

РАЖДАНЕТО НА AI ВИДЕОПОКОЛЕНИЕТО

Десислава Бакърджиева

Университет по библиотекознание и информационни технологии

Резюме: Настоящата статия разглежда създаването и развитието на AI (ИИ, изкуствен интелект) видеопоколение – ИИВП или AI ВП. ИИ (AI) не е формат или функция, той е процес и способност за обработка на информация и анализ на данни.¹ ИИ не е предназначен да замени хората. Целта му е да даде възможност на индивида за по-успешно реализиране на даден проект или задача. Само преди няколко години теорията да се създаде видеоклип само с изписване на ключови думи, е немислима, също така идеята да се разговаря с виртуален приятел, изглежда като научна фантастика. Работещите с изкуствен интелект инструменти за генериране на видео са не само реалност, те са част от всяка платформа и кинематографични визуализации.

Ключови думи: AI видеопоколение; AI кинематографични визуализации; AI – ИИ изкуствен интелект; чатбот; deepfake

ВЪВЕДЕНИЕ

ИИ е набор от технологии, които съчетават данни, алгоритми и изчислителна мощ.² Съвременните IT специалисти използват изкуствен интелект, или т.нар. „чатботове“, за по-ефективно изпълнение на задачи, свързани с комуникацията им с клиентите. Чрез предварително заложи идентифициращи модели на общуване ИИ чатботът може да проведе разговор, да създаде виртуалното ви изображение и да ви насочи към търсения артикул или информация, без общо да участвате в процеса. Всяка компания се нуждае от специалист в съответната сфера с математически познания и способност да използва алгоритми, за да създаде AI чатбот, отговарящ на съответните специфики и отрасли. Работещите с AI инструменти за генериране на видео са не само реалност, но и променят творческия пейзаж – от кратки, динамични клипове до зашеметяващи кинематографични визуализации. Изкуственият интелект дава на създателите му безпрецедентен контрол върху разказването на истории и изображения.

МЕТОДОЛОГИЯ НА ИЗСЛЕДВАНЕТО

Метод на наблюдение, системен анализ, приложение на метода на аналогията за идентифициране на връзките и взаимодействията в съвременната информационна и технологична среда.

Целта на статията е да се изследват създаването и развитието на ИИ видеопоколение, рисковете и силните страни при използването и внедряването на ИИ във всички медии и платформи. Да се проследят и онагледят последните разработки и приложенията на ИИ в съвременната информационна среда, трансформациите и взаимодействията. Да се зададат бъдещи модели на развитие на ИИВП.

ИЗЛОЖЕНИЕ

Изразът „изкуствен интелект“ (ИИ), или на английски Artificial Intelligence (AI), се използва за първи път официално през 1955 г.³ Европейската комисия (ЕК) създава регулаторен орган, който има за цел да въведе правила при използването на ИИ – AI. Европейският съюз регулира използването на изкуствен интелект и едно от правилата, които са въведени от ЕК, е създаденото от изкуствен интелект да бъде разпознаваемо – да има отличителен знак или маркировка на точно определено място. В списъка е и забраната за използването на изкуствения интелект за биометрично наблюдение, класифицирана информация, използване на лични данни, с изключение на прогнозна полицейска дейност.⁴ Ако в средата на миналия век изкуственият интелект е само тема на теоретични научни статии, то днес системи, базирани на ИИ, са широко разпространени във всички аспекти на живота ни. С помощта на чатботовете не само се търси информация, а се оптимизира селското стопанство, разчитат се лабораторни изследвания на човешка кръв, диагностицират се заболявания и се начертават схеми за лечение, разпознават се образи за рентгенов анализ, на базата на термовизия се диагностицира рак и т.н. Чрез ИИ се управляват коли без наша намеса, създава се виртуален приятел – партньор, и чрез няколко команди се създава видеоклип. Първоначалните експерименти с движещи се кадри сега се превръщат в надпревара между технологичните гиганти за изграждане на най-сложните, реалистични и достъпни AI видеоинструменти. Само през 2024 г. се отбелязват революционни иновации в създаването чрез AI видеоизображение от OpenAI, Google, Tencent и Imagine Art. Инструменти като Sora, Veo 2, Hunyuan и Videocraft са водещи, като всеки от тях разширява границите на възможното (Kurzweil 2025).

Терминът „AI видеопоколение“ не обозначава демографска група по възраст (както е при поколенията Z, X или Millennials), а технологично-социална генерация, която се формира около възхода на генеративния изкуствен интелект и неговото приложение във видеосъдържанието. Това поколение не е биологично, а функционално и културно – съставено от технологии, практики и потребители, които създават, консумират и взаимодействат с видеосъдържание, генерирано от ИИ. Не става дума за възраст, а за технология. Докато поколение Z например включва хора, родени между края на 90-те и 2010-те, AI видеопоколението обединява всички участници в една нова медийна екосистема – от разработчици до потребители, без оглед на възраст, пол или местоположение. Терминът „видеопоколение“ подчертава доминантния медиен формат.⁵ Ако за предишни поколения основният комуникационен формат е бил текстът (уеб 1.0), а след това – изображението (уеб 2.0 и социалните мрежи), то видеосъдържанието, особено създаденото от ИИ, се превръща в доминираща форма на изразяване, комуникация и творчество в уеб 3.0 и отвъд. AI е не просто инструмент, а партньор в креативния процес. Появата на платформи като Sora (OpenAI), Runway, Pika Labs и др. води до радикално нов модел: не просто потребление на видео, а създаване на синтетична реалност, в която ИИ участва

като съавтор. Затова говорим за поколение, което не само гледа, но и „генерира“ реалност чрез видео.⁶

През последните две години усъвършенстването на технологията и програмния код позволява създаването на графични и видеоизображения, като Chat GPT свързва с някои видеопрограми. AI генерирането на текст към видео позволява на потребителите да създават видеоклипове просто като описват дадена сцена (само текстово).⁵ На базата на този текст или команда, зададена от потребителя, AI интерпретира и генерира съответстващо видео, с движение, осветление и максимално близко до естественото човешко поведение. Това е най-популярният тип генериране на AI видео (Luger et al. 2004).

Модел	Година	Методология	FVD↓	IS↑	Вход	Изход	Източник
Sora (OpenAI)	2024	Latent Diffusion	18.2	N/A	Текст	60s 1080p	OpenAI Blog
VideoCrafter2	2023	Diffusion	21.5	9.8	Текст	16s 720p	Tencent AI
Gen-2 (Runway)	2023	GAN+Transformer	25.0	8.9	Текст/ Видео	4s 720p	Runway
Make-A-Video	2022	Diffusion+CLIP	28.3	8.5	Текст	5s 480p	Meta AI

Метрики: FVD = Fréchet Video Distance (по-ниско е по-добре), IS = Inception Score (по-високо е по-добре). Това е сравнителна таблица на някои от най-известните модели за текст-към-видео (text-to-video) генерация, базирана на различни технологии и с различни показатели за качество.

При генерирането на видео – от видео към видео, се използва AI за подобряване, модифициране или трансформиране на съществуващи видеоклипове, вместо да се създават нови от нулата. Създава се структура от кодове – скеле, запълват се съответните файлове; това е така познатият архитектурен модел. Това може да включва подобряване на качеството на видеоклипа, промяна на стилове, добавяне на специални ефекти или дори промяна на елементи в кадрите – като премахване на обекти или замяна на фонове.⁸

Synthesia е водещ AI видеогенератор, който позволява да се създават видеоклипове със студийно качество, с реалистични говорещи AI аватари, като дублажът може да бъде синхронизиран с говора. Този тип програми помагат за синхрона при озвучаване на различни езици. Той включва над 60 видеа и над 140 езика и акцента.⁹

Deepfake е съставна дума от deep learning (дълбоко обучение) и fake (фалшив). Терминът се използва за означаване на синтетични медийни материали – обикновено видео или аудио, създадени чрез дълбоки невронни мрежи, които могат да имитират човешко лице, глас, изражения и поведение с изключителна убедителност. Дълбоката невронна мрежа е изкуствена невронна мрежа с множество слоеве между входния и изходния слой. Тези слоеве се наричат скрити слоеве и са отговорни за прогресивното извличане на по-сложни представяния

от данните. Първоначално deepfake системите използват автоенкодери, които научават как да компресират (encode) и възстановяват (decode) изображения на лица. При смяна на декодера може да се възпроизведе лице на друг човек върху оригиналното видео.¹⁰

Generative Adversarial Networks (GANs). Generative Adversarial Networks, или съкратено GANs, представляват иновационна технология, която се използва масово в сферата на изкуствения интелект. Този вид изкуствени невронни мрежи, използвани за създаване на ново, синтетично съдържание, което изглежда изключително реалистично – изображения, видео, текст, аудио и др., намира широко приложение при създаването на синтетично съдържание. Въведени за първи път от Иън Гудфелоу и неговия екип през 2014 г., GANs бележат нов етап в развитието на машинното обучение, базиран на концепцията за съревнование между две невронни мрежи (Goodfellow, Bengio, Courville 2016).

Този модел функционира чрез т.нар. „състезателен обучителен процес“, в който участват две основни мрежи: генератор и дискриминатор. Генераторът има за цел да създава нови, изкуствени примери, наподобяващи реални данни, докато дискриминаторът се опитва да различи кои от входните данни са истински и кои – генерирани. С всяка итерация и двете мрежи се усъвършенстват, като генераторът става все по-добър в създаването на фотореалистично съдържание, а дискриминаторът – в разпознаването на измама (Creswell, White, Dumoulin, Arulkumaran, Sengupta, Bharath 2018).

Технологията на GANs намира приложение в множество области. В медийната индустрия тя се използва за създаване на deepfake видеа – високореалистични фалшификати (Mirsky, Lee 2021).

В изкуството и дизайна GANs се използват за създаване на оригинални произведения, които често надминават човешките възможности по отношение на детайли и креативност. В медицината се прилага за генериране на синтетични образи, използвани при обучение на специалисти или за симулации в клинични среди.

Наред с позитивните приложения GANs пораждат и сериозни етични, правни и социални въпроси. Потенциалът за злоупотреба с технологията – особено при разпространяването на дезинформация или създаването на фалшиво компрометиращо съдържание, налага необходимост от регулация и развитие на надеждни методи за откриване и маркиране на синтетични медии. Сред основните проблеми на младото поколение е приемането за вярна на всяка информация в интернет пространството без проверка на източници, компетентни автори и капацитети в съответната област (Goodfellow, Pouget-Abadie, Mirza, Warde-Farley, Courville, Bengio 2021).

Очаква се до 2030 г. AI видеопоколенията да постигне:

▷ пълна фотореалистичност с контрол върху осветление, камера, движения и физически взаимодействия;

▷ интерактивни AI аватари и герои, способни на диалог в реално време;

▷ демократизация на продукцията, при която всеки с минимални умения

ще може да създава анимационни филми, реклами и документални видеа с мощта на AI инструменти;

▷хиперперсонализирани видеа – базирани на индивидуални предпочитания, емоции или данни в реално време.

Перспективи за регулация

Развитието на генеративния AI поставя спешна нужда от правни и етични рамки. Очакваните посоки за регулация включват:

▷ясно етикетирание на синтетично съдържание – законово изискване генерирани видеа (особено *deepfake*) да бъдат маркирани като такива;

▷авторски права и собственост – нови законови категории, които дефинират кой е носителят на права при AI генерирани произведения (потребител, разработчик, платформа);

▷лицензиране на AI инструменти – възможно задължително лицензиране и сертификация на генеративни системи, използвани в медии, реклама и изкуство;

▷етика и защита на човешкия образ – забрана за използване на лица, гласове или идентичности без изрично съгласие;

▷регулаторни органи – Европейският закон за AI (AI Act), както и инициативи в САЩ и Япония ще създадат надзорни структури и стандарти за безопасна употреба.¹²

Ключовото предизвикателство ще бъде намирането на баланс между свободата на творческо изразяване и защитата на обществения интерес. Гъвкавите, но стриктни регулации ще бъдат от съществено значение, за да се избегнат злоупотреби, без да се потиска иновацията в изкуствата.

ЗАКЛЮЧЕНИЕ

Раждането на AI видеопоколение обхваща еволюцията на технологиите за генериране на видео чрез изкуствен интелект. Тази област е изключително динамична и напредва с огромни темпове благодарение на дълбокото обучение и генеративните модели. Все повече компании приемат конкурентните предимства на AI за бизнеса и превръщат използването на тази технология в приоритет. AI предлага много инструменти и възможности за намаляване на разходите и рисковете. Generative Adversarial Network е сред най-мощните и трансформиращи инструменти в съвременния ИИ. Способността да симулират реалността в дигитална форма, е както обещаваща, така и предизвикателна – съчетаваща научен прогрес с необходимост от отговорна и етична употреба.

БЕЛЕЖКИ

1. LUGER, George et al., 2004. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. 5th. The Benjamin/Cummings Publishing Company, Inc. Available from: <<https://www.cs.unm.edu/~luger/ai-final/tocfull.html>>.
2. ARTIFICIAL intelligence for Europe, Legislative Observatory (OEIL).

- European Parliament. Ref.: 20200918STO87404, 2020. Available from: <Artificial intelligence: threats and opportunities | News | European Parliament (europa.eu).
3. EUROPEAN coordinated plan on artificial intelligence, press release. Council of the EU, 18 February 2019. Available from: <Texts adopted - Civil Law Rules on Robotics - Thursday, 16 February 2017 (europa.eu).
4. SYNTHESIA Ai Alt – text to video Generator, 2024. Available from: <<https://www.synthesia.io/futures/text-to-video>.
5. https://www.capital.bg/biznes/finansii/2023/05/29/4488848_tehnologijite_bazirani_na_izkustven_intelekt_sa_dobri_i/.
6. YANG, G. et al., 2023. *VideoCrafter2: A strong video diffusion model with motion-aware guidance*. Available from: arXiv:2309.11472.
7. MCKINSEY & Company. Power up: How Southeast Asia's largest bank is becoming AI-fueled, 2022. Available from: <<https://www.mckinsey.com/capabilities/quantumblack/our-insights/power-up-how-southeast-asias-largest-bank-is-becoming-ai-fueled>.
8. GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A., 2016. *Deep Learning*. MIT Press. Available from: <http://www.deeplearningbook.org>.
9. LECUN, Y.; BENGIO, Y.; HINTON, G., 2015. Deep learning. *Nature*, 521(7553), pp. 436 – 444. Available from: <https://doi.org/10.1038/nature14539>.
10. SCHMIDHUBER, J., 2015. Deep learning in neural networks: An overview. *Neural Networks*, 61, pp. 85 – 117. Available from: <https://doi.org/10.1016/j.neunet.2014.09.003>.
11. SCHMIDHUBER, J., 2015. Deep learning in neural networks: An overview. *Neural Networks*, 61, pp. 85 – 117. Available from: <https://doi.org/10.1016/j.neunet.2014.09.003>.
12. HU, C.; TIAN, Q.; LI, T.; WANG, Y.; WANG, Y.; ZHAO, H., 2021. Neural Dubber: Dubbing for videos according to scripts. *arXiv preprint arXiv:2105.13284*. Available from: <https://arxiv.org/abs/2105.13284>.
- Представя модел за автоматичен дублаж, съчетаващ визуална синхронизация с TTS технологии, използван в AI видеосинтез.

REFERENCES

- KURZWEIL, Ray, 2005. *The Singularity is Near*. Penguin Books. ISBN 0-670-03384-7.
- LUGER, George et al., 2004. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. 5th. The Benjamin/Cummings Publishing Company, Inc. ISBN 0-8053-4780-1.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A., 2016. *Deep Learning*. MIT Press.
- CRESWELL, A.; WHITE, T.; DUMOULIN, V.; ARULKUMARAN, K.; SENGUPTA, B.; BHARATH, A. A., 2018. Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35(1), pp. 53 – 65.

MIRSKY, Y.; LEE, W., 2021. The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, 54(1), pp. 1 – 41.

GOODFELLOW, I.; POUGET-ABADIE, J.; MIRZA, M.; XU, B.; WARDEFARLEY, D.; OZAIR, S.; COURVILLE, A.; BENGIO, Y., 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems (NeurIPS)*.

THE BIRTH OF THE AI VIDEO GENERATION

Abstract: *This article examines the creation and development of the AI Video Generation – AIVG. AI is not a format or a function; it is a process and a capability for information processing and data analysis. AI is not intended to replace humans. Its purpose is to enable individuals to complete projects or tasks more successfully. Just a few years ago, the idea of creating a video clip by merely typing keywords seemed unthinkable, and talking to a virtual companion sounded like science fiction. Today, AI-powered video generation tools are not only a reality – they are an integral part of every platform and cinematic visualization.*

Keywords: AI video generation; AI cinematic visualizations; artificial intelligence (AI); chatbot; deepfake

Desislava Bakardzhieva, PhD

University of Library Studies and Information Technologies
Sofia, Bulgaria

E-mail: bakardjievad@yahoo.com