AI in Creativity

Sami Kamal

Rutgers University

Abstract

Humans have long been termed a creative species capable of conceiving novel ideas that are useful and serve a specific purpose. Through these efforts, creativity has been sparked in technological devices through AI, leading to whether AI can augment creativity. Subsequently, AI has demonstrated a sustained competence owing to creativity that assists in strengthening human creativity. Human cognitive domains such as reasoning and decision-making are often subject to knowledge and computational capabilities. Conversely, AI portrays manifold creativeness in problem-solving, and functionalities are ever-improving through learning patterns. Consequently, AI in creativity is a mandatory topic of discussion in the contemporary world because it helps in understanding our bounded rationality and embracing alternative approaches through the adoption of AI creativity. This paper will explore the application of AI in creativity through computer-generated images and natural language processing to provide insights into understanding human creativity. This study seeks to answer research questions, including: What can AI teach us about how humans use creativity for natural language processing and computer-generated images, and what insight does it offer when studying human creativity? I will explore literature relating to text-to-image generation, AI writing bots, language models using GPT-3, a generative model for music, and image generation through CLIP Latents to address the research questions.

Keywords: Creativity, Language models, Human creativity, Natural language processing, Computer-generated images

AI in Creativity

1 Introduction.

1.1 Background

The advancement of technology has left scientists with an urge to define creativity in a contemporary sense due to the cutting-edge cognitive science applications invented and discovered. The need to express creativity in unexpected, novel, and valuable attributes places artificial intelligence in the spotlight. However, as noted by modern science, creativity began in 1950 following Gilford's paper "Creativity," which became acclaimed by the American Psychological Association (Still & d'Inverno, 2016). According to Guilford, creativity measures the psychological propensity of power. Subsequently, creativity was to describe individuals who possessed novel ideas that provided something of value, satisfying, tenable, or usefulness. Consequently, this definition became widely adopted in Artificial Intelligence and Psychology as encompassing the ability to generate valuable and novel ideas (Boden, 1998). The use of creativity in both fields emphasizes the old mystery of how the human mind could produce extraordinary achievements of "creative genius" (Still & d'Inverno, 2016). As a result, Artificial Intelligence and Psychology attempt to provide a scientific understanding of the diverse mechanisms that underly these achievements.

Furthermore, advancement in AI technology has highlighted the limits of human rationality, thus indicating that computers can be creative. Artificial intelligence algorithms used in developing new programs and data structures augment novelty in creativity (Gobet & Sala, 2019). Conceivably, AI illuminates new pathways for understanding human cognition in psychology. Whereas psychology evaluates creativity using simple tasks such as alternative uses tests, AI designs new classes of experiments that make creativity its moonshot. For instance, IBM's Watson comprehensively grasped natural language, thereby outclassing human players (Gobet & Sala, 2019). AlphaGo, developed by DeepMind, quickly transcended human cognition by outsmarting grandmasters in Go. In addition, automated assistants and self-driving cars continue to raise profound issues regarding the nature of human creativity and rationality. Generally, AI provides novel prospects for understanding human creativity by

developing original modules of experiments and theories of creativity. The use of AI in understanding creativity continues to gather momentum in modern science, emphasizing the approach of problem-solving while enlightening human creativity. There are several applications of AI in creativity in contemporary science ranging from robot science, color painting, and deep learning algorithm. These developments are essential in understanding the human cognitive process and approaches to advancing technological intelligence.

1.2 Statement of the Problem

Human creativity remains a controversial topic because of the inability of assessment criteria. Whereas psychology has taken up creativity, it cannot stimulate artistic approaches vital for facilitating human cognition. Conversely, AI application in invention promises a wide variety of applicability and the ability to promote human reasoning and thinking. Subsequently, human understanding can be enhanced through AI in problem-solving and the development of human creativity. Previous studies have demonstrated the importance of creativity in developing AI by emphasizing embracing new classes of experiments of the invention (Still & d'Inverno, 2016). The pace of AI advancement in creativity is astounding compared to the best human minds.

Recent studies in AI continue to demonstrate the inability of humans to understand or perceive the range of creativeness of AI in problem-solving (Gobet & Sala, 2019). Consequently, AI has sustainably outpaced human cognition when employed in practical gaming and deep learning programs. These developments provide an opportunity for exploiting human imagination that inevitably contributes to the development of novel ideas, thus enhancing human creativity. As a result, AI enables humans to communicate with machines through natural language processing, thus exhibiting unique intelligence from the human. In addition, AI can offer a vital understanding of how humans utilize artistic and visual creativity by developing computer-generated images. These applications help acquire a deeper understanding of human creativity through the new experimental classes of creativity and theory.

1.3 Significance of the Study

Technological evolution has offered a plurality of novel tools that facilitate creative functionalities. Subsequently, the application of AI can assist humans in understanding the use of creativity for computer-generated images and natural language processing. The recent advent of AI through deep learning techniques has contributed to success in various applications. Conceivably, the use of these technologies for creative purposes manifests in natural continuity embraced within the artistic trend of this era. In addition, this study explores the use of AI in studying human creativity. The study of AI in creativity transcends the traditional psychological approaches utilizing alternative services task or insight problems, as illustrated by previous studies on human creativity (Gobet & Sala, 2019). In addition, the study offers an opportunity to understand human creativity and cognitive development. AI gives human creativeness more options by enlarging insights into their work, mainly through the awareness of audience engagement.

Whereas human creativeness might be challenged by limited computational and knowledge capacities, thus constraining their decision-making processes, AI can learn and continuously perfect its ingenuity. As a result, humans may rely on AI creativity because of the bounded reality theory that beats the best of us (Gobet & Sala, 2019). Since AI creativity is unlimited compared to humans, it is essential to understand the creative aspects of AI to increase human creativity (Miroshnichenko, 2018). Subsequently, AI enables designers to construct more comprehensive designs by facilitating their thought processes and creative options. This study challenges creation possibilities and offers a deep understanding of creativity. As a result, human creativity will be enhanced to understand AI processes such as computer-generated images and natural language processing.

2. Literature Review.

2.1 Text-to-Image Generation.

Ramesh et al. study on text-to-image generation provide insights into approaches to determining improved modeling expectations for training based on a secure dataset (2021). The authors discovered unprecedented outcomes where the model demonstrated the ability of generalization in ways that they did not anticipate. The study, therefore, suggests that the model has developed an elementary ability entailing

composing unusual concepts through heightened abstraction levels. In addition, the researchers determined that the model able to undertake zero-shot image-to-image transformation using natural language. The study found that the model worked with multiple modifications such as image operations, including grayscale conversion and image color transformation. Moreover, style transfer involving drawing on a postage stamp, greeting card, and cellphone case was realizable. Ramesh et al. found that the model could perform the primary item segmentation in some transformations, such as transforming the object's color (2021).

The study aimed to train a transformer to autoregressively model the image and text tokens as a sole data stream. Custom text-to-image generation was centered on developing better model conventions involving complex architectures and side information. Conversely, the study aimed to demonstrate the capability of a single generative model. The researchers created and developed a dataset of the same scale as JFT-300M by gathering 250 million text-image pairs via the internet (Ramesh et al., 2021). The datasets comprised conceptual captions, YFCC100M filtered subset, and text-image pairs from Wikipedia. The researchers used a subset of the image, text, and joint image and text filters to construct the dataset.

The study's main findings were that scale can improve generalization concerning the range of abilities emerging from a sole generative model and regarding zero-shot performance relating to the previous domain-specific techniques. The study's strength involved using a qualitative comparison of samples of the research model to those of previous work. In addition, the study used a manual selection design to reduce the rate of false-negative images. However, the study included findings with low reliability and failed to mention the need for further studies on the subject area.

2.2 AI Writing Bots

Tatalovic's study suggested how AI provides insights into human creativity by revolutionizing journalism (2018). The study found that several journals apply AI algorithms to summarize scientific research papers, turning them into important news stories and press releases. Mainstream media organizations such as Washington Post and Associated Press were found to have increased creativity by

applying AI bots. Conceivably, Tatalovic found that AI and algorithms are progressively playing a significant role in all sectors of journalism and society. The researcher found that AI creativity could summarize complex text and learn new ways of improving, thus posing a potential new business model.

The researcher's purpose was to explore creative AI bots with a high propensity to shape science journalism. Tatalovic examines how AI and algorithms have a creative dimension that could significantly change writing news stories, thus necessitating algorithmic accountability. The study used archival data and online news articles to collect data. Tatalovic's significant finding was a discovery that the society was fast moving towards AI-assisted content generation where all research papers would feature an automatically-generated news story and press release. The study's strengths include providing an in-depth understanding of AI creativity in a neglected field and utilizing relevant data sources to undertake a explore exploratory case study in the field. However, researcher bias is not addressed in the study.

2.3 Language Models.

Brown et al. study on language models found that GPT-3 is proficient in generating news article samples that are challenging for human evaluators to distinguish from human writings. GPT-3 demonstrated high performance on NLP datasets, such as undertaking 3-digit arithmetic and unscrambling words. The study aimed to illustrate how scaling up language models significantly improves few-shot performance and task-agnostic, achieving competitiveness with previous fine-tuning techniques.

The procedure for data collection used in this study included a pre-training approach involving training, data, and a model. The researchers then used GPT-2 architecture and models such as pre-normalization, reversible tokenization, and modified initialization. Subsequently, Brown et al. trained datasets by downloading and filtering a version of CommonCrawl according to the similarity of a range of high-quality corpora (2020). Next, the researchers undertook fuzzy deduplication within and across datasets at a document level. The final training step involved adding high-quality reference corpora to increase diversity and augment CommonCrawl.

The study's primary finding was the observation that large language models might be a critical ingredient in developing adaptable general language systems. The researchers strengthened the

examination by using task-specific fine-tuning and task-specific datasets, thus attaining high performance of desired tasks. Conversely, the study demonstrated weaknesses through GPT-3, which has a poor text synthesis than GPT-2. In addition, the use of GPT-3 presented algorithmic and structural limitations to the study. This study retained biases in data trained, leading the model to generate prejudiced or stereotyped content (Brown et al., 2020). Conceivably, shared language models present a poor sample during pre-training, thus weakening the study.

2.4 Generative model for music.

Dhariwal et al. study on the jukebox as having the capacity to capture creative processes found that the combined jukebox model at scale could produce a variety of high-fidelity songs with a multiple minute coherence (2020). The study found that the model used could generate songs from diverse genres such as jazz, rock, and hip-hop. The research aimed to investigate the creative processing power of AI through the jukebox's capability to generate raw audio music by imitating various artists and styles. Consequently, the study presents a creative AI by applying generative models in music generation tasks to challenge the non-symbolic approach of producing music as an audio piece.

The data collection procedure used in the study involved an experiment with a 1.2 million songs dataset that was paired alongside metadata and lyrics from LyricWiki (Dhariwal et al., 2020). Subsequently, the researchers performed training on 32-bit, 44.1 kHz raw audio while augmenting the dataset through a random downmixing of the left and right channels to provide mono channel audio. The significant finding in the study showed the generative model is capable of conditioning genre and artists to steer the vocal and musical style, including unaligned lyrics, thus controlling singing. The study strengths are derived from VQ-VAE architecture that accommodates audio compression at high levels and a loss function that facilitates retaining the maximum amount of musical information. However, the approach does not support melodies, choruses, and long-term musical patterns since the top-level lack the entire song's context.

2.5 Image Classification through Generative Pretraining.

Chen et al. found that GPT-2 model is capable of learning strong image presentations when measured through low data, fine-tuning, and linear probing classification (2020). Moreover, the research suggests that a generative image model can produce high-quality unsupervised image presentations by predicting pixels. The researchers purposed to investigate whether generative models are capable of learning valuable representation for images as depicted by the use of natural language. As a result, Chen et al. examined two training objectives, including denoising objective and auto-regressive prediction similar to contemporary neural sequence models (2020). The procedure for collecting data involved a pre-training phase that was subsequently tracked by fine-tuning steps. In the pre-training, the researchers explored both BERT and auto-regressive objectives. Furthermore, sequence transformer architecture was used in predicting pixels in place of language tokens.

The study's primary outcomes included the observation that similar to unsupervised representation for natural language learning, a creative dimension of AI facilitates the valuable representation of images. Conceivably, predicting pixels enables learning state-of-the-art representations in low-resolution datasets. The study's approach enhances the reliability of the findings because of its competitiveness with various self-supervised results on Image Net. However, the study indicated weakness through self-attention to model low resolution instead of CNN encoders.

2.6 Image Generation with CLIP Latents.

Ramesh et al. found that image representation through generative models enhances its diversity with minor alterations in the caption similarity and photorealism (2022). In addition, the study found that language-guided image adjustments in the context of zero-shot are reinforced through the CLIP's joint-embedding space. The researchers determined that guiding improved the aesthetic quality for unCLIP and GLIDE by inducing a trade-off between diversity and fidelity.

The study purposed to leverage the representation of images comprising of style and semantics by a contrastive model such as CLIP. To achieve the goal, the researchers performed a data collection method entailing generating a CLIP image embedding a specific text caption and processing the image embedding through a decoder. Consequently, the authors were able to come up with a primal finding that generating

image representations enhance its diversity with minimal alterations. The researchers used two approaches, including training diffusion models, to develop a creative AI capable of solving text-conditional image generation. On the other hand, the approach supports unCLIP, which is poor at binding features to objects compared to the GLIDE model.

3. Discussion.

Human creativity has been stretched through the advent of artificial intelligence, attempting to beat humans in novelty. Whereas traditional forms of AI creativity have majored in exploratory approaches, modern scientific techniques have necessitated the emergence of transformational AI applications that gear toward facilitating creativity. The recent demonstrations of AI capabilities in self-driving, the IBM Watson project, and the gaming AI programs persuade humans that AI creativity is valuable. Furthermore, the value of AI creativity has been shown through the ability of the models to learn and advance their skills, knowledge, and creativity in unprecedented dimensions. Conceivably, most creative AI has progressively shown a distinct feature of scalability using diverse applications and approaches. AI has led to a hurtful revelation concerning human creativity by unveiling its limits through applications in gaming programs such as AlphaGo. Consequently, AI creativity is a clarion call to human cognition, challenging their reasoning and novelty in creativity. Over the years, natural language processing and computer-generated images have facilitated human creativity.

Many documented gains in natural language processing tasks have allowed fine-tuning of given tasks to develop creative models. Subsequently, GPT-3 models can generate news articles that are challenging for humans to distinguish. Similarly, studies involving writing bots suggest that journalism is advancing towards using AI algorithms that can automatically convert various research articles into news stories. Conceivably, human creativity is challenged by these AI applications that indicate the limits of human cognition can be advanced by creativity in AI. The future for AI in creativity is unchallenged due to the accuracy that technology such as writing bots and AI algorithms has indicated.

Moreover, text-to-image generation provides insight into human creativity through contemporary machine learning techniques. Generative novel visual scenes utilizing machine learning approaches are

gradually improving to move past the performance of elementary tasks. Apart from using one design for generating an image, researchers have determined that pixels can be a better alternative in demonstrating the creativity of AI. Generative pre-training models for images are showing a significant improvement in the state-of-the-art low-resolution representation learning settings. Consequently, generative image modeling provides a promising framework for learning high-quality image representation that does not require supervision. Similarly, using diffusion models that leverage guiding approaches, text-conditional image generation has contributed to improved fidelity, thereby unveiling state-of-the-art video and image generation tasks.

Previous and current studies are in a consensus regarding the state of AI creativity, where both point toward the potential for the technology to move past the human creativity limitations. Generally, there is a perception of unsatisfactory among researchers who boldly claim that AI can achieve higher performance. The majority of the studies on AI creatives have equally expressed similar sentiments emphasizing the need for further development of the technology and merging with creative ideas to bring forth the invincibility of creativity.

References

Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial Intelligence*, *103*(1-2), 347-356. https://doi.org/10.1016/S0004-3702(98)00055-1

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877-1901.

Chen, M., Radford, A., Child, R., Wu, J., Jun, H., Dhariwal, P., Luan, D., & Sutskever, I. (2020, November). Generative pre-training from pixels. In *International Conference on Machine Learning* (pp. 1691-1703). PMLR.

Dhariwal, P., Jun, H., Payne, C., Kim, J. W., Radford, A., & Sutskever, I. (2020). Jukebox: A generative model for music. *arXiv:2005.00341*. 1-20. https://doi.org/10.48550/arXiv.2005.00341

Gobet, F., & Sala, G. (2019). How artificial intelligence can help us understand human creativity. *Frontiers in Psychology*, *10*(1401), 1-6. https://doi.org/10.3389/fpsyg.2019.01401

Miroshnichenko, A. (2018). AI to bypass creativity. Will robots replace journalists? (the answer is "yes"). *Information*, *9*(7), 183. https://doi.org/10.3390/info9070183

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditional image generation with clip latents. *arXiv:2204.06125*.

Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., & Sutskever, I. (2021, July). Zero-shot text-to-image generation. In *International Conference on Machine Learning* (pp. 8821-8831). PMLR.

Still, A., & d'Inverno, M. (2016). A history of creativity for future AI research. In *Proceedings of the Seventh International Conference on Computational Creativity,* 6, 147-154.

Tatalovic, M. (2018). AI writing bots are about to revolutionize science journalism: we must shape how this is done. *Journal of Science Communication*, *17*(1), 1-7. https://doi.org/10.22323/2.17010501