「ディープフェイク」による権利侵害についての考察

潘敬瑩

目次

人工知能を使い、既存映像素材における合成された音声・写真・動画、ディープフェイク。その合成されたニセモノと本人の区別がほとんどつかない場合がある。素材として使用された肖像の権利者の肖像権を侵害しかねない。不当利用によって、名誉権などの権利を侵害する可能性もある。本稿では、ディープフェイクの運用メカニズム、関連している法益と可能の権利侵害を検討する。そこで、ディープフェイクに関する実際の事例において、従業者と専門家の意見を踏まえ、ディープフェイクは規制されるべきかなどの問題についての一助となることを目標とする。

キーワード：ディープフェイク　権利侵害　肖像権　名誉権　著作権

1　問題意識

人工知能技術の多くの分野の中で、AI による音声・画像などの合成技術は一般に最も身近な分野の 1 つであり、人工知能関連業界で最も成熟したアプリケーションの 1 つである。AI ディープラーニング運用の一つとして、主に「顔」情報の収集に基づいて、シミュレーション、再構築、さらには「動かす」ためにアルゴリズムが使用される。ディープフェイク（Deep Fake）と呼ばれるこうした映像は、アルゴリズム関連のトレーニングを受けず、一般人でも簡単に習得できる。「フェイク動画」が本物に見え、情報の信憑性と正確性を危うくしかねない。偽画像の拡散は、SNS 時代の虚偽情報の特徴をも言える[*1]。

ある新しい技術に対して、その発展を促進、保証、規制する方法は、関連業界内の肝心な問題であり、立法機関および司法機関が注意を払う必要があるだろう。したがって、ディープフェイクによる可能の権利被害について議論をする。

2　「ディープフェイク」とは

2017 年、「deepfakes」という名前の Reddit ユーザーは、ポルノスターの顔と有名人の顔を入れ替える方法による新しいタイプのポルノ動画をアップロードした。動画で使われたアルゴリズム、「Deep Fake」はすぐにインターネットで公表された。その後、フェイク動画と画像をたやすく作るためのソフト・ウェブサイト・スマートフォンアプリが登場した。

新しいタイプのポルノ動画を作成することから公表されたディープフェイクは、最初から「技術濫用」による悪名高い問題である。その後、スマートフォンアプリなどで動画の合成が広く行われるようになったことで、「ディープフェイク」という言葉に対する悪印象が薄くなる。しかし、2019 年 9 月まで、インターネット上に流通している 1 万 4678 件のディープフェイク（動画のみ）のうち、96 %はポルノコンテンツ[*2]であった。影響力が最も高い４つのフェイクポルノウェブサイトの総閲覧数は、１億 3436 万を超えた。

ディープを作る技術は４つのパターンに分けられている。Faceswap（人の顔を、別人の顔に入れ替えること）、Lip sync（映像に映っている人の唇の動きと別人の唇の動きをシンクロさせる。映像を作った製作者の望む通り、実際に発言してない発言をしているように見える）、Facial reenactment（映像に映っている人の表情を、別人の表情に移し替えることができる。特定のシーンと音声を加え、その人の特定情感と態度を表すことができる）と Motion transfer（体の動きを別人のものに移し替えること）[*3]。

[*1] 福長秀彦「SNS 時代の誤情報・虚偽情報とマスメディアの打ち消し報道～留意すべき事柄を考える」放送研究と調査 8 月号（2019 年）100 頁。

[*2] Henry Ajder, Giorgio Patrini, Francesco Cavalli, and Laurence Cullen（2019）,The State of Deepfakes: Landscape, Threats, and Impact（https://sensity.ai/mapping-the-deepfake-landscape/, 2020 年 11 月 15 日最終閲覧）。

[*3] Francesco Marconi & Till Daldrup(2018), How The Wall Street Journal is preparing its journalists

現在、ディープフェイクの悪影響のうち、「デジタル性犯罪」が最も広く疑問視されている。わいせつ物の合成から始まるディープフェイクに対して、「その存在自体が厳然とした性犯罪である」という見方もある。わいせつ合成物の生産・流通は、権利者の個人的な尊厳を低下させ、権利者の社会的評価を著しく低下させ、深刻な名誉権損害を与える可能性が非常に高くなる。さらに、権利所有者の許可なしに肖像を使用すること、または営利目的で権利所有者の肖像を含む合成物を販売することは、肖像権に重大な侵害となりうる。

知的財産権の視点では、ディープフェイクその合成プロセスで使用された大量の素材は、著作権を侵害する危険性がある。

政治的観点から見ると、AI 合成物の真偽を肉眼で識別することは難しいため、ディープフェイクはマスコミと国民を混乱させる政治目的にも使用され、信頼の社会的危機を引き起こし、ないし「ネット動画は全て信じない」というような過剰反応が拡散などの悪影響を与え、という恐れがある。

3　考えられる権利侵害

3.1　肖像権

権利者の許可なしに肖像を使用すること、または営利目的で権利所有者の肖像を含む合成物を販売することは、肖像権に重大な侵害となりうる。特に、電子決済などの分野での顔認証技術の活用により、情報漏えいや不正利用を防止するため、顔情報の取り扱いにもっと注意を払うべきである。一つの例として、ディープフェイク作るアプリ「ZAO」のユーザー規約で、ユーザーがアプリで自分自身または他の人の顔を使用した場合、肖像の権利所有者が「ZAO」とその関連会社がその顔情報を世界中での使用を許可したとみなされていることが記載されていた。アプリ方の権利は「完全に無料、取消不能、永続的、ライセンス可能、サブライセンス可能な権利」であり、非常に広い権利である。ユーザーが自分の肖像権とユーザー規約への注意の欠如により、肖像の不当利用をもたらす可能性が非常に高いであり、肖像権侵害によって訴えることは難しい。

インターネット上の肖像権侵害について、裁判所は必ずしも異なる判断基準を設けていないが、（ア）被写体がその人物であると特定できるか、（イ）被写体から公表の同意を得ているか、（ウ）被写体の人格利益侵害が社会生活上受忍の限度を超えるか、という判断アプローチについての見解がある[*4]。

3.2　名誉権

皮肉とユーモアなどの効果のために、ディープフェイクの作成者は意図的にキャラクターのイメージを歪め、悪質にする場合がある。有名人であれ、一般人であれ、社会的イメージや人格尊厳を損害し、深刻な悪影響を引き起こすだろう。

日本では、芸能人の顔の差し替えポルノ動画を公開したことによって、名誉毀損と著作権違反両方の疑いで逮捕した男性 2 人の事例がある。新しい技術に運用する侵害手段でも、現行法による権利侵害行為に当たると、規制すべきだ、という立場を明らかにする。

to detect deepfakes,（https://www.niemanlab.org/2018/11/how-the-wall-street-journal-is-preparing-its-journalists-to-detect-deepfakes/，2020 年 11 月 15 日最終閲覧）。

[*4] 数藤雅彦「インターネットにおける肖像権の諸問題：裁判例の分析を通じて」情報の科学と技術 70 巻 5 号（2020 年）233 頁。

3.3 著作権

一般論として、著作権法で保護された「作品」は、「革新的な知的成果」の実質的な要件と「再現性」の正式な要件を満たす必要があると考えられている。著作権者の許可なしに、素材として合法的に公開された他人の作品を使用する行為より、著作権侵害に該当する可能性がある。あるプラットフォーム上の音画素材は主に 2 つのソースがある。プラットフォーム自体によってアップロードされ、あるいはユーザーによってアップロードされるもの。ユーザー自身が著作権者を確認することができないため、侵害のリスクがあるかどうかを判断することは困難だろう。前述のように、ほとんどすべてのプラットフォームはユーザー規約で、著作権侵害の責任はユーザーが負うことを規定している。ユーザーは動画の著作権問題を調査することができず、しかし、プラットフォームは審査の義務をユーザーに譲渡したいと考えている。著作権紛争が発生すると、ユーザーは重大な不利益を被るだろう。

アーティスト JAY-Z の著作権紛争を例にとると、アルゴリズム合成物についての著作権主張は困難である。JAY-Z の声を合成し、音楽スタイルを模倣して作成されたディープフェイクが YouTube に公開された。彼はそれらの動画に著作権侵害の申し立てをした。しかし、技術的な観点から見ると、元の音声が直接に使用されたわけではなく、アルゴリズムによって、声を「シミュレート」しただけである[*5]。また、特定の音楽スタイル自体は、著作権法の保護対象とされない。

3.4 現行法の枠外の権利

現行法の下で、侵害が発生した相当な証拠があれば、如何なる侵害手段にもかかわらず、法律で適用することができる。ディープフェイクによる侵害を着目し、一つの疑問点は前述のように「現行法上の権利侵害に当たるかどうか」の問題であり、もう一つは、現行法の条文と異なり、「新たな権利を明確すべきか」の問題であろう。例えば、プライバシー権・名誉権・肖像権といった隣接する個別的人格権を整合し、「自己像の同一性」に対する、まだ未成熟な権利が議論される[*6]。または、アメリカの判例における、「公衆の誤認」による権利侵害は、以下のように一般的な名誉毀損と異なる立場が見られる。

（ア）名誉毀損の程度に当たるものでなくてもよい。

（イ）公衆の誤認による侵害の訴訟の要件として、対象物が相当数の公衆に公表されなければならない。

（ウ）誤認による侵害は名誉毀損より精神的苦痛を救済するものであるから、社会的名誉が毀損されたという立証の重要性が低い。[*7]

これらの権利は、人の人格が誤って社会的に表象されることからの保護に関わる権利であり、ディープフェイクから生じる法律問題と類似性があるだろう。

[*5] Marc Hogan(2020), What Does JAY-Z's Fight Over Audio Deepfakes Mean for the Future of AI Music? (https://pitchfork.com/thepitch/what-does-jay-zs-fight-over-audio-deepfakes-mean-for-the-future-of-ai-music/, 2020 年 11 月 15 日最終閲覧)。

[*6] 曽我部真裕「「自己像の同一性」に対する権利について」法学論叢 167 巻 6 号 (2010 年)1 頁。

[*7] ジョン・ミドルトン「アメリカにおける虚報とプライバシー侵害の成否」一橋法学 7 巻第 3 号 (2008 年)661 頁。

4　プラットフォーム責任

オンラインプラットフォーム内で、サードパーティによって公開されたコンテンツから生じる責任問題についての紛争では、プラットフォームが直接責任を負うべきか、共同責任を負うべきかについて、国によって異なる規制がある。たとえば、アメリカでは、1996 年に制定された通信品位法（Communication Decency Act of 1996）で、ネットワークサービスプロバイダーは、サードパーティによって提供されたコンテンツに対して、一部の例外を除き発行者と見なすことはできないとしている[*8]。つまりネットワークサービスプロバイダーは、ディープフェイクによる不法行為責任に問われても、法的責任はない。表示されたコンテンツについて、プラットフォームは実質的に制御、編集などのことができない。オンラインプラットフォームには第三者が公開したコンテンツに対して法的に必要な審査義務がないため、権利者は、オンラインプラットフォームに権利を主張することができない。

一方、起こり得る権利侵害責任を回避するために、プラットフォームは多くの場合、事前にユーザー規約などの方法を使用し、自身の責任を完全に回避しようとする。侵害紛争から、プラットフォームの責任を完全な免除するのは不合理だろう。紛争が発生すると、ユーザーは重大な不利益を被る可能性がある。アメリカ通信品位法 230 条が規定されているように、プラットフォームは、コンテンツによって引き起こされた権利侵害について、その責任が法で免除されることは、ネット技術の発展に伴い、改めて議論する必要があるだろう。一般に、国際社会で広く認められている「ノーティスアンドテイクダウン」（Notice and Take Down）ルールに従い、侵害者からコンテンツの削除要求の通知を受け取った後、プラットフォームが必要な措置を取らない場合、侵害行為による損害の拡大部分に、侵害者と共同責任を負うべきだとされている。

プラットフォームは、損害の防止に大きな役割を果たす可能性があるため、アップロードされたすべてのディープフェイクに対して、プラットフォームは、電子標識または内容審査などの手段をしなければならないという見解がある。しかしながら、プラットフォームがコンテンツの内容審査を担当する場合、スクリーニングのプロセスに人為的なバイアスを導入することが容易になり、中立性が損なわれる危険性がある。

5　裁判側における問題点

5.1　権利侵害かどうかについての判断

権利侵害が発生したかどうかを判断する上で重要な問題は、ディープフェイクが「ニセモノ」として明確な目印をつけられ、娯楽目的でのみ使用されていると主張する場合、権利者は、侵害が発生したと主張できるか、および訴訟の対象を選択することができるか。名誉権についての紛争を例に取ると、少なくとも 2 つの問題があるはずだろう。まず、ディープフェイクが「ニセモノ」であることを明らかにした場合、実際の人の社会的評価を損なうと言えるか。２つ目は、伝播の範囲が拡大し、結果として予期せぬ名誉毀損が実際に発生し、その責任を誰が負うべきか。元の作成者か、あるいはその影響を拡大する者か。

さらに、ディープフェイクの合成は膨大な量の素材に依存した可能性があるため、最終の合成物は、一見したところ、使用された素材とは何の関係もない可能性がある。ディープフェイクの制作で使われたアルゴリズ

[*8] 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. 通信品位法 230 条 (c) (1) である。(`https://www.law.cornell.edu/uscode/text/47/230`, 2020 年 11 月 15 日最終閲覧)。

ムが「識別可能」であるという前提で、使用された素材の権利者が、本人の本意でない発表されたことにより、権利侵害であると主張すると、裁判機関の視点から、十分な根拠を持っているかどうかことは、争点になってしまう。合成アルゴリズムが悪用できる法律の「抜け穴」はあるのだろうか。

5.2　偽りのエビデンス

偽造された証拠を識別し、判決の結果が左右されないことを保証し、それは裁判所が直面している問題である。証拠の信憑性を確認するための手順は、通常、裁判前に専門の司法鑑定機関によって完了される。ディープフェイクの存在とその使いやすさのため、偽りのエビデンスを識別するプレッシャーが高まっている。更に、偽造された証拠を識別することは困難であるため、その状況が長期的に続くと、電子証拠に対する信頼が低下する可能性もある。

「百聞は一見に如かず」という観念に基づいて、証人の証言と比べると、ヒトはしばしば音声やビデオなどの証拠を信じる傾向がある。しかし、もし裁判所が合成や改ざんされた偽りのエビデンスを識別することを確保できないとすれば、司法の信頼性への悪影響は計り知れないだろう。

現在、鑑定機関は、素材ソースを識別し、合成の痕跡を探すことにより、証拠が意図的に合成されたかどうかを弁別している。「AI に対抗するため、AI を使う」はもう 1 つの実行可能な方法である。例えば、人物のまばたき頻度を検出することで、アルゴリズムは実際の人物であるかどうかを判断できる。フェイク動画を検出する AI の研究開発とインターネット業界におけるその普及は深刻な価値を持っている。しかし、フェイク動画をつくる技術と比べると、検出する AI は常に受動的で不利な立場に置かれるである。将来の結果は、両方の AI その自身の学習能力次第になるかもしれない。作る技術を持つ者が増加していることを前提として、フェイク動画を見破るため、技術の進化に直面しなければならない。政府または関連機関が適切に対応できるか、この懸念は残る。

6　結びにかえて

ある新しい技術は、その出現の初期段階では、規制の欠如により一定の混乱を引き起こす状況がしばしば見える。ある新しい技術の実際の効果は、使用される方法と手段次第だろう。技術は特定の社会的矛盾を生み出すのではなく、その矛盾の影響を強化し、特定の社会的矛盾を悪化させることにすぎない。逆に、技術を合理的に使用することで、既存の対立を緩和し、積極的な社会的役割を果たすことができる。

まず、ディープフェイクに伴う大きなリスクを認識し、技術そのものではなく、起こり得るリスクを厳格に管理し、権利侵害が発生した場合に使用すべき対応策を明確にする必要があるだろう。既存の法規制に従い、民事損害補償、刑事罰、および行政罰などの対応策と救済措置を採用すべきである。一方で、それは法律で保護される権利者の正当な権利と利益を確保するための要求である。また、医療、教育、エンターテイメントの分野で積極的な役割を果たすべきディープフェイクという新技術が悪用されないように、法的な対応が求められる。

ディープフェイクなどの新技術の登場により、人間認知の領域で新たな課題が生じることが明らかにする。科学技術の支援で、認知範囲が大幅に拡大したことにも関わらず、生物として、人間は常に目や耳など自分の感覚器官に依存している。しかし、今の時代では、すべての情報は、技術によって偽造されたものである可能性が存在している。したがって、感覚への依存度を減らし、感官による判断への信頼感を弱め、「真」と「偽」の複雑な関係を再検討する必要があるかもしれないだろう。

7　参考文献

・福長秀彦「SNS 時代の誤情報・虚偽情報とマスメディアの打ち消し報道～留意すべき事柄を考える」放送研究と調査 8 月号（2019 年）100 頁。

・数藤雅彦「インターネットにおける肖像権の諸問題：裁判例の分析を通じて」情報の科学と技術 70 巻 5 号（2020 年）233 頁。

・曽我部真裕「「自己像の同一性」に対する権利について」法学論叢 167 巻 6 号 (2010 年)1 頁。

・ジョン・ミドルトン「アメリカにおける虚報とプライバシー侵害の成否」一橋法学 7 巻第 3 号 (2008 年)661 頁。

・Henry Ajder, Giorgio Patrini, Francesco Cavalli, and Laurence Cullen (2019), The State of Deepfakes: Landscape, Threats, and Impact（https://sensity.ai/mapping-the-deepfake-landscape/，2020 年 11 月 15 日最終閲覧)。

・Francesco Marconi & Till Daldrup(2018), How The Wall Street Journal is preparing its journalists to detect deepfakes,（https://www.niemanlab.org/2018/11/how-the-wall-street-journal-is-preparing-its-journalists-to-detect-deepfakes/，2020 年 11 月 15 日最終閲覧)。

・Marc Hogan(2020), What Does JAY-Z's Fight Over Audio Deepfakes Mean for the Future of AI Music?（https://pitchfork.com/thepitch/what-does-jay-zs-fight-over-audio-deepfakes-mean-for-the-future-of-ai-music/，2020 年 11 月 15 日最終閲覧)。

・47 U.S. Code § 230 - Protection for private blocking and screening of offensive material (https://www.law.cornell.edu/uscode/text/47/230，2020 年 11 月 15 日最終閲覧)。