

Analyzing Copyright Infringement by Artificial Intelligence: The Case of the Diffusion Model

Shumin Wang^{1,a,*}

¹School of Civil and Commercial Law, Northwest University of Political Science and Law, Xi'an, China

^aShuminwang2021@outlook.com

*Corresponding author

Abstract: As the use of artificial intelligence technology, particularly deep learning models like the diffusion model, becomes more prevalent in the creation of art, concerns regarding the use of unauthorized work samples in their training have emerged. The lack of supervision and relevant laws has contributed to the problem. This study examines the potential infringement issues that may arise during the training of the diffusion model and explores the legality of using unauthorized samples for deep learning model training. While some scholars argue that copyright law only protects expression and not painting style, therefore, using unauthorized works in model training is not considered infringement, we propose a different viewpoint. By considering the essence of artificial intelligence from an information theory perspective, we highlight that it is still a deterministic algorithm and that data processing does not bring about an increase in information entropy without the input of additional information. Thus, the painting created by the diffusion model is essentially a mash-up of paintings in its training space, and as such, it is a copy or adaptation of the original work and should be licensed by the creator. We highlight the critical difference between human learning and AI "learning," emphasizing the need for effective protection and encouragement of human artistic innovation while embracing the wave of AI.

Keywords: Copyright Law; Infringement; Artificial intelligence; Information theory

1. Introduction

Artificial intelligence has made significant progress in recent years and is now applied in almost every field [1][2]. Legal risks associated with artificial intelligence were first identified by Solum as early as 1991 [3]. However, it was not until the emergence of the diffusion model and large-scale language models that the legal issues surrounding artificial intelligence received widespread attention. These neural networks possess strong creative abilities, which pose a threat to the interests of creators. This paper uses the diffusion model as an example to analyze the potential infringement issues associated with it.



Figure 1: Image generated by stable diffusion model. The input statements are "The flowers are in full bloom on the grass" and "A flying bird in dark sky".

The diffusion model is a neural network model that converts an image into pure Gaussian noise by adding Gaussian noise continuously during the forward process. In the generation process, the model can restore pure Gaussian noise to an image. The diffusion model's impressive creation abilities enable it to draw a picture conforming to any text description based on previous training data (See Fig.1). However, the high performance of artificial intelligence is reliant on extensive training data. To perform well, the diffusion model must use a considerable amount of data, which may lead to hidden infringement risks for the training data [4]. Carlini's research reveals that the diffusion model imitates the painting style of the training data, and may even directly copy it [5]. Identifying and defining the infringement act of the diffusion model, as well as assessing whether existing laws can address it, are crucial issues that have yet to be resolved.

2. Potential Infringement Risk in Diffusion Model

The training of the diffusion model requires a vast amount of data samples, and these samples' sources are often complex, categorized into three primary aspects: (1) existing databases; (2) web crawling technology to collect data from the internet [6]; and (3) digitization of non-digital samples via scanning or photography. Here, we will only discuss the first two sources of data, as the third collection mode is often costly at large data scales.

Legal risks associated with using existing datasets primarily involve violations of public dataset licenses. Some licenses, such as the Creative Commons Attribution-NonCommercial-ShareAlike Unported License [7] adopted by the PhosphoSite dataset [8], explicitly prohibit commercial activities. However, several public datasets such as Cifar10 [9] do not indicate the license, exposing hidden legal risks. Crawler images, on the other hand, are the primary offenders. The founder of Midjourney, a well-known company, has personally admitted that his company used hundreds of millions of unauthorized images to train image generators. The internet is inherently weak in terms of copyright [10], as copyright holders cannot determine whether their images have been posted online unless they are explicitly informed. However, the advent of the diffusion model has expanded the harms of such infringement by training on numerous unlicensed images that enable the model to imitate a creator's painting style without their permission [11]. Nonetheless, whether the imitation of painting style by a deep learning model constitutes infringement remains a topic of controversy.

3. Infringement or Non-infringement? When Public Paintings are Used for Training without Permission

A crucial question arises: does using a creator's painting without permission to train a diffusion model to imitate their style constitute copyright infringement? Currently, there is no legal regulation on the use of data for model training in the field of AI or machine learning in China, Europe, or the United States. Therefore, the issue of determining whether the unauthorized use of public works for training constitutes infringement should be evaluated based on the provisions of copyright law.

According to China's Copyright Law, copyright owners have the right to publication, authorship, modification, integrity protection, reproduction, distribution, and adaptation of their works [12]. The object of protection under copyright law is original intellectual achievements expressed in certain forms in the fields of literature, art, and science. Copyright law protects expressions, not ideas [13]. From this perspective, using published paintings for machine learning training appears reasonable. Copyright owners have the right to prohibit others from exploiting their works for profit, but not the right to prevent their appreciation and learning. Therefore, using published works in the diffusion model's training does not infringe the author's copyright and does not require authorization or permission.

However, from the essence of the diffusion model, the most complex neural network is also an algorithm with deterministic properties. From the perspective of information theory, the processing of the diffusion model is a type of information processing process, and the information entropy will not increase in the process of data processing without the input of additional information [14]. Therefore, the artificial intelligence model without the ability to perceive the outside world does not have an increase in information entropy in the process of training and generating works, so it is not original. All of its generated works are mapped from the works of the existing training set in a nonlinear way. Hence, the "learning" of the diffusion model and artificial intelligence is different from that of human beings. Human learning combines observation and practice of the natural world, resulting in additional input of information. This is the source of human innovation. AI learning, on the other hand, is a nonlinear

combination. It learns how to select an appropriate combination that can mix different works together and still make sense from the perspective of human cognition. Therefore, when the diffusion model uses unauthorized paintings to learn, it infringes on the author's right to copy and adaptation.

4. The Future of Copyright Law: Embrace the AI Revolution

The rapid advancement of AI technology has presented novel legal challenges and issues. The use of generative models, such as the diffusion model, has resulted in a significant impact on the creative output of human artists. From a judicial standpoint, the high-dimensionality of machine learning makes it more difficult to detect and prove instances of copyright infringement. The infringement of works by machine learning models can be implicit and harmful [15], as they imitate the unique style of creators without permission, potentially reducing their audience and economic opportunities. The black box nature of machine learning further complicates the ability to identify the source of the generated work.

While AI drawing may appear to offer infinite possibilities for human art creation, it ultimately limits artists' ability to explore and innovate in the creative process. The space for AI art creation is confined to the creative style included in its training data, whereas human art has an infinite range of inspirations, emotions, and experiences to draw upon. We can feel the wind, listen to the rain, smell on the flowers, look at the birds, smile in the love and cry in the pain. All these are the source of our artistic creation and the inexhaustible spring of our innovation. Therefore, it is crucial for copyright law to protect the value of human innovation and creativity by clearly defining infringement behaviors and standards in the training of artistic works by AI. This would not only encourage human innovation but also foster new artistic expressions and styles.

References

- [1] van Dis E. A., Bollen J., Zuidema W., van Rooij R., & Bockting C. L. (2023). *ChatGPT: five priorities for research*. *Nature*, 614(7947), 224-226.
- [2] Li Q., Cai W., Wang X., Zhou Y., Feng D. D., & Chen M. (2014, December). *Medical image classification with convolutional neural network*. In *2014 13th international conference on control automation robotics & vision (ICARCV)* (pp. 844-848). IEEE.
- [3] Solum L. B. (1991). *Legal personhood for artificial intelligences*. *NCL Rev.*, 70, 1231.
- [4] Naqvi Z. (2020). *Artificial intelligence, copyright, and copyright infringement*. *Marq. Intell. Prop. L. Rev.*, 24, 15.
- [5] Carlini N., Hayes J., Nasr M., Jagielski M., Sehwag V., Tramèr F., & Wallace E. (2023). *Extracting training data from diffusion models*. *arXiv preprint arXiv:2301.13188*.
- [6] Thomas D. R., Pastrana S., Hutchings A., Clayton R., & Beresford A. R. (2017, November). *Ethical issues in research using datasets of illicit origin*. In *Proceedings of the 2017 Internet Measurement Conference* (pp. 445-462).
- [7] Smith L. N., & Chambers R. *Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Public License*.
- [8] Hornbeck P. V., Chabra I., Kornhauser J. M., Skrzypek E., & Zhang B. (2004). *PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation*. *Proteomics*, 4(6), 1551-1561.
- [9] Krizhevsky A., & Hinton G. (2009). *Learning multiple layers of features from tiny images*.
- [10] Schlachter E. (1997). *The intellectual property renaissance in cyberspace: why copyright law could be unimportant on the Internet*. *Berkeley Tech. LJ*, 12, 15.
- [11] Ruipérez C., Gutiérrez E., Puente C., & Olivás J. A. (2017). *New Challenges of Copyright Authorship in AI*. In *Proceedings on the International Conference on Artificial Intelligence (ICAI)* (pp. 291-296). *The Steering Committee of the World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp)*.
- [12] Lazar J. C. (1995). *Protecting ideas and ideals: Copyright law in the People's Republic of China*. *Law & Pol'y Int'l Bus.*, 27, 1185.
- [13] Jones, R. H. (1990). *The myth of the idea/expression dichotomy in copyright law*. *Pace L. Rev.*, 10, 551.
- [14] Kullback S. (1997). *Information theory and statistics*. *Courier Corporation*.
- [15] Gilotte J. L. (2019). *Copyright infringement in ai-generated artworks*. *UC Davis L. Rev.*, 53, 2655