Face/Off: "DeepFake" Face Swaps and Privacy Laws

By: Erik Gerstner

Erik Gerstner is an associate at David, Kamp & Frank, L.L.C. in Newport News, Virginia. Erik received his JD from William & Mary Law School in 2018. He focuses his practice on civil litigation, including business litigation, real estate, personal injury, and reputational injury. This article has been expanded from one the author published in For The Defense.

IN 2018, a curious trend spread rapidly across the Internet: people posting videos of Nicholas Cage's performances in various Hollywood films.[1] To the uninitiated viewer, these videos might appear to be nothing special, just various facets of Cage's prolific career. However, closer inspection would reveal a subtler thread running throughout these clips: none of these performances actually involved Cage. Rather, thanks to relatively new artificial intelligence (AI)-powered software programs colloquially known as DeepFakes, Internet users had seamlessly inserted Cage's face over the faces of the original actors in these scenes, making it appear as though Cage had always portrayed those characters.[2]

Nicholas Cage's central role in DeepFake videos is fitting, given his starring role alongside John Travolta in 1997's *Face/Off*, a film in which his and Travolta's characters both end up wearing the other's faces throughout the film. Although it was only a fanciful Hollywood invention in 1997, face swapping technology entered the mainstream in 2017. In August of that year, University of Washington researchers released a video, seemingly of Barack Obama, discussing topics such as terrorism, fatherhood, and job creation, which had been created using machine

[1] John Maher, *This was the year of the deepfake Nicolas Cage meme*, THE DAILY DOT (Dec. 27, 2018), available at https://www.dailydot.com/unclick/nicolas-cage-memes-deepfakes-2018.

[2] *Id.*

learning algorithms.[3] By 2018, similar tools became publicly available, with the most popular, called FakeApp, available for free online. FakeApp was developed using Google's open-source deep learning software. In its first two months of being publicly available, it was downloaded more than 120,000 times.[4]

The many-faceted ramifications stemming from the widespread availability of this and other similar software are staggering. The unprecedented ability to create fabricated messages from politicians and other celebrities, or "fake news" in the parlance of our current political climate, is a major concern - the Pentagon alone has already spent tens of millions of dollars in an effort to research and combat DeepFakes.[5] However, although the political and cultural ramifications of DeepFakes are significant, and worthy of considerable attention across the spectrum of areas that they affect, this article will be limited to primarily examining the legal issues likely to arise from these programs, including privacy, the right to one's own likeness, and defamation/false light claims. The first section will discuss the software and the technology behind it, including a brief introduction to how it technically functions. Next, this article will discuss the state of the relevant law and examine how face swaps have and will continue to intersect with applicable statutory and case law. Finally, it will discuss potential judicial and legislative solutions to present and future problems arising from these sorts of AI technologies.

I. Fakeapp and Machine Learning

The influx of fake videos stems largely from the widespread availability of simple yet powerful software tools such as FakeApp. Utilizing machine learning to train AI, it condenses what would be an exceedingly complex operation for even the most experienced digital artists into a single button press to create a face swapped video. While having a moderately powerful computer is a slight barrier to the effective usage of the program, it otherwise is relatively uncomplicated to create fake media

[3] Jennifer Langston, *Lip-syncing Obama: New tools turn audio clips into realistic video*, UNIVERSITY OF WASHINGTON NEWS (July 11, 2017), available at http://www.washington.edu/news/2017/07/11/lip-syncing-obama-new-tools-turn-audio-clips-into-realistic-video.

[4] Kevin Roose, *It Was Only a Matter of Time: Here Comes an App for Fake Videos*, NEW YORK TIMES (Mar. 4, 2018), at A1, available at https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html.

[5] Dan Robitzski, *Pentagon's AI Director Calls for Stronger Deepfake Protections*, FUTURISM, (Aug, 30, 2019), available at https://futurism.com/the-byte/pentagon-ai-director-deepfake-protections.

with it.[6] In its most basic form, all a user needs is a "base" video and a number of source images of the face of the person being pasted into the video. The more source images input into the program, the more seamless the final video will appear.[7]

After creating the datasets, FakeApp then trains the deep learning algorithm, a process that can take hours or even days, depending on how powerful a computer is used and the quality sought for the final video. Thereafter, the user needs only to click one more button to create the resulting video. A more experienced creator may be able to achieve a higher degree of realism through more involved interaction with the FakeApp software, but by following the basic steps, even a novice can fairly easily create a face swap using the program.[8]

While this process is straightforward for the front-end user, it is anything but for the computer running FakeApp.[9] The software utilizes Google's open source TensorFlow machine learning algorithm to power its operations and requires considerable computing power from any hardware on which it runs. The final video quality is determined by a combination of factors, including the similarity of the faces and poses among the base video and the source images, and the amount of time spent and quality of the AI training. What is not a factor, however, is the software itself – computer-generated faces were once strictly the domain of big-budget studios with deep pockets, proprietary software tools, and considerable amounts of time. For example, the much-discussed appearance of a computer-generated young Carrie Fisher in *Star Wars: Rogue One* in 2016 was the product of a $200 million production budget, and, according to the visual effects supervisor, "a super high-tech and labor-intensive version of doing makeup."[10] Now, private individuals are able to create videos equaling or even surpassing those created by these studios for a tiny fraction of the time and expense.[11] Machine learning is the great equalizer: thanks to the powerful AI algorithms powering

---

[6] Roose, *supra* note 4.

[7] *Id*.

[8] *Id.*

[9] Note that other face swap AI programs do not necessarily operate in the same way. For example, the University of Washington algorithm is considerably more in-depth, learning what shapes mouths make when vocalizing certain sounds, then creating video from whole cloth to match a given audio track, rather than simply superimposing one face over another in an existing video. *See* Langston, *supra* note 3.

[10] B.J. Murphy, *Reddit user outperforms Disney with AI-generated Princess Leia*, GRAY SCOTT (Jan. 25, 2018), available at https://www.grayscott.com/seriouswonder-//reddit-user-outperforms-disney-with-ai-generated-princess-leia.

[11] *See*, *e.g.*, *id.*

FakeApp and similar software, those able to make full use of it have had the power at their fingertips enhanced exponentially, a process of growth that is likely to continue as this technology continues to progress.[12]

II. Privacy, Defamation, and Fake News: The Present State of the Law

There are potential ramifications flowing from the creation and use of the resulting media that span the legal spectrum, including ramifications in election law, criminal law, evidence, and intellectual property. This article, however, will focus on potential privacy issues, including defamation, false light, and the right of publicity, also sometimes known as personality rights.[13] While case law and statutory law regarding deep fakes currently is scant or nonexistent, there are analogues which may provide some guidance as to how the law in the United States will address issues arising from these new technological advancements.

One potential legal concern flowing from these fake images is defamation. A defamation cause of action could arise from an individual using FakeApp or similar software to create a fake video of an individual saying or doing something that would injure the individual's reputation if it were true. For example, in the aforementioned video the University of Washington created of President Obama, the audio could be any recording the creator wants to use, literally putting words of the creator's choosing into Obama's mouth, including statements that could be highly offensive to an unsuspecting viewer.[14] In states that recognize a difference between slander and libel, a face swapped video could easily give rise to both of these causes of action. For example, if someone creates a video purportedly showing Person A saying defamatory things about Person B, then Person B might have a claim for slander (as the defamatory statements were verbal), while Person A might have a cause of action for libel.

---

[12] Joe McKendrick, *More artificial intelligence, fewer screens: the future of computing unfolds*, ZDNET (Sept. 9, 2017), available at http://www.zdnet.com/article/artificial-intelligence-the-new-user-interface-and-experience.

[13] *See*, *e.g.*, Benny Evangelista, *If you think fake news is bad, fake video is coming*, SAN FRANCISCO CHRONICLE (March 14, 2018), available at https://www.sfchronicle.com/business/article/If-you-think-fake-news-is-bad-fake-video-is- 12751052.php; Ari Breland, *Lawmakers worry about rise of fake video technology*, THE HILL (Feb. 19, 2018), available at http://thehillcome/policy/technology/374320-lawmakers-worry-about-rise-of-fake-video-technology.

[14] Langston, *supra* note 3.

While fake media of the sort described above would seem to satisfy the requirements for a defamation claim, there is no common law jurisprudence either way for a claim resulting from such a created video. There is some precedent, though: in some states, the tort of defamation explicitly applies to altered still images.[15] Generally speaking, videos are treated the same as still images under the law.[16] Thus, a defamatory video should be fully actionable if a plaintiff attempts to bring suit. However, there are powerful affirmative defenses for defamation claims that apply in many cases and make it difficult for plaintiffs to win these lawsuits.[17] The primary defense for a content creator is a claim that their face swapped video is parody, which may be an absolute defense in defamation suits.[18]

Defamation is by its nature mutually exclusive of parody. By definition, defamation requires a false statement of fact; parody, to the degree that it is perceived as parody by its intended audience, conveys the message that it is not the original and, therefore, cannot constitute a false statement of fact.[19]

As the above passage states, for something to be considered parody, it must be perceived as parody (and thus not as a statement of fact) by its audience.[20] In other words, it must not "*reasonably* be understood as describing actual facts. . . or events" (emphasis added).[21] This analysis would be applied on a case-by-case basis, and different finders of fact may come out with dramatically different results based on the facts of each individual case.

Face swapped videos are also likely to trigger the common law tort of intentional infliction of emotional

---

[15] *See*, *e.g.*, Kiesau v. Bantz, 686 N.W.2d 164, 178 (Iowa 2004) (holding that an altered image depicting a female police officer in uniform, standing in front of her official vehicle, with her breasts exposed, was libelous per se); Morsette v. "The Final Call," 764 N.Y.S.2d 416 (N.Y. App. Div. 2003) (upholding jury verdict of libel stemming from altered image of woman in newspaper making it appear that she was a convict).

[16] *See*, *e.g.*, Arizona v. Steinle, 372 P.3d 939, 945 (Ariz. 2016) (stating that for the purposes of evidence, requirements for admission of video evidence should be the same as for a photo).

[17] One common defense in modern defamation cases is the *New York Times v. Sullivan* standard, which applies in cases involving the defamation of public figures, who must prove that publishers of defamatory content did so with a mens rea of "actual malice." New York Times Co. v. Sullivan, 376 U.S. 254 (1964). However, even for public figures, this defense would offer no sanctuary because as face swapped videos are intentionally created to be fake representations of real people, their creators have actual knowledge of the falsity, with intent going well beyond the actual malice standard.

[18] Hustler v. Falwell, 485 U.S. 46, 46 (1988) (holding that trial court properly dismissed plaintiff's defamation claim because content in question was ruled to be parody).

[19] 50 AM. JUR.2d, *Libel and Slander* § 156 (2018).

[20] *Id.*

[21] *Hustler*, 485 U.S. at 46.

distress (IIED) - conduct that causes severe emotional trauma in a victim. Frequently a defamation claim is accompanied by an IIED claim, and they are often adjudicated similarly, though with certain key differences.[22] Unlike with defamation, when the victim of an IIED claim is a private figure, there is no need to analyze whether a faked video would qualify as a false statement, for the only concern with regard to IIED is conduct, false or not. Parody is not an absolute defense against IIED either, as parody, even if demonstrably not a false statement of fact, may still rise to the level of being patently offensive.[23] This is not a blanket rule, however. Similar to defamation, IIED claims against public figures do require that there be a false statement of fact, made with "actual malice" – that is, with knowledge that it was false or with reckless disregard as to whether or not it was true.[24] With that said, any creator of a faked video would by definition have actual knowledge that the statements made in the video and attributed to the depicted individual are false, meeting this *mens rea* requirement even for public figures. Note too that a faked video could potentially give rise to IIED suits from multiple parties stemming from the same video: both the purported individual depicted in the video, and any recipients who may be shocked by the things the person in the video is saying or doing.[25]

On the other hand, for any defense to an IIED claim should focus first on the intent element. While there will clearly be intent in the creation of the media itself, in many cases it is unlikely that a court will find actual intent *to cause emotional distress*.[26] In states where this is the sole element, this may prove to be a bar to IIED claims. A court could very easily hold, however, that the very creation of a face swapped video, particularly an unflattering one, is by its very nature likely to result in emotional distress if published, and thus it is per se reckless. This uncertainty makes an IIED claim less favorable for a plaintiff in the context of face swaps, but still a valid option in some cases for those victimized by this technology.

---

[22] *See*, *e.g.*, *Hustler*, 485 U.S. at 46; Rykowsky v. Dickinson Public School Dist. No. 1, 508 N.W.2d 348 (N.D. 1993); Barker v. Huang, 610 A.2d 1341 (Del. 1992); Lewis v. Benson, 701 P.2d 751 (Nev. 1985).

[23] *Hustler*, 485 U.S. at 46.

[24] *Id.*; the *Sullivan* standard was enumerated in *Sullivan*, 376 U.S. at 279-280.

[25] *See*, *e.g.*, Dzamko v. Dossantos, 2013 WL 5969531 (Conn. Super. Oct. 23, 2013) (holding that both subject and recipient of false images had an actionable claim for IIED).

[26] This is especially true when it comes to faked pornography. *See* Emma Grey Ellis, *People Can Put Your Face On Porn—And The Law Can't Help You*, WIRED (Jan. 26, 2018), available at https://www.wired.com/story/face-swap-porn-legal-limbo.

False light is another tort claim, though many states do not recognize it as a separate cause of action.[27] Like defamation, false light claims arise from the spread of falsehoods about a plaintiff that would be considered objectionable by a reasonable person. However, unlike with defamation, false light claims award damages based on the emotional harm they suffered from the spread of the falsehoods.[28] One of the four causes of action outlined by William Prosser in his seminal 1960 article, false light differs from defamation and IIED primarily because it is a privacy tort, in that it seeks to protect an individual from claims about them which are released to the public.[29] Note a key difference from defamation claims: for false light, truth is not an affirmative defense; rather, the burden is on the plaintiff from the outset to establish the false or misleading nature of the statement, making false light a more difficult cause of action compared to defamation, at least in some jurisdictions. Despite this, however, the analysis does not change much compared to the other two torts already discussed – because these videos are, by definition false, what might otherwise be a difficult hurdle for a plaintiff is met. The other elements are much easier to establish before a jury, and thus a false light claim, in those states which recognize it, would be a powerful potential avenue for a plaintiff harmed by appearing in a fake video.

Defamation, IIED, and false light are quite similar as far as their respective elements and the types of harms they seek to redress. Returning to Prosser's four categories of privacy torts, another of these that is likely to give rise to litigation stemming from face swapped videos: "[a]ppropriation, for the defendant's advantage, of the plaintiff's name or likeness," commonly referred to in the United States as the right of publicity.[30] As with false light, not every state recognizes the right of publicity, though it is more widespread across the country than false light. Because there is no federal scheme protecting this right, it varies by state. Nonetheless, in its most basic form, the tort applies when one "appropriates the commercial value of a person's identity by using without consent the person's name, likeness, or other indicia of identity

---

[27] *Defamation vs. False Light: What is the Difference?*, FINDLAW, available at http://injury.findlaw.com/torts-and-personal-injuries/defamation-vs--false-light--what-is-the-difference-.html.

[28] *See False Light*, LEGAL INFORMATION INSTITUTE, available at https://www.law.cornell.edu/wex/false_light.

[29] William L. Prosser, *Privacy*, 48 CALIF. L. REV. 383, 389 (1960).

[30] *Id*. at 389.

for purposes of trade."[31] In many states, such as California, this applies not only to an individual's actual image, but also voice, signature or other types of likenesses. In Indiana, these protections are extended to, among other things, distinctive appearance, gestures or mannerisms as well.[32]

In those states which do recognize the right, there are considerable differences in how they approach it. Some, like California, treat the right of publicity similar to a property right. Others, such as New York, address it as a privacy right, more along the lines of Prosser's original incarnation.[33] Thus, depending on where the suit is filed, the results may vary considerably. With that said, the right of publicity is one very likely to be invoked by victims of face swap exploitation, so long as there is some commercial value involved in the end usage of the media.[34] It is irrefutable that a face swap, by its very nature, captures the likeness of its subject. If, for example, the video of Obama previously discussed were changed to endorse a particular fast food restaurant, or car model or clothing brand, this would clearly meet the elements of a right of publicity claim. Depending on the level of a person's celebrity, as well as how their likeness is being used, a right of publicity suit can be quite lucrative. For example, in 2015 NBA legend Michael Jordan was awarded $8.9 million after an Illinois court found that a local supermarket chain had violated his right of publicity. [35] Because of these potential judgments against face swap creators, the right of publicity is a strong deterrent against media creators with deep pockets, such as companies that might be tempted to use a celebrity's likeness for their own commercial ends. It is a powerful tool in any individual's (and his/her attorney's) arsenal should they become the subject of a face swap gone viral.

The right of publicity has another, even stronger application as well: a potential cause of action against platforms hosting face swapped videos.[36] In those states

---

[31] RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 46 (AM. LAW INST. 2006).

[32] CAL. CIV. CODE § 3344 (West 1984); IND. CODE ANN. § 32-36-1-1 (West 2012).

[33] Jonathan Faber, *A Brief History of the Right of Publicity*, RIGHT OF PUBLICITY (July 21, 2015), available at http://rightofpublicity.com/brief-history-of-rop.

[34] Jesse Lempel, *Combatting Deep Fakes through the Right of Publicity*, LAWFARE BLOG (Mar. 30, 2018), available at https://www.lawfareblog.com/combatting-deep-fakes-through-right-publicity.

[35] Darren Rovell, *Supermarket chain must pay Michael Jordan $8.9 million for use of name*, ESPN (Aug. 21, 2015), available at http://www.espn.com/nba/story/_/id/13486052/supermarket-chain-pay-michael-jordan-89-million-use-name. For the court's discussion of this case, and how the right of publicity applies to the First Amendment, *see* Jordan v. Jewel Food Stores, Inc., 743 F.3d 509 (7th Cir. 2014).

[36] Lempel, *supra* note 34.

that recognize the right of publicity as an intellectual property right, plaintiffs potentially could sue Twitter, Facebook, YouTube, Reddit, or other websites hosting this content in addition to the creators themselves.[37] For the other potential causes of action discussed, any potential liability to platforms would be curtailed by the Communications Decency Act, which states that "no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."[38] However, the Act also creates an exception: it does not "limit or expand any law pertaining to intellectual property."[39] The application of this so-called "intellectual property exception" to the right of publicity has not yet been tested in courts, and there are a number of obstacles preventing it from becoming a viable option for plaintiffs moving forward.[40] The multibillion dollar companies that these lawsuits might target could argue persuasively that it is not fair for them to have to police the millions of posts their users create each day, especially given how seamless these fakes can appear.[41] A solution might be one similar to the Digital Millennium Copyright Act (DMCA), which allows the owner of intellectual property to request that hosting platforms remove it; if platforms do so, it absolves them of liability.[42] However, this would be a solution to be created through legislative means, rather than judicial.

III. FUTURE LEGAL APPROACHES TO FACE SWAPS

For many emerging technologies, the law has no immediate solution to the unique problems they pose. As a result, courts and attorneys must create solutions for these problems as they come across their respective dockets and desks. The efficacy of this process varies considerably, based on the field of law, type of technology, the specific court hearing the case, and a number of other factors as well. With copyright

---

[37] *Id.*

[38] 47 U.S.C. § 230(c)(1) (1998).

[39] *Id.* at § 230(e)(2).

[40] Lempel, *supra* note 34. Lempel describes three basic hurdles: (1) fitting the right of publicity into the intellectual property exception to begin with; (2) meeting the "commercial use" requirement for right of publicity claims; and (3) overcoming First Amendment hurdles. He concludes that there is significant precedent for right of publicity claims to succeed given the first two issues, and that while the First Amendment is a more difficult one, case law provides plaintiffs with some arguments here as well.

[41] Facebook alone sees 300 million photos uploaded per day. *The Top 20 Valuable Facebook Statistics*, ZEPHORIA DIGITAL MARKETING, available at https://zephoria.com/top-15-valuable-facebook-statistics.

[42] 17 U.S.C. § 512 (2010).

law, for example, many statutes are written for specific technologies, and are poorly suited to deal with newer emerging technologies in which copyrights may be held.[43] Cable television and other paid TV broadcasts are a useful case study: in the mid-1970s, FCC regulations for cable were based on the traditional transmission of signals, via ground wires or microwaves.[44] The advent of satellite transmissions were not addressed by these specifically-written regulations, however, and as they became more prevalent, courts struck down many of the prior regulations. This forced the FCC to scramble to create an entirely new regime for cable companies by the 1980s.[45] In many cases, courts may be better-equipped to be on the forefront of emerging technologies, as they are able to examine these technologies on a case-by-case basis, in real-time, as necessary. There is an argument that for face swaps the best solution is to permit courts and attorneys to create common law precedents in order to handle the risks and issues with the technology as it and its harms become more widespread across society and culture.

Courts, however, are not infallible, especially when it comes to dealing with new, unfamiliar subjects. In some cases, the common law approach results in inconsistent rulings across different jurisdictions or even within the same one. Copyright provides a useful example: peer-to-peer (P2P) file sharing sites such as Napster, Grokster, and Kazaa, which allow users to share files (including but not limited to music, videos, and other media) with other individuals, usually for free, via the internet. In 2001, the Ninth Circuit found Napster liable for contributory and vicarious copyright infringement for enabling this behavior.[46] Three years later, however, the same court found other P2P services, including Grokster and Kazaa, not liable for their users doing essentially the same thing as Napster's users had been doing – that is, sharing copyrighted material with each other.[47] These issues are exacerbated when different courts come to radically different conclusions about how to handle technology-

---

[43] Jessica D. Litman, *Copyright Legislation and Technological Change*, 68 OR. L. REV. 275, 277 (1989).

[44] *Id*. at 343.

[45] *Id.* at 343-346.

[46] A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004 (9th Cir. 2001).

[47] Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 380 F.3d 1154 (9th Cir. 2004). This decision was overturned a year later by the Supreme Court, which held that any individual who promotes their product as a way to infringe copyrights is liable for the resulting acts of infringements. Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 545 U.S. 913 (2005). The comparison, however, holds – the Ninth Circuit was unable to properly resolve this technology issue, even though it had already done so once in the past.

related issues. A recent example of this came from conflicting decisions regarding the Fourth Amendment and its applications to data stored on foreign servers, as the Second Circuit and the Eastern District of Pennsylvania came to opposite conclusions regarding the United States' jurisdiction over this data.[48] Practically speaking, this means that for more complex issues, potential plaintiffs and defendants alike may be left wondering exactly what the law is, depending on the venue in which their case is proceeding.

On top of circuit splits, courts can at times have difficulty grasping the implications of newer technologies, and prior precedent may not provide an adequate framework for evaluating new technologies. The Supreme Court has grappled with these issues a number of times in the realm of emerging technologies and the Fourth Amendment. In *US v. Jones*, for example, the Court ruled that law enforcement officials must obtain a warrant to track a private individual's movements via Global Positional System (GPS) devices.[49] In her concurrence, Justice Sotomayor noted that the majority decision revolved around the constitutional maxim that "when the Government physically invades personal property to gather information, a search occurs."[50] However, "physical intrusion is now unnecessary to many forms of surveillance," and therefore the majority's "approach is ill suited to the digital age."[51] This is just one example. In cases where technology is advancing rapidly, the judicial solution may not be practical as the traditional methods of dealing with an issue may not be able to keep pace with the problems created by these advances.

While there are arguments for allowing the courts to handle face swaps, Congress and state legislative bodies can also pass proactive laws to deter abuses of the technology and have a process already in place to deal with them if and when they do become a more serious issue. Indeed, bills regulating DeepFakes were introduced into both the U.S. Senate and House in 2019, although neither have been enacted into federal law as of this writing. Multiple states, including Virginia and Texas, have enacted laws addressing DeepFakes,

---

[48] Lucy Bertino, *Courts Continue to Split on the Fourth Amendment in Cyberspace*, NORTH CAROLINA JOURNAL OF LAW & TECHNOLOGY (Feb. 22, 2017), available at http://ncjolt.org/circuit-split-4th-amendment-cyberspace/. This split was shortly thereafter resolved by the CLOUD Act, H.R.4943, which grants the United States and foreign parties access to data stored around the globe. David Ruiz, *Responsibility Deflected, the CLOUD Act Passes*, ELECTRONIC FRONTIER FOUNDATION (Mar. 22, 2018), available at https://www.eff.org/deeplinks/2018/03/responsibility-deflected-cloud-act-passes.

[49] U.S. v. Jones, 565 U.S. 400 (2012).

[50] *Id.* at 414.

[51] *Id.* at 414, 417.

albeit in different contexts.[52] As discussed briefly earlier, these can target two different but equally important groups: private video creators and the websites hosting the content themselves.

For individual video creators, the easiest step is to simply modify existing laws to explicitly address face swaps and other fake media. Unfortunately, one of the common uses for DeepFakes is the creation of fake pornography. Many states already have "revenge porn" statutes, which criminalize publishing intimate media of another person which they wish to remain private (or, as it is sometimes called, nonconsensual pornography).[53] In some states, these offenses are misdemeanors, in others, felonies. However, in most of these cases, the existing statutes do not address face swaps.[54] As these statutes demonstrate, there are some basic legal elemental and First Amendment issues stemming from the nature of these types of media: the people being supposedly depicted in a DeepFake are not the ones whose bodies are actually in the videos.[55] As discussed in the previous section, there is some question about this; certain types of claims may or may not be viable for plaintiffs victimized by face swap media.

A better approach, regardless of the issue being addressed, might be to draft new laws altogether which specifically regulate face swaps, regardless of who it depicts or for which end purpose it is created. For such a law to be effective, it may make more sense to adopt a strict liability standard for content creators instead one that makes it a crime to create *any* fake media depicting an individual without their consent. This would run into a bevy of First Amendment issues, and ultimately may not stand if

---

[52] Adi Robertson, *Virginia's 'revenge porn' laws now officially cover deepfakes*, THE VERGE, (July 1, 2019) available at https://www.theverge.com/2019/7/1/20677800/virginia-revenge-porn-deepfakes-nonconsensual-photos-videos-ban-goes-into-effect. Virginia's updated law, addressing DeepFakes in the context of nonconsensual pornography, can be found at VA. CODE. § 18.2-386.2 (2019). Texas' law, addressing DeepFakes in the context of election law, can be found at TEX. ELECTION CODE § 255.004 (2019). Note that these two exemplars differ dramatically, in both their approaches to the law and the problems they seek to address. This is an excellent demonstration of the shortfalls of legislative solutions: drafting comprehensive laws to address issues with wide-ranging effects is very difficult, especially if the causes and/or effects are not yet well-understood by lawmakers.

[53] Liz Crampton, *Taking New Steps to Put an End to "Revenge Porn"*, THE TEXAS TRIBUNE (Aug. 21, 2015) available at https://www.texastribune.org/2015/08/21/texas-law-criminalizing-revenge-porn-goes-into-effect/. 38 states, as well as Washington, DC, currently have revenge porn statutes in effect. *46 States + DC + One Territory have Revenge Porn Laws*, CYBER CIVIL RIGHTS INITIATIVE (last accessed Dec. 10, 2019), available at https://www.cybercivilrights.org/revenge-porn-laws.

[54] Ellis, *supra* note 26.

[55] *Id.*

challenged in court.[56] However, if properly narrowly tailored, such an approach would be the most effective option to combat creators of face swaps used for offensive purposes.

Another option is to target and make liable platforms hosting falsified media. These platforms are where the vast majority of falsified media would be shared and viewed. While the applications of such liability to the right of publicity have been touched on already, explicit, broad laws opening up these companies to liability for hosting fake media would likely cause a sharp decline in the availability of these videos. There are concerns, however: strict liability for media platforms based on content posted by unaffiliated users would essentially make them "arbiters of truth," a position of power in which legislators and judges alike are understandably wary to place private companies.[57] With that said, there are already tools – themselves ironically utilizing machine learning and artificial intelligence – available to detect face swapped media.[58] For content hosts like YouTube, which already deploy AI algorithms to detect various prohibited content such as copyright violations, it would not be too complicated to add face swap detection as well to the bots which scan each and every post uploaded to their servers.[59]

The final potential target of legislation would be the software used to create these programs.[60] By holding creators of these apps liable for their misuse, legislators would force the authors of these programs to severely restrict their distribution or monitor their use. This too is potentially problematic – moviegoers, such as those who saw and enjoyed *Rogue One*, would be the first to acknowledge the positive uses of face swap technology.[61] It is, in theory, not fair to restrict the spread of powerful technology simply out of fear of what malefactors may do with it, and it may very well be unconstitutional, due to the same First Amendment concerns that arise from criminalizing face swaps in general.[62] As with pursuing content hosts, there would likely be sweeping results. These would need to be balanced against a clearly compelling governmental interest in preventing American citizens from

---

[56] *Id.*

[57] Lempel, *supra* note 34.

[58] Ellis, *supra* note 26.

[59] *Id.*

[60] *Id.*

[61] Murphy, *supra* note 10. Indeed, in response to a proposed New York law which would criminalize creating "digital replicas" of people without their consent, the Motion Picture Association of America stated that such a law would restrict the ability of filmmakers to depict real people and events. Robertson, *supra* note 52.

[62] Damon Beres and Marcus Gilmer, *A guide to 'deepfakes,' the internet's latest moral crisis*, MASHABLE (Feb. 2, 2018) available at https://mashable.com/2018/02/02/what-are-deepfakes.

being victimized by face swap software. Whether a law could be narrowly tailored enough to overcome the First Amendment concerns would depend on both the legislative body attempting to do so and the court reviewing the resulting legislation, but it seems like a valid approach for lawmakers to attempt.

IV. CONCLUSION

Face swap software utilizing machine learning, whether FakeApp or otherwise, is a powerful and exciting tool, with valuable applications in entertainment and elsewhere. However, along with its potential for simple entertainment, such as ensuring that Nicholas Cage appears in every film ever created, there are considerable avenues for misuse of this technology as well. As the availability of the software has spread, so too have concerns resulting from its proliferation, especially videos exploiting the images of others. As the technology continues to improve, these issues will only continue to rise in profile and frequency. Ideally, legislatures and judiciary will attempt to confront these problems before they become major ones, but it is not a matter of if, but when face swaps will become too big an issue to ignore. And it is only a matter of time before attorneys are called upon to defend clients—whether content creators, content hosts, or others—from claims stemming from this technology.