

Seton Hall University

eRepository @ Seton Hall

Student Works

Seton Hall Law

2024

Caught in the Web: Balancing First Amendment Considerations in the Face of Deepfake Child Pornography

Iqra A. Niazi

Follow this and additional works at: https://scholarship.shu.edu/student_scholarship



Part of the Law Commons

I. Introduction

It is 10:38am. You are fourteen years old and in your first year of high school. Walking down the hallways surrounded by upperclassmen has always been intimidating, but something is different today. As you walk by, with a heavy bag of books on your back and a pacing heart, you notice the snickers as they look down at their phone and look back up at you. You feel the knot in your stomach grow tighter as the anxiety increases and the collective laughter of teenagers starts to ring in your ears.

Buzz. A notification from Snapchat. A friend from your math class shared a snap. As you innocently open the video, you fail to realize that your life is about to change forever. On your screen is a high quality video with your face, your body, and your voice depicting you engaging in sexual activity with someone you have never seen before.

What is this? When did this happen? Was I drunk? Was I assaulted? Where did this video come from?

Suddenly it clicks what everyone else has been giggling at on their phones. In a moment of pain, embarrassment, and shame, you fight the lump in your throat and call your mom. She picks up worried, asking what happened, but you fall silent. How do you defend yourself when the video is real? How do you explain that something is not real when *it appears to be so real?*

This paper provides an overview of deepfake technology and its contemporary uses in society. Next, it discusses current federal law concerning the intersection of child pornography and the First Amendment and the current circuit court split over the issue of morphed child pornography. Finally, this paper concludes by exploring potential methods of combatting pornographic deepfakes of children. This includes an analysis of current

state laws pertaining to deepfakes and a reconsideration of Supreme Court precedent in light of an evolving technological landscape. The overarching theme emphasizes the need for heightened federal regulation to protect children from the lasting harms of sexual exploitation while maintaining the delicate balancing between the protection of free speech and society’s most vulnerable individuals.

II. The Deepfake Problem

A. What is a deepfake?

A deepfake is a video, image, or other form of media that has been created through deep generative approaches in order to produce a realistic-looking face that does not belong to a preexisting identity, or to perform highly realistic facial manipulations and swaps.¹ In order to produce such material, a user must input a set of genuine photos, videos and audio files of various people – called a “faceset” – into a machine learning program.² Deepfake technology first went viral in 2017 when a Reddit user posted a falsified video to the website depicting actress Gal Gadot’s face superimposed onto the body of an actual adult pornography actress’s body.³

B. How are deepfakes created?

There are two main methods by which deepfakes are generated. Under the first and most popular method, a pair of general adversarial networks (“GANs”) are utilized.⁴ The first network, called the generator, inputs a latent sample and generates an image to

¹ Juefei-Xu, F., Wang, R., Huang, Y. *et al.*, *Countering Malicious DeepFakes: Survey, Battleground, and Horizon*, INT J COMPUT VIS 130 (2022). <https://doi.org/10.1007/s11263-022-01606-8>

² Moncarol Y. Wang, *Don’t Believe Your Eyes: Fighting Deepfaked Nonconsensual Pornography with Tort Law*, 2022 U. Chi. Legal F. 415 (2022).

³ *Id.*

⁴ Raina Davis, Chris Wiggins & Joan Donovan, *Deepfakes*, Harv. Kennedy Sch. Belfer Ctr. for Sci. and Int’l Aff. 3, https://www.belfercenter.org/sites/default/files/files/publication/Deepfakes_2.pdf [<https://perma.cc/2AY8-KM47>].

feed into the second network, called the discriminator, which then identifies if the generated image is authentic or fake based on its training on real data.⁵ The generator utilizes this feedback until the discriminator is no longer able to distinguish between a real image and the generator’s artificially-created image.⁶ The larger the set of training data for the networks, the more believe the generated image will be.⁷

The second method of deepfake creation is through variational autoencoders (“VAEs”).⁸ VAEs are generative models that work together to “encode an image into a low-dimensional representation and then subsequently decode the representation back into an image” while configuring hundreds of stored images of a desired subject until the input and output images from the VAEs match.⁹ The system is also trained on a wide variety of other curated subjects, allowing the decoder to configure artificially created content of its original subject partaking in virtually any of the other frames, posing, and lighting that it has also been trained on.¹⁰

The result of either of these two methods is one: hyper realistic content depicting an individual saying or doing something that never actually happened.

C. Use and Abuse

Despite the relative infancy of the technology, its use has expanded quickly. A museum can now generate a life-size depiction of Salvador Dali to greet visitors and

⁵ *Id.*

⁶ *Id.*

⁷ *Id.*

⁸ *Id.*

⁹ Molly Mullen, *A New Reality: Deepfake Technology and the World Around Us*, 48 Mitchell Hamline L. Rev. 210, 213 (2022).

¹⁰ *Id.*

share his famous quotes as if he was alive and present today.¹¹ Lovers of history can listen to President John F. Kennedy share the speech he was meant to deliver the day he was assassinated.¹² The underlying technology behind deepfakes could help medical researchers use realistic data, such as fake brain MRI scans, instead of waiting on real patient data, to help train machines to better identify tumors in scans.¹³ Hip hops fans can see their favorite deceased rappers from the 90s seemingly come back to life in collaborative music videos and interviews with today's top artists.¹⁴ Celebrities looking to engage with fans across the world or partake in international charity efforts can now share their message seamlessly in any language.¹⁵ Amyotrophic lateral sclerosis (“ALS”) patients can still preserve their voice to speak with their friends and family despite the detrimental impact of the neurodegenerative disease on the patient's ability to speak.¹⁶ There is no doubt that the emerging technology has shown tremendous potential for positive impact. Despite opportunities for positive use, however, many remain conflicted on the ultimate effect of deepfake technology.

Some have identified that the concern with deepfakes is not the creation of fake content or media manipulation, but rather “the democratization of Hollywood-style

¹¹ Dami Lee, *Deepfake Salvador Dali takes selfies with museum visitors*, THE VERGE (May 10, 2019).

<https://www.theverge.com/2019/5/10/18540953/salvador-dali-lives-deepfake-museum>

¹² Simon Chandler, *Why Deepfakes Are A Net Positive For Humanity*, FORBES (Mar. 9, 2020).

<https://www.forbes.com/sites/simonchandler/2020/03/09/why-deepfakes-are-a-net-positive-for-humanity/?sh=3c718b842f84>

¹³ *Id.*

¹⁴ Alyse Stanley, *This VFX Team Made a Tupac Deepfake to Rap with Snoop Dog*, GIZMODO (Mar. 7, 2020).

<https://gizmodo.com/these-vfx-artists-made-a-tupac-deepfake-to-rap-with-sno-1842185314>

¹⁵ Sam Forsdick, *How deepfake technology enables David Beckham to speak nine languages in malaria charity ad*, NS BUSINESS (Apr. 9, 2019). <https://www.ns-businesshub.com/technology/david-beckham-deepfake-malaria-ad/>

¹⁶ Amanda Morris, Alexa Juliana Ard & Szu Yu Chen, *Patients were told their voices could disappear. They turned to AI to save them.*, THE WASHINGTON POST (Apr. 20, 2023).

<https://www.washingtonpost.com/wellness/interactive/2023/voice-banking-artificial-intelligence/>

technology that can create really compelling fake content”.¹⁷ Internet users worldwide now have access to technology that requires no real skills to use but can generate hyper realistic images.¹⁸ This access, coupled with practically unlimited distribution ability via social media and the Internet as a whole, leaves almost anyone – celebrities, politicians, private citizens, and children alike – exposed to the possibility of having a deepfake image created of them.¹⁹ Practically unlimited access to deepfake creation and distribution becomes of extra concern, however, when considering the harsh reality of the situation: 96% of all deepfakes are non-consensual pornography.²⁰ As stated by one writer, “Google gave the world powerful AI tools, and the world made porn with them.”²¹

The use of this non-consensual pornography is varied. For some, deepfake applications and websites have become an alternative to their regularly consumed adult content, allowing the average user to customize graphic pornographic material of his favorite actress.²² An Indian investigative journalist was targeted by political opponents for her reporting on corruption in Hindu nationalist politics with a video depicting her face on the body of a naked woman, exposing her to doxing and threats of gang rape.²³ Controversy arose in the online streaming community when viewers noticed a popular

¹⁷ Waldrop, M. Mitchell. *Synthetic media: The real trouble with deepfakes*. KNOWABLE MAGAZINE (Mar. 16, 2020). <https://knowablemagazine.org/article/technology/2020/synthetic-media-real-trouble-deepfakes>

¹⁸ *Id.*

¹⁹ *Id.*

²⁰ Chen, Angela, *Forget fake news – nearly all deepfakes are being made for porn*, MIT TECHNOLOGY REVIEW, (Oct. 7, 2019), <https://www.technologyreview.com/2019/10/07/132735/deepfake-porn-deeptrace-legislation-california-election-disinformation/>

²¹ Dave Gershgor, *Google gave the world powerful AI tools, and the world made porn with them*, QUARTZ (Feb. 7, 2018), <https://qz.com/1199850/google-gave-the-world-powerful-open-source-ai-tools-and-the-world-made-porn-with-them>

²² Drew Harwell, *Scarlett Johansson on fake AI-generated sex videos: ‘Nothing can stop someone from cutting and pasting my image’*, THE WASHINGTON POST (Dec. 31, 2018), <https://www.washingtonpost.com/technology/2018/12/31/scarlett-johansson-fake-ai-generated-sex-videos-nothing-can-stop-someone-cutting-pasting-my-image/>

²³ Danielle Keats Citron, *Sexual Privacy*, 128 Yale L.J. 1870 (2019).

married streamer's purchase of nonconsensual deepfake pornography depicting female streamers, including some of his own colleagues and friends.²⁴ Given the obsession of today's society with sharing personal details and photos online, it was no surprise that private individuals would become victims of this same content.²⁵ Some victims have been facing battles since they were minors, such as Noelle Martin, who has been fighting for years after anonymous predators stole non-explicit images from her social media and uploaded them to porn sites and threads, leading to the creation of doctored pornographic videos of Martin performing sexual acts.²⁶

More recently, several high schools in the United States have been victims of these deepfake incidents. In New Jersey, a mother received a text from her fourteen year old daughter, Francesca Mani, stating that naked photos of her were being distributed at school.²⁷ Upon the school's further investigation, it became evident that other students had used artificial intelligence to generate fake nude images of Mani and other female students.²⁸ Despite efforts to delete the photos and stop the distribution, the damage had been done instantly: several victims and their friends, all girls, were seen crying around the hallways after their classmates had now seen what they perceived to be naked photos of them, creating a new form of sexual exploitation with seemingly no substantive form

²⁴ Megan Farokhmanesh, *The Debate on Deepfake Porn Misses the Point*, WIRED (Mar. 1, 2023)

<https://www.wired.com/story/deepfakes-twitch-streamers-qtcinderella-atric-pokimane/>

²⁵ Danielle S. Van Lier, *The People v. Deepfakes California AB 1903 Provides Criminal Charges for Deepfakes Activity To Guard Against Falsified Defaming Celebrity Online Content*, L.A. Law., May 2020, at 21.

²⁶ Cara Curtis, *Deepfakes are being weaponized to silence women – but this woman is fighting back*, TNW (Oct. 5, 2018), <https://thenextweb.com/news/deepfakes-are-being-weaponized-to-silence-women-but-this-woman-is-fighting-back>

²⁷ Tim McNicholas, *New Jersey high school students accused of making AI-generated pornographic images of classmates*, CBS NEW YORK (Nov. 2, 2023). <https://www.cbsnews.com/newyork/news/westfield-high-school-ai-pornographic-images-students/>

²⁸ *Id.*

of regulation.²⁹ The school, unsure of how to punish a rather novel issue, temporarily suspended the student who created the images until he was allowed to return a few days later.³⁰ Many parents have found this punishment to be inadequate and are pushing school officials and lawmakers alike to implement better safeguards to protect children from the gruesome risks of such technology.³¹

Victims like Martin and Mani raise a pressing and important question for the government: what happens when a child’s likeness is used in the production of these images?³² As it stands, most victims of deepfakes may only rely on general legal principles to seek relief due to an absence of current federal legislation that would adequately penalize synthetic pornographic content of children.³³ The next section examines the current circuit split over the issue of morphed child pornographic material.

III. The First Amendment and Child Pornography

A. *Miller v. California* (1973)

In *Miller*, the Court was confronted with the task of defining obscenity. Prior to *Miller*, the Supreme Court had found that obscenity is not protected speech under the First Amendment and that States have a legitimate interest in preventing the distribution of obscene materials.³⁴ In *Miller*, the Appellant carried out a mass mailing campaign to advertise illustrated books that were also referred to as “adult” material.³⁵ The Appellant

²⁹ *Id.*

³⁰ Teresa Priolo and Amanda Geffner, *Westfield High School student accused of creating AI nude images of classmates*, FOX 5 NEW YORK (Nov. 2, 2023), <https://www.fox5ny.com/news/westfield-high-school-new-jersey-artificial-intelligence-pornographic-images-incident>

³¹ McNicholas, *supra* note 27

³² Deepfake technology poses a heightened risk of harm to children as Congress has concluded that most pedophiles prefer to customize pornographic material to cater to their specific sexual deviation. S. REP. NO. 95-438, at 6 (1977).

³³ Ryken Kreps, *I (Don’t) Know It When I See It: The Dangers of Deepfakes*, 107 Minn. L. Rev. (2023).

³⁴ *Roth v. U.S.*, 354 U.S. 476, 486 (1957)

³⁵ *Miller v. California*, 413 U.S. 15, 16 (1973).

had distributed the advertisement brochures to a restaurant where they were opened by the manager and his mother.³⁶ While the brochures did include some descriptive material in print, most of it consisted of pictures and drawings explicitly showing men and women partaking in sexual activity.³⁷ The restaurant manager and his mother had not requested the brochures and complained about them to the police as they had not indicated any interest in receiving such material.³⁸

The *Miller* court laid out a framework for analyzing obscenity. To test whether a work is obscene, the *Miller* test asks (a) whether the average person, applying contemporary community standards, would find that the work, taken as a whole, appeals to the prurient interest; (b) whether the work depicts or describes, in a patently offensive way, sexual conduct specifically defined by the applicable state law; and (c) whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value.³⁹ Interestingly, the Court recognized a general shift toward seemingly progressive views regarding obscene material, but failed to recognize this trend as sufficient to warrant unregulated access to pornographic materials.⁴⁰

The Court also clarified that under the first part of the test, the State is not required to apply a national standard to the material alleged to be obscene but rather the relevant community standards of the State regulating such material.⁴¹ As such, the Court

³⁶ *Id.* at 18

³⁷ *Id.*

³⁸ *Id.*

³⁹ *Id.* at 24

⁴⁰ “The ‘sexual revolution’ of recent years may have had useful byproducts in striking layers of prudery from a subject long irrationally kept from needed ventilation. But it does not follow that no regulation of patently offensive ‘hard core’ materials is needed or permissible; civilized people do not allow unregulated access to heroin because it is a derivative of medicinal morphine.” *Id.* at 36

⁴¹ *Id.* at 31

reaffirmed that obscene materials are not protected speech under the First Amendment, held that such material may be constitutionally regulated by state law, and that obscenity is analyzed using local community standards as opposed to a national standard.⁴²

B. *New York v. Ferber* (1982)

In *Ferber*, the Court clarified another exception to First Amendment protection: depictions of actual minors engaging in sexual conduct. In *Ferber*, the Defendant owned a New York bookstore focusing on selling “sexually oriented products”.⁴³ The Defendant sold two films depicting young boys masturbating to an undercover police officer.⁴⁴

The Defendant was indicted on two counts of violating § 263.10 and two counts of violating § 263.15, both of which are New York state laws regulating the distribution of child pornography.⁴⁵ The Defendant was acquitted for the “two counts of promoting an obscene sexual performance” but was found guilty for the two counts under § 263.15.⁴⁶ Under § 263.15, the State was not required to prove that the films were obscene.⁴⁷

On appeal, the Appellate Division of the New York State Supreme Court affirmed Ferber’s convictions.⁴⁸ The New York Court of Appeals reversed his convictions, however.⁴⁹ The New York Court of Appeals found that § 263.15 was unconstitutional as it was underinclusive because although it regulated visual depictions of minors engaged in sexual activity, the law did not prohibit any other categories of dangerous activity

⁴² *Id.* at 36-37

⁴³ *New York v. Ferber*, 458 U.S. 747, 752 (1982).

⁴⁴ *Id.*

⁴⁵ *Id.*

⁴⁶ *Id.*

⁴⁷ *Id.*

⁴⁸ *Id.*

⁴⁹ *Id.*

portrayed in films.⁵⁰ It was underinclusive “because it discriminated against visual portrayals of children engaged in sexual activity by not also prohibiting the distribution of films of other dangerous activity.”⁵¹ The court found the same law to be “overbroad because it prohibited the distribution of materials produced outside the State, as well as materials, such as medical books and educational resources, which deal with adolescent sex in a realistic but nonobscene manner.”⁵² The Supreme Court then granted the State’s petition for cert and considered whether a state could prohibit the dissemination of materials that depict minors engaged in sexual activity regardless of whether such material is obscene.⁵³

The Court reaffirmed its prior holdings that prevention of the abuse and exploitation of children is an important governmental interest.⁵⁴ The Court also reasoned that the dissemination of child pornography is inherently connected to the abuse of children in multiple ways.⁵⁵ “First, the materials produced are a permanent record of the children's participation and the harm to the child is exacerbated by their circulation.... Second, the distribution network for child pornography must be closed if the production of material which requires the sexual exploitation of children is to be effectively controlled.”⁵⁶ “Third. The advertising and selling of child pornography provide an economic motive for and are thus an integral part of the production of such materials, an activity illegal throughout the Nation.”⁵⁷ “Fourth. The value of permitting live

⁵⁰ *Id.*

⁵¹ *Id.*

⁵² *Id.*

⁵³ *Id.* at 753

⁵⁴ *Id.* at 756-57

⁵⁵ *Id.* at 759

⁵⁶ *Id.*

⁵⁷ *Id.* at 761

performances and photographic reproductions of children engaged in lewd sexual conduct is exceedingly modest, if not *de minimis*.”⁵⁸ “Fifth. Recognizing and classifying child pornography as a category of material outside the protection of the First Amendment is not incompatible with our earlier decisions.”⁵⁹

Under *Ferber*, only visual depictions of minors engaging in sexual activity are unprotected under the First Amendment.⁶⁰ The Court also adjusted the *Miller* test for obscenity to create a test for child pornography.⁶¹ Specifically, under this test, the material: (1) does not need to appeal to the prurient interest of the average person; (2) does not need to display sexual conduct in a patently offensive manner; and (3) does not need to be considered as a whole.⁶²

The *Ferber* court repeatedly emphasized the importance of government regulation of child pornography, stating that “the prevention of sexual exploitation and abuse of children constitutes a government objective of surpassing importance.”⁶³ “It is evident beyond the need for elaboration that a State’s interest in ‘safeguarding the physical and psychological well-being of a minor’ is ‘compelling.’⁶⁴ ‘A democratic society rests, for its continuance, upon the healthy, well-rounded growth of young people into full maturity as citizens.’⁶⁵ Accordingly, we have sustained legislation aimed at protecting the physical

⁵⁸ *Id.* at 762

⁵⁹ *Id.* at 763

⁶⁰ *Id.* at 764

⁶¹ *Id.*

⁶² *Id.*

⁶³ *Id.* at 757

⁶⁴ *Globe Newspaper Co. v. Superior Court*, 457 U.S. 596, 607 (1982).

⁶⁵ *Prince v. Mass.*, 321 U.S. 158, 168 (1944).

and emotional well-being of youth even when the laws have operated in the sensitive area of constitutionally protected rights.”⁶⁶

C. Ashcroft v. Free Speech Coalition (2002)

In *Ashcroft v. Free Speech Coalition*, the Court toyed with the idea of synthetic or simulated child pornography. The Free Speech Coalition, an adult entertainment trade association, challenged the CPPA for being overbroad.⁶⁷ The Free Speech Coalition felt that although its members did not use children in the production of their sexually explicit work, they could still be found guilty due to overbreadth and vagueness in the language of the legislation.⁶⁸ The organization argued that language like “appears to be” and “conveys the impression” found in the provisions of the CPPA are overbroad and vague and could criminalize their material that is protected under the First Amendment.⁶⁹ The District Court disagreed and granted summary judgment.⁷⁰ Free Speech Coalition appealed and the 9th Circuit Court of Appeals reversed, holding that the CPPA was overbroad for banning material that is neither obscene nor involves the exploitation of minors.⁷¹ The Supreme Court granted cert over the circuit split.⁷²

The Court analyzed whether the First Amendment allowed Congress to prohibit sexually explicit virtual images of minors if no actual minors are visible.⁷³ “The CPPA cannot be read to prohibit obscenity, because it lacks the required link between its

⁶⁶ *Ferber* at 757

⁶⁷ *Ashcroft v. Free Speech Coalition*, 535 U.S. 234, 243 (2002).

⁶⁸ *Id.*

⁶⁹ *Id.*

⁷⁰ *Id.*

⁷¹ *Id.*

⁷² Various circuit courts had previously upheld the validity of the CPPA: *U.S. v. Fox*, 248 F.3d 394 (5th Cir. 2001); *U.S. v. Mento*, 231 F.3d 912 (4th Cir. 2000); *U.S. v. Acheson*, 195 F.3d 645 (11th Cir. 1999); *U.S. v. Hilton*, 167 F.3d 61 (1st Cir. 1999).

⁷³ *Ashcroft* at 240

prohibitions and the affront to community standards prohibited by the definition of obscenity.”⁷⁴ The statute was upheld in *Ferber* because of the intrinsic connection between the speech at issue and the crime it results from i.e. you cannot create child pornography without exploiting and abusing a child in the process.⁷⁵

The Court rejected the Government’s argument that this imagery can actually lead to real life instances of child sexual abuse, finding that causal link to be too indirect.⁷⁶ The Court also rejected the Government’s argument that the *Ferber* court found all child pornography by definition to be illegal.⁷⁷ The Court instead found that *Ferber* found some works that fall into this category of material to still carry some significant value, but that this can be remedied through alternative methods such as virtually produced images.⁷⁸ The Court rejected the Government’s argument “that its objective of eliminating the market for pornography produced using real children necessitates a prohibition on virtual images as well”, finding that photographers would not risk prosecution by creating real child pornography if virtual images were actually identical to real images.⁷⁹

Ultimately, the Court refused to create a categorical ban on the material despite the potential for harm to children because the crux of the criminalization of child pornography stems from the inherent harm posed to children in its production which is not comparable to the harm posed to children by simulated materials. “The Government cannot ban speech fit for adults simply because it may fall into the hands of children. The

⁷⁴ *Id.* at 248

⁷⁵ *Id.* at 249-250

⁷⁶ *Id.* at 250

⁷⁷ *Id.* at 251

⁷⁸ *Id.*

⁷⁹ *Id.* at 254

evil in question depends upon the actor's unlawful conduct, conduct defined as criminal quite apart from any link to the speech in question.... The objective is to prohibit illegal conduct, but this restriction goes well beyond that interest by restricting the speech available to law-abiding adults.”⁸⁰ Although the Court was willing to admit that such material may be “closer” to real child pornography as these images appear to portray an identifiable minor partaking in sexual activities, the Court was unwilling to establish that such material is close *enough* to also constitute unprotected speech under the First Amendment.⁸¹

Justice Thomas wrote a separate concurrence highlighting the problem that technological innovations may pose on this opinion:

“At this time, however, the Government asserts only that defendants *raise* such defenses, not that they have done so successfully. In fact, the Government points to no case in which a defendant has been acquitted based on a “computer-generated images” defense.... this speculative interest cannot support the broad reach of the CPPA, technology may evolve to the point where it becomes impossible to enforce actual child pornography laws because the Government cannot prove that certain pornographic images are of real children. In the event this occurs, the Government should not be foreclosed from enacting a regulation of virtual child pornography that contains an appropriate affirmative defense or some other narrowly drawn restriction.”⁸²

D. CPPA

There were two provisions of the CPPA at dispute in *Ashcroft*: Section 2256(8)(B) and Section 2256(8)(D).

§ 2256(8)(B) prohibits: “any visual depiction, including any photograph, film, video, picture, or computer or computer-generated image or picture, whether made or

⁸⁰ *Id.* at 252-53

⁸¹ *Id.* at 242

⁸² *Id.* at 259

produced by electronic, mechanical, or other means, of sexually explicit conduct” that “is, or appears to be, of a minor engaging in sexually explicit conduct.”⁸³

§ 2256(8)(D) prohibits any sexually explicit image that was “advertised, promoted, presented, described, or distributed in such a manner that conveys the impression” it depicts “a minor engaging in sexually explicit conduct”.⁸⁴

IV. Circuit Split

Following *Ashcroft*, several circuit courts remain split on the issue of synthetic child pornography, with a majority leaning toward eliminating First Amendment protection of such content.

A. Second Circuit

In 2011, the Second Circuit found that a defendant who superimposed images of the faces of minor female children onto nude and partially nude adult female bodies engaging in sexually explicit conduct was not protected by the First Amendment.⁸⁵ The court again highlighted that the interests of a real child are implicated because even if the bodies in the images do not belong to the children, the result is still such that the final product is only recognizable as a depiction of a minor partaking in sexual activity.⁸⁶ This harm was further exacerbated because the actual names of the minors were added to the media, further implicating the child’s image and reputation.⁸⁷ The court also recognized “the psychological harm of knowing that their images were exploited and prepared for

⁸³ 18 U.S.C.A. § 2256(8)(B)

⁸⁴ 18 U.S.C.A. § 2256(8)(D)

⁸⁵ *U.S. v. Hotaling*, 634 F.3d 725, 729-30 (2nd Cir. 2011) holding that, “sexually explicit images that use the faces of actual minors are not protected expressive speech under the First Amendment.”

⁸⁶ *Id.* at 730.

⁸⁷ *Id.*

distribution by a trusted adult” as another definite harm resulting from this type of media.⁸⁸ The Second Circuit’s holding suggests that the court is more likely to recognize a definite harm resulting from these synthetic images when the identity of the child whose image has been used is more likely to be recognized.⁸⁹

B. Sixth Circuit

The Sixth Circuit joined the majority of circuit courts in a 2012 decision.⁹⁰ In *Boland*, the court considered a case involving a lawyer who downloaded stock images of children and superimposed their faces onto the bodies of adults engaging in sexual activity.⁹¹ The attorney attempted to use before and after versions of these images in the defense of his clients on trial for child pornography, testifying that it would not be possible for someone who did not partake in the production of these images to know that they include actual children.⁹² The court rejected this argument, finding that the images resulted in many of the same reputational, emotional and privacy injuries that are caused by actual child pornography.⁹³ The court also highlighted the difference between morphed media of actual children and visual media that involves adults but appears to depict children partaking in sexual activities to achieve an artistic goal as mentioned by *Ashcroft*, such as *Romeo and Juliet*.⁹⁴ The court firmly declared that morphed images of actual children offer a weak expressive value and are not similarly protected by the First Amendment because they implicate the interests of identifiable children.⁹⁵

⁸⁸ *Id.*

⁸⁹ *Id.*

⁹⁰ *Doe v. Boland*, 698 F.3d 877, 882 (6th Cir. 2012).

⁹¹ *Id.*

⁹² *Id.* at 880

⁹³ *Id.* at 881

⁹⁴ *Id.* at 884

⁹⁵ *Id.*

C. Fifth Circuit

The Fifth Circuit was the latest to join the majority of the circuit courts in 2020.⁹⁶ In *Mecham*, the defendant was indicted after over 30,000 pornographic files were seized from his home.⁹⁷ All of the images featured the defendant's four granddaughters as he admitted to superimposing their faces onto other media depicting actual adults partaking in sexual activity.⁹⁸ The court rejected his argument that such content was protected by the First Amendment, finding that because the morphed images feature an identifiable child, they pose the same psychological, reputational and emotional harm to children as real child pornography.⁹⁹

D. Eighth Circuit

One circuit court stands alone in its protection of such material. In 2014, the Eighth Circuit found that a defendant who superimposed a minor's face onto an image of two adults was afforded protection by the First Amendment.¹⁰⁰ The court rejected the Government's argument that the children were still exposed to harm because the images created a lasting record of an identifiable minor that appears to be engaged in sexually explicit behavior.¹⁰¹ The court reasoned that even though the depicted minor was identifiable, the speech was protected under precedent because the child's face was superimposed onto the nude bodies of adults.¹⁰² The court reasoned that such material

⁹⁶ *U.S. v. Mecham*, 950 F.3d 257, 267 (5th Cir. 2020).

⁹⁷ *Id.* at 260.

⁹⁸ *Id.*

⁹⁹ *Id.* at 267

¹⁰⁰ *U.S. v. Anderson*, 759 F.3d 891, 896 (8th Cir. 2014).

¹⁰¹ *Id.* at 894

¹⁰² *Id.* at 895

may only be prosecuted as child pornography when the underlying sexual activity depicted is itself a crime, more specifically the sexual abuse of a child.¹⁰³

V. Potential Solutions

As the reality begins to set in that artificial intelligence and machine learning are here to stay and infiltrate societies worldwide, there is a pressing need for the federal government to keep up with the increased risk of negative usage of these emerging technologies.¹⁰⁴ The next section aims to explore potential solutions for the federal government in order to combat the issue, including the enactment of a private right of action, criminalizing the production and distribution of such materials, and revisiting deepfakes through the lens of *Miller* and *Ashcroft*.¹⁰⁵

A. Federal Law Regulating Deepfakes

One solution is the enactment of federal legislation to regulate nonconsensual deepfakes of children. This section will analyze the laws of two states on this matter: California and Virginia, the former of which provides a private right of action against creators of deepfake media and the latter of which categorizes such an offense as a misdemeanor.

1. California

¹⁰³ *Id.*

¹⁰⁴ “As imaging technology improves... it becomes more difficult to prove that a particular picture was produced using actual children.” *Ashcroft* at 242.

¹⁰⁵ Some scholars have proposed the use of tortious claims, such as intentional infliction of emotional distress, as one venue for remedying the harm to victims of deepfakes. Mullen, *supra* note 9, at 232. An individual may succeed on a claim for IIED if they can prove: (1) intentional or reckless conduct; (2) extreme or outrageous conduct; (3) a causal connection between the wrongful conduct and the emotional distress; and (4) severe emotional distress. *Id.*

Section 1708.86(b) provides private right of action against any person who: (1) creates and intentionally discloses¹⁰⁶ sexually explicit material and the person knows or reasonably should have known the depicted individual in that material did not consent to its creation or disclosure; or (2) intentionally discloses sexually explicit material that the person did not create and the person knows the depicted individual in that material did not consent to the creation of the sexually explicit material.¹⁰⁷ § 1708.86 does not allow a defendant to exert a defense that the material included a disclaimer indicating that the inclusion of the depicted individual was not authorized or that the depicted individual did not partake in the creation of the material.¹⁰⁸ Notable is the definition of “consent” under § 1708.86(3)(A) under which a child by definition is unable to consent to their exploitation.¹⁰⁹

Thus, under the California law, an individual depicted in a pornographic deepfake can sue the creator or intentional distributor of the deepfake.¹¹⁰ The depicted individual is able to receive money damages for: (1) an amount equal to any monetary gain made by the defendant by creating or distributing the content, (2) damages for harm caused by the disclosure of the content, including emotional distress, (3) punitive damages, and (4) attorney’s fees and costs.¹¹¹

¹⁰⁶ “‘Disclose’ means to publish, make available, or distribute to the public” thus depicted individuals have a right of action against not just the individuals creating this material but also against websites hosting such material to dispel the harm caused by distribution. Cal. Civ. Code § 1708.86(a)(7)

¹⁰⁷ Cal. Civ. Code § 1708.86(b)

¹⁰⁸ Cal. Civ. Code § 1708.86(d)

¹⁰⁹ “Child pornography.... is child sexual exploitation.... The act creating the pornography was by force... and there was unequivocally no consent, legally or commonsensically speaking.” Whitney J. Gregory, *Honeypots: Not for Winnie the Pooh but for Winnie the Pedo-Law Enforcement’s Lawful Use of Technology to Catch Perpetrators and Help Victims of Child Exploitation on the Dark Web*, 26 Geo. Mason L. Rev. 259, 263 (2018).

¹¹⁰ *Id.*

¹¹¹ Cal. Civ. Code § 1708.86(e)

This California law sets out a framework that could prove to be a useful model for Congress to mimic in order to create federal protections for children and adults nationwide. Because of the Supreme Court’s precedent in *Ashcroft* that fails to prohibit synthetic child sexual assault material, a federal private right of action can serve as a powerful tool for victims of deepfakes across the country. The most pressing concern for the federal government should be the potential for detrimental impacts on children if they become the victims of such deepfakes especially considering the Court’s precedent acknowledging the Government’s right to limit civil liberties in order to protect children.¹¹²

2. Virginia

Under § 18.2-386.2, an individual commits a misdemeanor if he: “with the intent to coerce, harass, or intimidate, maliciously disseminates or sells any videographic or still image created by any means whatsoever that depicts another person who is totally nude, or in a state of undress so as to expose the genitals, pubic area, buttocks, or female breast, where such person knows or has reason to know that he is not licensed or authorized to disseminate or sell such videographic or still image”.¹¹³ “‘Another person’ includes a person whose image was used in creating, adapting, or modifying a videographic or still image with the intent to depict an actual person and who is recognizable as an actual person by the person's face, likeness, or other distinguishing characteristic.”¹¹⁴

Under the Virginia law, individuals who share nonconsensual deepfake pornography are subject to a misdemeanor and face up to twelve months in jail and a \$2,500

¹¹² “The power of the state to control the conduct of children reaches beyond the scope of its authority over adults.” *Prince v. Mass* at 170

¹¹³ Va. Code Ann. § 18.2-386.2(A)

¹¹⁴ *Id.*

fine.¹¹⁵ Section 18.2-386.2 functions as an amendment to preexisting Virginia laws criminalizing revenge porn.¹¹⁶ Although Virginia’s law serves as another potential model for the federal government to consider in its development of anti-deepfake legislation, it appears to fall short when considering the severity and longevity of negative consequences to victims of sexual abuse and exploitation, especially children.¹¹⁷

3. *Ashcroft* Revisited

One of the primary reasons why the Supreme Court found virtual child pornography to be protected by the First Amendment but the opposite for actual child pornography is because the latter inherently involves the abuse and exploitation of children.¹¹⁸ Child pornography poses the risk that the child may be followed by the material for an indefinite period of time, including adulthood.¹¹⁹

Despite a failure to categorically create an exception, the *Ashcroft* Court was willing to acknowledge that morphed child pornography causes harm to the real interests of children that is closer to the harm caused by actual child sexual assault material.¹²⁰

Several circuits courts have since recognized these harms – including emotional, mental,

¹¹⁵ Wang, *supra* note 52, at 427.

¹¹⁶ Davis, Wiggins & Donovan, *supra* note 4, at 8

¹¹⁷ “The permanent record of a child’s sexual abuse can alter his or her [life] forever. Many victims of child pornography suffer from feelings of helplessness, fear, humiliation, and lack of control.” *Child Pornography*, U.S. Dep’t of Just., <https://www.justice.gov/criminal/criminal-ceos/child-pornography>

¹¹⁸ *Ferber* at 249-250

¹¹⁹ “Pornography poses an even greater threat to the child victim than does sexual abuse or prostitution. Because the child’s actions are reduced to a recording, the pornography may haunt him in future years, long after the original misdeed took place. A child who has posed for a camera must go through life knowing that the recording is circulating within the mass distribution system for child pornography.” *Shouvlín, Preventing the Sexual Exploitation of Children: A Model Act*, 17 Wake Forest L.Rev. 535, 545 (1981).

¹²⁰ *Ashcroft* at 242

physiological, and reputational harm – as sufficient to deny First Amendment protection to morphed child pornography.¹²¹

Deepfakes pose a unique challenge to Supreme Court precedent as they appear to create a new and almost closer category to actual child pornography, beyond even the harms of morphed child pornography due to the rise of easily accessible deepfake technology.¹²² High schools across the country are already experiencing the challenges of this technology as underage students continue to become victims to this rising form of sexual exploitation.¹²³ Some scholars suggest that even though the depicted children in altered images may not have to reexperience an actual sexual assault, they still are exposed to the ongoing risk of the circulation of their images to an unlimited amount of people online, generating another form of victimization.¹²⁴

One of the issues that the Court grappled with in *Ashcroft* was the overbreadth and vagueness of the CPPA language.¹²⁵ To remedy this, the CPPA, or similar legislation, could be amended to include “deepfake” images explicitly as opposed to what “appears to be” and what “conveys the impression”.¹²⁶ Considering their easy production and stunning accuracy, it seems more likely that the Supreme Court, when faced with a case of child nonconsensual deepfake pornography, would find that such material has a

¹²¹ See *U.S. v. Hotaling*, 634 F.3d 725, 729-30 (2nd Cir. 2011); *Doe v. Boland*, 698 F.3d 877, 882 (6th Cir. 2012); *U.S. v. Mecham*, 950 F.3d 257, 267 (5th Cir. 2020).

¹²² Davis, Wiggins & Donovan, *supra* note 4, at 1: “even technologically unsophisticated actors are now able to create and distribute deepfakes.”

¹²³ Tate Ryan-Mosley, *A high school’s deepfake porn scandal is pushing US lawmakers into action*, MIT TECHNOLOGY REVIEW (Dec. 1, 2023).

¹²⁴ Taylor Comerford, *No Child Was Harmed in the Making of This Video: Morphed Child Pornography and the First Amendment*, 62 B.C.L. Rev. E-Supplement II. -323, II. -343 (2021).

¹²⁵ *Ashcroft* at 243

¹²⁶ *Id.*

detrimental impact to children with no other meaningful purpose besides the sexual exploitation of children and thus would fail to be covered by First Amendment protection.

Justice Thomas' concurrence in *Ashcroft* also seemingly becomes more relevant within the context of deepfake technology. Justice Thomas highlighted that technology may evolve to the point where even the Government is not able to prove that a pornographic image depicts an identifiable child.¹²⁷ Some scholars have suggested that the existence of deepfakes creates a new issue for courtrooms nationwide as they raise questions of authentication of evidence and detection of such materials in the courtroom.¹²⁸ While some technology has been trained to detect signs of deepfakes using biometric data and anomaly detection models, avid creators of deepfake material continue to subvert such efforts through the use of other technology created to evade detection.¹²⁹ While deep-learning experts continue to expand the use of the technology, the American legal system unsurprisingly falls behind in its regulation, leaving children the most vulnerable.

4. Child Porn Deepfakes Under the *Miller* Obscenity Test

One potential solution is to challenge the deepfakes and other similar simulations of child pornography as obscene under the *Miller* test. To test whether a work is obscene, the *Miller* test asks (a) whether the average person, applying contemporary community standards, would find that the work, taken as a whole, appeals to the prurient interest; (b) whether the work depicts or describes, in a patently offensive way, sexual conduct

¹²⁷ *Ashcroft* at 259

¹²⁸ Mullen, *supra* note 9, at 224 . “Deepfakes can be quite difficult to identify as inauthentic because they generally use genuine footage, crisp audio, and are popularly shared to hundreds of thousands, if not millions of people via social networks.” *Id.* at 214

¹²⁹ *Id.* at 229

specifically defined by the applicable state law; and (c) whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value.¹³⁰

Although the *Miller* Court was unwilling to find adult pornography obscene, it is possible that the Government can challenge deepfake depictions of children engaging in sexual acts on the basis that they are obscene.¹³¹ Considering how close and accurate the depictions shown in deepfakes can be while also recognizing the preexisting free speech exception for child pornography, a jury may find that deepfakes depicting children engaging in sexual activity are obscene as defined in *Miller*.¹³² Considering the mechanisms involved in producing a deepfake, it is likely that the average person would not be able to distinguish between an image of a virtually-created child partaking in sexually explicit activity, an image of an identifiable child superimposed onto the bodies of actual adults engaging in sexual activity, or an actual child whose sexual abuse has been documented.¹³³

Under the second prong of the obscenity analysis, the *Ashcroft* court focused on the CPPA's potential prohibition on creative and literary works that depicted themes regularly accepted by society, such as teenage lovers consummating their relationship in a

¹³⁰ *Miller* at 24

¹³¹ "By using the standards set down in *Miller v. California*, the government should be able to prosecute the producer of hard core depictions of child pornography even if such portrayals are virtual. A depiction of a child, particularly a young one, performing a sexual act or a lewd exhibition of their genitals would most likely never be found to have serious literary, scientific, or artistic value. A jury may very well find that this type of hard-core child pornography, when taken as a whole, appeals to the prurient interest. It is arguably material that plays no part in the public debate due to the fact that it is only viewed by pedophiles. Such a strategy is most likely to succeed when the materials at issue involve explicit sex acts and contain virtual depictions of very young children." Ryan P. Kennedy, *Ashcroft v. Free Speech Coalition: Can We Roast the Pig Without Burning Down the House in Regulating "Virtual" Child Pornography?*, 37 Akron L. Rev. 379 (2004).

¹³² The *Ferber* court created one of the exceptions for free speech protection when it found that visual depictions of children engaging in sexual activity are not protected under the First Amendment. *Ferber* at 764

¹³³ Davis, Wiggins & Donovan, *supra* note 4, at 8

juvenile love story.¹³⁴ It is difficult to imagine, however, that a jury would find a deepfake video depicting extremely accurately visual representations of actual minors engaging in sexually explicit conduct to be as socially acceptable as a movie like *Romeo and Juliet*.¹³⁵

Even if this type of material was considered to be a creative liberty by production, a prosecutor would likely be successful in convincing a jury that machined-generated images of six-year-old children engaged in hardcore sexual activity do not fall within the scope of pervasive themes that are accepted by the larger American society, especially as most courts have demonstrated reluctance to identify any artistic value to this type of media.¹³⁶

Further, various circuit courts have already grappled with this question, and are of the shifting opinion that these modern types of synthetic child pornography indeed carry many of the harms that justified the First Amendment exception for actual child pornography in the first place.¹³⁷

VI. Conclusion

In the ever-evolving landscape of deepfake technology, the path toward safe use will require heightened regulation by the federal government, especially to protect children from sexual exploitation. As society becomes more dependent on technology, lawmakers, technology experts and parents alike have an increased responsibility to

¹³⁴ *Ashcroft* at 247

¹³⁵ *Id.*

¹³⁶ The *Ferber* court recognized that “the value of permitting live performances and photographic reproductions of children engaged in lewd sexual conduct is exceedingly modest, if not de minimis. We consider it unlikely that visual depictions of children performing sexual acts or lewdly exhibiting their genitals would often constitute an important and necessary part of a literary performance or scientific or educational work.” *Ferber* at 762-63.

¹³⁷ “By using identifiable images of real children.... morphed child pornography implicates the reputational and emotional harm to children that has long been a justification for excluding real child pornography from the First Amendment.” *U.S. v. Machem*, 950 F.3d 257, 265 (5th Cir. 2020).

ensure the safe upbringing of children online. Deepfake nonconsensual child pornography is one of many growing risks to children online and the issue requires proactivity from lawmakers to carefully balance First Amendment protections with a need to prevent child victims from the harm of a permanent digital footprint that can subject them to a lifetime of victimization.