Romanian Journal of Petroleum & Gas Technology
VOL. V (LXXVI) • No. 1/2024

ARTIFICIAL INTELLIGENCE – POWERED VIDEO CONTENT GENERATION TOOLS

Cosmina – Mihaela Roșca [1]

Ionuț Adrian Gortoescu [2]

Marius Radu Tănase [3]

[1] Petroleum-Gas University of Ploiesti, Romania
[2] Petroleum-Gas University of Ploiesti, Romania
[3] Petroleum-Gas University of Ploiesti, Romania
e-mail: cosmina.rosca@upg-ploiesti.ro

DOI: 10.51865/JPGT.2024.01.10

ABSTRACT

This article discusses the considerations of artificial intelligence-powered video content generation tools, exploring their applications, ethical considerations, and evaluation criteria. Through discussions of various artificial intelligence (AI) tools, including features, limitations, and implications, the authors analyze the evolving landscape of video creation in the digital age. Key themes include the ethical implications of deep fake technology and the importance of responsible AI principles, exemplified by Microsoft's guidelines.

This paper identifies five of the most promoted free social media tools. Evaluation criteria for these tools, such as visual quality, relevance, coherence, authenticity, and transparency, are examined to assess the suitability of AI-generated videos. While AI offers promising opportunities, the discussion underscores the continued need for human oversight and ethical considerations to ensure the responsible use of AI technologies in video content generation.

Keywords: AI-powered video content generation, ethical considerations, deep fake technology, responsible AI principles, evaluation criteria

INTRODUCTION

Artificial Intelligence (AI) has undoubtedly transformed various aspects of our lives, from how we interact with technology to how businesses operate. However, despite the significant advancements in AI technologies, there remains a notable gap between the capabilities of these tools and the complexities of human intelligence.

AI is revolutionizing the video content landscape across various domains, offering unparalleled capabilities. AI-powered video surveillance enhances security by detecting anomalies and identifying threats in real-time. With the rapid expansion of network technology, there's a surge in audio and video content creation, demanding adequate data protection in the online sphere. Extracting targeted information from vast audio and video datasets presents a significant challenge for surveillance purposes. Paper [1] introduces

an automatic identification system tailored for audio content, blending machine vision and sound analysis in deep learning frameworks. This method's evaluation revealed the highest identification accuracy at 98%, while recognition performance on the dataset was achieved at a rate of 97%. These results underscore the significance of employing AI technologies, particularly in computer systems, for enhanced data processing and analysis [1].

In sports, tennis motion precision video monitoring simulation, based on AI algorithms, continuously enhances with algorithmic advancements. This technology offers precise detection and processing of sports items, even at smaller dimensions, through advanced video detection techniques [2]. Through data analysis, the study [2] demonstrates the advantages of tennis motion precision video detection based on AI algorithms, achieving high efficiency in image precision, resolution, fluency, and fidelity, with efficiencies of 85.98%, 90.14%, 93.06%, and 98.25%, respectively. Paper [3] evaluates the efficacy of an AI video feedback system in college tennis teaching. Integrating this methodology in college physical education enhances students' acquisition of learning knowledge more intuitively, facilitates faster establishment of action models and spatial thinking concepts, and offers personalized and timely correction and consolidation of technical actions. These advancements are not aimed at replacing teachers but at enhancing teaching methods to make education more accessible. Paper [4] reviews recent research on attitudes towards AI in education and its practicality. Building upon previous studies, a new investigation focused on creating video content using innovative technologies. The increased integration of intelligent technology and AI in education may lead to an improved understanding of current and future challenges in the educational process.

The rise of diverse broadcasting platforms has shifted web content consumption towards multimodal formats, sparking interest in integrating multimedia into AI research. The study [5] introduces a comprehensive dataset by analyzing various multimodal information in video clips. Enriched with semantic labels, this dataset aids AI research in classifying information, covering contextual, intentional, and emotional aspects depicted through vision, speech, and language. This dataset fills a gap in public data for multimodal interaction research, especially in Korea, and is expected to support the development of AI services like Korean dialogue processing and visual information extraction.

In Indonesia, integrating AI into education is an important factor for society's evolution due to challenges in learning system management, human resource shortages, and flawed incentives. The paper [6] investigates the role of AI in creating digital learning content to enhance knowledge acquisition and learning empowerment. Through qualitative analysis of AI-generated video content and feedback from respondents, it's found that AI-driven videos positively impact student learning.

Paper [7] presents a computer-based AI system for generating short videos. It features an optimization algorithm for efficient large-angle offset calculation and prediction model selection based on complexity. The system integrates various AI capabilities like face and speech recognition, text recognition, and natural language processing to streamline content aggregation and improve production efficiency. Results indicate the system's robust learning and processing capabilities, facilitating rapid text retrieval and organization to enhance writing and publishing efficiency.

Paper [8] addresses content shortages encountered by museums, art galleries, cinemas, and cultural complexes, attributing them to high production costs associated with traditional production methods. The paper proposes a solution to integrate online and offline media. By utilizing software technologies such as video understanding and enhancement, a diverse range of audio and video programs from the internet are collected, integrated, optimized, and recreated. This process aims to overcome fragmentation and poor quality, achieving high-resolution, high-color-depth, and high-frame-rate cultural products at a more affordable cost. This approach meets the sustainable demand of modern cultural complexes for high-quality cultural content at a lower price point. Video streaming efficiency remains a significant challenge, particularly with the increasing demand for high-resolution content and minimal buffering times. The study [9] introduces a novel solution to enhance user experiences by combining adaptive streaming, AI algorithms predicting network conditions, and edge computing for optimized task allocation. These technologies aim to reduce latency, decrease network traffic, and enable seamless transitions in video quality. The results of this research improve streaming services and facilitate real-time video applications, particularly in remote or mobile environments, presenting significant implications for the future of video streaming.

This article explores the current landscape of AI technologies and tools, highlighting their diverse applications and discussing the limitations of replicating human expertise. Furthermore, we utilize generated video content from various AI applications as illustrations to underscore the evolving capabilities and challenges in the field.

Additionally, AI algorithms optimize video production workflows, automating tasks such as editing and captioning, saving time and resources. As AI continues to advance, its integration into video content creation and analysis promises transformative impacts across diverse sectors, shaping the future of digital media.

The following sections will provide an overview of video content across all branches of artificial intelligence, followed by a focus on content generation. Additionally, the study will delve into the importance of responsible usage of generated video content. Subsequently, the most popular AI services for video content generation will be identified, along with various quality aspects examined to compare and determine the best tool.

VIDEO CONTENT IN THE AI CONTEXT

Processing video content can be utilized across all components of AI, as this type of content is prevalent in most industries. From machine learning algorithms to recommendation systems, video content is a rich data source for training and inference. AI tools can be categorized based on their primary objectives and applications. Here are some key classifications [10]:

- *Machine Learning Algorithms*:
    - **Supervised Learning** represents algorithms trained on labeled data to make predictions or decisions, such as a recommendation system that provides movie suggestions to users depending on their past ratings of movies.

- **Unsupervised Learning** uses algorithms to analyze unlabeled data to uncover hidden patterns or structures - for instance, an anomaly detection system that spots irregular patterns in CCTV footage to identify potential security risks.
- **Reinforcement Learning** benefits of algorithms learn through trial and error by interacting with an environment to achieve specific goals. For example, a self-driving car algorithm learns to navigate through traffic by interacting with its environment and receiving feedback on its driving decisions.

- *Natural Language Processing (NLP) Tools*:
  - **Sentiment Analysis** determines the sentiment or emotional tone in text data, e.g. an AI tool analyzes customer reviews in video comments to gauge the overall sentiment toward a product or service.
  - **Named Entity Recognition** (NER) identifies and categorizes named entities (such as names of people, organizations, or locations) in text. For example, a news video transcript is analyzed to identify and categorize named entities, such as names of politicians, companies, or locations.
  - **Language Translation** translates text from one language to another while retaining its original meaning. For instance, an AI system translates subtitles in a video from English to Romanian while preserving the original meaning and context.
- *Computer Vision Systems*:
  - **Object Detection** involves identifying and pinpointing objects in images or videos. Example: A security camera can detect and highlight suspicious items or packages in a busy airport terminal.
  - **Image Classification** categorizes images into categories or labels that are predefined. An example could be a wildlife conservation organization that uses AI to classify camera trap images, distinguishing between animal species.
  - **Facial Recognition** verifies and recognizes individuals based on facial features, e,g., a video surveillance system identifies and tracks individuals entering a high-security facility based on facial features.
- *Generative Models*:
  - **Generative Adversarial Networks** (GANs) models consist of dual neural networks, a generator, and a discriminator, employed to produce authentic synthetic data like images or text. For instance, an AI artist employs a GAN trained on a dataset of natural landscape images to craft photorealistic scenery paintings.
  - **Variational Autoencoders** (VAEs) are methods for generating new points as data, instrumental in image generation and anomaly detection, such as a video editing tool that generates new video effects by learning from a dataset of existing special effects and applying variations.

- *Robotics and Autonomous Systems*:
    - **Autonomous Vehicles** are AI-powered vehicles that can navigate and operate without human intervention. For example, a self-driving delivery van navigates through urban streets to deliver packages to customers without human intervention.
    - **Industrial Robots** are equipped with AI for manufacturing, assembly, and logistics tasks. An example might be an AI-powered robotic arm that precisely assembles electronic components in a factory assembly line.
- *Expert Systems*:
    - **Rule-based systems** utilize predefined rules to make decisions or provide recommendations within specific domains. For instance, an AI tutor provides personalized video lessons based on a student's learning progress and performance.
    - **Knowledge-based systems** incorporate expert knowledge to solve complex problems or assist decision-making processes, e.g., a medical diagnosis system analyzes symptoms presented in a video consultation and recommends treatment options based on expert medical knowledge.
- *Recommendation Systems*:
    - **Collaborative Filtering** recommends items or content based on user preferences and similarities with other users. For example, a video streaming platform suggests social-media content to users based on their viewing history preferences and similarities with other users.
    - **Content-based filtering** recommends different items based on their attributes and similarities to items previously liked by the user, such as a cooking video platform that recommends recipes to users based on their past recipe likes and preferences.

These classifications represent a fraction of the diverse applications of AI across various industries, each with its unique set of challenges and opportunities. As we delve deeper into AI's capabilities, it becomes evident that while these technologies continue to advance, they still fall short of replicating the multifaceted nature of human intelligence. Through AI-generated images, we aim to shed light on the progress and inherent limitations in achieving human-like intelligence in machines.

In May 2024, social media platforms are inundated with posts advocating that individuals no longer need to engage in traditional studies to become experts in various fields. Instead, the narrative promotes that proficiency can be achieved simply by mastering AI tools. Furthermore, there's an intense push towards effortlessly making money using these AI-powered tools with minimal effort.

The authors of this article aim to debunk these notions and advocate for the continued pursuit of in-depth studies to foster expertise in diverse domains. While acknowledging the utility of AI tools, the article stresses that they are adjuncts rather than substitutes for human expertise. It is imperative to emphasize that these technologies augment decision-

making processes but cannot supplant the final judgment, which should always rest with domain experts.

In practice, while AI tools have undoubtedly revolutionized various aspects of industry and everyday life, they are not a panacea for skill acquisition or financial success. They can assist in data analysis, automation, and optimization. However, true mastery and innovation still stem from a deep understanding of the subject matter coupled with human creativity and critical thinking. Thus, the narrative advocating for effortless expertise through AI tools must be tempered with recognizing the irreplaceable value of dedicated study and expertise cultivation.

USING AI VIDEO CONTENT RESPONSIBLY

While the ability to generate video content using AI offers significant advantages, there are also instances of malicious use. One notable example is the phenomenon known as "deepfake." Deepfake technology entails using AI algorithms to produce exceedingly realistic videos, typically by overlaying the faces and voices of individuals onto preexisting video content [11].

The implications of deepfake technology are concerning, as it can be used to create convincing but entirely fabricated content, such as counterfeit celebrity endorsements, political propaganda, or even fake news reports. These videos can be incredibly deceptive, making it difficult for viewers to discern between genuine and manipulated footage.

On 2024 January 22nd, Energy Minister of Romania, Sebastian Burduja, announced that he filed a criminal complaint with the Prosecutor's Office after a deepfake video surfaced on Facebook. The video, created using AI, utilized Burduja's image and voice to falsely promote a nonexistent investment scheme by Enel company [12]. In the manipulated clip, Burduja purportedly endorsed an innovative investment platform by Enel, promising investors an incredible monthly sum of 9,000 RON. Similarly, Prime Minister Marcel Ciolacu was targeted by online scammers, appearing in a manipulated video on social media presiding over a government meeting. His altered voice claimed that shareholders would receive monthly payments of at least 10,300 RON on their cards for a minimum investment of 1,200 RON. These instances highlight the misuse of deepfake technology for fraudulent purposes, prompting legal action and public awareness campaigns [12].

The National Bank of Romania (NBR) issued a warning on 2024, February 5th, regarding the circulation of AI-generated videos on social media involving Mugur Isărescu, the governor of the NBR. These videos falsely promote an investment platform, urging people to spend money on a nonexistent platform. The bank advised against falling for these online scams, emphasizing that the NBR does not make investment recommendations. In the video, an AI-generated voice attributed to Mugur Isărescu promotes an automated platform with a minimum investment of 1,100 RON and potential earnings of 8,200 RON. NBR clarified that Isărescu's image and voice fraudulently promoted this scam [12].

Detecting deep fakes can be challenging as technology continues to evolve and improve. However, researchers are developing methods to identify and combat deep fake videos, including forensic analysis, machine learning algorithms, and blockchain technology to verify the authenticity of video content.

This case violates the fundamental principles of AI service development. It represents a breach of ethical standards and undermines the core values that guide responsible AI implementation. Such instances compromise the integrity of AI technologies and erode trust in their usage and deployment.

Microsoft's Responsible AI principles outline the company's commitment to developing and deploying AI technologies ethically and responsibly. These principles guide Microsoft's efforts to ensure that AI systems are designed and used in ways that prioritize fairness, reliability and safety, privacy and security, transparency, and inclusiveness [10]. The principles have been established to ensure that AI technologies are built and used responsibly and ethically. One of these essential principles is transparency, which emphasizes informing users when they interact with AI-generated content. This entails explicitly stating that certain content or experiences have been generated with the help of AI algorithms so that users are aware of their origin and nature.

While AI-generated video content offers numerous benefits, the misuse of technologies like deepfake underscores the importance of vigilance and skepticism when consuming digital media. It also highlights the need for ongoing research and development of countermeasures to mitigate the potential risks associated with malicious uses of AI.

Analyzing hundreds of social media posts and exploring websites, along with the services promoted on these sites, the authors of this work have identified the following categories of services promoted as AI associated with video content:

- Video Generator are services that create videos automatically;
- Video Editing represents platforms for video editing.

In the following sections, we analyze the first category of AI services (tools) and identify the most popular ones promoted online. Additionally, the authors conduct a case study highlighting the seamless integration of AI-generated content without users knowing whether it was generated with AI technologies. This exploration examines prevalent AI tools across online platforms. It illustrates their effective utilization through case studies, demonstrating the seamless integration of AI-generated content into various contexts while maintaining authenticity and user engagement.

AI VIDEO CONTENT-GENERATED TOOLS

AI-powered video generation services leverage sophisticated algorithms and large datasets to automatically create content based on user input or predefined parameters.

The authors of this study have synthesized the top five most promoted tools on social media according to source [13]. These tools were tested and analyzed from the perspective of future software engineers who approach the evaluation qualitatively, distinct from influencers who typically endorse such tools. The research excludes voice analysis and focuses exclusively on analyzing images from the video content. This means that any examination or consideration related to the audio component of the videos will not be included in the study. Instead, the analysis will solely concentrate on visual elements, such as frames, scenes, objects, and other visual cues in the video content. This decision allows for a more targeted investigation into the visual aspects of the videos, providing insights into their quality, relevance, and other relevant factors without being

influenced by audio-related considerations. This approach seeks to provide a more objective and rigorous assessment of the capabilities of tools, focusing on their functionality, performance, and suitability for software development. By evaluating the tools through the lens of engineering principles rather than marketing hype, the study aims to offer insights more aligned with the practical needs and standards of software development practices.

The main AI video generators for video creation are the following:

- *Runway* facilitates the creation of short videos from ideas with intuitive text-to-image and text-to-video conversion. The user can explore endless variations and styles with image and video transformation tools. The application is available at https://app.runwayml.com/
- *Visla* offers a smart AI-powered video creation and editing tool tailored for businesses. Visla AI allows users to create professional videos with storyboards, scripts, B-rolls, voice-overs, subtitles, and more. It is reachable at https://app.visla.us/
- *Chromox* is an advanced AI video generator specializing in transforming text into visual content. It emphasizes the expansion of creative possibilities and novel options for video creation. It is offered at https://chromox.alkaidvision.com/
- *Genmo* is an AI-driven tool for generating compelling videos from textual input. It harnesses the power of NLP algorithms to produce engaging visual content. It is accessible at: https://www.genmo.ai/
- *Invideo*.ai is a video creation platform powered by AI, enabling users to produce professional-quality videos effortlessly. It is delivered at https://ai.invideo.io/

To conduct a comparative analysis, five characteristics of the generated video content are assessed using a 5-point scale (1 - very dissatisfied, 2 – dissatisfied, 3 – neutral, 4 - satisfied, and 5 - very satisfied). This will determine which tool is best and whether it fully meets the essential characteristics required for AI-generated video content. The five characteristics are as follows:

- **Visual quality** assesses the clarity, details, and resolution of the images and animations in the video. This characteristic ensures the video is clear and easy to view, without visual artifacts or distortions;
- **Relevance** of content ensures that the video's content is relevant to its description. It verifies whether the message and content suit the context in which the video was generated and will be used;
- **Consistency** follows the logical structure of the video. It ensures that the scenes and transitions are smooth and that the content unfolds logically and consistently;
- **Authenticity** verifies whether the video maintains an authentic appearance. This characteristic ensures that animated characters or voice generation does not appear forced or artificial;
- **User experience** evaluates the video's generated process experience. This includes response time and loading speed.

By assessing these characteristics, it is possible to establish which tool is most suitable for generating AI-driven video content for scientific purposes and whether it meets the quality, relevance, and daily use criteria.

The same description was used for all tests, requesting the generation of video content: "*A serene park with birds chirping and flying, with calming background music and text on the screen saying 'Exploring AI Video Generation.' After that, a computer screen with text saying 'Testing capabilities' shows a calming beach with waves gently crashing*."

Table 1 synthesizes the scores of the five characteristics for the five tools and presents the total score for each tool. This table provides an overview of how each tool performs across different aspects of video content generation, allowing for comparative analysis. Aggregating the scores for each characteristic offers insight into each tool's overall performance and suitability for generating AI-driven video content. The authors outlined a methodology for assessing the quality of AI-generated video content using human expertise:

- The five parameters identified by the authors were individually analyzed in sequence;
- The authors assigned a rating between 1 and 5 to each parameter, facilitating comparative analysis of the tools.

For the description text mentioned above, the authors generated video content for each AI tool tested. Additionally, the results will include private links to access the generated content, allowing for verification of the visual quality of the frames.

Table 1. AI video generator tests

Tool name	Visual quality	Relevance	Consistency	Authenticity	User experience	Total
Runway	2p	3p	1p	5p	3p	14p
Visla	4p	4p	3p	3p	4p	18p
Chromox	2p	3p	1p	5p	3p	14p
Genmo	3p	3p	1p	5p	4p	16p
Invideo.ai	4p	3p	5p	2p	5p	19p

Runway—The graphical output can range from cartoon-like to realistic styles. However, the objects often contain numerous artifacts and fail to interact correctly with one another. While it meets some specified requirements, it struggles with processing more complex tasks. Like other generators, it completely ignores the remaining provisions. Additionally, it does not generate multiple scenes within a single video. The demonstrative video generated by the Runway tool can be accessed at the private address https://youtu.be/iUuFrHQDoSg;

Visla - The video clips are realistic and generally meet the requirements, as seen in the demonstrative generated video https://youtu.be/MRSxgk78Xpo. However, some scenes may lack proper image focus or may not match the input precisely. The generator meets the requirements and is the only one that includes music and voiceover. It can generate

multiple scenes related to our request, though irrelevant scenes, such as those containing text, may also appear. Most materials used are sourced from various locations and are typically watermark-free. The tool is easy to use, allowing the input of an idea for which a ChatGPT API rewrites the script. The resulting video content will not be relevant if we input the generated script into our own ChatGPT. It loses one point because the export duration is very long for the free version despite having all these tools for requirement description. When presenting an idea to be rewritten by the AI, we do not need to provide instructions about on-screen text, specific sounds, or similar elements, as it does not accurately interpret these instructions;

Chromox - The scene from the input is distinguishable, indicating a "park" with birds, but the image quality is relatively poor, and many artifacts are present. The video generated by the AI tool can be accessed at a private address: https://youtu.be/CfP2RfyH78w. The generator meets some specified requirements but struggles with processing more complex tasks. It completely ignores or combines the remaining provisions, resulting in unsatisfactory output. It cannot create multiple scenes and has difficulty understanding the full scope of the requests, processing a maximum of one sentence. The interface is similar to ChatGPT, but the waiting time increases depending on the number of user requests. This tool is more suited for generating GIFs rather than complex videos;

Genmo - The graphics of the video content are anime-like, and objects are distinguishable despite some artifacts, as presented in generated content: https://youtu.be/hkwQrwDdWNc. The generator meets some specified requirements but struggles with processing more complex tasks. It either completely ignores or combines the remaining provisions, similar to Chromox. It does not generate multiple scenes. The format resembles ChatGPT, featuring minimal settings and a low waiting time;

Invideo.ai - The generated video clips are realistic and generally meet the requirements (https://youtu.be/X8bvuAAnUuM). However, some scenes may lack proper image focus or may not match the input precisely. The generator mostly meets the requirements and, along with Visla, is one of the few that includes music and voiceover. It can generate multiple scenes related to our request, though irrelevant scenes, such as those containing text, may also appear. Most of the materials used are sourced from iStock. The tool offers numerous settings, including script options, different output types, and music. However, the waiting time can be longer than that of the previously presented tools.

According to Table 1, Runway and Chromox both score lower overall, with 14 points each. While they exhibit acceptable visual quality and relevance, they struggle with consistency. Visla demonstrates solid performance across most categories, scoring 18 points. It excels particularly in visual quality and relevance, but consistency and authenticity could be improved. Genmo shows promising results with 16 points, boasting good visual quality and user experience. However, it falls short in consistency. Invideo.ai is the top performer with 19 points, excelling in consistency and user experience. It also maintains high visual quality and relevance scores, though authenticity is an area for potential enhancement. While each tool has strengths and weaknesses, Invideo.ai stands out as the most well-rounded option, offering a balance of quality, relevance, consistency, authenticity, and user experience.

Based on the total score obtained, Invideo.ai has achieved the best results. The next issue to be addressed concerns the originality of the content. To assess the quality of the content, frames were extracted from the video content and subsequently searched using Google Images. Figure 1 illustrates a frame extracted from the video content.

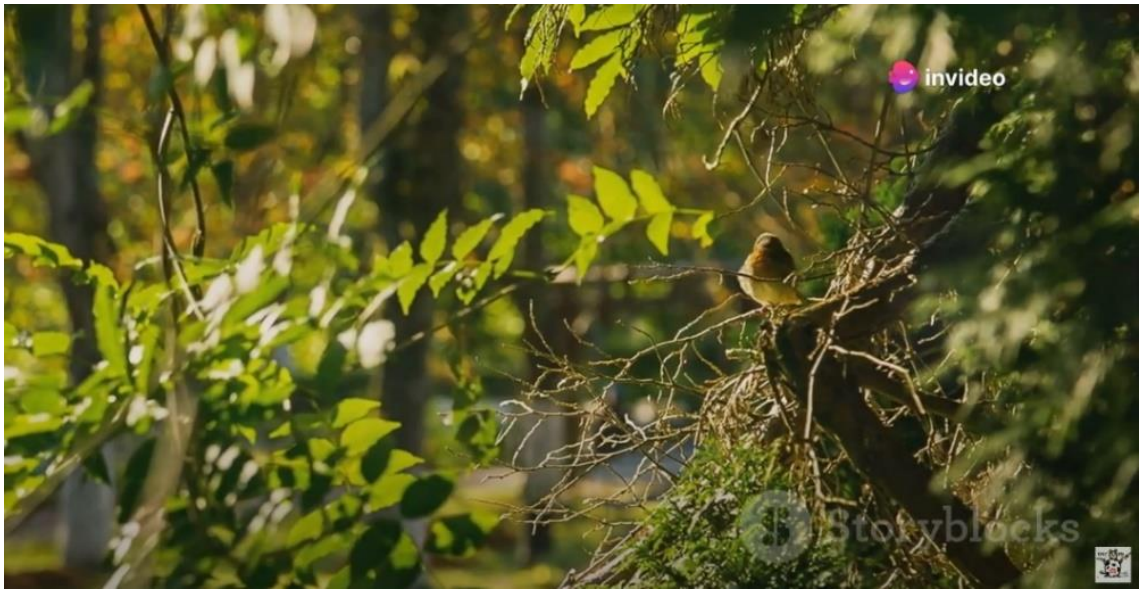

Figure 1. *Extracted frame from video content generated by Invideo AI service (source: www.invideo.ai)*

By conducting reverse image searches on extracted frames, it becomes possible to identify any potential plagiarism or unauthorized use of existing content. This process helps ensure the generated content is novel and does not infringe on intellectual property rights. Additionally, examining the extracted frames provides insights into the visual quality of the content. It allows for identifying any artifacts or distortions that may affect the viewing experience.

No relevant results were generated in Google Images to identify the frame in question. This negative outcome suggests that the content is original and has not been sourced from other visual materials online. This finding strengthens the credibility and authenticity of the content generated using the tool Invideo.ai.

It is important to note that the absence of results in Google Images does not exclude the possibility that the frame may be used in other contexts or present in sources not indexed by the search engine. However, the lack of results suggests that the frame is not part of an online recurrent or widely distributed image.

This evaluation stage confirms that the generated content is original and not directly associated with existing materials. Ensuring originality promotes copyright compliance and reinforces the integrity of the entire AI content generation process.

AI video generators provide a convenient solution for automating video content creation, but one of the common challenges they face is maintaining logical continuity between frames. The algorithm implemented by each AI tool, unknown to the authors, must ensure the logical continuation of motion and eliminate stuttering.

The transition between scenes is important in video storytelling, as it helps to guide the viewer through the narrative smoothly. However, AI algorithms often struggle to replicate human editors' understanding of context and pacing. As a result, the transitions generated by AI video generators may sometimes feel abrupt or disconnected, disrupting the narrative flow. Traditional editing techniques, such as cuts, fades, or wipes, can be replicated by AI, but the challenge lies in applying them to maintain coherence and enhance the viewer's understanding of the story.

AI video generators must improve user experience by customizing and fine-tuning scene transitions. This could include options to adjust the duration of transitions, select different transition effects, or manually override the algorithm's decisions.

In Figure 2, a frame is highlighted where the scene is generated incorrectly according to the initial description "generating emotion" suggesting that achieving performance identical to a custom-made video by a human professional may not be feasible in these AI-generated videos. This observation underscores the limitations of AI-generated content compared to human-created content. This discrepancy underlines the importance of understanding the capabilities and limitations of AI technology when using it for content creation purposes. Additionally, it emphasizes the continued need for human intervention and oversight to ensure the quality and accuracy of generated content, particularly in scenarios where precision and attention to detail are paramount.

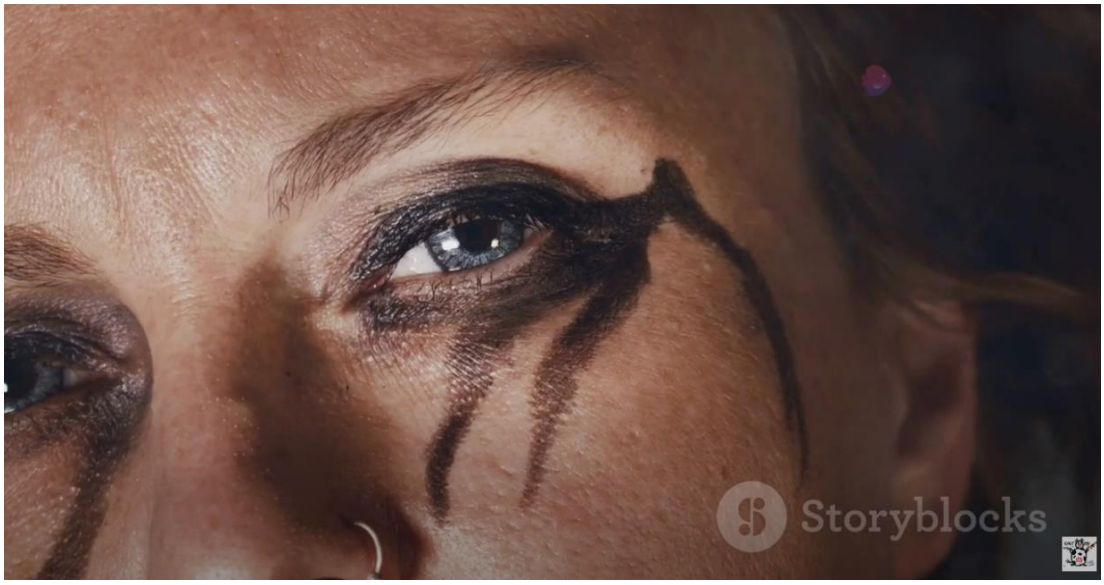

Figure 2. *Incorrectly generated object in the video content (source: www.invideo.ai)*

While AI tools offer automation in the video generation content process, they may not always match the quality and precision of content created by human professionals, especially in complex scenarios. Despite AI advancements, human intervention and oversight remain crucial to ensuring AI-generated content's quality, accuracy, and ethical use.

CONCLUSIONS

In this article, we illustrated video content across specific applications within all branches of AI, highlighting the significance of video content that must continuously align with the latest advancements in the field.

While AI-driven video content offers significant advantages, there are also instances of misuse, as exemplified by deep fake technology, which poses risks to authenticity and trust in digital media. Responsible advancement and deployment of AI technologies are crucial to effectively addressing these risks and upholding ethical and transparent utilization.

In the online environment, AI tools for video content generation and editing have been notable for promotion. However, this study exclusively analyzed the former category. Analyzing AI video generation tools reveals the importance of assessing visual quality, content relevance, consistency, authenticity, and user experience. By evaluating these characteristics proposed by the authors as a comparative measure between the analyzed tools, it becomes possible to identify the most suitable one for generating AI-driven video content, particularly for scientific purposes.

The study underscores the need for continued human oversight and intervention to maintain the quality of AI-generated content, especially in complex scenarios.

In summary, AI presents promising opportunities for video content generation, but it is essential to approach its utilization with critical evaluation, ethical considerations, and responsible practices. Through ongoing research and development, alongside ethical guidelines and responsible usage, AI-driven video content can contribute positively to various domains while upholding moral standards and user trust.

REFERENCES

[1] Gong Z., Evaluation of the Application of Artificial Intelligence Technology in Computers, 2nd International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI), Zakopane, Poland, 2023, pp. 114-118.

[2] Zhou X., Tennis Motion Precision Video Monitoring Simulation Based on Artificial Intelligence Algorithm, 2022 International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS), Bristol, United Kingdom, 2022, pp. 211-215.

[3] Lin J., Song J., Sun L., The Application of Artificial Intelligence Video Feedback System in Tennis Teaching in Colleges and Universities, 2020 International Conference on Artificial Intelligence and Education (ICAIE), Tianjin, China, 2020, pp. 28-31.

[4] Rakovac Bekeš E., Galzina V., Utilizing smart digital technology and artificial intelligence in education for transforming the way content is delivered, 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 573-578.

[5] Shin S., Jang J., Jung M., Kim J., Jung Y., Jung H., Construction of a machine learning dataset for multiple AI tasks using Korean commercial multimodal video clips, 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2020, pp. 1264-1266.

[6] Kurniawan Y., Jordan, Ongko J. G., Julianto, Anwar N., Public Interest Research on Digital Learning Content Produced by Artificial Intelligence, 2023 7th International Conference on New Media Studies (CONMEDIA), Bali, Indonesia, 2023, pp. 326-331.

[7] Lin H., Lin T., Research on the Application of Computer Artificial Intelligence Animation Technology in Short Video Production, 2024 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China, 2024, pp. 122-126.

[8] Song J., Deng J., Xin X., Wu X., Liu H., Exploration of Online and Offline Media Convergence Based on Artificial Intelligence Technology, 2021 IEEE/ACIS 19th International Conference on Computer and Information Science (ICIS), Shanghai, China, 2021, pp. 102-107.

[9] Darwich M., Khalil K., Ismail Y., Bayoumi M., Adaptive Video Streaming: An AI-Driven Approach Leveraging Cloud and Edge Computing, 2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings), Mount Pleasant, MI, USA, 2023, pp. 1-5.

[10] Sharp J., Exam Ref AI-900 Microsoft Azure AI Fundamentals, Pearson Education, 2021.

[11] Chadha A., Kumar V., Kashyap S., Gupta, M., Deepfake: An Overview. In: Singh, P.K., Wierzchoń, S.T., Tanwar, S., Ganzha, M., Rodrigues, J.J.P.C. (Eds) Proceedings of Second International Conference on Computing, Communications, and Cyber-Security. Lecture Notes in Networks and Systems, vol 203, pp 557–566, 2021, Springer, Singapore.

[12] Otopeanu C., Imaginea și vocea guvernatorului BNR, folosite într-o escrocherie financiară tip deepfake. Reacția Băncii Centrale, Libertatea, 2024. https://www.libertatea.ro/stiri/imagine-voce-guvernator-bnr-mugur-isarescu-escrocherie-financiara-deepfake-reactie-bnr-4795520

[13] Power AI, Video Generator AI Tools, 2024. https://www.powerfulai.tools/tool/category/video-generator

Received: May 2024; Revised: June 2024; Accepted: June 2024; Published: June 2024