

ARTIFICIAL INTELLIGENCE AND GRAPHIC DESIGN

Boyan Blazhev

South-West University "Neofit Rilski", Bulgaria

Abstract: *The connection between artificial intelligence (AI) and graphic design is close and increasingly significant in the world of technology and creativity. Graphic design involves the use of various elements, such as images, text, and color, to create visual communication. This process is often slow and laborious. Artificial intelligence can be useful to automate these processes and improve the efficiency of design work. Artificial intelligence in the modern era is a powerful weapon for creating new graphic elements, improving the quality of designs and analyzing user behavior. In the future, it is expected that these capabilities will develop and new ways to use artificial intelligence in graphic design will appear, but one should not forget the dangers that this spanless new tool hides.*

Keywords: *Artificial Intelligence, Graphic Design, Digital, Art, Collaboration*

ИЗКУСТВЕН ИНТЕЛЕКТ И ГРАФИЧЕН ДИЗАЙН

Боян Блажев

Югозападен университет „Неофит Рилски“, България

Резюме: *Връзката между изкуствения интелект (ИИ) и графичния дизайн е тясна и все по-значима в света на технологиите и креативността. Графичният дизайн включва използването на различни елементи, като изображения, текстове и цветове, за да се създаде визуална комуникация. Този процес често пъти е бавен и трудоемък. Изкуственият интелект може да бъде полезен за автоматизиране на тези процеси и да подобри ефективността на дизайнерската работа. В съвременната епоха е мощно оръжие, за създаване на нови графични елементи, подобряване на качеството на дизайните и анализ на потребителското поведение. В бъдеще, се очаква тези възможности да се развият и да се появят нови начини за използване на изкуствения интелект в графичния дизайн, но не трябва да се забравят опасностите, които крие този необятен нов инструмент.*

Ключови думи: *Изкуствен Интелект, Графичен Дизайн, Цифров, Изкуство, Сътрудничество*

Историята на развитие на изкуствения интелект (**ИИ**; на английски; artificial intelligence, **AI**) започва в далечната 1956 г. Терминът е въведен от Джон Маккарти, Марвин Мински, Натаниел Рочестър и Клод Шенън по време на първата международна конференция по проблемите на изкуствения интелект известна като Дартмутска конференция. През годините, изкуственият интелект се развива с изключително бързи темпове и в последните години има много значителен напредък в тази област. Нови техники, модели и алгоритми се разработват и използват в широка гама от приложения, като обработка на естествени езици, разпознаване и генериране на изображения и звук, роботика и много други. Тези постижения имат

голям потенциал да променят начина, по който работим, живеем и комуникираме помежду си.

Един от ранните успехи в ИИ е експертната система Dendral, която е разработена през 1960-те и 1970-те години за анализ на молекули. В началото на 1980-те години, машинното обучение става основен подход за ИИ и тогава се появяват невронните мрежи и генетичните алгоритми. Благодарение на големите масиви от данни, които се събират, обработват и анализират, от началото на 21-ви век има значителен напредък в развитието на ИИ. Машинното обучение става все по-мощен инструмент и се използва в много сфери, като изкуство, здравеопазване, автомобилната индустрия, финансовия сектор, роботиката и други.

Изкуственият интелект е област, която се занимава с разработването на системи и програми, които могат да изпълняват задачи, които обикновено изискват човешки интелект. ИИ включва машинно обучение (machine learning), експертни системи (expert systems), невронни мрежи (neural networks), генетични алгоритми (genetic algorithms) и други.

Изкуственият интелект е особено силно популяризиран след пускането на GPT-3 на 30 ноември 2022 г. Разработен от OpenAI¹ базиран на архитектура разработена през 2020 година. GPT-3 е голям езиков модел, който използва машинно самообучение за да генерира текст с висока степен на автономност и сходство с човешкия език. Нейросетев чат-бот² ChatGPT е изключително разпространен инструмент за комуникация и автоматизация на различни видове услуги, благодарение на своята способност да генерира автономни, но все още естествено звучащи отговори.

През последните години изкуственият интелект навлиза все по-осезаемо в ежедневието на човека. Постепенно той започва да изпълнява рутинни функции и незабелязано става неразделна част от съвременния наратив. Животът рязко и фундаментално се променя. В пост-постмодерната епоха появата на ИИ е епохално събитие съизмеримо с навлизането на персоналните компютрите и Интернет, което ще бъде обект на множество изследвания през годините. В областта на развитието на визуалната комуникация ИИ има голямо влияние. В по-тесен смисъл графичният дизайн е област, която е сериозно засегната от развитието на изкуствения интелект. Новите технологии и методи позволяват на дизайнерите да автоматизират и оптимизират много от задачите, свързани с проектирането на графични елементи и съдържание.

¹ OpenAI е американска технологична компания, която е основана през 2015 година с цел да разработи изкуствен интелект с широко приложение. Целта на компанията е да разработи технология, която да бъде полезна за хората и да не нанася вреда на обществото.

² Изкуствено невронна мрежа

Могат да се разграничат три етапи в историята на развитие на ИИ за генериране на визуално изкуство

- **Първи модели и невронни мрежи 1960-2000 г.**
- **Ерата на изкуствения интелект, невронните мрежи и системите за deep machine learning - 2000-2020 г.**
- **Отвъд „Закона на Мур“ 2020 – 2023 г.**

Първи модели и невронни мрежи 1960-2000 г.

В периода 1957-1962 г. Франк Розенблат (Frank Rosenblatt) разработва вид изкуствена невронна мрежа - Персептрон, която може да бъде обучена да разпознава модели и изображения, която за съжаления е прекалено ограничена. Въпреки това Персептрон е една от основите за развитието на изкуствените невронни мрежи. През 1950-те и 1960-те години, компютърната графика е използвана за създаване на прости модели и форми. По късно през 1960-те години на миналия век компютърните специалисти започват да изследват концепцията за машинно обучение (machine learning), при генериране на изображения чрез ИИ. Немският математик и учен Фридер Наке през 1967 г. представя поредицата "Matrix Multiplications". В нея се разглеждат различни матрични алгоритми за компютърната графика. Наке също така представя много примери за приложения на матричните алгоритми в компютърната графика, включително рисуване на криви и повърхности. Той изследва как матричните алгоритми могат да се използват за анимация и представя някои алгоритми за симулация на физически явления като светлина и сенки. "Matrix Multiplications" има важна роля в историята на компютърната графика и изкуствената интелигентност. Тя разглежда как матричните алгоритми могат да се използват за създаване на изображения, включително примитивни геометрични форми и генериране на цветове. По късно през 70-те и 80-те ИИ генераторите се развиват и могат да правят комплексни фигури, фрактали и някои от методите се прилагат в софтуери за дигитална обработка като Adobe Photoshop, Adobe Illustrator и др. Паралелно се развиват и системите за компютърно подпомаганото проектиране (CAD). Това е процес на използване на компютърни системи и софтуер за създаване, моделиране, анализ и оптимизация на дизайни и технически чертежи. Благодарение на CAD системите архитектите и дизайнерите работят все по-ефективно и ефикасно върху проекти, като се намалят времето и разходите за разработка и производство. Чрез CAD системите потребителите могат да създават по-сложни и реалистични изображения, тримерни модели, да използват различни материали и размери, да променят форми и пропорции, да правят визуализации и анимации, да използват автоматизирани инструменти за проектиране, да генерират технически чертежи и други.

През 1973 г. е разработена компютърна програма за рисуване AARON от американския художник Харолд Коен. Програмата използва невронни мрежи, за да

генерира изображения, като има вградена евристика³, която позволява на програмата да прави творчески решения за изображенията, които генерира. AARON е използвана за създаването на множество изображения, включително портрети, пейзажи и абстрактни композиции (фигура 1). Програмата използва многослойни невронни мрежи, за да анализира и реплицира стилове на други художници, като например Пикасо и Ван Гог. Програмата AARON е един от първите примери за използване на изкуствен интелект в изкуството (Garcia, 2016). Тя продължава да бъде активно развивана и се използва все още в днешно време.



Фигура 1. Абстрактна композиция-Харолд Коен

По късно се появява и конволюционната невронна мрежа (Convolutional Neural Network - CNN). Това е вид невронна мрежа, която е специално проектирана да обработва данни със силно корелиращи се структури, като изображения и видео. Една от първите успешни реализации на CNN е моделът Neocognitron⁴, разработен от японския учен Кунихико Фукушима през 80-те години на миналия век (Fukushima, 1980). Този модел представлява значителен скок в областта на компютърната наука, но в сравнение със съвременните CNN има ограничени възможности и изисква прекалено много изчислителна мощ.

³ В контекста на изкуствен интелект, евристиката може да се използва за усъвършенстване на системи за генериране на изображения и други творчески процеси. Това включва програми, които използват машинно обучение и невронни мрежи, за да генерират нови изображения, музика и други творби. Евристиката може да бъде използвана, за да се гарантира, че тези системи генерират качествени и оригинални творения.

⁴ Neocognitron има две главни компоненти: първоначален слой за извличане на характеристики и каскадна мрежа от слоеве, които използват свръх позиционирана ретина за постигане на инвариантност по отношение на трансляция. Първият слой използва свръх позициониран филтър, за да извлича ниско ниво на характеристики от входните изображения. След това каскадната мрежа от слоеве извлича все по-абстрактни характеристики от изображението, като последните слоеве се използват за класификация на обектите в изображението.

С развитието на компютърните системи и технологиите в последното десетилетие на 20-ти век, става възможно и разработването на първите невронни модели на базата на машинното самообучение (machine learning). Появява се LeNet-5 това е първата успешна конволюционна невронна мрежа, разработена е от Ян Лекун, Леон Боту, Йошуа Бенгио и Патрик Хафнер през 1998 година. Това е 5-слойна мрежа, която е била използвана за разпознаване на ръкописни цифри. LeNet-5 използва конволюционни слоеве за извличане на характеристики от изображенията и слоеве за приемане на решения (decision-making), като последният слой е изпълнен със софтвакс функция⁵ за класификация на изображенията в няколко класа. Тази мрежа се смята за революционна в областта на разпознаването на изображенията и довежда до значителен напредък в областта на компютърното зрение, като помага за въвеждането на deep machine learning⁶. LeNet-5 е в основата на много от последващите модели на конволюционни невронни мрежи, които са създадени за различни задачи, като например обработка на изображения и видео, разпознаване на реч, превод на езици и други.

Ерата на изкуствения интелект, невронните мрежи и системите за deep machine learning - 2000-2020 г.

Първото десетилетие на 21-ви век се характеризира като години на глобализация и бурно развитие на изкуствения интелект. През този период невронните мрежи търпят огромен напредък. Появяват се различни разработки и софтуери свързани с генерирането и обработката на изображения. Появява се ImageNet, това е една от най-големите бази данни от изображения, създадена през 2009 г. от професор Фей Фей Ли и неговия екип (Deng et al., 2009). Базата данни включва над 14 милиона изображения, разделени в над 20 000 категории (Krizhevsky et al., 2017). ImageNet е станала известна в света на изкуствения интелект през 2012 г., когато екип от изследователи, водени от професор Андрю Йанг и професор Фей Фей Ли, използват ImageNet за обучение на невронна мрежа с име AlexNet. Тази мрежа се смята за преломен момент в историята на компютърното зрение, тъй като постигна значителен напредък в точността на класификацията на изображенията в ImageNet. Една от ключовите характеристики на мрежата е, че използва техники за увеличаване на данните като обръща и изрязва входните изображения, за да създаде допълнителни изображения за обучение. Това помага за намаляване на пренастройването и подобрява ефективността на модела. AlexNet се използва като основа за развитието на следващите модели за deep machine learning.

⁵ Софтвакс функцията е математическа функция, която се използва често в машинното обучение, за да се нормализират резултатите от изчисления, които са изведени чрез някакъв алгоритъм или модел. Тя приема вектор от произволни стойности като вход и връща нов вектор със същия брой стойности, като всички стойности в новия вектор са неотрицателни и сумират до 1. Това означава, че със софтвакс функцията можем да превърнем произволен вектор в вероятностен разпределение. Софтвакс функцията се използва често като последен слой в много класификационни модели, където изходите на модела трябва да бъдат интерпретирани като вероятности за всяка класа.

⁶ Deep machine learning - дълбоко машинно обучение

Иън Гудфелоу на 10.06.2014 г. предлага невронна мрежа, която се използва за неконтролиран *deep machine learning* наречена General Adversarial Network (GAN). Тези мрежи са едни от най-важните постижения в областта на машинното обучение и изкуствения интелект. Те работят паралелно при изпълнението на определена задача чрез създаване на две невронни мрежи. Едната е генератор, който произвежда изображения, а другата е дискриминатор⁷, който определя дали изображенията, създадени от генератора, са реалистични или не. Тези мрежи правят революция в областта на генеративното моделиране. Благодарение на тях се създават изключително реалистични изображения, които са трудно различими от изработените чрез конвенционалните методи и технологии. Предишните модели са много по-ограничени във възможностите си за изграждане на разнообразни висококачествени данни. Приложението на GAN мрежата е в различни области, компютърна графика, медицинска диагностика, синтез на текст и музика, включително генериране на изображения, редактирането им и синтез на данни. То подобрява значително производителността на други модели за *deep machine learning*, като модели за класификация на изображения и сегментиране. GAN мрежите бързо се превърнаха в един от най-мощните инструменти за генериране на изключително реалистични изображения. В напредъка на дълбокото машинно самообучение се наблюдава експоненциален ръст, като на всеки няколко месеца се разработват все по-бързи модели и реалистични (Artbeeder, n. d.).

Разнообразието и приложението на моделите е голямо, но от гледна точка на последователност и оказване на влияние най-значимите са:

GoogleNet е невронна мрежа, която е известна още като Inception v1 и е разработена от изследователи от Google през 2014 година. Тя е известна със своята дълбочина и сравнително малък брой параметри, което я прави ефективна за използване на мобилни устройства. GoogleNet включва множество модули на Inception, които използват различни размери на ядрата на филтрите и позволяват на мрежата да избира оптималния филтър в зависимост от характеристиките на изображението. Това води до по-добра ефективност на мрежата и намаляване на изискванията за памет. GoogleNet е първата мрежа, която постига точност в класификацията, по-висока от 70% в ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Тази невронна мрежа открива широко приложение в различни области, включително медицинските изображения и автономните превозни средства.

Друг проект на Google е DeepDream създаден през 2015 г. Този софтуер за генериране на изображения с ИИ невронни мрежи, по-специално с *deep learning CNN* е техника за обработка на изображения, която използва невронни мрежи за генериране на хиперреалистични и психеделични изображения. Техниката е базирана на обратно подаване на сигнали в мрежата, като се максимизира активацията на определени неврони. Това довежда до усилване на някои детайли в

⁷ Дискриминаторът е вид невронна мрежа, която е обучена да прави разлика между реални и изкуствено генерирани данни

изображението и създаване на нови, което води до необичайни и забележителни ефекти. За да се използва DeepDream, се избира изображение и се подава през невронната мрежа. След това, мрежата извършва редица итерации, при които се променят пикселите на изображението, за да се усилят някои характеристики на изображението, които мрежата е научила да разпознава. Резултатът е изображение, което изглежда като комбинация от оригиналното изображение и тези характеристики, които мрежата е усилила.

Следващата година се появява техниката Styletransfer. Тя се използва в компютърното зрение и дълбокото обучение, която включва прилагането на стила на едно изображение към съдържанието на друго (Dumoulin et al., 2016). Този процес включва извличането на характеристики за съдържанието и стила на двете изображения и тяхното комбиниране по начин, който дава ново изображение със съдържанието на оригиналната картинка и стила на изображението-образец. Първият генератор, който използва техниката на style transfer, се нарича "neural style transfer" и е представен в статията на Gatys et al. от 2015 година, "A Neural Algorithm of Artistic Style". Техниката се развива и днес се използва от множество други генератори за създаване на изображения. Някои от тези генератори включват Fast Neural Style Transfer, Adaptive Style Transfer, Arbitrary Style Transfer и други.

Процесът на styletransfer обикновено използва конволюционни невронни мрежи (CNN), обучени на големи набори от изображения. Характеристиките за съдържание и стил на изображенията се извличат от различни слоеве на CNN, като например средните представяния на характеристики в слоевете, които изобразяват високо ниво на абстракция. След извличането на характеристиките, процесът на styletransfer използва алгоритъм, който оптимизира новото изображение, като променя пикселите му, така че да съчетава съдържанието на едната картина и стила на другата. Този процес може да бъде изпълнен с различни алгоритми, като например gradient descent, което води до оптимално изображение, което е комбинация от съдържанието и стила на двете изходни изображения. Styletransfer има множество приложения в изображенията, например в рекламата, дизайна и други. Техниката може да бъде използвана за създаване на уникални и впечатляващи рекламни банери, създаване на нови и креативни изображения в дизайнерските проекти, създаване на нови картини в изобразителното изкуство и т.н.

Емблематичен пример е проектът "The Next Rembrandt" на ING Bank и J. Walter Thompson Amsterdam, който използва ИИ, за да създаде ново портретно изображение, базирано на стиловете и техниките на Рембранд. Проектът е разработен от сътрудници на различни институции, включително Техническият университет в Делфт, Музея на науките в Амстердам и компанията за машинно обучение J. Walter Thompson Amsterdam. Използвайки анализ на данни и машинно обучение, проектът "The Next Rembrandt" създава изображение, което наподобява стила на Рембранд. Използвани са данни от 346 портретни картини на Рембранд,

които са сканирани и анализирани от ИИ алгоритми. След анализа, ИИ моделът генерира нова портретна картина, която прилича на тези на Рембранд (фигура 2).



Фигура 2. Фотография на процеса на принтиране на портрета създаден от ИИ проектът “The Next Rembrandt”

Проектът “The Next Rembrandt” има голямо значение за художниците и дизайнерите, тъй като демонстрира възможностите на изкуствения интелект да генерира изображения, които са не само реалистични, но и могат да бъдат уникални и креативни. Това може да предизвика нови начини за изкуственото създаване на изображения и да промени начина, по който тези професионалисти работят. Освен това, проектът може да бъде полезен за рекламната и маркетинговата индустрия, като дава възможност за създаване на персонализирани изображения и кампании. В крайна сметка, “The Next Rembrandt” открива нови хоризонти за технологиите на изкуствения интелект и техните възможности за създаване на изкуство и дизайн.

Един от главните ефекти на проекта “The Next Rembrandt” е предоставянето на нова инструментална възможност за художници и дизайнери. Този проект демонстрира възможностите на машинното обучение и изкуствен интелект в областта на изобразителното изкуство и дизайна. Той дава възможност на творци да използват ИИ инструменти, за да създават нови видове изкуство и да откриват нови техники. От една страна, този проект може да предизвика дискусии относно автентичността на творбите, създадени от машини, и дали те могат да бъдат считани за истински изкуствени произведения. От друга страна, този проект може да помогне на творци да преодолеят технически ограничения и да създадат неща, които не биха били възможни с традиционните методи на рисуване и дизайн. Като цяло, “The Next Rembrandt” демонстрира потенциала на машинното обучение и изкуствения интелект в изобразителното изкуство и дизайна и открива нови възможности за творческия процес.

Отзивите свързани с авторските права на проекти като „The Next Rembrandt“ и други подобни проекти, които генерират произведения на изкуството, са разнообразни. Някои хора смятат, че тези системи нарушават авторските права на художника, чиито творби са използвани като основа за обучението на алгоритъма, докато други смятат, че това е просто нов начин за изразяване на творческата идея. Някои също така смятат, че тези системи могат да бъдат използвани като инструмент за създаване на нови, уникални произведения на изкуството, като същевременно запазват идеята и стила на оригиналния автор. Все още не съществува ясно законодателство относно авторските права върху произведенията, генерирани от изкуствени интелигентни системи, и това е обект на активни дискусии и дебати.

Резултатът при прилагането на styletransfer е толкова добър, че отключва голям дебат сред артистите по отношение на авторските права и дали така създаденото изображение е оригинал или е копие.

До този момент повечето ИИ генератори на изображения и алгоритми работят на принципа изображение към изображение (image-to-image) и изображение към текст (image-to-text) (Parmar et al., 2023), (Zhu et al., 2017). С други думи алгоритмите превръщат изображенията в текстови и математически анализи, тогава се ражда идеята процеса да се обърне, като се създаде изцяло ново изображение. През 2015 г. е публикувана статията на Елман Мансимов, Емилио Парисото, Джими Лей Ба и Руслан Салахутдинов, в която те описват първите си опити с модел за генериране на изображения от текст, наречен “text-to-image” (Mansimov et al., 2016). Този модел използва невронна мрежа, която е обучена да генерира изображения, които са съответствие на даден текст. Екипът прилага условен генеративен модел (Conditional Generative Model), който включва две невронни мрежи: генератор и дискриминатор. Генераторът е обучен да създава изображения, които са съответствие на даден текст, докато дискриминаторът е обучен да разпознава, дали изображението е реално или генерирано. За да обучат модела, колективът извлича набор от данни, наречен COCO (Common Objects in Context), който включва над 330,000 изображения с над 2,5 милиона анотации. За всеки текст в набора от данни, генераторът създава съответстващо изображение, като дискриминаторът оценява, дали това изображение е реално или генерирано. Резултатите от експеримента показват, че моделът може да генерира реалистични изображения, които са съответствие на даден текст. Например, когато на моделът е бил даден текст да „генерира червен училищен автобус на паркинг“, той е генерирал изображение, което е много близко до това, което се описва в текста, и въпреки, че качеството на изображенията е с ниска резолюция, резултатът е вълнуващ. Този експеримент е отличен пример за използването на генеративни модели за създаване на нови изображения. Такива модели имат много приложения в различни области, като например в създаването на герой на игри, графичен дизайн и медицинските изображения (Absolutely Ai, n. d.).

През 2016 г. Isola et al. разработва модела Pix2Pix Той използва дълбока конволюционна генеративна мрежа (GAN), за да превърне изображения от един домейн в изображения на друг домейн. За разлика от други модели за генериране на изображения, като например техниката на style transfer, която използва едно изображение като "стил" за преобразуване на друго изображение, Pix2Pix работи с две изображения - едно, което е входното изображение, и друго, което е целевото изображение. Той може да се използва за решаване на различни задачи, като например за превръщане на черно-бели изображения в цветни, за превръщане на графични изображения в реалистични фотографии, за превръщане на линии в рисунки и други. На следващата година е пуснат алгоритъма алгоритъмът CycleGAN, той може да променя вида, стила, цвета или дори обектите в едно изображение, както и във видео. CycleGAN използва две условни генеративни невронни мрежи (GAN) за превръщане на изображения от един домейн в друг, като при това се запазват важните характеристики на изображенията. Този модел се отличава от други алгоритми за трансфер на стил, защото не изисква съпоставяне на изображения от двата домейна в даден момент. Вместо това, той използва две противоположни GAN мрежи, които работят заедно, за да генерират изображения, които са едновременно автентични в двата домейна. Това означава, че може да бъде постигната двупосочна трансформация между двата домейна. Алгоритъмът на моделът е базиран на принципа на циклична консистентност, която гарантира, че след превръщането на изображенията от един домейн в друг и обратно, резултатните изображения са максимално близки до оригиналните изображения. Този принцип помага за избягване на проблема с информационната загуба и запазва важните детайли на изображенията. CycleGAN има множество приложения, като например превръщане на черно-бели изображения в цветни, превръщане на конските изображения в зебри, превръщане на стиловете на картини и много други.

BigGAN е генеративен модел, базиран на изкуствен интелект, който е създаден от OpenAI. Той е един от най-мощните генеративни модели до момента и е способен да генерира изображения с висока резолюция и качество. BigGAN е разработен като подобрение на предишния генеративен модел GAN, и за разлика от тях, BigGAN има способността да генерира изображения с висока разделителна способност – 512x512 пиксела. В процеса на обучение на BigGAN, той използва много големи данни и невронни мрежи с дълбочина до 512 слоя, което позволява на модела да създава изображения с висока детайлност и качество. Освен това, моделът е обучен с техника на условното генериране, която му позволява да генерира изображения в специфичен стил или в избран контекст. BigGAN се използва за създаване на изображения за видео игри, филми и други. В последно време, той е бил използван за генериране на изображения, които могат да бъдат използвани за обучение на други генеративни модели и за подобряване на техните резултати.

По късно NVIDIA показват генеративен модел StyleGAN. Той предлага редица предимства в сравнение с други генеративни модели, като може да генерира изображения с висока резолюция и качество. StyleGAN предоставя възможност на

потребителя в управлението на стила на генерираните изображения, може да се избират различни аспекти на изображението, като цветовете, текстури, форми и други. Също така, потребителят може да контролира процеса на генериране на изображението, като например променя детайлите, формите или текстурите, докато генерацията е в процес. Този модел е обучен с голям брой изображения, което му позволява да генерира изображения в различни стилове и контексти. Освен това, този модел може да се изучи на големи набори от данни. Позволява на потребителите да прехвърлят стилове между изображения, като например прилагат текстури или цветовете от едно изображение към друго. Използване на графични процесори за ускоряване на работата: StyleGAN може да се изпълнява на графични процесори (GPU), което позволява на потребителите да генерират изображения по-бързо и ефективно.

Приложението на модела е в различни сфери на визуалното изкуството, дизайна, рекламата, изображения за видео игри, филми и други. Тези предимства правят StyleGAN много полезен и гъвкав инструмент за създаване на реалистични и креативни изображения. Моделите също са използвани за създаване на изключително реалистични deep fake⁸ видеоклипове, които поражда опасения относно потенциалната им злоупотреба.

GauGAN е генеративен модел, който използва техниката на Conditional Generative Adversarial Networks (GANs), за да генерира реалистични изображения на ландшафти чрез текстово описание. Този модел е създаден от NVIDIA през 2019 година и има много приложения в областта на визуалните изкуства, дизайна, архитектурата и други. Името му е вдъхновено от френския постимпресионист Пол Гоген. Алгоритъмът се използва в различни области в архитектурата и интериорния дизайн, игрите и филмовата продукция. Помага на дизайнерите и художниците да визуализират своите идеи по-бързо, като им позволява да генерират фотореалистични изображения от прости скици или груби чертежи. Използва се и за образователни цели, тъй като „преподава“ на ученици за перспектива, осветление и теория на цветовете.

NVIDIA Canvas е софтуер, който използва технологиите на изкуствен интелект и машинно обучение, за да позволи на потребителите да създават изображения и илюстрации чрез рисуване с помощта на графичния таблет или мишката (NVIDIA Canvas, n. d.). Този софтуер е разработен от NVIDIA и е част от тяхната серия от продукти, свързани с изкуствен интелект. Потребителите могат да създават изображения във високо качество и реалистичен стил, като могат да избират от множество различни стилове, като например маслени бои, въглен, молив и други. Софтуерът използва технологията на GAN (Generative Adversarial Networks), за да

⁸ „Deepfake“ са реалистични видеоклипове, аудио записи или изображения, които са създадени или модифицирани с помощта на алгоритми AI. Тази технология може да създава изключително убедително съдържание, което изглежда автентично, но всъщност е изфабрикувано или манипулирано. Известен пример за „deepfake“ е експеримент направен през 2018 г., където Доналд Тръмп и други известни личности, говорят за интимния си живот. Видеото е фалшиво и накрая системата се изключва за да се покаже актьора.

генерира изображения, които отговарят на стиловете, които потребителите избират. Основното предимство на NVIDIA Canvas е, че с помощта на изкуствения интелект и машинното обучение, той може да помогне на художниците и дизайнерите да създават изображения по-бързо и по-ефективно, като същевременно предоставя възможност за творчество и експериментиране с различни стилове и техники на рисуване. Потребителите могат да използват този софтуер за създаване на илюстрации, рекламни материали, компютърни игри, филми и други.

Приложението на ИИ в генерирането на изображения, които могат да се класифицират като изкуство бележи своята повратна точка през 2018 г. ,когато портретът „Edmond de Belamy, from La Famille de Belamy“ стана първата в света картина генерирана от изкуствен интелект, продадено на търг на Christie's. Това произведение на изкуството е резултат на колаборацията между френската група Obvious и технологичната компания Art AI, която използва технологията на генеративните мрежи (GANs) за създаване на изображения.

Портретът на Едмонд де Белами е създаден чрез изкуствения интелект, който е бил обучен с помощта на набор от стари портрети. Процесът на генериране на изображението е станал възможен благодарение на технологията на GANs, която използва две мрежи - генеративна и дискриминативна, за да генерира нови изображения, които приличат на обучаващия набор. "La Famille de Belamy" е серия от 11 портрета, създадени чрез технологията на GANs. Тези произведения на изкуството представляват едни от първите успешни опити за използване на изкуствения интелект във визуалните изкуства и се считат за важен момент в историята на изкуството. Друг известен художник, който работи с GAN системите и създава AI изкуство, което продава, е Марио Клингеман. Неговите портрети са излагани по музеите в цял свят.

StyleGAN2 е подобрена версия на алгоритъма за генериране на изображения StyleGAN. Той е първо публикуван през 2019 година от НВИДИА, като е основно насочен към подобряване на качеството на генерираните изображения и увеличаване на разнообразието им. Сравнено с оригиналния StyleGAN, StyleGAN2 използва нова архитектура, наречена "Skip Connections", която позволява по-лесно споделяне на информация между различните слоеве на мрежата. Това помага за подобряване на качеството на изображенията и за постигане на по-добра реализъм. Освен това, StyleGAN2 включва нови техники за регулиране на сложността на мрежата и насърчаване на разнообразието на генерираните изображения. Например, въведен метод на "пресичане на каналите" (channel-wise crossing), който позволява на модела да контролира нивото на детайлност на изображенията в различните области на изображението. В резултат на тези подобрения, StyleGAN2 създава изображения с по-висока резолюция, по-голям детайл и по-голямо разнообразие в сравнение с оригиналния StyleGAN. Той е бил използван в множество приложения, включително в създаването на реалистични лица, карти, мебели и други обекти.

Отвъд „Закона на Мур“ 2020 г.

След 2020 г. се наблюдава огромен скок на технологично ниво в развитието на невронните мрежи. Събитията засягащи ИИ и изкуството следват едно след друго, компаниите се надпреварват да изкарват все по-добър софтуер, който има за цел да генерира по-бързо и по-качествени изображения. Според закона на Мур „броят на транзисторите върху чиповете се удвоява на всеки две години, което означава, че те стават по-мощни и могат да обработват по-големи данни по-бързо и ефективно. Това от своя страна позволява разработването на по-сложни модели за машинно обучение и изкуствен интелект, които могат да обработват огромни количества на данни и да създават по-добри резултати. На базата на което компютърни специалисти предвиждат до 2075 г, а по-смелите изказват мнение за 2045 г. да се случи компютърната сингулярност – зависимост, когато ИИ ще е „по-умен“ от всички хора на Земята.

В днешно време всеки сериозен уважаващ себе си технологичен бранд се включва в надпреварата за разработване на все по-добри ИИ системи. Една от водещите компании в това отношение е OpenAI. През 2021 г. тя пуска първият си модел text-to-image⁹ DALL-E (DALL·E, n. d.). Това е генеративен модел, който може да генерира изображения от текстово описание. Моделът е кръстен на художника Салвадор Дали и името на робота WALL-E, който е герой в анимационен филм. Алгоритъмът на модел може да генерира сложни изображения, като например животни, предмети, хора и дори сцени в резолюция от 512x512 пиксела. За да изгради тези изображения, моделът използва текстово описание на изображението, което е дадено от потребителя. След това DALL-E използва генеративна мрежа, за да създаде изображението, което отговаря на описанието. На 7 април 2022 г. OpenAI пускат следващата версия на генератор за изображения – Dall-E 2. Новият алгоритъм може да създава оригинални и реалистични изображения с 4 пъти по-голяма резолюция от предишната версия, редакции на съществуващи изображения добавя и премахва елементи, надписи, като взема предвид сенки, отражения и текстури. DALL·E 2 е нова голяма стъпка в развитието, защото позволява създаването на вариации, вдъхновени от оригинални изображения.

OpenAI разработват и публикуват през 2020 г. системата за генериране на изображения Noise-conditional Score Network (NCSN). Тя използва авто енкодери, за да създаде реалистични изображения, като при това се използват случайни входове за да се генерират множество различни изображения. NCSN добавя допълнителна функционалност, като също така включва и система за оценка на качеството на генерираните изображения. Така нареченият (оценъчен механизъм) - т.е. оценява дали дадено изображение е "добро" или "лошо". NCSN има редица предимства в сравнение с други системи за генериране на изображения. Той е в състояние да създаде много качествени реалистични изображения, които са трудно различими от

⁹ Най-широко използваният вид модели за генериране на изображения. Термина се превежда „текст към изображение“ и на практика превръща думите в снимки.

реални. Генерираните изображения са с по-висока резолюция, включително и 1024x1024 пиксела, което е по-висока резолюция в сравнение с други системи за генериране на изображения. Голяма част от всички нови ИИ генератори са базирани на процеса дифузия. OpenAI не спират до тук и продължават да разработват различни модели през следващата година пускат CLIP (Contrastive Language-Image Pre-Training). Това е модел за обработка на естествен език и изображения. Той използва големи количества данни, за да научи обща репрезентация за езика и изображенията, която позволява да се измерва сходството между текст и изображения в един и същ пространствен модел. CLIP е предвиден да работи за множество задачи, като класификация на изображения, решаване на задачи на мултимодално сходство (както например свързване на изображения с описателни думи) и генериране на изображения от текст. Моделът може да се обучава на големи данни, включително от Интернет, и след това да се използва за решаване на различни задачи без нужда от допълнително обучение. Системата CLIP бързо се популяризира и започва да се използва от повечето модели създадени през последните години, като DeepDaze; Big Sleep; Aphantasia; VQGAN+CLIP; CLIP Guided Diffusion (първия алгоритъм работещ с CLIP и модела дифузия); LAION-400M, Disco Diffusion (популярен AI генератор за живописни изображения).

Същата година NVIDIA пускат продължението на GauGAN под името NVIDIA Canvas. Canvas използва генеративна моделна мрежа, за да анализира чертежите, рисунките или скицираните бележки и да генерира цифрово изображение, което прилича на тях. Моделът на Canvas може да разпознава и различни стилове, като например акварел, мастило, молив и много други, и да генерира изображения, които приличат на тези стилове. Canvas е създаден с цел да бъде интуитивен и да бъде използван от криейтив професионалисти като художници, дизайнери и аниматори, за да ускорят процеса на създаване на цифрови изображения.

Художникът и програмист Марио Клингеман (Mario Klingemann), който има голям опит в използването на невронните мрежи за създаване на интерактивни изкуства, представя на 13 март 2022 г. генератора Midjourney. Това е невронна мрежа за генериране на изображения, която използва архитектура, наречена StyleGAN2. Това е подобрена версия на оригиналния модел StyleGAN, който първоначално е представен от ИИ на NVIDIA. Моделът е обучен с голям брой изображения от интернет, които са проучени и анализирани, за да се извлекат определени характеристики. Така наречените "latent codes" или "style codes", се използват след това от модела за генериране на нови изображения. Midjourney може да генерира фотореалистични изображения с висока резолюция и ярки цветове трудно различаващи се от традиционно създадените, като същевременно предоставя голям контрол върху процеса на генериране. Моделът позволява на потребителя да въведе определени параметри, като цветове, форми и текстури, за да се генерират изображения, които отговарят на тези характеристики. Midjourney е също така стабилен и надежден модел, който може да генерира качествени изображения без големи разходи за изчислителна мощност. Това го прави подходящ за различни приложения, включително изкуство, дизайн, реклама и много други.

Популярността на модела нараства лавинообразно, както сред професионалистите, така и сред по-широката общественост. Това помага за непрекъснато засилващия се интерес и популяризирането на генерирането на изображения и произведения на изкуството, чрез изкуствен интелект.

По късно през юли месец се появява модел за генериране на изображения Stable Diffusion от Google Research. Той е базиран на концепцията за "diffusion processes" (процеси на дифузия). Този процес се извършва чрез последователно размазване на изображението с шум, като след всяко размазване шумът се намалява. Това позволява на модела да генерира изображения, които са по-прецизни и по-детайлни, тъй като информацията от предишния етап все още се запазва в последващите етапи. Тази концепция се основава на различни подобрения върху предишни модели, като DeepDream, StyleGAN и VQ-VAE. Основното предимство на Stable Diffusion е, че той може да генерира изображения с висока резолюция и с много детайли, като в същото време запазва и реалистичността на изображенията. Това го прави изключително полезен инструмент за генериране на изображения в области като компютърна графика, изкуствен интелект, дизайн, медии и други.

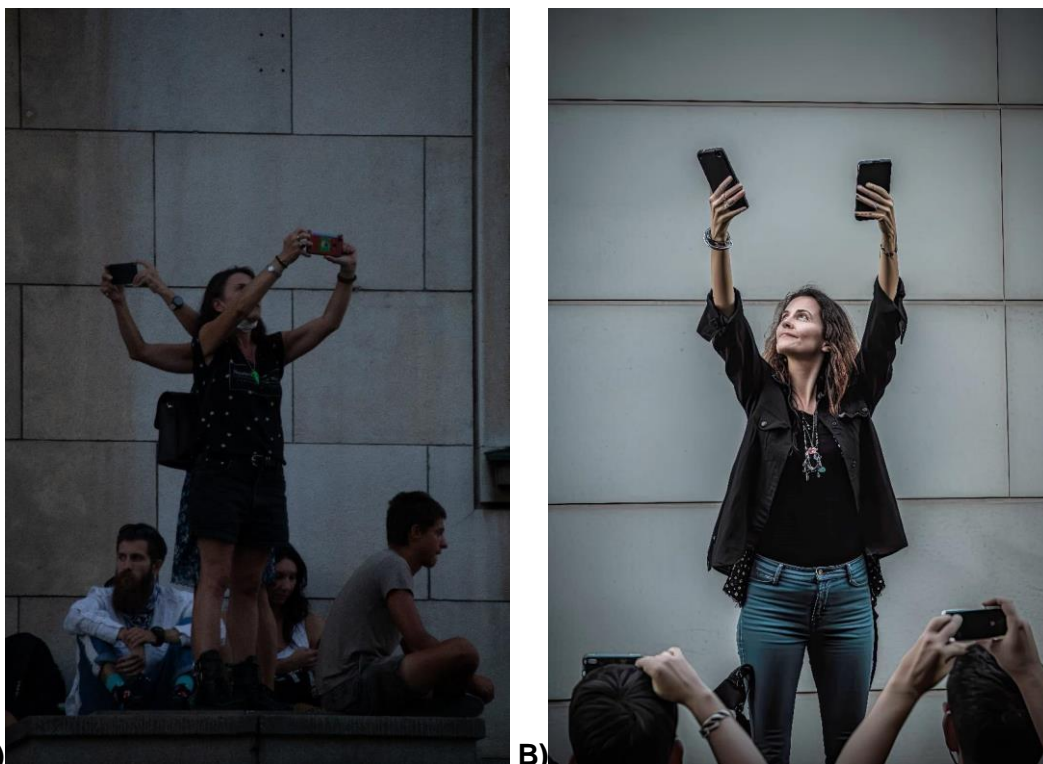
Този кратък исторически дискурс свързан с развитието на изкуствения интелект отваря много въпроси и показва, каква динамична област от науката и изкуството представлява изкуственият интелект и в частност генерирането и коригирането на изображения. Все пак трябва да се каже, че живеем в нова ера ерата на изкуствения интелект. Със сигурност той ще промени начина на живот, както компютрите и Интернет го направиха.

Изкуственият интелект има потенциала да промени графичния дизайн, визуалната комуникация и визуалното изкуство по много начини. Една от основните области, в които ИИ има влияние, е автоматизирането на процеса на създаване на дизайн, като се използват алгоритми и машинно обучение, за да се генерират автоматично дизайн елементи като цветове, форми, композиции и т.н. Това може да се използва за създаване на голям брой вариации на един и същи дизайн, като се ускори процесът на тестване и подобрение на дизайн решенията. Изкуственият интелект може да бъде използван и за оптимизиране на производствения процес на дизайн и визуалното изкуство, като се намали времето и разходите, свързани с ръчното създаване на дизайн елементи. Друга област, в която ИИ има влияние, е персонализацията на дизайн, като се използват алгоритми за анализ на данните за потребителското поведение и предпочитанията, за да се генерират уникални дизайн решения, персонализирани за конкретния потребител. Изкуственият интелект има влияние и върху визуалната комуникация като се използва за автоматизиране на процеса на анализ на социални медии и други данни за комуникационните тенденции, за да се създават по-ефективни комуникационни стратегии и съобщения. Въпреки, че ИИ може да има голям потенциал за автоматизация на процесите в графичния дизайн и визуалната комуникация, все още е важно да се запази човешката креативност и умение за оценка и интерпретация на естетическия опит.

Изкуственият интелект има потенциал за създаване на изображения, но все още съществуват някои предизвикателства и недостатъци. Един от главните му недостатъци за графичен дизайн, е липсата на креативност и интуиция, които често са необходими за създаването на успешен дизайн. На този етап е трудно да научиш машините, какво е красиво. Това е нещо, което изисква интуиция и много опит. Също така, изкуственият интелект може да бъде склонен към повтаряне на шаблони, което може да доведе до монотонни или предвидими дизайни. Ако алгоритъмът е проектиран да създава определен тип графичен елемент, тогава той може да попадне в капан на еднообразие. Също така, изкуственият интелект може да се бори да разбере контекста на определени думи или изображения, което може да доведе до неправилни или неуместни решения. В проучвания, изследователите разглеждат проблема със създаването на „коректни“ изображения за определени думи, като например „добродетел“ или „грях“. Изследователите заключават, че изкуственият интелект все още има трудности в разбирането на тези абстрактни концепции и правенето на подходящи решения. Някои критици се тревожат за етичните проблеми, свързани с технологията, включително въпросите за авторски права, защита на личните данни и злоупотреба. Трябва да се има в предвид, че въпреки постиженията, съществуват и някои недостатъци в използването на изкуствен интелект за създаване на изображения. Един от най-големите недостатъци е, че изображенията, генерирани от ИИ, все още не могат да бъдат напълно разграничени от тези, създадени чрез конвенционални техники от човек (Фигура 4)¹⁰ автор на фотографиите (Самуил Младенов). При използването на ИИ, възниква опасността да се стигне до генериране на изображения, които могат да бъдат използвани за манипулиране на общественото мнение или за други нежелани цели.

Също така, ИИ може да създава изображения, които са изключително реалистични, но не и оригинални или иновативни. Това е поради факта, че ИИ може да използва само информация, която вече е била предоставена на машината. Ако няма съществуващи изображения или информация за определен обект или концепция, ИИ може да създаде само ограничен брой възможности за това изображение. В допълнение към това, някои учени смятат, че използването на ИИ за генериране на изображения, може да доведе до загуба на природата и същността на изобразявания обект. Например, може да генерира изображение на кон, което изглежда напълно реалистично, но не отразява същността на това какво е кон. Като цяло, използването на изкуствен интелект за създаване на изображения е все още в начален етап на развитие и със сигурност има нужда от повече изследвания и разработки, за да се намалят недостатъците и да се увеличат предимствата.

¹⁰ Параметри на изображенията: в ляво - Canon EOS 750D Sigma 18-300mm F3.5-6.3 DC MACRO OS HSM ISO: 800 f/6.3 Скорост на затвора: 1/160 sec. Фокусно разстояние: 230 mm. UV филтър. В дясно - Midjourney Версия: 5 Въведена команда: <<https://s.mj.run/Kv3icmMUsBw>> Standing woman with raised hands, taking picture with phone. She has four arms in both directions. All four hands are holding a one phone Параметри: --ar 2:3; --v 5; --s 250; --q 2; --iw 1 Image seed: 2067008541 Допълнителен софтуер: Topaz Gigapixel AI; Topaz Sharpen AI



Фигура 4. В ляво - фотография на Самуил Младенов създадена чрез конвенционална техника В дясно - изображение генерирано с помощта на ИИ

Появата на ИИ, с който да се генерират изображения ще промени значително представите ни в областта на изкуството. Може да се каже, че всяка форма на медия ще претърпи промени от всички гледни точки, а изкуството може би ще бъде най-засегнато. Ще бъде все по трудно да осъзнаем, какво наистина е изкуството. Тепърва ще излизат все повече нови въпроси, като дали изкуствения интелект ще подпомогне за създаването на изкуство или ще се стигне до унищожаване и унифициране по своето безкрайно разнообразие, едва ли някой може да каже. Все пак може да се заключи че ИИ може да бъде използван като мощен инструмент, който ако попадне в ръцете на истински творец, би дал завидни резултати.

Научна литература

Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, Kai; and Fei-Fei, Li (2009). ImageNet: A large-scale hierarchical image database, 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: <https://doi.org/10.1109/CVPR.2009.5206848>

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybernetics* 36, 193–202 (1980). <https://doi.org/10.1007/BF00344251>

Krizhevsky, A.; Sutskever, I.; E. Hinton, G. (2017). ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 6 (June 2017), 84–90. doi: <https://doi.org/10.1145/3065386>

- Mansimov, E.; Parisotto, E.; Lei Ba, J.; Salakhutdinov, R. (2016).** Generating Images from Captions with Attention. 4th International Conference on Learning Representations (ICLR 2016), Conference Track Proceedings, doi: <https://doi.org/10.48550/arXiv.1511.02793>
- Parmar, G.; Kumar Singh, K.; Zhang, R.; Li, Y.; Lu, J.; Zhu, J. -Y.; (2023).** Zero-shot Image-to-Image Translation, doi: <https://doi.org/10.48550/arXiv.2302.03027>
- Zhu, J. -Y.; Park, T.; Isola, P.; Efros, A. A. (2017).** Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2242-2251, doi: 10.1109/ICCV.2017.244.
- Artbreeder. (n. d.).** <https://www.artbreeder.com/> (last view: 27.04.2023)
- Absolutely Ai Creative content agency. (n. d.).** <https://www.absolutelyai.com.au/> (last view: 27.04.2023)
- DALL·E: Creating images from text. (n. d.).** <https://openai.com/dall-e-2/> (last view: 27.04.2023)
- Dumoulin, V.; Shlens, J.; Kudlur, M.; Google Brain Team, (2016).** Supercharging Style Transfer. <https://ai.googleblog.com/2016/10/supercharging-style-transfer.html> (last view: 27.04.2023)
- Garcia, C. (2016).** Harold Cohen and AARON - a 40-year collaboration, Computer History Museum. <https://computerhistory.org/blog/harold-cohenand-aaron-a-40-year-collaboration/> (last view: 27.04.2023)
- NVIDIA Canvas. (n. d.).** <https://www.nvidia.com/en-us/studio/canvas/> (last view: 27.04.2023)

Asst. Dr. Boyan, Blazhev

South-West University "Neofit Rilski", Blagoevgrad, 2700, Bulgaria

ORCID  <https://orcid.org/0000-0002-3376-8249>

Boian78@gmail.com; boian_lb@abv.bg

boyanblazhev.com

ас. др. Боян Блажев

Юозападен университет „Неофит Рилски“, ул. "Иван Михайлов" №66
Благоевград, 2700, България

AUTHOR'S DATA WERE PUBLISHED ACCORDING GDPR RULES AND PUBLICATION ETHICS OF THE JOURNAL (<http://www.math.bas.bg/vt/kin/>)

Received: 14 May 2023

Accepted: 20 June 2023

Published: 30 June 2023

DOI: <https://doi.org/10.55630/KINJ.2023.090109>

KIN Journal, 2023, Volume 09, Issue 1

Science Series Cultural and Historical Heritage: Preservation, Presentation, Digitalization

Научна поредица Културно-историческо наследство: опазване, представяне, дигитализация

Научная серия Культурное и историческое наследие: сохранение, презентация, оцифровка

Editors

Prof. PhD. Petko St. Petkov

Prof. PhD. Galina Bogdanova

Редактори/съставители

проф. д-р Петко Ст. Петков

проф. д-р Галина Богданова

Copy editors

Assist. prof. PhD. Nikolay Noev

Assist. prof. PhD. Kalina Sotirova-Valkova

PhD. Paskal Piperkov

Технически редактори

гл. ас. д-р Николай Ноев

ас. д-р Калина Сотирова-Вълкова

д-р Паскал Пиперков

© Editors, Authors of Papers, 2023

© Редколегия, Авторски колектив, 2023

Published by

Institute of Mathematics and Informatics

at the Bulgarian Academy of Sciences,

Sofia, Bulgaria

Издание на

Институт по математика и

информатика при Българска академия на

науките, София, България

<http://www.math.bas.bg/vt/kin/>

ISSN: 2367-8038