세상에 없던 사람을 만든다, 딥페이크

김원배 정보통신용어표준화위원회(WORDSTD) 위원, 전자신문 부장

2018년 4월 유튜브에 게재된 한 동영상에서 버락 오바마 전 미국 대통령이 “트럼프 대통령은 쓸모없는 사람”이라고 태연한 표정과 목소리로 말하는 모습을 볼 수 있다. 충격적 내용에 많은 사람이 당황했다. 하지만 동영상은 실제가 아니었다.

동영상은 콘텐츠 제작사 버즈피드가 만든 것이다. 영화감독 조던 필이 성대모사한 목소리를 실제 오바마 전 대통령 영상에 입히고, 입 모양을 바꾼 것이다. 동영상 제작에는 '딥페이크' 기술을 사용했다.

딥페이크(deepfake)는 인공지능(AI) 핵심기술 ‘딥러닝(deep learning)’과 가짜를 의미하는 ‘페이크(fake)’의 합성어다. AI의 한 분야인 딥러닝(Deep learning)을 활용해 사람의 얼굴이나 특정 부위를 합성해 제작된 가짜 사진과 이미지, 음성, 영상 또는 제작 프로세스 자체를 의미한다.

적대관계생성신경망(GAN) 기계학습 기술을 사용해 사진이나 영상, 음성 등을 원본에 겹치는 방식이다. GAN은 두 개의 알고리즘이 작동한다. 하나는 이미지 식별 기술을 이용해 정교한 이미지를 만들고, 다른 알고리즘은 제작된 이미지 진위 여부를 판단한다. 두 알고리즘이 서로 대립하며 차이점을 분석하고, 스스로 학습하며 진짜 같은 가짜 영상을 만든다. AI가 사진이나 영상 등을 공부하면 할수록 특정 단어를 얘기할 때 입꼬리가 올라가는 정교한 부분까지 구현이 가능하다. 과거에도 사진 합성, 영상 합성 등 가짜 콘텐츠가 있었지만 부자연스러웠다. GAN 기계학습 기술이 발전하면서 이제 가짜인지 진짜인지 구분하지 못할 정도의 정교한 영상을 구현한 것이다.

2017년 미국 대형 온라인 커뮤니티 ‘레딧’에 올라온 합성 포르노 영상을 딥페이크의 시초로 본다. ‘Deepfakes’라는 아이디를 쓰는 이용자가



오픈소스 소프트웨어(SW) 텐서플로를 활용해 유명 연예인과 포르노를 합성한 영상을 올렸다. 이후 'FakeApp'이라는 무료 SW가 배포되며 초보자도 쉽게 딥페이크 영상을 만들어 올리기 시작했다.

AI 기술은 어느새 진짜와 가짜를 구분할 수 없는 수준에 이르렀다. 딥페이크는 AI 기술의 무궁한 발전과 기술의 가능성을 보여주지만, 그만큼 악용되면 매우 위험한 기술이다. 가짜와 진짜를 구분할 수 없는 딥페이크 기술은 많은 범죄에 악용되고 있어 딥페이크 기술 존재 자체에 대한 찬반 논쟁까지 이어지고 있다.

가장 심각한 범죄는 합성 포르노이다. 우리나라는 물론 세계적으로 연예인 뿐만 아니라 일반인까지 얼굴 합성 포르노 피해자 사례가 증가하고 있다. 일반인의 얼굴을 합성하고 음성을 위조해 일명 '지인능욕'이라는 딥페이크 포르노 영상까지 유포되고 있다.

네덜란드 딥페이크 탐지 기술업체 '딥트레이스'가 2019년 발간한 보고서 '더 스테이트 오브 딥페이크'에 따르면 인터넷에 유포된 딥페이크 영상 96%가 음란물이다.

가짜뉴스 또한 심각한 문제다. 정치인 얼굴을 합성한 가짜 뉴스 등 정치인이나 연예인 등 특정인이 하지 않은 발언이나 행위를 묘사하는 영상을 만들어 여론을 호도할 가능성을 배제할 수 없다.

이뿐만 아니다. 딥페이크 악용 사례는 다양하다. 다른 사람 음성이나 영상을 허락없이 사용할 경우에 개인정보는 물론 초상권, 명예권, 성명권을 침해할 수 있다.

그렇다고 딥페이크 기술이 부작용만을 초래하는 필요악은 아니다. 원자력 기술을 발전용으로 사용할 수 있고, 폭탄 제작을 위해 사용할 수 있듯이 딥페이크 기술도 다른 용도로 사용될 수 있다. 좋은 목적으로 활용하면 획기적 AI 기술이지만, 범죄에 악용하면 무서운 기술이 되는 것이다.

딥페이크 기술 발전은 다양한 산업에 긍정적 영향을 준다. 독일 뤼벡대 연구진은 GAN 기반 딥러닝 알고리즘을 암 진단에 활용해 이전보다 정확한 진단이 가능함을 입증했다. 엔터테인먼트 산업에도 딥페이크 기술을 적용, 실제 촬영 없이 진짜 같은 영상을 쉽게 만든다.

딥페이크를 활용하면 쉽고 간편하게 특수효과를 만들어 내거나 증강현실(AR) 영상을 제작할 수 있다. 2016년 개봉한 영화 'Rogue One'에는 1977년 'Star Wars: A New Hope'에 출현한 배우가 당시의 모습 그대로 등장했다. 외형적으로 유사한 대역 배우를 섭외한 후, 대역배우의 얼굴에 과거 배우의 얼굴을 합성하는 모션 캡처 기법과 딥페이크를 활용한 결과다.

딥페이크는 신원보호 영역에서도 활발하게 사용되고 있다. 표정과 얼굴을 가리는 모자이크 대신, 익명성을 보장하면서 감정과 표정을 전달할 수 있는 딥페이크 기술은 여러 분야에서 활용될 전망이다.

딥페이크는 양날의 검처럼 이로운 부분과 해로운 부분이 명확하게 구분된다. 딥페이크의 산업적 가치, 잠재력은 무궁무진하지만 부작용 대처는 숙제다. 딥페이크 기술은 사람이 분간하기 어려울 정도의 수준에 이르렀다. 유통되는 수많은 콘텐츠 진위 여부를 한층 판단하기 어렵게 만들고, 새로운 문제를 만들 개연성이 상당하다.

앞으로 발전할 AI·딥페이크 기술을 바르게 사용하기 위해 악용에 대한 대처·규제가 필요하다. 급속도로 발전하는 딥페이크 변화에 제대로 대처하려면 법적으로, 제도적으로 일정한 규제는 필수다.