WHAT CAN THE LAW DO ABOUT 'DEEPFAKE'?

Posted on March 30, 2018

Categories: Insights, Publications

Introduction

In November 2017, videos began appearing on www.reddit.com featuring celebrities' faces superimposed onto actors and actresses in pornographic videos (and a surprising number of random Nicolas Cage appearances, most of which are thankfully not pornographic). Such face-swapping videos were made possible by software technology that uses machine-learning algorithms to analyze photographs of a target subject, "learning" the map of the target subject's face (the more photographs, the more accurate the portrayal), and then superimposing the target subject's "face" onto a video, sometimes producing extremely convincing results. Such open source technology is commonly referred to as "deepfake technology", the term "deepfake" being a portmanteau of the concept of machine "deep learning" and the word "fake." In January 2018, a free, user-friendly app was launched that made deepfake technology accessible to the general public and to those without technological expertise.

Deepfake technology has useful applications, such as in creating satirical, humorous videos or aiding with motion picture special effects and video content production. However, the creation of realistic, non-consensual videos (e.g. causing a person's likeness to be "transposed" or "transplanted" onto the speech and/or motion of others) and the ease with which such videos may be made has set off alarm bells in the digital world. In response, several large social platforms have revised their user policies to forbid users from uploading or sharing such non-consensual videos created by deepfake technology.

Legal Recourse: Causes of Action

The law generally does not evolve as quickly as technology does. That being said, in the case of videos created by deepfake technology, several causes of action existing in our current laws may be applicable (or extended to be applicable) in addressing the wrongs committed by a person's misuse or abuse of deepfake technology. Below is a summary of some of such causes of action:

1.*Copyright Infringement*

Under the Copyright Act, R.S.C., 1985, c. C-42, a copyright owner has the sole right to produce or reproduce a work in any material form. In addition, the author of a work has moral rights, or the right to the integrity of the

work, and this right is infringed where the work is distorted, mutilated, or otherwise modified.[1] Undoubtedly, some deepfake videos, particularly those that are derived from copyrighted videos and photos, will infringe copyright laws based on the modification and republication of such videos and photos. Therefore, the owner of a video (onto which a "deepfake transplant" is done) and the owner of the photos (from which the "deepfake transplant" is sourced) may have remedies in copyright law which include injunctions, damages, and court orders to have the copies of the modified videos and photos destroyed. In addition, while the Copyright Act only provides copyright protection within Canada, similar laws exist in many other countries through those countries' ratification of international treaties such as the Berne Convention and the Agreement on Trade Related Aspects of Intellectual Property Rights.

2. Defamation

Words, acts, and audible or visible matters that would tend to lower or sully a person in the eyes of a reasonable person are known at law as acts of defamation. Some deepfake videos may create false statements of fact about a person's presence and actions that lead to a loss of reputation of that person. Persons defamed by such videos may be entitled to damage awards and in some cases injunctive relief to prevent the defamatory material from being further disseminated.[2] However, where there is a disclaimer that the video is fake or manufactured, an action for defamation may not succeed since it is no longer purporting the truth of a false statement (or action).[3] In addition, a Canadian court may not be the appropriate forum to hear a case, particularly if the publisher of the fake or manufactured video is located overseas.[4]

3. Violation of Privacy

Privacy is protected under common law torts and privacy legislation in Canada. Under s. 1(1) of the *BC Privacy Act,*[5] "it is a tort, actionable without proof of damage, for a person, willfully and without a claim of right, to violate the privacy of another." Other privacy legislation, such as PIPEDA and PIPA[6], protect information about an identifiable individual. However, deepfake videos, due to their false nature, may not technically create privacy issues, since the videos are not actually exposing any part of the victim's life or personal information.[7] That being said, victims in the province of Quebec may have a remedy under the Quebec Civil Code, which states that it is an invasion of privacy to use a person's name, image, likeness, or voice for a purpose other than the legitimate information of the public.

4. Appropriation of Personality

The tort of appropriation of personality arises where a person attempts to gain an economic advantage by using some aspect of another person's name, likeness, or personality without that person's consent. To be successful, the plaintiff must establish that the defendant appropriated his or her persona for economic gain. This is also enshrined in legislation such as the *BC Privacy Act*, which states that it is a tort to use a likeness, still

or moving, including a likeness deliberately disguised to resemble a person "for the purpose of advertising or promoting the sale of, or other trading in, property or services." In either case, this tort may be relevant depending on the context of how the deepfake video is used. For example, deepfake videos may be posted as part of commercial operations (i.e. on a commercial or social media site that itself makes money from advertisements or that allows the user to monetize from traffic or advertising on the post). Such use of deepfake videos for commercial or economic gain may therefore invite legal claims under this cause of action.

5. *Criminal Code*

There are several ways that Canadian criminal law may apply. First, there are broad, robust child pornography laws that would directly capture any pornographic deepfake videos depicting any person under the age of 18. In past criminal cases where an accused superimposed the face of a child onto pornographic pictures, the accused was charged with making and possessing child pornography.[8] This would be no different for deepfake videos involving minors, and if posted, there may be an additional charge of distribution of child pornography.

Second, the *Criminal Code* was amended in March 2015 to criminalize the publication of an intimate image without consent, sometimes known as "revenge pornography." Under this law, an "intimate image" is defined as a visual recording of a person in which the person is nude, exposing intimate regions, or engaged in explicit sexual activity. It will, however, remain to be seen whether a face that is swapped into such a visual recording could fall under this definition.[9]

In addition, where deepfake videos are used for illicit purposes such as blackmailing or fraudulent purposes, this would clearly fall under criminal provisions against extortion or fraud. Where the videos are used to harass a person in such a way as to cause the person to reasonably fear for their safety, this would constitute criminal harassment.

6. *Human Rights Complaint*

Where deepfake videos arise in the workplace, it may be possible to file a discrimination complaint under provincial or federal human rights legislation. In the province of British Columbia, any discrimination, harassment, or bullying related to a prohibited ground[10] in the *Human Rights Code* is a claim that can be filed with the BC Human Rights Tribunal (the "BCHRT"). For example, a pornographic deepfake video would constitute sexual harassment, which is a form of discrimination related to the prohibited ground of sex. The BCHRT has the authority to provide a compensation order for any lost wages or expenses incurred as a result of the contravention, and to provide a compensation award for "injury to dignity, feelings and self respect."

7. *Intentional Infliction of Mental Suffering*

While this tort is typically seen in the employment law context, the tort of intentional infliction of mental suffering arises out of flagrant conduct that is calculated to produce harm and that does produce harm, often in the form of a visible and provable psychiatric illness or mental distress. The obvious difficulty is proving that the creation of the deepfake video was indeed calculated to produce harm, as the videos may have been created for the pleasure of the creator rather than to cause emotional distress. In addition, the victim may not be able to pursue a legal remedy in the absence of incurring such illness or mental distress.

8. *Harassment*

In 2017, an Ontario court affirmed the existence of the tort of harassment, which was previously controversial in Canada.[11] The test requires that 1) the conduct be outrageous, 2) the perpetrator intended to cause emotional stress or had reckless disregard for causing the emotional stress, 3) the person did suffer from severe emotional stress, and 4) the conduct was the actual and proximate cause of the emotional stress. Since this test is a lower bar than the test for intentional infliction of mental suffering, harassment may be a viable avenue for legal recourse if subsequent case law upholds this tort.

Who is liable and what can be done?

When a harmful deepfake video is posted that gives rise to an action, a question will naturally arise: who is liable for it? Clearly, the person who posts the video may be liable under at least one of the headings above. However, it may be difficult to locate or identify that person, given the anonymity offered by many Internet sites. For this reason, one must consider also pursuing the service provider: that is, the forum, site or platform where the video is hosted.

In the United States, the *Communications Decency Act* provides a liability shield to a service provider for the actions of its users, stating "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."[12] However, Canada does not have similar legislation (except, perhaps, to a limited degree for copyright infringement), which means that online providers of forums, social media, comments sections, or other interactive websites may in fact be liable for the acts of their users under at least some of the causes of action described above. That is, the potential in Canada of proceeding against both the poster and the site does provide additional avenues for a victim to achieve a remedy. Thus, sites that operate in Canada should be conservative in approaching the moderation of deepfake videos and other defamatory, illegal, or harmful content.

However, policing the Internet has proven to be difficult. In the case of *Google Inc v Equustek Solutions Inc*, 2017 SCC 34, the Supreme Court of Canada affirmed the issuance of a takedown order requiring Google to remove harmful results from its search index for users worldwide.[13] A United States court subsequently blocked enforcement of the Canadian order, citing Google's immunity under the *Communications Decency*

Act and stating that the Canadian order undermines its policy goals and threatens free speech on the Internet. Thus, while an action may succeed in Canada, the multi-jurisdictional nature of the Internet may make such judgment difficult to enforce.

This area of law is evolving, and gives rise to legitimate debates about inter-jurisdictional independence, free speech, privacy, and judicial or legislative overreach. As the saying goes, "the Internet never forgets." Still, while taking legal action may not effectively take down a harmful deepfake video in all jurisdictions, it may nevertheless provide some measure of redress, particularly if damages can be obtained.

Conclusion

News articles published about pornographic deepfake videos have identified the potential dangers of misusing and/or abusing deepfake technology, including the ability to create fake news, cause political havoc, and target specific demographics such as but not limited to women, minorities, and vulnerable persons.

While the area of law surrounding deepfake videos is evolving and there are no direct laws per se that specifically address deepfake technology, there is an existing body of law in Canada that may already be able to address or may be extended to address the potential harms resulting from the publication of deepfake videos. Of course, Parliament and Legislative Assemblies may always introduce new statutory laws. In the meantime, case law will continue to develop to take these "techno-torts" into account, and there are some precedents of similar fact patterns already available in other jurisdictions. For example, in the United States, a male pilot superimposed a female pilot's face onto pornographic pictures and posted them to a computer bulletin board. The female pilot was awarded damages for defamation, intentional infliction of emotional distress, invasion of privacy, and punitive damages for acting with malice.[14]

Legal remedies, however, are not the only solution, and as any lawyer who practices in the area of online infringement or defamation can attest, playing "whack-a-mole" to take down content from the Internet is a never-ending battle. As such, other practical solutions should also be employed. Such solutions include having Internet platforms institute policies to ban harmful deepfake videos and creating better mechanisms for detecting, flagging, and removing such videos. Boosting AI fraud detection systems may become increasingly important as the laws in this area continue to evolve.

by Ryan Black, Pablo Tseng and Sally Wong (Articled Student)

[1] Please note that there are fair dealing exceptions in Canadian copyright law such as research, private study, parody, and satire.
[2] We note that this has been quite difficult to do owing to the multi-jurisdictional characteristic of the Internet.

[3] In the UK case *Charleston v News Group Newspapers Ltd*, [1995] 2 WLR 450 (cited in *Simpson v Mair,* 2008 SCC 40), a photo was published of two actors' faces superimposed onto nude bodies in a pornographic pose with supporting text explaining that these photos were manufactured without the actors' consent. The House of Lords held that the text and picture must be viewed as a whole, and was therefore not defamatory.

[4] In *Goldhar v Haaretz,* 2016 ONCA 515, an Israeli newspaper with no subscribers or business presence in Canada posted defamatory material online about a Canadian businessman, which was viewed by 200-300 viewers in Canada. The defendant argued that the court lacked jurisdiction as only a minor element of the tort was committed in Ontario (i.e. minimal readership in Canada). The Court of Appeal found that the subject matter of the article did have a significant connection in Ontario and thus the court had jurisdiction. This case has been granted leave to the Supreme Court of Canada, and will provide guidance for future defamation cases involving online activity.

[5] *Privacy Act*, [RSBC 1996] Chapter 373.

[6] PIPEDA refers to the *Personal Information Protection and Electronic Documents Act,* (S.C. 2000, c.5) and PIPA refers to the *Personal Information Protection Act*, [SBC 2003] Chapter 63.

[7] Though there may be an argument that the photographs used to source the deepfake video and the data extracted therefrom are personal information.

[8] See *R v H*(C), 2010 ONCJ 270, *R v Butler,* 2011 NLTD(G)5, D(R) v S(G), 2011 BCSC 1118.

[9] In *R v Craig*, 2016 BCCA 154, the judge states: "The purpose of the provision is to protect one's privacy interest in intimate images sent to and entrusted with another," an interpretation that, if correct, may mean that deepfake videos (like traditional privacy law) may not apply.

[10] Prohibited grounds in the *Human Rights Code* [RSBC 1996] Chapter 210 are race, colour, ancestry, place of origin, political belief, religion, marital status, family status, physical or mental disability, sex, sexual orientation, gender identity or expression, or age of that person or because that person has been convicted of a criminal or summary conviction offence that is unrelated to the employment or to the intended employment of that person.

[11] *Merrifield v The Attorney General*, 2017 ONSC 1333.

[12] Section 230. This section is currently in the process of being amended pursuant to the *Stop Enabling Sex Traffickers Act* though the amendments do not appear to apply to deepfake videos.

[13] See "[The New Frontier of Jurisdiction: Supreme Court of Canada Upholds Worldwide Injunction Against Google](#)".

[14] *Butler v Continental Express, Inc.*, No. 96-1204091, Montgomery County District Court, Texas.

A Cautionary Note

The foregoing provides only an overview and does not constitute legal advice. Readers are cautioned against

making any decisions based on this material alone. Rather, specific legal advice should be obtained.