Knowledge E
enriching | engaging | empowering

Conference Paper

Legal Protection for Victims of Artificial Intelligence-based Pornography in the Form of Deepfakes According to Indonesian Law

Cindy Monique[1], Tongat*[2], Siti Wulandari[2], Aprilia Bhirini Slamet[2]

[1]Master of Law, Universitas Brawijaya, Malang, Indonesia
[2]Faculty of Law, University of Muhammadiyah Malang, Malang, Indonesia

ORCID
Cindy Monique: https://orcid.org/0000-0001-8735-0676
Tongat: https://orcid.org/0000-0002-1981-0412
Siti Wulandari: https://orcid.org/0009-0007-4937-762X
Aprilia Bhirini Slamet: https://orcid.org/0000-0002-7997-6860

Corresponding Author: Tongat; email: tongat@umm.ac.id

Published 5 January 2024

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the 4th INCLAR Conference Committee.

Abstract.
Deepfake is an artificial intelligence-based technique for synthesizing an image of a person, using a special method to combine images or videos to make the result look realistic (POLRI 2020). Deepfake is a relatively new type of technology that allows you to download deepfake apps for free. Anyone can access the Deepfake app to create freely edited videos and images. The original purpose of using deepfakes was entertainment on TV and social media. But over time, technology is used as a tool to mislead people and spread misinformation. Deepfakes can undermine public trust, especially when it comes to big and famous people. Not only fake videos but reputation is also easily damaged by this technique.

Keywords: deepfake, artificial intelligence, information and transaction electronic

1. INTRODUCTION

Deepfake is an artificial intelligence-based technique for synthesizing an image of a person, using a special method to combine images or videos to make the result look realistic (Polri 2020). Deepfake is a relatively new type of technology that allows you to download deepfake apps for free. Deepfakes are the result of artificial intelligence (AI). Anyone can access the Deepfake app to create freely edited videos and images. The original purpose of using deepfakes was entertainment on TV and social media. But over time, technology is used as a tool to mislead people and spread misinformation.

Deepfakes can undermine public trust, especially when it comes to big and famous people. Not only fake videos, reputation is also easily damaged by this technique. Many people use deepfake applications to spread negative content, such as spreading

How to cite this article: Cindy Monique, Tongat*, Siti Wulandari, Aprilia Bhirini Slamet, (2024), "Legal Protection for Victims of Artificial Intelligence-based Pornography in the Form of Deepfakes According to Indonesian Law " in *4th International Conference on Law Reform*, KnE Social Sciences, pages 265–275. DOI 10.18502/kss.v8i21.14724

hoax messages or manipulating data and this is very easy to do. Especially in 2018, deepfake applications like FakeApp have started to appear. This application allows users to edit and swap faces and the result is a video (Qualitiva 2021). This app uses special algorithms and techniques. Other similar apps include DeepFaceLab, FaceSwap and myFakeApp.

The functionality of these apps is more or less the same. You can create accurate facial reconstructions and apply them to your videos and videos. The effect of deepfakes on victims is the manipulation of images and videos by malicious artificial intelligence. This is how Deepfake Pornography tries to humiliate its targets. Apart from that, deepfakes are also used in targeted scams and revenge porn (Vit 2020). Deepfakes can also damage a person's reputation, image and credibility. Especially if the resulting deepfake looks genuine and resembles the original. Widespread deepfakes can threaten a person's status and job. Crimes that utilize deepfakes continue to run rampant, not only celebrities but also public figures such as politicians can become victims, causing harm to many people.

Initially, celebrities were the main targets for deepfake pornography victims[1]. Because their photos and videos are very easy to get. However, the victims of deepfake pornography are no longer limited to celebrities and public figures. Anyone can become a victim of deepfake porn in cyberspace, because it is easy for perpetrators to steal photos of victims, especially through social media. There is a short 61-second video showing an obscene scene in which the perpetrator has a face similar to socialite and artist Nagita Slavina. Responding to the crowd, the police investigated the video and confirmed that it appeared to be edited by someone to look like Nagita Slavina. Police have previously established that the technique uses deepfake technology

Elements of articles that meet the criteria for the example cases above are found in Article 27 paragraph (1) of the ITE Law, namely:

1. Everyone
2. Deliberately and without rights distributing and/or transmitting and/or making Electronic Information and/or Electronic Documents containing illegal contents accessible.

Apart from local artists, deepfakes also take victims in the form of international artists. On the Internet, you can find erotic videos with a face similar to actress Gal Gadot. Finally it is known that the video also uses AI technology. Common tools or applications used to create deepfakes are After Effects CC, FakeApps, and DeepFaceLab. Most deepfake victims are celebrities, politicians and artists, especially female artists, who become pornographic video material.

The formulation of the problems found are as follows:

1. What is the legal basis for Deepfake-based pornography cases in Indonesia so far?
2. What are the legal remedies against deepfake-based pornography around the world?

2. METHODOLOGY/ MATERIALS

The type of research used in this paper uses normative legal research based on the statute approach and conceptual approach. The technique of collecting legal materials used in this research is library research. If all the data has been collected, it will be processed using analytical methods using qualitative prescriptive which will be presented and describe and explain what this research finds.

3. RESULTS AND DISCUSSIONS

3.1. Legal Basis for Deepfake-Based Pornography Cases in Indonesia

According to S. R. Sianturi, in summary the elements of crime are: the presence of a subject, the presence of an element of error, an act against the law (PMH), an act that is stopped or prohibited by laws/regulations and others. those who violate it are subject to criminal sanctions, at certain times, places & circumstances[2]. Referring to the elements of the crime earlier, S. R. Sianturi formulates the meaning according to crime is an act at a certain place, time & condition, which is stopped (or violates obligations) & is threatened by using statutory penalties and contrary to the rules. & contains elements of errors that can be made by the person in charge.[3]

Determination of an act to be a criminal act is carried out through an analysis of whether the act has fulfilled the elements regulated in a particular criminal law article[4]. For this reason, an analysis must be carried out according to an act and whether the act fulfills the elements according to the articles of criminal law[5]. If it fulfills, then it can be said/categorized that the act is a criminal act and is in sync with one of the articles that was violated. However, if one of these elements is not proven, then it is concluded that the act is not a crime.[6]

Regarding the use of deepfake applications with bad intentions/abuse of deepfakes, it can be classified as a crime if it meets the elements of the criminal budget that regulate it. Therefore, in an effort to ensnare perpetrators of deepfake abuse, it is necessary to know in advance the influence of the abuse of deepfake implementation that they are doing. Below we will describe the articles that can be applied to perpetrators of abuse of deepfake implementation according to the impact they have.

The misuse of deepfake applications cannot be separated from cybercrime because its spread is based on the results of deepfake photo/video edits carried out through social media which also use the internet network for its operation, therefore the misuse of deepfake applications can also be classified as cybercrime. So, the criminal act of misusing the application of deepfakes is analyzed according to the law governing cybercrime which is related to the law governing the impact arising from the misuse of the application of deepfake.

The impact of deepfake technology itself on Indonesia's national privacy law has not been fully enforced as Indonesia still has limited regulations on personal data protection. In general, article 27 of the ITE Law is used to involve perpetrators in cases where information is not or should be disseminated:

Everyone intentionally and without rights distributes and/or transmits and/or makes accessible

Electronic Information and/or Electronic Documents that contain content that violates decency.

Everyone intentionally and without rights distributes and/or transmits and/or makes Electronic Information and/or Electronic Documents that contain gambling content accessible.

Everyone intentionally and without rights distributes and/or transmits and/or makes accessible Electronic Information and/or Electronic Documents that contain extortion and/or threats

In that article there is a diction that anyone who intentionally and without rights distributes and/or transmits and/or provides electronic information. In our opinion, this article cannot prosecute photorealistic video makers, but the law can only prosecute video makers that contain dirty words, insults, slander, extortion and threats. So that legal protection for victims and sanctions against perpetrators of deepfake pornography crimes have not been achieved through this article alone.

So, in this case, the author is of the opinion that there is no standard and clear legal regulation regarding how sanctions are imposed on perpetrators of deepfake-based crimes. That crimes such as the use of deepfakes themselves can be a threat that will occur in the future, therefore this is an appropriate urgency why legal regulations regarding the abuse of deepfakes must be immediately regulated or implemented in Indonesia.

3.2. Legal Enforcement of Deepfake-Based Pornography Cases in the World

Next, we will discuss what legal remedies have been implemented by several countries in dealing with deepfake-based pornography cases. We consider this discussion important, as material for legal comparison if in the future Indonesia wants to regulate specifically the rules regarding sanctions imposed on perpetrators of crimes in the form of creators/spreaders of deep hoax-based pornographic content. Actually, the legal rules regarding the use of AI in the form of deepfakes have not been regulated by international law. However, the United Nations provides the widest possible discretion for countries to determine which rules they want to apply regarding the use of deepfakes in a particular case.[7] Even so, the United Nations has made several efforts to align regulations regarding the use of this deepfake technology, one of which is by forming the Telecommunications Union. However, the UN does not give the obligation to use the union as a legal basis in settling international cases related to the use of AI, but the UN gives freedom to countries in the world to use it as a legal basis or not (in a case).

The first countries to be discussed in this journal are countries in Europe. In dealing with cases of pornography based on deep fabrication, European countries have implemented a personal right known as the Right to be Forgotten, or RTBF for short. RTBF is a right to be forgotten that can be demanded by citizens to protect their personal data that has been spread in the media, with the aim that negative things related to it can be slowly forgotten by the wider community. RTBF began to be widely used after the case that befell Mario Costeja Gonzales in 2014. The idea of implementing RTBF as a way of resolving deepfake-based pornography cases is an awareness idea that the right to be forgotten is a very important moral asset, because basically every human being's moral rights must be respected by all levels of society in carrying out their daily activities. and is a form of personal rights. However, in fact, in Indonesia itself there are legal instruments that have a 'similar' concept to the RTBF that applies in the European Union, more precisely contained in Article 26 paragraph (3) of the ITE Law which focuses on granting rights to individuals. request deletion of information or data regarding individuals. , where the core of this article has the essence that is quite the same as RTBF, namely that electronic system operators are expected to be able to remove excessive information and electronics under their influence according to their individual wishes.

The second country to be discussed in this journal is the United States. The United States itself has long echoed the urgency of the need to make special regulations governing deepfake abuse, because deepfake abuse itself is one of the most frequent cases in the United States, especially in 2019, especially deepfake pornography. So that since 2020, several states in the United States have formulated regulations regarding deepfake abuse, and states that are quite early in regulating this include:[8] California, Texas and Virginia. Virginia and Texas have actually banned the use of deepfakes during the presidential election as an effort to reduce slander during the presidential campaign period (the use of deepfakes on the face & body of US presidential candidates). Meanwhile, California just banned the use of deepfakes 60 days before the presidential election day with the following sound:

“In the U.S., injunctions against deepfakes are likely to face First Amendment challenges. Even if such injunctions survive a First Amendment challenge, lack of jurisdiction over extraterritorial creators of deepfakes would inhibit their effectiveness. Therefore, injunctions against deepfakes may only be granted under few specific circumstances, including obscenity and copyright infringement.”

In a case regarding deepfake pornography that occurred in the United States in 2020[9] Article 230 of the Communications Decency Act applies, whereby disseminators of pornographic content created by deepfake applications can be fined around US$2,500 or around Rp.35,000,000.00 (thirty-five million rupiah) or imprisonment for a maximum of 12 months. But it cannot be denied, even though there are regulations that formulate sanctions against perpetrators of deepfake-based pornography crimes, this still cannot prevent the spread of non-consensual pornographic content, including those created by deepfake applications. Judging from the sanctions given to the perpetrators, the crime of spreading deepfake-based pornographic content is also not considered a serious crime in the United States, because ultimately posting deepfake content or miss-information is not considered illegal. act in the United States.

The third country to be discussed in this journal is South Korea. South Korea is a country capable of technology, so it’s no wonder that deepfakes have become popular in this country since 2018, one year after the existence of deepfake technology emerged as a phenomenon of technological progress in the world. Unfortunately, just like the problems in previous countries, many irresponsible persons take advantage of deepfakes as an opportunity to create pornographic content, and the majority of victims are young women aged 14-21 years. Cases of spreading deepfake-based pornographic content in South Korea reached their peak in mid-2021[10], when the police at that time managed to arrest 94 suspects in just the last 5 months.

Spreading pornographic content is a serious crime in this country, because South Korea is a country that upholds the privacy of its citizens, even though at that time there were no specific regulations that contained sanctions against the perpetrators of the crime of spreading deepfake-based content. As a result, in the same year an anonymous petition (titled deepfake_strong punishment) was created to urge the government to immediately make legal regulations regarding sanctions against related cases, which had been signed by more than 330,000 signatures in just one day since the petition was released. Finally, the Korea Communications Standards Commission (KOCSC)[11] took action by blocking the general public's access to deepfake websites. Revisions are made in 2021, whereby under the revision, offenders who have created deepfake content with someone's consent by including sexual insults in it, can be subject to a penalty of five years in prison, or a fine of ₩50,000,000 (or approximately six hundred and forty five million rupiah), provided that if the content is distributed for commercial purposes, the prison sentence will be increased to 7 years.

The last country discussed in this journal is England. Similar to previous countries, this deepfake abuse has the potential to occur in developed countries, including the UK. Actually, legal regulations regarding the use of AI technology have been regulated in the Online Security Bill which was passed in May 2021, but this law does not yet regulate the abuse of deepfakes in the spread of pornographic content, so that in March 2022 a revision was made[12] against the law. In the revision of the Online Safety Bill, the UK officially criminalizes the dissemination of deepfake (non-consensual) pornographic content.

3.3. Efforts to Realize Legal Protection for Victims of Deepfake Pornography in Indonesia

Basically the existence of the emergence of audio/visual manipulation is not something that is really new and has never been predicted before. In 1998, American scientist David Brin warned that photo-checking technology would soon be abandoned and we would soon reach a society where expertly controlled computers would be able to adjust images microscopically pixel by pixel without leaving a trace of clues. Currently in 2022, the concerns described by David Brin are finally starting to occur slowly, where technological developments are increasingly developing until it is possible to manipulate video into audio using a technological discovery phenomenon known as deepfake, and the worst is the potential for its application. to pornographic content which is then disseminated for the benefit of public consumption.

The urgency that arises is how to avoid deepfake abuse. Education should be provided to identify the types of deepfake videos circulating in the media. Identification is an original video where another person's face is "glued" to the face of a natural person in the video, as if the video were actually created from a new face inserted into the original video. It is feared that it will be abused in several crimes, one example of which is revenge porn. A criminologist named Anastasia Powell and a researcher named Nicola Henry further explained that there are at least three types[13] :

The creation of nude or sexual images without consent;

The distribution or sharing of nude or sexual images without consent (including images that were self-created by the victim or consensually created with another person); and

The threat of distribution of nude or sexual images.

In the second form, it is described as 'manipulating' the speech and facial expressions in the video as if the person in the video said those words. This is a form of deepfake video. Current problems, according to a survey by the Association of Indonesian Internet Service Providers (APJII)[14] in 2020, around 24.7% of people use the internet to communicate via messages, while 18.9% use the internet to share social media. The high activity of the Indonesian people towards internet use certainly increases the risk of the Indonesian people becoming victims of deepfake abuse by irresponsible persons.

Peter Fleischer, a consultant who works at Google, then gave his views/questions[15] regarding the protection of victims in the misuse of all forms of technology that can be accessed by the wider community, namely:

If something online, do I have the right to delet it again?

If I posting something an someone else copies it and re-post it on their own site, do I have the right to delete it?; and

If someone else post something about me, do I have a right to delete it?

Peter's first view, of course, would be easy given that the data subject has access control to the data source, but this is out of sync with views 2 & 3, given that the data subject has no authority. over data. The data subject needs to submit a request to the party who has re-uploaded the data and published the data which still has the data subject (victim) in it. Problems will arise when requests for data deletion are rejected, thus creating a potential threat to someone who is a victim of a deepfake video that has been spread. This happens because basically a person's self-portrait is direct data that must be legally protected.

Although it is not easy to define privacy, in general privacy can be divided into 3 basic types, namely; Physical Privacy, Information Privacy, and Organizational Privacy. Meanwhile, to ensure that data privacy often conflicts with public data, you can see the division, namely: our Communication Privacy, our Behavioral Privacy, and our Personal Privacy. So in this case Indonesia can apply the RTBF that has been described in the previous chapter, bearing in mind that several elements in the RTBF actually already exist in Indonesian law, more precisely in Article 26 paragraph (3) and (4) UU ITE which reads as follows:

Article 26

Every Electronic System Operator is obliged to delete irrelevant Information and Electronics under their control at the request of the person concerned based on a court order;

Every Electronic System Operator is required to provide a mechanism for deleting electronic information that is no longer relevant.

What distinguishes it from the RTBF concept that applies in Europe is the result of fulfilling this right.[16] In the RTBF concept, this information disappears from search engine results but can be found in the original link where the data information is stored. Therefore, RTBF is often referred to as a measure to make it difficult for anyone to access information about someone on search engines/websites. Therefore, a ministerial regulation on the right to be forgotten must be issued immediately to provide legal protection for victims of this deepfake abuse.

4. CONCLUSION AND RECOMMENDATION

Deepfake, is an example of a form of technological progress that unfortunately many people misuse to become a suggestion for committing a crime, including deepfake pornography. Indeed, in Indonesia, there have been several articles in the ITE Law to the Criminal Code to deal with deepfake pornography cases that have occurred in Indonesia so far, but unfortunately, special regulations regarding the misuse of deepfake technology have not been regulated in Indonesia so far. Article 26 of the ITE Law, which is predicted to be able to handle cases of deepfake abuse in pornographic content, has not been able to provide protection for victims. The concept of RTBF applied in European countries should have been fully implemented in that article, or even better if a ministerial regulation on the right to be forgotten which has the same essence as

RTBF is immediately ratified and enforced in Indonesia, so as to be able to make laws protecting against victims of deepfake pornography, including revenge pornography.

Supposedly, efforts to prevent media manipulation, whether video, photo or sound by using deepfake techniques, can be done by limiting the excessive publication of personal documentation, both in the form of photos and videos. If the data has been published, then the step that can be taken is to carry out the data selection process, so that the information available on the internet is not exposed to excess.

References

[1] Marria Saimima J, Liminanto E, Wasia Z. Edukasi Hukum tentang Kekerasan Seksual Pada Perempuan Dan Anak Di Kelurahan Lateri Kota Ambon. J Dedik Huk. 2022 Apr;2(1):75–84.

[2] Pujinoto S, Mashdurohatun A, Sulchan A. Juridical analysis of application of forgiveness (Rechterlijk Pardon) as a basis of judge consideration in deciding the criminal. J Daulat Huk. 2020 Jun;3(2):307–12.

[3] S. Siantur, Asas-Asas Hukum Pidana Di Indonesia Dan Penerapan. 3rd ed. Jakarta: Storia Grafika; 2002.

[4] Tongat T, Prasetyo SN, Aunuh N, Fajrin YA. Hukum yang Hidup dalam Masyarakat dalam Pembaharuan Hukum Pidana Nasional. J. Konstitusi. 2020 May;17(1):157–77.

[5] Tongat. “Restorative Justice Dan Prospek Kebijakan Idealnya Dalam Hukum Pidana Indonesia,”. Restor. Justice Dan Prospek Kebijak. Idealnya Dalam Huk. Pidana Indones. 2013;42(4):542–8.

[6] Langa J. Deepfakes, real consequences: Crafting legislation to combat threats posed by deepfakes. New York: Hein Online; 2021.

[7] Stevens D. Regulating deepfake technology. Tilbg Law Sch J. 2020.

[8] Briscoe S. U.S. Laws Address Deepfakes. United States of America: Security Management Magazine; 2021.

[9] Dauer F. Law enforcement in the era of deepfakes. Police Chief Magazines; 2022.

[10] “Police arrest 94 suspects over deepfake crimes in 5 months.” All News, 2021. .

[11] Ryall J. “‘Deepfakes’ rattle South Korea’s tech culture.” DW News, Nov-2021.

[12] Lomas N. UK to criminalize deepfake porn sharing without consent. United Kingdom; 2022.

[13] Kasita Dewi I. “Deepfake Pornografi: Tren Kekerasan Gender Berbasis Online (KGBO) di Era Pandemi Covid-19.” Wan dan Kel. 2022;3(1).

[14] “Survei Profil Pengguna Internet Indonesia.”

[15] Khusna Hidayatul I. “Deepfake Tantangan Baru untuk Netizen.” Promedia. 2019;5(2).

[16] Abir Jufri MA, Kurnia Putra A. “Aspek Hukum Internasional dalam Pemanfaatan Deepfake Technology terhadap Perlindungan Data Pribadi.” Promedia. 2021;2(1).