или магнитно-резонансная томография²) [7, 8].

Машинный интеллект в совершенстве владеет текстом: на текущем уровне развития технологий искусственного интеллекта способность предсказывать следующий элемент текста важна для понимания его смысла и создания новых содержательных текстов. Справедливо отметить, что алгоритмы создания визуальных изображений также используют «предсказание следующего пикселя», но в отличие от текстовых моделей (GPT-3 и др.) и генерируемых ими текстовых фраз, при диалоговом взаимодействии синхронизация между людьми происходит на уровне нервно-психических функций [9], например, увеличивается с подключением общего эмоционального поля [10]. Положительные эффекты этой нейронной синхронизации применяются в коммуникационных экспериментах [11, 12].

ВИЗУАЛЬНОЕ ВОСПРИЯТИЕ И ЭМОЦИИ

Разработчики непрерывно стремятся усовершенствовать функциональность и результативность нейронных сетей (приложения DALL-E 2, ruDALL-E, Stable Diffusion, Midjourney и др.), а их появление вдохновляет учёных на изучение визуального восприятия смыслов, заложенных в художественные объекты, с помощью генеративного искусства³ [6, 13]. В связи с этим возникает логичный вопрос, связано ли восприятие объектов цифрового искусства с личностными характеристиками смотрящего? В частности, Р. Achlioptas и соавт. [14] провели исследование эмоций, сопровождающих зрительное восприятие произведений искусства, связанных с ними объяснений собственных эмоций. В этом эксперименте произведения изобразительного искусства использовались в качестве стимульного материала, для того чтобы вызвать сильный эмоциональный отклик. Как подчеркнули авторы [14], аффективный компонент часто недооценивается при разработке систем искусственного интеллекта.

Проведём небольшой эксперимент, ответив на вопрос: «Как вы считаете, какое из двух изображений, представленных на рис. 2, создано нейронной сетью?»

Ответ простой: оба изображения (см. рис. 2) созданы с использованием искусственного интеллекта [15].

Благодаря развитию генеративных моделей text2image представляется реальным быстрое создание тематического ряда уникальных цифровых изображений с помощью нейронной сети. Уже сейчас почти любой исследователь

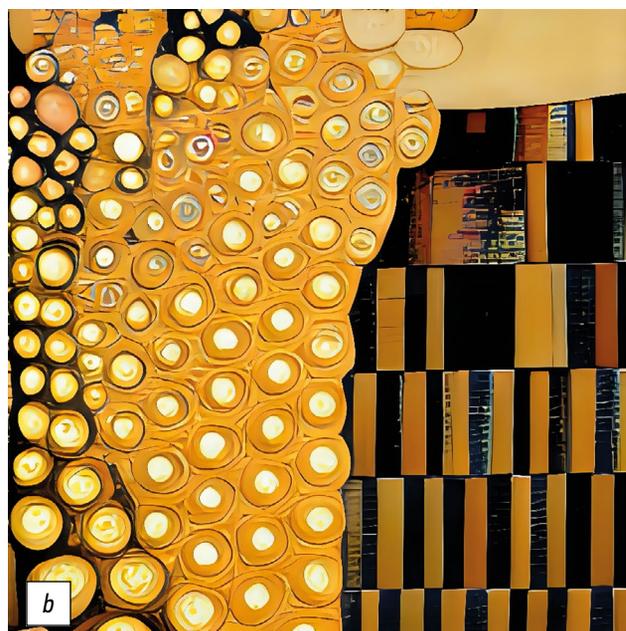


Рис. 2. Изображения (а, б), созданные нейронной сетью.

² Подробнее см. обзоры о роли генеративных состязательных нейросетей в анализе медицинских изображений.

³ Под генеративным искусством понимаются художественные объекты, созданные с помощью информационных технологий, в частности GAN или DM [6, 13].

может воспользоваться подобным инструментарием, сгенерировать новые контекстуальные изображения и спланировать собственный дизайн эксперимента.

Экологически валидно использовать визуальное искусство в качестве стимула для организации исследований. Человек в своих реакциях испытывает множество переживаний, в том числе эмоции и саморефлексию. Данный опыт в высшей степени индивидуален, а реакции смотрящих на один и тот же объект существенно различаются. Подтверждением данных индивидуальных различий служат паттерны нейронной активности различных подсетей головного мозга [16].

ЗАКЛЮЧЕНИЕ

Положительно отвечая на поставленный в письме вопрос (возможно ли использовать машинный интеллект для создания генеративных изображений и применения в экспериментальной диагностике), необходимо акцентировать внимание на то, что развитие исследований на стыке психологии и генеративного искусства, где машинный интеллект создаёт полноценные художественные работы, способствует появлению интеллектуальных систем, поддерживающих эмоциональное человеко-машинное взаимодействие. В свою очередь, подобные системы будут встраиваться в роботов, которые в роли социального партнёра будут помогать человеку адаптивно управлять и регулировать собственные эмоции, а в роли врача-ассистента — организовывать терапевтическую деятельность.

Подобный подход будет реализовываться не только как интерактивный и интеллектуальный инструмент

на рабочем столе психолога и врача, например, для целей экспериментальной диагностики аффективных процессов у пациентов, но как более сложная система⁴, обеспечивающая контролируемое взаимодействие врача-машинного интеллекта и пациента для целей практической медицины.

ДОПОЛНИТЕЛЬНО

Источник финансирования. Автор заявляет об отсутствии внешнего финансирования при проведении поисково-аналитической работы.

Конфликт интересов. Автор декларирует отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Вклад автора. Автор подтверждает соответствие своего авторства международным критериям ICMJE. (Автор внёс существенный вклад в разработку концепции, проведение поисково-аналитической работы и подготовку статьи, прочёл и одобрил финальную версию перед публикацией).

ADDITIONAL INFORMATION

Funding source. This article was not supported by any external sources of funding.

Competing interests. The author declare that he has no competing interests.

Author's contribution. The author made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work.

СПИСОК ЛИТЕРАТУРЫ

1. Tanveer M., Richhariya B., Khan R.U., et al. Machine learning techniques for the diagnosis of alzheimer's disease: a review // *ACM Transactions Multimedia Computing Communications Applications*. 2020. Vol. 16, N 1. P. 35. doi: 10.1145/3344998
2. Sharma S., Mandal P.K. A comprehensive report on machine learning-based early detection of alzheimer's disease using multimodal neuroimaging data // *ACM Computing Surveys*. 2023. Vol. 55, N 2. P. 1–44. doi: 10.1145/3492865
3. Koich M.F., Pessotto F. Projective aspects on cognitive performance: distortions in emotional perception correlate with personality // *Psicologia Reflexão Crítica*. 2016. Vol. 29, N 17. P. 1–8. doi: 10.1186/s41155-016-0036-6
4. Адаскина А.А. Терапевтические возможности цифрового художественного творчества // *Современная зарубежная психология*. 2021. Т. 10, № 4. С. 107–116. doi: 10.17759/jmfp.2021100410
5. Paladines-Jaramillo F., Egas-Reyes V., Ordonez-Camacho D., et al. Using virtual reality to detect, assess, and treat frustration. In: Morales R.G., Fonseca C., Salgado E.R., et al. (eds.) *Information and communication technologies. TICEC 2020*. Vol. 1307. Communications in Computer and Information Science. Springer, Cham, 2020. doi: 10.1007/978-3-030-62833-8_28
6. Cetinic E., She J. Understanding and creating art with ai: review and outlook // *ACM Trans. Multimedia Comput Commun*. 2022. Vol. 18, N 2. P. 1–22. doi: 10.1145/3475799
7. AlAmir M., AlGhamdi M. The role of generative adversarial network in medical image analysis: an in-depth survey // *ACM Computing Surveys*. 2022. doi: 10.1145/3527849
8. Ali H., Biswas R., Ali F., et al. The role of generative adversarial networks in brain MRI: a scoping review // *Insights into Imaging*. 2022. Vol. 13, N 98. P. 1–15. doi: 10.1186/s13244-022-01237-0
9. Lankinen K., Saari J., Hari R., et al. Intersubject consistency of cortical MEG signals during movie viewing // *NeuroImage*. 2014. N 92. P. 217–224. doi: 10.1016/j.neuroimage.2014.02.004
10. Nummenmaa L., Glerean E., Viinikainen M., et al. Emotions promote social interaction by synchronizing brain activity across individuals // *Proceedings Nat Academy Sci*. 2012. Vol. 109, N 24. P. 9599–9604. doi: 10.1073/pnas.120609510

⁴ Медицинское изделие.

11. Tseng P.H., Rajangam S., Lehew G., et al. Interbrain cortical synchronization encodes multiple aspects of social interactions in monkey pairs // *Sci Rep*. 2018. Vol. 8, N 1. P. 4699. doi: 10.1038/s41598-018-22679-x
12. Shanechi M.M. Brain-machine interfaces from motor to mood // *Nat Neurosci*. 2019. Vol. 22, N 10. P. 1554–1564. doi: 10.1038/s41593-019-0488-y
13. Vlasov A. GALA Inspired by Neo Klimt: 2D images processing with implementation for interaction and perception studies (preprint). 2022. doi: 10.13140/RG.2.2.10806.57928

14. Achlioptas P., Ovsjanikov M., Haydarov K., et al. ArtEmis: affective language for visual art // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, October 6, 2021. P. 11569–11579. doi: 10.48550/arXiv.2101.07396
15. Gala Klimt. Digital art collection of pictorial poems. Ridero. 2022. Режим доступа: <https://www.researchgate.net/project/GALA-KLIMT>. Дата обращения: 15.07.2022.
16. Vessel E.A., Starr G.G., Rubin N. The brain on art: intense aesthetic experience activates the default mode network // *Front Hum Neurosci*. 2012. N 6. P. 66. doi: 10.3389/fnhum.2012.00066

REFERENCES

1. Tanveer M, Richhariya B, Khan RU, et al. Machine learning techniques for the diagnosis of alzheimer's disease: a review. *ACM Transactions Multimedia Computing Communications Applications*. 2020;16(1):35. doi: 10.1145/3344998
2. Sharma S, Mandal PK. A comprehensive report on machine learning-based early detection of alzheimer's disease using multimodal neuroimaging data. *ACM Computing Surveys*. 2023;55(2):1–44. doi: 10.1145/3492865
3. Koich MF, Pessotto F. Projective aspects on cognitive performance: distortions in emotional perception correlate with personality. *Psicologia Reflexão Crítica*. 2016;29(17):1–8. doi: 10.1186/s41155-016-0036-6
4. Adaskina AA. Therapeutic possibilities of digital artistic creativity. *Modern Foreign Psychology*. 2021;10(4):107–116. (In Russ). doi: 10.17759/jmfp.2021100410
5. Paladines-Jaramillo F, Egas-Reyes V, Ordonez-Camacho D, et al. Using virtual reality to detect, assess, and treat frustration. In: Morales R.G., Fonseca C., Salgado E.R., et al. (eds.) *Information and communication technologies. TICEC 2020*. Vol. 1307. Communications in Computer and Information Science. Springer, Cham, 2020. doi: 10.1007/978-3-030-62833-8_28
6. Cetinic E, She J. Understanding and creating art with ai: review and outlook. *ACM Trans Multimedia Comput Commun Applications*. 2022;18(2):1–22. doi: 10.1145/3475799
7. AlAmir M, AlGhamdi M. The Role of generative adversarial network in medical image analysis: an in-depth survey. *ACM Computing Surveys*. 2022. doi: 10.1145/3527849
8. Ali H, Biswas R, Ali F, et al. The role of generative adversarial networks in brain MRI: a scoping review. *Insights Into Imaging*. 2022;13(8):1–15. doi: 10.1186/s13244-022-01237-0
9. Lankinen K, Saari J, R Hari, et al. 2014. Intersubject consistency of cortical MEG signals during movie viewing. *NeuroImage*. 2014;92:217–224. doi: 10.1016/j.neuroimage.2014.02.004
10. Nummenmaa L, Glerean E, Viinikainen M, et al. Emotions promote social interaction by synchronizing brain activity across individuals. *Proceedings Nat Academy Sci*. 2012;109(24):9599–9604. doi: 10.1073/pnas.120609510
11. Tseng PH, Rajangam S, Lehew G, et al. Interbrain cortical synchronization encodes multiple aspects of social interactions in monkey pairs. *Sci Rep*. 2018;8(1):4699. doi: 10.1038/s41598-018-22679-x
12. Shanechi MM. Brain-machine interfaces from motor to mood. *Nat Neurosci*. 2019;22(10):1554–1564. doi: 10.1038/s41593-019-0488-y
13. Vlasov A. GALA Inspired by Neo Klimt: 2D images processing with implementation for interaction and perception studies (preprint). 2022. doi: 10.13140/RG.2.2.10806.57928
14. Achlioptas P, Ovsjanikov M, Haydarov K, et al. ArtEmis: affective language for visual art. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, October 6, 2021:11569–11579. doi: 10.48550/arXiv.2101.07396
15. Gala Klimt. Digital art collection of pictorial poems. Ridero. 2022. Available from: <https://www.researchgate.net/project/GALA-KLIMT>. Accessed: 15.08.2022.
16. Vessel EA, Starr GG, Rubin N. The brain on art: intense aesthetic experience activates the default mode network. *Front Hum Neurosci*. 2012;6:66. doi: 10.3389/fnhum.2012.00066

ОБ АВТОРЕ

* Власов Андрей Васильевич;

адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1;
ORCID: <https://orcid.org/0000-0001-9227-1892>;
eLibrary SPIN: 3378-8650; e-mail: a.vlasov@npcmr.ru

AUTHOR'S INFO

* Andrey V. Vlasov;

address: 24-1 Petrovka street, 127051 Moscow, Russia;
ORCID: <https://orcid.org/0000-0001-9227-1892>;
eLibrary SPIN: 3378-8650; e-mail: a.vlasov@npcmr.ru