Deep Synthesis Works Compliance Governance and Copyright Protection: Chinese Perspective and Practices

Hang Liu*

Law School, Beijing Normal University, Beijing 100875, China

*Corresponding author: Hang Liu, liuhang@bnu.edu.cn

Abstract: Deep synthesis works are creative works produced by applying deep synthesis technology. Depending on whether a performer’s image is entirely generated by deep synthesis technology, deep synthesis works can be classified as native synthesized works or regenerated synthesized works. The major legal issues of the deep synthesis works include the classification of the works, copyrightability, rights competition, and potential rights conflicts. This article argues that the originality of deep synthesis works is determined by objective standards. In the event of concurrence of rights, the performer should choose the most appropriate way to protect their rights in specific circumstances. Furthermore, conflicts can be resolved by clarifying the boundaries of the rights of the work. Deep synthesis technology and its application shall be regulated to reduce its negative effects, so as to transform “deepfake” to “deep synthesis” and maximize the benefit of this technology.

Keywords: Deep fake; Compliance governance; Copyrightability; Portraiture right; Performers’ rights

Online publication: December 27, 2023

1. Introduction

Deep synthesis technology is an image synthesis technology based on artificial intelligence that utilizes technologies like Generative Adversarial Networks (GAN), Recurrent Neural Networks (RNN), and Autoencoders (AE) to generate highly realistic images and human voices [1]. The precursor to deep synthesis technology is deepfake technology. Deepfake technology is often involved in alleged violations of reputational and portrait rights, such as the unauthorized use of other people’s portraits in adult videos via “AI face-swapping.” While prohibiting the use of deep learning and virtual reality technologies to engage in illegal activities, the Provisions on the Governance of the Online Information Content Ecosystem promulgated by the Cyberspace Administration of China also expressed concern and encouragement regarding technological developments. Distinguishing the transition of deep synthesis from deepfake is the result of incorporating emerging technologies into regulatory frameworks [1]. Works produced by using deep synthesis technology
belong to AI-generated content (AIGC). Academic discussions on AIGC, including works and creations, have focused mainly on copyrightability and copyright ownership. Thus far, there has been a general consensus that raises the question of whether AI-generated content qualifies as a work under objective standards. An output is considered a work if it is unique [2]. At the same time, before artificial intelligence’s status as a civil subject is confirmed, the copyright of AI-generated content belongs to its owner [3]. AI developers also agree on this form of copyright ownership [4]. Compared to the legal attributes and ownership rights of conventional AI-generated content, copyright issues in deep synthesis works are more complex. Discussions on the legal entitlement arising from deep synthesis works could influence the evolution of the laws that protect AI-generated content [2].

2. Regulatory framework of deep synthesis technology

A set of clear legal rules is necessary for the “rehabilitation” of deep synthesis technology. Currently, the laws that are involved in regulating this technology are the Civil Code, the Regulations on the Administration of Network Audiovisual Program Services, and the Provisions on the Governance of the Online Information Content Ecosystem. Among them, Article 1019 of the Civil Code prohibits the violation of portrait rights using deep synthesis technology. Article 11 of the Regulations on the Network Audiovisual Program Services stipulates the identification obligations of network audiovisual information service providers and users. It also prohibits the fabrication of news information using deep synthesis technology. Article 23 of the Provisions on the Online Information Content Ecosystem prohibits the use of deep synthesis technology to engage in illegal activities. However, these above provisions and regulations are still too general to cover all scenarios of deep synthesis technology misuse, providing guidance for technological developments. An improved regulatory model for deep synthesis technology should be developed from the following three aspects: scenario-based awareness, protection of legal interests, and cross-field synergy.

2.1. Scenario-based awareness

Scenario-based awareness means adopting differentiated governing methods and a combination of lenient and strict regulatory concepts according to the different application scenarios of deep synthesis technology. In scenarios where the authenticity of information is extremely important, such as political activities and news reporting, the use of deep synthesis technology should be prohibited to ensure the authenticity of information. In application scenarios such as film and television entertainment, communications and social networking, education, and healthcare, the application of deep synthesis technology should fully be based upon the stakeholder’s consent.

2.1.1. Political activities

There have not yet been incidents of deep synthesis technology applications relating to political activities in China. In other countries, the use of deep synthesis technology to interfere with political activities might exist. For example, former US President Trump once posted a forged video on Twitter using deep synthesis technology. In the video, when Trump honored McGee, an honorary veteran of World War II, the Speaker of the House Pelosi tore up the State of the Union address [5]. The deepfakes of such political activities, especially those of national leaders, do not only infringe on individuals’ civil rights but may also have an impact on national security and social public interests [6]. Therefore, the application of deep synthesis technology in political activities should be prohibited.
2.1.2. News reporting

News reporting usually bears the stamp of authority in China. If deep synthesis technology is allowed to fabricate news, it will disrupt the social order of the nation and severely undermine the government’s credibility [7]. Therefore, the Regulations on Network Audiovisual Program Services clearly stipulate the prohibition of using deep synthesis technology to fabricate news. It should be noted that whether it is in political activities or news reporting, the legitimate reason for prohibiting deep synthesis technology lies in preserving the authenticity of content. This technology is allowed provided that the authenticity of the content is not affected [8]. For example, virtual anchors generated by deep synthesis technology can report news 24/7, freeing news anchors from arduous physical labor [9], while ensuring visual quality.

2.1.3. Film and television entertainment

Deep synthesis technology is currently most extensively used in film and television entertainment. For example, in *Furious 7* released in 2015, the film crew used deep synthesis and CGI motion capture to recreate the on-screen image of the late actor Paul Walker [10]. Similarly, *Rogue One: A Star Wars Story* in 2016 and *Star Wars: The Last Jedi* in 2017 also used deep synthesis technology. In addition to applications in audiovisual works, deep synthesis technology can also enhance the gaming experience for players. For example, Tencent YouTu's DittoGAN technology can create avatars of the players through deep synthesis, providing an immersive experience for players. In general, applications in film and television entertainment do not involve national security and public interests. Therefore, as long as the generated portraits are authorized, such application is lawful and shall be permitted.

2.2. Lawful interests protection

Lawful interests refer to interests protected by law before they are formally recognized as legal rights. Protecting lawful interests requires targeted measures and special sanctions according to different types of interests that may be infringed upon by deepfakes.

2.2.1. National security

According to the ASI Science Data Centre (ASDC), a two-hour video of former US President Trump declaring war on foreign countries can be created using deep synthesis technology [11]. If application as such is not regulated, it will amount to a serious national security menace. In addition to forging videos of political figures in speeches, deep synthesis technology may also be used to interfere with elections and other illegal activities. Serious criminal penalties should be applied for regulatory purposes if deep synthesis technology violates legitimate interests, including national security.

2.2.2. Public order

The use of deep synthesis technology should not disturb public order. For example, deep synthesis technology may be abused to fabricate evidence in judicial proceedings. In the British TV series, *The Capture*, the protagonist Shaun Emery was imprisoned due to a forged video using deepfake technology [3]. Although this plot is fictional, the potential abusive practices do exist. The abuse of deep synthesis technology may make it difficult to determine the evidentiary value of video evidence in trials, thus undermining judicial efficiency.

2.2.3. Civil rights

The materials required for deep synthesis technology come from physical persons’ biometric information, mainly portraits. In practice, in almost all scenarios where deep synthesis technology is abused, physical
persons’ civil rights are infringed. In the case where individuals’ civil rights are infringed, tort liability is applicable. For example, using individuals’ portraits to produce audiovisual works without permission constitutes an infringement of portrait rights. If the produced audiovisual work is an adult film, it also infringes the victims’ reputation rights. The key requirement in applying the deep synthesis technology is the prior awareness and consent of the person involved.

2.3. Cross-field synergy

From the regulatory perspective, deep synthesis technology is constrained by the synergistic effects of four factors: market, architecture, community norms, and laws \[^{12}\]. Among them, the diverse market demand is the main incentive mechanism promoting the development of deep synthesis technology. The factor of “architecture” refers to technical means that may affect the person subjected to regulation. In the context of deep synthesis technology governance, the architecture refers to its detection technology and tracking technology. Community norms refer to artificial intelligence codes of ethics that regulate deep synthesis technology, while laws are mandatory norms that regulate deep synthesis technology.

2.3.1. Technology as regulatory means

Deepfake video detection technology is an internal solution to technology abuse. Mainstream deepfake video detection technologies are divided into frame-based detection methods and video integrity-based detection methods. The former uses the discontinuities in the video’s spatiotemporal state caused by “visual artifacts” and “visual noise” in each frame of deepfakes to detect the falsity. The latter embeds watermarks into the original video and use data integrity verification and other means to trace the source of the video and detect tampering \[^{13}\]. The detection technologies mentioned could curb the abuse of deep synthesis technology to a certain extent. Detection technology can identify deepfake videos and thus ensure the reliability of video content. However, the development of deep synthesis technology far outpaces that of its detection technology \[^{14}\]. Therefore, other regulatory means are required to fill in the gap.

2.3.2. Ethical guidance

Technology is essentially value-neutral, so the application of technology needs ethical guidance. Although ethical norms do not have the same legally binding force as laws, they have value-orientation and self-restraint functions. On November 25, 2021, the United Nations Educational, Scientific and Cultural Organization (UNESCO) adopted the Recommendation on the Ethics of Artificial Intelligence (hereinafter referred to as the “Recommendation”), calling on member states and all stakeholders to take feasible steps to implement its policy recommendations. In the name of “respect, protection, and fulfillment of human rights and fundamental freedoms as well as human dignity,” the Recommendation proposes the basic principle of attributing ethical and legal responsibilities to natural persons or existing legal entities, in the application of artificial intelligence at any stage. When deep synthesis technology runs against ethics and violates the law, its users or developers should bear the responsibility.

2.3.3. Legal regulations

Among various factors that produce synergistic effects, the law is the most direct and powerful regulatory tool. The legal governance of deep synthesis technology can be broadly divided into the US model and the European Union model \[^{16}\]. The former has introduced a series of special laws to regulate deep synthesis technology. For example, Virginia’s “Unlawful Dissemination or Sale of Images of Another” clause stipulates that disseminating deepfake videos without permission constitutes a crime. California has also passed a bill “Elections: Deceptive
Audio or Visual Media” to regulate deep synthesis technology. The EU model uses existing laws such as the General Data Protection Regulation (GDPR) to regulate deepfakes by targeting the “abuse of personal biometric information.” The governance model adopted in China is a combination of the above two models: it prohibits the use of deep synthesis technology to infringe on individuals’ portrait rights protected by the Civil Code; at the same time, separate normative documents have been formulated to prohibit the abuse of deep synthesis technology. Since existing regulations are broad and some of them lack operability, the legislature should take further improvement steps. Legislative improvement measures should take into account China’s specific conditions, as well as learn from the experiences and lessons of foreign countries.

3. Copyright substantiation of deep synthesis works

Deep synthesis technology has wide applications. Deep synthesis technology is mainly applied in the creation of audiovisual works. The manner of production affects the definition and classification of works produced by deep synthesis technology. Audiovisual works include cinematographic works and works created by cinematographic methods. The characteristics of audiovisual works are highly compatible with the features of deep synthesis technology: audiovisual works consist of a series of related images with or without accompanying sound, or with a related series of images corresponding to deep synthesis videos, or with their accompanying sound that corresponds to deep synthesis audio. Therefore, the copyright-related issue of deep synthesis technology mainly concerns audiovisual works. Audiovisual works using deep synthesis technology can be qualified as “deep synthesis works.” Copyright law closely follows technological development\(^{[17]}\). New legal problems will arise due to the application of deep synthesis technology. From the perspective of the nature of works, the new problems revolve around two main issues: work types and copyrightability.

3.1. Analysis of deep synthesis work types

3.1.1. Basis for deep synthesis works categorization

Deep synthesis works can be divided into native synthesized works and regenerated synthesized works depending on whether the performers’ images are completely generated by deep synthesis technology. In the aforementioned application examples such as *Furious 7* and other films, the on-screen images of actors like Paul Walker are completely generated by deep synthesis technology. Their actions and expressions in the film are independent creations to which the creators have nothing or little to contribute in addition. Thus, they belong to native synthesized works. The situation for regenerated synthesized works becomes complicated when the work has been completed before deep synthesis technology is involved, deep synthesis technology is required for post-production. For example, in the TV series *The Surrounded*, the character Lü Dekuang was originally played by Zhao Lixin, who was banned due to inappropriate remarks. The crew used deep synthesis technology to replace his face with that of actor Li Zhiqiang. This type of deep synthesis work belongs to regenerated synthesized works.

3.1.2. Legal implications of deep synthesis work categorization

The categorization of deep synthesis works has important legal implications for the following reasons: First, the originality of different types of works varies. For regenerated synthesized works, the original performer’s actions and expressions need to be accurately recreated to maintain the integrity of the original work, which is similar to the “precise tracing” of the original\(^{[18]}\). Therefore, this part of the content lacks originality. The originality of the deep synthesis content in native synthesized works needs to be identified according to different circumstances. Second, not all deep synthesis works have competing performer rights. In regenerated
synthesized works, the new “actor” only provides his or her portrait and is not considered a performer, so there is no conflict between portrait rights and performer rights. In contrast, native synthesized works involve competing performers’ rights. However, different types of works involve different kinds of rights conflicts. The rights conflicts in native synthesized works are mainly manifested as conflicts between the rights of creators and AI owners. Since there is no replacement of performers in native synthesized works, there will be no conflict between their rights. At the same time, because the deep synthesis content in regenerated synthesized works lacks originality, the owner of deep synthesis technology does not enjoy copyright, so there will be no conflict on copyright entitlement. The rights conflicts in regenerated synthesized works are mainly manifested as conflicts between the rights of current performers and original performers.

3.2. Copyrightability of deep synthesis works

In terms of the copyrightability of either deep synthesis works or traditional work types, the key to determining whether a “work” is protected by copyright law lies in its originality. The originality of a work means that the work is independently created by the author, rather than plagiarized from another work [19]. Specifically, the requirement of “independence” means that the expression of the work is formed independently, rather than originating from others. The requirement of “creativity” means that the expression of the work can reflect the author’s personal intellectual choices. The originality of deep synthesis works should be examined from the perspectives of the work both as a whole and synthesized parts as well.

3.2.1. Originality of the entire work

Deep synthesis works are audiovisual works under statutory work classification. Determining the originality of audiovisual works is a difficult issue in copyright law [20]. It is generally believed that the originality of audiovisual works is premised on the overall selection and arrangement of continuous images. Specifically, the originality of audiovisual works is mainly reflected in the visualization of plot content, image shooting, and post-production [21]. To determine the originality of deep synthesis works, an innovative thinking method is required, that is, the deep synthesis work may be distinguished into two different scenarios to judge their originality. The first scenario is that all the images in the work are produced by deep synthesis technology and the second scenario is that a small proportion of the work is completed by deep synthesis technology. In the first case, the originality of deep synthesis content determines whether the work as a whole is protected by copyright law. Therefore, it should be determined by the above criteria, that is, whether the whole work is independently completed and whether it reflects the author’s personal intellectual choices. In the second case, since the proportion of deep synthesis content in the work is manifestly reduced, its impact on the originality of the whole work becomes insignificant. For intermediate cases that lie in between the above two “radical” scenarios, originality should be assessed based on the extent of deep synthesis participation in the work.

3.2.2. Originality of synthesized parts

According to the standard of self-awareness, artificial intelligence can be divided into weak artificial intelligence and strong artificial intelligence [4]. Weak artificial intelligence is artificial intelligence that is applied by human beings as a tool, while strong artificial intelligence may acquire in some sense, personality. The difficulty in determining the originality of AI-generated content is that its content may originate from the AI’s selections and arrangements, that is, the content remains completely under the control of its owner (developer). In that case, the premise of discussing the personality of artificial intelligence is debatable. Nevertheless, since the purpose of using deep synthesis technology is to realize the creator’s personal intellectual choice, the determination of the originality of deep synthesis content is not affected by AI factors. That is, artificial intelligence is
always applied as a tool in deep synthesis works, and its sole purpose is to meet the creative needs of creators. Therefore, unlike other AI-generated content involving subjective factors, the originality assessment of the synthesized parts should only depend on whether their content is formed independently and whether it reflects the intellectual choice of the creators. By far, the development status of deep synthesis technology, in particular, whether AI has self-awareness is still irrelevant to the originality of the synthesized work.

4. Rights analysis of deep synthesis works

Rights competition and rights conflicts are the much-debated new issues brought about by deep synthesis works. The various problems can be generalized as “competition and conflict of rights” [22]. Among them, rights competition refers to several rights arising from the same legal relationship at the same time. Rights conflicts refer to several rights holders claiming mutually opposing rights to the same object. Precisely, rights competition in deep synthesis works refers to the competition between portrait rights and performers’ rights in deep synthesis works. Rights conflicts include conflicts between creators (producers) and AI owners, as well as conflicts between current performers’ rights and original performers’ rights. Giving clear solutions to the legal issues in the above situations will not only help clarify the copyright ownership in deep synthesis works but also help improve the intellectual property rights protection rules for AI-generated content.

4.1. Rights competition in deep synthesis works

4.1.1. Competition between portrait rights and performers’ rights

According to the definition in paragraph 2 of Article 1018 of the Civil Code, a portrait is the external image of an actual person like sculptures and paintings. In the Beijing Treaty on Audiovisual Performances (hereinafter referred to as the “Beijing Treaty”), a performer refers to an actor, singer, musician, dancer, or other person who performs, sings, delivers, declaims, plays, or otherwise performs literary or artistic works or expressions of folklore. When natural persons perform in a work, they not only enjoy portrait rights but also performers’ rights. Therefore, when infringement occurs, performers should freely choose from multiple grounds of claims [23]. In the context of deep synthesis works, the use of others’ portraits to create works is controlled by portrait rights, but whether the “performer” who only provides a portrait is a performer in the sense of copyright law remains an open question. In the aforementioned examples, this issue is whether actor Paul Walker (his heirs) in Furious 7 and actor Li Zhiqiang in The Surrounded enjoy performers’ rights.

The copyright law protects performers’ rights because performers facilitate the dissemination of works. Performers who only provide portraits and have other content generated by deep synthesis technology can be defined as performers by analogy. They may also be analyzed as “supporting performers.” During the Beijing Treaty negotiations, there were major disputes over whether supporting performers (such as stunt doubles) should enjoy performers’ rights. The Beijing Treaty provides that supporting performers do not have identifiable individual contributions and that the standard of protection is determined by member states [24]. Therefore, identifiable individual contribution is the key to determining whether performers’ rights are recognized by those providing portraits. Objectively, the portrait of a specific performer facilitates the dissemination of the work. For example, works synthesized using celebrities’ portraits are more attractive to the general public. Therefore, natural persons who only provide portraits can enjoy performers’ rights.

4.1.2. Comparing the two solutions to rights competition

In deep synthesis works, performers enjoy both portrait rights and performers’ rights over their performances (images). Although both are civil rights, portrait rights, and performers’ rights differ in the following ways:
First, due to the publicity of intellectual property rights, the law imposes more restrictions on performers’ rights. The Civil Code contains rules on the fair use of portrait rights. However, a natural person’s portrait right is subject to less derogation in the Civil Code when compared to the exceptions to the performers’ rights. For example, the reasonable use of exhibiting or preserving a copy of the work is an exception that specifically applies to the performers’ rights. Secondly, the economic content of performers’ rights has a clear protection period, which ends on December 31 of the 50th year after the performance takes place. There is currently no clear time limit for the protection of portrait rights in China, and the mainstream view is that the portrait rights of the deceased should also be protected. Finally, the protection of performers’ rights is narrowed to certain specific purposes, such as permitting others to disseminate their performances through the cyber net. The Beijing Treaty provides performers with more comprehensive and full protection on the basis of the Rome Convention and WPPT, in particular confirming performers’ rights in audiovisual fixations (audiovisual works and videograms). The protection of portrait rights is not limited to specific ways of use, and its scope of protection is broader than that of performers’ rights. For example, in the “Jinlei Zhang Case,” the court of appeal expanded the scope of protection of portrait rights to the image of Sun Wukong in the game. However, the above differences do not mean that choosing to protect portrait rights is always an advantageous strategy for performers. If there are discrepancies between the screen image and the natural portrait of the performer in the audiovisual work, it would be unreasonable for the performer to enjoy portrait rights over their screen image. Consequently, in the precedent case, the actors could only choose to claim the protection of their performers’ rights in the audiovisual works. In addition, unlike portrait rights, the scale of damage for infringement of performers’ rights is easier to determine. Therefore, performers are suggested to choose the mode of rights protection according to their respective situations.

4.2. Rights conflicts in deep synthesis works

4.2.1. Conflicts between creator and other co-authors

In terms of allocation of rights, audiovisual works can be divided into two categories. According to paragraph 2 of Article 17 of the Copyright Law, for cinematographic and television works, the copyright is enjoyed by the producer. For audiovisual works other than cinematographic and television works, the copyright ownership is determined by agreement between the parties. In cases of civil rights conflicts, the best way to resolve rights conflicts is for the parties to clarify the boundaries of rights through agreements. The most important commercial application of deep synthesis works is the creation of movies and TV shows. Consequently, a valid agreement between the parties regarding copyright ownership is the best way to avoid rights conflicts. Paragraph 2 of Article 17 of the Copyright Law is formally a mandatory provision that cannot be derogated through agreement. However, according to Article 153 of the Civil Code, a civil act that derogates or breaches a mandatory provision will not necessarily be invalid. Under Chinese law, mandatory provisions are divided into two types: one is validity mandatory provisions, which are related to state and social public interests, the violation of which will result in the invalidity of a civil act; the other is qualified as “administrative” mandatory provisions, the legal consequence of whose violation shall be nuanced case by case. China’s judicial practices indicate that the consequence of invalidity has been rare and is strictly applied to specific situations. Article 17 of the Copyright Law is deemed as an administrative mandatory provision. Moreover, this article does not apply to content requirements like “market access” restriction. Thus, the violation of Article 17 will not lead to the invalidity of an agreement by the parties. The legislative purpose of this provision is to protect transaction security and promote the dissemination of audiovisual works. Therefore, an agreement between the parties regarding copyright ownership in movies and TV works remains legally binding. However, the agreement is not
opposable to third parties.

4.2.2. Rights conflicts between primary and secondary performers

In regenerated works, conflicts may arise between the rights of the primary performer and the secondary performer: the secondary performer’s performance is based upon that of the primary performer with some modification. Such rights conflicts should be resolved by prioritizing the parties’ autonomy. In the absence of agreement between the primary and secondary performers, the identifiable individual contributor criteria are applied to determine who is entitled to the performers’ rights. The criteria are based on the purpose and common practice of international copyright treaties on performers’ rights. In addition, in cases where neither the primary performer’s consent nor the secondary performer’s is obtained, and the creator uses the secondary performer’s portrait to replace the primary performer, the legal relationship should be clarified as follows: Firstly, the creator’s act of replacing image without consent constitutes a breach of the primary performer’s moral rights to preserve the integrity of the work. Secondly, since the primary performer’s performance image has been replaced, the creator’s acts of reproducing, distributing, broadcasting, and disseminating the new work on the cyber net do not infringe on the primary performer’s property rights. Finally, the rights of the secondary performer should be determined according to objective standards, that is, he or she is entitled to the performer’s rights because his or her image objectively promotes the dissemination of the work. Moreover, recognizing the performer’s rights in the above case has the following two advantages: firstly, it can strengthen the protection of the secondary performer’s economic interests on the basis of portrait rights, and conduce the secondary performer and the creator to reach an agreement; secondly, as copyright is an exclusive right, the secondary performer has the right to control the dissemination of the deep synthesis work, thus in line to protect his or her rights and interests.

5. Conclusion

The development of science and technology has always been bringing out the problems and opportunities as well. This article supports the principle of technological neutrality in that risks created by new technologies shall be objectively assessed and controlled and that new technologies shall be developed with the goal of benefiting society as a whole. Legal issues in the application of deep synthesis technology should be treated in the spirit of balancing risk management with technology development. Deep synthesis technology has broad application prospects in film and television entertainment, communications and social networking, education and medicine, etc. In the field of film and television entertainment, deep synthesis technology is mainly applied in the creation of audiovisual works. Clarifying the rights relationships in deep synthesis works will help improve the legal rules of intellectual property for AI-generated content and fully protect the rights and interests of various stakeholders. In the future, the benefits of deep synthesis technology can be further deployed through an integrated approach of regulation taking into account both legal and ethical considerations.

Disclosure statement

The authors declare no conflict of interest.

References


Publisher's note
Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.