



欧米各国で、ディープフェイクという巧妙なフェイク動画が、衝撃をもたらしている。政治家や映画女優の顔を入れ替えた本物そっくりの動画が出回り、被害も出ている。このディープフェイクは、ポストトゥルース時代を深刻なものにするかもしれない。逆に輝く未来を切り開いてくれるかもしれない。そこで注目されている新技術がGANsである。本稿は、これらの話題を取り上げて紹介することにしたい。

欧米で問題視されるディープフェイク

オバマ前大統領が「トランプ大統領は救いようのないマヌケ」だと批判するフェイク動画が、2018年4月にYouTubeで流され、大きな注目を集めた。人工知能(AI)を利用した巧妙な偽動画であった。

このフェイク動画は、米のオンラインメディア「バズフィード((BuzzFeed))」と映画監督のジョーダン・ピール((Jordan Peele))による作品で、ディープフェイク(deepfake)の脅威を世界に警鐘するためのものであった。

NHKのビジネス特集でも、「その動画は本物? “ディープフェイク”の衝撃」(2018年10月4日放送)で放送され、かつて米CBSテレビの著名なニュースキャスターであったダン・ラザーも、新技術の影響について大きな危機感を表明していた。

ディープフェイクによるフェイススワップ(Face Swaps、顔のすげ替え)は、今後の欧米の選挙戦などに大きな影響が出る危険性がある。映像作家ニコラス・ガーディナーによる実験的短編映画『Proxy』では、「炎と怒り」という言葉を使って、北朝鮮を糾弾するドナルド・トランプの顔を、中国の国家主席である習近平の顔にすり替えていた。

ディープフェイクは、政治の世界だけではなく、映画の世界も揺るがしている。ハリウッド女優や有名女優が、ポルノ映画に出演しているように見せるフェイクポルノ動画が、ネットで流れ被害者も出ている。

フェイクポルノの最初は、2017年末とされている。多くの人々が作ったフェイクポルノ動画が、海外掲示板Redditに投稿されたからである。2018年2月時点で、このコミュニ

ティの参加者は 9 万人を超えていたという。

これは、海外掲示板の **Reddit** にネット上で「ディープフェイクス (Deepfakes)」というハンドルネームのプログラマー (正体不明) が、誰でも簡単にフェイクポルノを作るアプリを公開したからである。

日本でも同様の事件が起こっている。フェイクポルノ動画共有サイトで、日本のアイドルや女優などのフェイクポルノがアップされている。日本はネット上にポルノ動画 (本物) が溢れており、アイドルや女優のフェイク被害が今後拡大する可能性がある。

被害者は女性だけではない。男性が見知らぬ女性から身に覚えのないフェイクポルノで脅迫されるケースもある。政治家や政府要人や有名人であれば、スキャンダルにされかねない。欧米諸国で問題視される理由は、ここにある。

当然、掲示板や動画サイトではフェイクポルノの投稿を禁止する動きが広がっている。フェイクポルノの誕生サイトである海外掲示板 **Reddit** は、この主の投稿を禁止する方針を発表している。

ガル・ガドットなど「**FakeApp**」を公開するサブレディット「**deepfakes**」は閉鎖されている。しかし、ネットでの拡散を防止することは、難しいと言ってよい。

ポストトゥルース時代におけるディープフェイクの影響は

最初に紹介した前オバマ大統領のフェイク動画が注目された背景には、2016 年から 2017 年にかけて、欧米各国で大統領や首相を選ぶ選挙戦において、フェイクニュースの氾濫が、大きな影響を及ぼしていたからである。

米「バズフィード」による「2016 年にフェイスブック上でヒットした捏造記事」で人気を集めたベスト 5 (米政治関連) の一つが、「オバマ大統領が学校での『忠誠の誓い』の朗読を禁じた」であった (出所:「ポストトゥルース」の起源とは? そして最もヒットした「フェイクニュース」とは? courrier.jp/news)。

この 2016 年を代表する最も注目すべき言葉として、オックスフォード英語辞典が選んだのが「ポストトゥルース」(Post-truth、ポスト真実) であった。フェイクニュースの氾濫が、「ポストトゥルース」という新語を誕生させたといっている。

ポストトゥルース時代の到来は、英国の欧州連合 (EU) 離脱 (2016 年 6 月の選挙) と米国のトランプ大統領の就任 (2017 年 1 月) によって、誰もが実感させられた。この 2 つの選挙で、マスコミ等による事前の予想が、全く覆されてしまったからのである。

ポストトゥルースとは、「世論形成において、客観的な事実より、虚偽であっても個人の感情に訴えるものの方が強い影響力を持つ状況」(知恵蔵) である。事実より、個人の感情にアピールする人物の発言や行為が、強い影響力を持つ。

当然、SNS でのフェイクニュースより、投稿サイトでのディープフェイクによる影響力は、強くなると言ってよい。ディープフェイクにより、世論の形成に大きな影響力を持つ人物の印象を歪曲した偽映像が流されれば、その人物の印象が変わってしまう。

ちなみに、本物と偽物の区別がつかないフェイススワップ (顔すげ替え) により、多重人格者が勝手に創作されてしまうかもしれない。かつて実在した多重人格者ビリーミリガン (2014 年、59 歳で死亡) は、24 の人格を持つと言われていた。

しかし、暗い話題ばかりではない。次に紹介する GANs という技法は、本物と偽物を乗り越える新しい世界を、近い将来、登場させるかもしれないからである。

ディープフェイクで注目される技法 GANs

さて、ディープフェイクとは、そもそも何か。ウィキペディアは「ディープフェイクとは『深層学習（deep learning）』と『偽物（fake）』を組み合わせた混成語（かばん語）で、人工知能にもとづく人物画像合成の技術を指す」と記述している。

このディープフェイクで注目されている技法の一つが GANs である。ウィキペディアは「『敵対的生成ネットワーク（Generative adversarial networks、略称 GANs）』と呼ばれる機械学習技術を使用して、既存の画像と映像を、元となる画像または映像に重ね合わせて（スーパーインポーズ）、結合することで生成される」と説明している。

GANs は、イアン・グッドフェロー（Ian Goodfellow）らによって 2014 年に考案された技法で、教師なし学習（Unsupervised Learning）の一つである。学習対象のデータだけを与えることで、コンピュータ自身に特定のデータの特徴量を発見させられる。たとえば、犬と動物の大量の画像から、犬の特徴量をコンピュータ自身が見だし、犬の認識を獲得できるようになる。

GANs は、ジェネレータ（generator）とディスクリミネータ（discriminator）という 2 つのネットワークを、対抗させる仕組みになっている。

ジェネレータは、より本物に似せた偽物の生成に努め、ディスクリミネータは、偽物と本物かを識別することに努めることを繰り返し、互いに学習しあう。簡単に言えば、偽札作りと鑑識係が、偽札の判定を競い合い学習するモデルである。

この学習を繰り返すことによって、本物と偽物との判別が出来なくなるまで、作業は続く。本物と区別のつかない精度の高い偽物が創作される仕組みである。

ただし、この 2 つのネットワークを競わせて学習させるのは技術的に難しく、安定した学習をさせるのに創意工夫が必要であるという。この「学習の安定性」を得るのが大きな課題になっており、課題のクリアには、しばらく時間がかかるとみられている。

GANs は、現在も進化し続けている。これまでの主要な成果（2014 年から 2018 年）と今後の予想については、「次世代の AI 技術 Generative Adversarial Networks（GANs）のビジネス応用可能性を考えてみる」（AI スタートアップ note、2019 年 2 月 16 日）に詳しく紹介されており、関心のある方はこれを参照されたい。この GANs の進化によって将来変わるかもしれないビジネスについて、この「AI スタートアップ note」は、次のような項目を列挙し解説している。今後の成果を見守りたい。

- ・ストックフォト・ビジネスが成り立たなくなる。
- ・デザイナーの仕事内容が、よりクリエイティブになる。
- ・AI の学習に、大量のデータが必要なくなる。
- ・映画・動画業界で加工パターンを、大量かつ簡単に試せるようになる。
- ・過去の動画データを、全て簡単にリマスタリングできるようになる。
- ・フェイクニュースを、見分けるビジネスが成立する。などなど

（TadaakiNEMOTO）