

유튜브 딥페이크(deepfake) 영상과 허위사실공표

허 순 철*

경남대학교 법학과 교수, 법학박사

■ 국문 초록

인공지능 기술이 발달함에 따라 진짜 같은 가짜 동영상을 만들 수 있게 되면서 예전에는 없었던 새로운 법적인 문제들이 생겨나고 있다. 2020년 말 엘리자베스 2세 영국 여왕이 크리스마스 인사를 하고 춤까지 추는 가짜 동영상이 공개된 적이 있었다. 딥페이크의 위험성을 경고하기 위해 2018년에는 오바마 전 미국 대통령이 “트럼프 대통령은 정말 쓸모 없는 사람”이라고 말하는 가짜 동영상이 공개되기도 했다. 최근 우리나라에서도 선거가 얼마 남지 않은 시점에 어느 전직 대통령이 특정한 대선 후보를 지지한다는 내용의 딥페이크 동영상이 여당의 공식 유튜브 채널에 게시되었다가 논란이 일자 삭제되는 일이 있었다.

그런데 선거를 앞두고 후보자에 관한 가짜 동영상이 유튜브(YouTube)에 게시되면 유권자가 후보자를 선택하는 데에 부정적 혹은 긍정적 영향을 미치게 되고 이는 결국 선거의 공정성을 전제로 하는 대의민주주의에 중대한 위협이 된다. 이에 미국의 캘리포니아주와 텍사스주는 공직선거의 후보자나 선거결과에 영향을 미칠 수 있는 딥페이크 동영상을 규제하는 법률을 제정하여 시행하고 있다.

가짜 동영상을 가리키는 “딥페이크(deepfake)”라는 단어는 “딥러닝(deep learning)”과 “가짜(fake)”의 합성어로서, “진짜처럼 보이기 위해 조작된 이미지, 오디오 또는 비디오”를 가리킨다. 그런데 캘리포니아주의 딥페이

* juristice@kyungnam.ac.kr

크금지법은 선거의 공정이라는 ‘진정한 국가의 이익’을 달성하기 위하여 언론의 자유에 대해 최소한의 규제만을 하는 것이라고 하기 어렵고, 언론의 자유에 관한 연방대법원의 ‘현실적 악의’의 법리 때문에 동 법이 실제 사례에 적용되기 어렵다는 비판이 제기되고 있다. 따라서 이러한 비판적 견해에 의하면 딥페이크는 명예훼손법이나 저작권법 등을 통해 규율하는 것이 더 바람직하다고 한다. 한편 텍사스주의 딥페이크금지법은 ‘현실적 악의’의 법리를 규정하지 않고 있고, 페러디나 풍자를 동 법의 적용예외로 규정하고 있지 않으므로 연방 헌법에 위반된다는 주장이 제기되고 있다.

최근 우리나라의 중앙선거관리위원회는 ‘딥페이크 영상 관련 법규운용 기준’을 제정하여, 딥페이크 영상을 이용한 선거운동은 가능하지만 “딥페이크 영상(AI OOO)”임을 표시하지 않는 경우 「공직선거법」 제250조에 위반된다는 유권해석을 내놓았다. 그러나 딥페이크 영상임을 표시해야 하는 법적인 근거가 명확하지 아니할 뿐만 아니라, 이러한 유권해석은 언론의 자유에 대한 사전 제약으로서 위축효과를 가져올 것으로 보인다. 한편 딥페이크 선거운동을 금지하는 내용의 「공직선거법」 개정안이 국회에서 발의되었으나, 기존의 ‘허위사실공표죄’로도 딥페이크를 충분히 규제할 수 있으므로 새로운 처벌 규정을 신설할 필요는 없는 것으로 보인다.

주제어: 딥페이크, 유튜브, 허위사실공표, 선거, 공직선거법

목 차

- I. 머리말
- II. 유튜브와 딥페이크
 - 1. 유튜브
 - 2. 딥페이크
- III. 미국에서의 유튜브 딥페이크와 허위사실공표
 - 1. 서설
 - 2. 유튜브 딥페이크와 허위사실공표
 - 3. 소결
- IV. 우리나라에서의 유튜브 딥페이크와 허위사실공표
 - 1. 서설
 - 2. 거짓말과 표현의 자유
 - 3. 유튜브 딥페이크와 허위사실공표
- V. 맺음말

I. 머리말

흔히들 “백 번 듣는 것이 한 번 보는 것보다 못하다(百聞不如一見)”고 한다. 서양에서는 “보는 것이 믿는 것이다(Seeing is believing)”라는 말이 있다. 그런데 만약 우리가 본 것이 진짜가 아니라 ‘가짜’라면 어떻게 될까? 1994년에 개봉한 영화 ‘포레스트 검프(Forrest Gump)’에서는 주인공 톰 행크스가 케네디 대통령과 만나는 장면이 나온다.¹⁾ 이 장면은 기록 영화로 남아 있던 영상에서 케네디 대통령의 모습을 배경과 분리하고 새로운 배경을 촬영한 뒤에 톰 행크스가 혼자 연기한 장면과 합성하는 “삼겹 합성본(three-layer composite)” 방식을 통해 만들어졌다.²⁾ 당시에는 이러한 장면을 만들려면 상당한 시간과 기술력 그리고 비용이 필요했기 때문에 일반인들은 그런 장면을 만들기가 어려웠다. 그런데 영화에

1) 네이버영화, 포레스트 검프, URL: <https://movie.naver.com/movie/bi/mi/basic.naver?code=17159>.

2) 네이버TV, 포레스트 검프 CG메이킹 영상, URL: <https://tv.naver.com/v/1076460>.

나오는 이러한 ‘가짜’ 영상은 비록 거짓이었지만 영화의 완성도를 높여 주는 긍정적인 역할을 한다는 점에서 별다른 법적인 문제가 없었다. 그러한 가짜 영상이 과거에 있었던 어떤 사실을 왜곡하는 것은 아니었으며, 대부분의 관객들은 이미 그 영상이 가짜라는 사실을 알고 있었기 때문이기도 했다.

그런데 인공지능 기술이 발달함에 따라 진짜 같은 가짜 동영상을 만들 수 있게 되면서 예전에는 없었던 새로운 법적인 문제들이 생겨나고 있다. 예를 들어, 2020년 말에는 엘리자베스 2세 영국여왕이 크리스마스 인사를 하고 춤까지 추는 가짜 동영상이 공개된 적이 있었다.³⁾ 가짜 정보의 심각성을 알리기 위해 이 가짜 동영상을 만든 영국 방송 채널4는 “가짜 여왕 영상은 눈으로 본 것조차 믿을 수 없다는 점을 상기시켜 줄 것”이라고 하였다.⁴⁾ 한편 2018년 4월 영화감독 조던 필(Jordan Peele)은 오바마 전 미국 대통령이 “트럼프 대통령은 정말 쓸모없는 사람”이라고 말하는 가짜 동영상을 공개하여 딥페이크의 위험성을 경고하기도 했다.⁵⁾

그런데 진짜 같은 가짜 동영상을 만들어 낼 수 있는 기술이 악용되는 경우에는 심각한 사회적 문제를 야기할 수도 있다.⁶⁾ 가짜로 만든 테러나 미사일 공격 동영상은 심각한 경우 전쟁을 야기할 수도 있다.⁷⁾ 유명인의 얼굴을 합성해서 만든 가짜 포르노그래피의 유포는 심각한 명예훼손

3) 이세영 (2020, 12, 26). AI 기술로 만들어낸 춤추는 ‘가짜 영국 여왕’. <조선일보>, URL: https://www.chosun.com/international/2020/12/26/BL2DA2RS5JGMJCTFGAU6R76AVE/?utm_source=naver&utm_medium=referral&utm_campaign=naver-news.

4) 이세영 (2020, 12, 26). 위 신문기사 참조.

5) 엄민재 (2018, 7, 5). 목소리도 행동도 똑같아...진짜 같은 가짜 ‘딥페이크’. <SBS 뉴스>, URL: https://news.sbs.co.kr/news/endPage.do?news_id=N1004834758&plink=SEARCH&cooper=SBSNEWSSEARCH.

6) “딥페이크 범죄 기승, 타깃 안 되려면” 국정원의 경고. (2021, 5, 26). <JTBC 뉴스>, URL: https://news.jtbc.joins.com/article/article.aspx?news_id=NB12007096.

7) Chesney, Bobby & Citron, Danielle. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, *California Law Review* 107, 1776.

손과 인격권 침해를 야기한다.⁸⁾ 이에 우리나라는 「성폭력범죄의 처벌 등에 관한 특례법」을 개정하여 딥페이크 기술을 이용한 ‘허위영상물 등의 반포등’의 행위를 처벌하는 규정을 신설하였다.⁹⁾

특히 선거를 앞두고 후보자에 관한 가짜 동영상이 유튜브(YouTube)에 유포되어 유권자가 후보자를 선택하는 데에 부정적 혹은 긍정적 영향을 미친다면 대의민주주의에 중대한 위협이 된다. 선거에 있어서 공정성은 가장 중요한 요소이며 불공정한 선거는 당선자에게 민주적 정당성을 부여할 수 없기 때문이다. 최근 우리나라에서도 유튜브 가짜 동영상이 문제가 된 적이 있었다. 대통령 선거가 얼마 남지 않은 시점에 “고(故) 노무현 전 대통령이 이재명 대선후보를 지지 선언한다는 콘셉트의 영상을 더불어민주당이 공식 채널에 올렸다가 논란이 일자 삭제”하였다.¹⁰⁾ 이 동영상에서 “가상의 노 전 대통령은 ‘사랑하는 국민 여러분. 대한민국 20

8) 미국의 경우 2019년 9월을 기준으로 인터넷상의 딥페이크 동영상의 약 96 퍼센트는 포르노그래피였다고 한다. Ajder, Henry et al. (2019, September). The State of Deepfakes: Landscape, Threats, and Impact, *Deeptrace*, URL: https://regmedia.co.uk/2019/10/08/deepfake_report.pdf. 한편 딥페이크 누드사진을 만들 수 있는 애플리케이션은 50달러에 판매되고 있으며, 이 애플리케이션을 이용하면 사진 속 인물의 얼굴을 나체의 여성 이미지와 합성하여 나체사진을 만들 수 있다. Ajder, Henry et al. (2019, September). 위의 자료 참조.

9) 제14조의2(허위영상물 등의 반포등) ① 반포등을 할 목적으로 사람의 얼굴·신체 또는 음성을 대상으로 한 촬영물·영상물 또는 음성물(이하 이 조에서 “영상물등”이라 한다)을 영상물등의 대상자의 의사에 반하여 성적 욕망 또는 수치심을 유발할 수 있는 형태로 편집·합성 또는 가공(이하 이 조에서 “편집등”이라 한다)한 자는 5년 이하의 징역 또는 5천 만원 이하의 벌금에 처한다. ② 제1항에 따른 편집물·합성물·가공물(이하 이 항에서 “편집물등”이라 한다) 또는 복제물(복제물의 복제물을 포함한다. 이하 이 항에서 같다)을 반포등을 한 자 또는 제1항의 편집등을 할 당시에는 영상물등의 대상자의 의사에 반하지 아니한 경우에도 사후에 그 편집물등 또는 복제물을 영상물등의 대상자의 의사에 반하여 반포등을 한 자는 5년 이하의 징역 또는 5천만원 이하의 벌금에 처한다. ③ 영리를 목적으로 영상물등의 대상자의 의사에 반하여 정보통신망을 이용하여 제2항의 죄를 범한 자는 7년 이하의 징역에 처한다. ④ 상습으로 제1항부터 제3항까지의 죄를 범한 때에는 그 죄에 정한 형의 2분의 1까지 가중한다.

10) 김명일 (2022, 2, 6). 故노무현이 “정의로운 이재명 지지”...與 영상에 친문도 발끈했다. <조선일보>, URL: <https://www.chosun.com/politics/election2022/2022/02/06/TOFLQ5BNG5EXTF6HEYC5243TIU/>.

대 대통령 선거, 얼마 남지 않았습시다’라며 ‘저 노무현은 사람 사는 세상을 꿈꾸며 가난과 역경을 딛고 일어나 오직 국민만을 생각하며 불의와 타협하지 않고 기득권과 싸워 이겨내는 정의로운 이재명 후보를 지지합니다’라고 했”으며, 비판이 제기되자 이 동영상은 곧 삭제되었다.¹¹⁾

우리의 일상생활에서 유튜브와 같은 동영상 공유 플랫폼은 이제 생활의 일부가 되었다고 해도 과언이 아니다. 선거운동은 물론 무료 강의 수강, 뉴스 시청, 정보 수집 등 일상생활에서 유튜브가 활용되는 분야는 매우 다양하다. 그런데 유튜브에 선거와 관련한 가짜 동영상이 게시된다면 사람들은 혼란에 빠질 수 있다. 이에 유튜브는 “조작된 선거 관련 동영상을 삭제할 방침”을 밝힌 바 있다.¹²⁾ 유럽의 일부 국가에서는 딥페이크 기술을 이용한 선거운동이 “유권자의 올바른 판단을 저해”할 수 있다는 이유로 금지되어 있다고 한다.¹³⁾ 미국에서는 캘리포니아주¹⁴⁾와 텍사스주¹⁵⁾가 딥페이크 기술을 이용한 가짜 동영상을 규제하는 법률을 제정하였다.

그렇다면 우리나라도 미국의 캘리포니아주나 텍사스주처럼 「공직선거법」에 딥페이크 기술을 이용한 가짜 동영상을 게시하여 선거에 영향을 미치려는 행위를 규제하는 새로운 입법을 해야 할까? 나아가 딥페이크를 이용한 패러디나 풍자도 금지해야 할까?¹⁶⁾

11) 김명일 (2022, 2, 6). 위의 신문기사 참조.

12) 이상헌 (2020, 2, 4). 페이스북 이어 유튜브도 ‘선거관련 조작 동영상’ 삭제 방침. <연합뉴스>, URL: <https://www.yna.co.kr/view/AKR20200204051200009?input=1195m>.

13) 강문규·유오상 (2021, 12, 9). ‘AI 윤석열’ 논란…美·유럽선 “유권자 판단 왜곡” 딥페이크 활용 금지. <헤럴드경제>, URL: <http://news.heraldcorp.com/view.php?ud=20211209000455>.

14) Lecher, Colin. (2019, October 7). California has banned political deepfakes during election season, *The Verge*, URL: <https://www.theverge.com/2019/10/7/20902884/california-deepfake-political-ban-election-2020>.

15) Morris, Allie, Texas is first state to ban political ‘deepfake’ videos, *San Antonio Express-News*, Oct. 9, 2019, <https://www.expressnews.com/news/local/politics/article/Texas-is-first-state-to-ban-political-14504294.php>.

16) 캘리포니아는 딥페이크를 규제하는 법률을 제정하면서 동 범이 풍자나 패러디에

미국에서는 딥페이크 규제를 둘러싸고 찬성과 반대의 논의가 전개되고 있으며, 최근 우리나라 국회에서도 ‘딥페이크 선거운동 금지법’이 발의되었다.¹⁷⁾ 이에 본 연구에서는 먼저 유튜브와 딥페이크의 개념과 특성을 살펴보고, 이를 토대로 미국 캘리포니아주와 텍사스주의 선거 관련 딥페이크 규제 입법과 이에 관한 논의를 검토하여 시사점을 찾아본 후에, 우리나라의 「공직선거법」상 허위사실공표죄를 중심으로 유튜브 딥페이크 동영상을 규제하는 새로운 법률을 제정해야 할 필요성이 있는지에 대해 검토해 보기로 한다.

II. 유튜브와 딥페이크

1. 유튜브

유튜브는 이용자들이 동영상을 시청, 공유 및 게시할 수 있는 동영상 공유 서비스이다.¹⁸⁾ 유튜브 홈페이지의 고객센터 코너에서는 ‘잘못된 선거 정보 관련 정책’이라는 제목 아래 ‘유권자 투표 방해’, ‘후보자 자격 요건’, ‘민주적 절차를 방해하는 선동’, ‘해킹된 자료의 배포’, ‘선거 공정성’에 관련된 콘텐츠를 게시하지 말라고 하면서, “이 정책을 위반하는 콘텐츠는 삭제되고 이와 관련된 알림 이메일이 발송”된다고 설명하고 있다.¹⁹⁾ 유튜브를 운영하는 구글(Google)의 담당자인 클레멘트 울프

해당하는 딥페이크에는 적용되지 않는 것으로 규정하고 있다. Cal. Elec. Code § 20010(d)(5). “This section does not apply to materially deceptive audio or visual media that constitutes satire or parody.”

17) 이정우 (2021, 12, 20). 이해석 의원, 유권자 기만하는 ‘딥페이크 선거운동 금지법’ 대표발의. <The Public>, URL: <https://www.thepublic.kr/news/newsview.php?ncode=1065576005911254>.

18) Webwise.ie, Explained: What is YouTube?, URL: <https://www.webwise.ie/parents/what-is-youtube/>.

19) YouTube 고객센터, Misinformation policies, 잘못된 선거 정보 관련 정책, URL:

(Clement Wolf)는 한국 기자들과 가진 온라인 간담회에서 “정치적인 허위정보, 특히 선거 관련 허위정보 판별이 쉽지 않은 과제라고 생각한다”면서, “매우 디테일하고 구체적인 정책을 운영하고 있다. 콘텐츠의 퀄리티를 반영해 양질의 콘텐츠가 계속 높은 순위(노출이 잘 되는 위치)에 오르도록 부각하고 있고, 저질 콘텐츠는 계속해서 부각되지 않도록 노력한다”라고 말했다.²⁰⁾ 한편 지난 2020년 2월에 유튜브는 “어떤 정부 관료가 사망한 것처럼 보이게 하거나, ‘부정확한 투표 날짜를 알려주는 것처럼 투표·인구조사 과정에서 사람들을 오도할 목적을 가졌거나’, ‘후보자의 시민권 요건과 관련한 허위정보에 근거해 그가 공직 자격이 없다는 주장과 같은 후보자 자격요건에 대한 허위사실이 담긴’ 동영상 등을 예시”로 들면서, “조작된 선거 관련 동영상을 삭제할 방침이라고 밝”힌 바 있다.²¹⁾

한편 미국은 1996년에 IT 기업들을 보호하기 위해 「통신품위법」(Communication Decency Act) 제230조를 제정하여 인터넷에 게시된 내용으로 인해 명예훼손과 같은 문제가 발생한 경우에 인터넷서비스 제공자는 면책시키고 게시자에게 책임을 지우는 정책을 유지해 왔다.²²⁾ 동 조항은 “상호 연결된 컴퓨터 서비스(interactive computer service)는 다른 정보 콘텐츠 제공자가 게시한 어떤 정보의 발행자나 화자(speaker)로 간주되지 않는다”고 규정하고,²³⁾ 민사책임에 관하여 “서비스 제공자나 이

https://support.google.com/youtube/answer/10835034?hl=ko&ref_topic=10833358#zippy=%2C%EC%84%A0%EA%B1%B0-%EA%B3%B5%EC%A0%95%EC%84%B1.

20) 금준경 (2022, 1, 20). 극단적 유튜브 채널 방치 이유 있다 질문에 구글의 답은. <미디어오늘>, URL: <http://www.mediatoday.co.kr/news/articleView.html?idxno=301900>.

21) 이상헌 (2020, 2, 4). 페이스북 이어 유튜브도 ‘선거관련 조작 동영상’ 삭제 방침. <연합뉴스>, URL: <https://www.yna.co.kr/view/AKR20200204051200009?input=1195m>.

22) 김성민 (2020, 10, 17). 페이스북·유튜브 가짜뉴스 놔두면 법적 책임 묻는다. <조선일보>, URL: https://www.chosun.com/economy/tech_it/2020/10/17/FY7BQ4NUSZG67B5UPQ32623NFQ/.

23) 47 U.S.C. 230(c)(1). “(1) Treatment of publisher or speaker. No provider or user

용자가 헌법적으로 보호를 받는지 여부와 관계없이 선의로 음란하거나, 외설적이거나, 선정적이거나, 저속하거나(filthy), 과도하게 폭력적이거나, 괴롭히거나(harassing) 그 밖의 불쾌한 자료에 대한 접근이나 이용을 제한하는 조치를 하거나 또는 그러한 자료에 대한 접근을 제한하기 위하여 정보 제공자나 그 밖의 자가 취한 기술적 조치와 관련하여 상호 연결된 컴퓨터 서비스의 제공자나 이용자는 어떠한 책임도 지지 않는다”고 규정하고 있다.²⁴⁾ 그런데 지난 2020년 5월 “우편투표가 선거 조작으로 이어질 수 있다’는 트럼프 대통령의 트윗에 트위터(Twitter)가 ‘경고 딱지’를 붙이면서 동조항의 개정 논의가 불붙었”으며, “미국 공화당은 진보 성향인 실리콘밸리 IT 기업들이 통신법 230조를 방패 삼아 보수주의자의 메시지를 검열한다고 비난”해왔다.²⁵⁾

2. 딥페이크

오늘날에는 소위 인공지능을 이용한 첨단기술의 발달로 인해 누구나 쉽게 소프트웨어를 다운로드받아서 가짜 음성이나 동영상을 만들 수 있게 되었다.²⁶⁾ 가짜 동영상을 가리키는 딥페이크(deepfake)라는 말은 “딥러닝(deep learning)”과 “가짜(fake)”의 합성어로서, 2017년 12월에 ‘딥페

of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”

24) 47 U.S.C. 230(c)(2). “(2) Civil liability. (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or (B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1).”

25) 김성민 (2020, 10, 17). 앞의 신문기사 참조.

26) Porup, J. M. (2021, March 18). How and why deepfake videos work—and what is at risk, *CSO*, URL: <https://www.csoonline.com/article/3293002/deepfake-videos-how-and-why-they-work.html>.

이크'라는 닉네임을 사용하던 익명의 이용자가 소셜 뉴스 웹사이트인 '레딧(Reddit)'에 유명인의 얼굴을 합성해서 만든 포르노그래피를 게시하였던 데에서 유래한 것으로 알려져 있다.²⁷⁾ 그런데 '딥페이크²⁸⁾ 또는 '(인공지능 기반) 첨단 조작기술'의 개념에 관하여, 다양한 견해들이 제시되고 있다. 예를 들어, "방송·신문 등 기존 미디어 매체 또는 인터넷 기반의 새로운 유형의 미디어 매체뿐만 아니라 음성·이미지·영상·문자 등 정보 유형을 불문하고 허위의 사실을 진실인 것처럼 포장하는 내용의 정보"²⁹⁾ 또는 "진짜처럼 보이기 위해 조작된 이미지, 오디오 또는 비디오"를 말한다는 견해가 있다.³⁰⁾ 이와는 달리 딥페이크의 범위를 비디오에 한정하여 "인공지능을 이용한 디지털 도구를 사용하여 창조되거나 개조된 비디오"로 정의하는 견해도 있다.³¹⁾ 최근 우리나라에서 제출된 「공직선거법」 개정 법률안에서는 딥페이크를 "실제로 말한 것이나 행동한 것처럼 보이도록 인공지능 기술을 이용하여 만든 거짓의 음성·화상

27) Delfino, Rebecca A. (2019). Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act, *Fordham Law Review* 88; Vincent, James. (2018, May 22). Why we need a better definition of 'deep fake', *The Verge*, URL: <https://www.theverge.com/2018/5/22/17380306/deepfake-definition-ai-manipulation-fake-news>.

28) 문화체육관광부와 국립국어원은 "인공지능(AI)을 이용해 가상 이미지나 영상을 만들어 내는 기술인 '딥페이크(deepfake)'는 '(인공지능 기반) 첨단 조작기술'로 대체하여 부르기로 했다고 한다. 양정우 (2020, 10, 26). '프롭테크'→'부동산 정보기술'...쉬운 우리말 대체어 선정. <연합뉴스>, URL: <https://www.yna.co.kr/view/AKR20201026040200005>; 정형모 (2020, 6, 6). 비말 대신 침방울, 슬로푸드 대신 정성 음식 어때요. <중앙일보>, URL: <https://www.joongang.co.kr/article/23795039>; '뉴 노멀'은 '새 기준, 새 일상'으로, '웨비나'는 '화상 토론회'로-국립국어원 새말모임(4.20.~22.) 대체어 마련. (2020, 4, 28). 문화체육관광부, 3.

29) 최중선 (2020). 허위조작정보 규제 시도에 대한 비판적 고찰. <홍익법학>, 21권 1호, 367.

30) Tashman, Alexandra. (2021). "Malicious Deepfakes"—How California's A.B. 730 Tries (and Fails) to Address the Internet's Burgeoning Political Crisis, *Loyola of Los Angeles Law Review* 54, 1397.

31) Dixon Jr., Herbert B. (2019, August 12). Deepfakes: More Frightening Than Photoshop on Steroids, *American Bar Association*, URL: https://www.americanbar.org/groups/judicial/publications/judges_journal/2019/summer/deepfakes-more-frightening-photoshop-steroids/.

또는 영상 등의 정보”로 규정하고 있다.³²⁾ 한편 미국의 텍사스주 선거법은 딥페이크를 “속이려는 의도로 제작된 비디오로서 실존 인물(a real person)이 실제로는 하지 아니한 행위를 한 것처럼 보이게 하는 비디오”라고 규정하고 있다.³³⁾

딥페이크와 달리 셸로우페이크(shallow fake 또는 cheap fake, 저급 허위조작정보)는 동영상의 재생속도를 낮추거나 높이는 것과 같은 단순한 조작을 통해 만든 동영상을 가리킨다.³⁴⁾ 2019년 봄에 낸시 펠로시(Nancy Pelosi) 미국 연방하원의장이 마치 술에 취해 어눌하게 말을 하는 듯한 동영상이 유포된 적이 있었는데, 이 동영상이 대표적인 셸로우페이크에 해당한다.³⁵⁾ 이 동영상은 딥페이크 기술을 사용하지 않고 단지 동영상의 재생속도를 늦추는 방법으로 만들어진 것에 불과했지만 펠로시 연방하원의장이 진짜 술에 취한 듯 보였다는 점에서, “전문가들은 딥페이크 기술이 발전하고 더 일반적인 것이 될수록 가짜뉴스의 도전은 심화할 것”이라는 우려를 표시했다.³⁶⁾

그런데 이미지를 바꾸는 기술은 이미 오래전부터 있었다. 소련의 독재자 스탈린(Joseph Stalin)은 정적들(political opponents)에 관한 사진을 교

32) 공직선거법 일부 개정 법률안 (이해식, 2021, 12, 20), URL: http://likms.assembly.go.kr/bill/billDetail.do?billId=PRC_L2O1S1Y2C1V5B1R1X5Q0V4B1M6O8I1.

33) Tex. Elec. Code § 255.004(e). “In this section, ‘deep fake video’ means a video, created with the intent to deceive, that appears to depict a real person performing an action that did not occur in reality.”

34) Dixon Jr., Herbert B. (2019, August 12). Deepfakes: More Frightening Than Photoshop on Steroids, *American Bar Association*, URL: https://www.americanbar.org/groups/judicial/publications/judges_journal/2019/summer/deepfakes-more-frightening-photoshop-steroids/; Ajder, Henry et al. (2019, September). The State of Deepfakes: Landscape, Threats, and Impact, *Deeptrace*, URL: https://regmedia.co.uk/2019/10/08/deepfake_report.pdf.

35) Ajder, Henry et al. (2019, September). The State of Deepfakes: Landscape, Threats, and Impact, *Deeptrace*, URL: https://regmedia.co.uk/2019/10/08/deepfake_report.pdf.

36) 정성호 (2019, 5, 29). WP “펠로시 조작 동영상, 새 가짜뉴스 대응할 준비부족 드러내”. <연합뉴스>, URL: <https://www.yna.co.kr/view/AKR20190529027900091?nput=1195m>.

묘하게 조작하여 정치에 이용하였다.³⁷⁾ 북한 역시 김일성을 우상화하기 위하여 “‘항일빨치산’ 시절 김일성의 모습을 담은 사진”을 조작하였다고 한다.³⁸⁾ 얼마 전 우리 정부는 G7정상회담 관련 사진을 배포하면서 “사진 앞줄 맨 왼쪽에 있던 시릴 라마포마 남아프리카공화국 대통령”을 사진에서 제외했다가 ‘사진조작’을 했다는 비난을 받기도 했다.³⁹⁾

사진의 이미지를 바꾸는데 사용되는 장비들 중 가장 널리 알려진 것은 1987년에 출시된 포토샵(Adobe photoshop)이다.⁴⁰⁾ 그런데 요즘에는 인공지능을 이용하여 일반인들도 정교한 덤페이크 동영상을 손쉽게 만들 수 있게 되었다.⁴¹⁾ 초기의 덤페이크 기술은 짧은 비디오 동영상에 유명인의 얼굴을 붙이는 정도에 불과했지만, 인공지능의 한 가지 형태 또는 기계학습(machine learning)이라고 할 수 있는 심층학습(deep learning, 덤러닝)을 통해 정교한 덤페이크 동영상물의 제작이 가능해졌다. 심층학습이란 컴퓨터가 인간의 수치 입력(input)이나 지적능력을 필요로 했던 일들을 수행할 수 있도록 하는 기계학습을 가리킨다.⁴²⁾ 다시

37) Gessen, Masha. (2018, July 15). The Photo Book that Captured How the Soviet Regime Made the Truth Disappear, *The New Yorker*, URL: <https://www.newyorker.com/culture/photo-booth/the-photo-book-that-captured-how-the-soviet-regime-made-the-truth-disappear>.

38) 배진영 (2021, 6, 14). 지도자 우상화 위해 사진 조작하는 북한 닳아가나?. <월간조선>, URL: http://monthly.chosun.com/client/mdaily/daily_view.asp?idx=12711&Newsnumb=20210612711.

39) 강주리 (2021, 6, 14). 文 사진 중앙에 “대한민국의 위상”...남아공 대통령 잘라냈다. <서울신문>, URL: https://www.seoul.co.kr/news/newsView.php?id=20210614500158&wlog_tag3=naver.

40) Adobe, Photoshop, Britannica, URL: <https://www.britannica.com/technology/Adobe-Photoshop>.

41) Metz, Cade. (2019, December 24). Internet Companies Prepare to Fight the ‘Deepfake’ Future, *The New York Times*, URL: <https://www.nytimes.com/2019/11/24/technology/tech-companies-deepfakes.html>.

42) Marr, Bernard. (2018, October 1). What is deep Learning AI?, A Simple Guide with 8 Practical Examples, *Forbes*, URL: <https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/?sh=14e204a58d4b>.

말해서 심층학습은 인공신경망(artificial neural networks)이 대량의 데이터를 학습하도록 하는 것이고,⁴³⁾ 신경망(neural networks)은 컴퓨터가 훈련사례들(training examples)을 분석함으로써 특정한 임무를 수행하는 방법을 학습하는 기계학습의 한 가지 방식이다.⁴⁴⁾

컴퓨터는 인간이 경험을 통해 학습을 하는 것과 유사한 방식으로 신경망이 반복적으로 임무를 수행하게 함으로써 산출물의 가치를 향상시킨다.⁴⁵⁾ 이러한 심층학습 알고리즘(algorithms)이 원본 비디오와 바꾸고 싶은 얼굴 이미지를 함께 학습함으로써 딥페이크 동영상을 만들어 낸다.⁴⁶⁾ 인공지능 기반의 소프트웨어가 원본 이미지에서 대상자의 입이나 얼굴이 움직이는 특징들을 찾아낸 후에 다른 사람의 동영상에 그러한 특징들을 복사해 넣고,⁴⁷⁾ 컴퓨터는 진짜처럼 보이도록 만들기 위해서 이 두 가지 동영상을 가장 잘 합성할 수 있는 방법을 찾으려 스스로를 훈련시킨다.⁴⁸⁾ 원본 이미지가 많으면 많을수록 딥페이크를 만드는 제작자는 더 그럴듯한 딥페이크 동영상을 만들 수 있으므로, 딥페이크 동영상에 유명정치인이나 할리우드 배우들과 같이 공개된 동영상이 많은 사람들이 자주 이용된다.⁴⁹⁾

심층학습을 이용하던 딥페이크 기술은 이제 ‘생산적 적대 신경망(generative adversarial network)’⁵⁰⁾을 이용하여 이전보다 더 진짜 같은

43) Bernard. (2018, October 1). 위의 자료 참조.

44) Hardesty, Larry. (2017, April 14). Explained: Neural networks, *MIT News*, URL: <https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>.

45) Bernard. (2018, October 1). 앞의 자료 참조.

46) Binder, Matt. (2020, January 20). Deepfakes Are Getting Easier to Make and the Internet’s Just Not Ready. *Mashable*, URL: <https://in.mashable.com/tech/10633/deepfakes-are-getting-easier-to-make-and-the-internets-just-not-ready>.

47) Dixon Jr., Herbert B. (2019, August 12). Deepfakes: More Frightening Than Photoshop on Steroids, *American Bar Association*, URL: https://www.americanbar.org/groups/judicial/publications/judges_journal/2019/summer/deepfakes-more-frightening-photoshop-steroids/.

48) Binder. (2020, January 20). 앞의 자료 참조.

49) Dixon Jr. (2019, August 12). 앞의 자료 참조.

가짜 동영상을 만들어 내고 있다.⁵¹⁾ 생산적 적대 신경망⁵²⁾은 2개의 기계학습모델 중 첫 번째 기계학습모델이 데이터를 이용하여 가짜를 만들어 내도록 하고, 두 번째 기계학습모델은 진짜를 찾아내도록 함으로써 두 모델이 끝까지 겨루게 하는 기술을 말한다.⁵³⁾ 이러한 생산적 적대 신경망 기술을 이용하여 만들어 낸 데이터가 진짜 데이터와 구별할 수 없을 때까지 최적화된 기술을 사용하여 진짜 같은 가짜를 만들 수 있게 된다.⁵⁴⁾

한편 딥페이크 동영상을 만드는 기술이 발달함에 따라 딥페이크 동영상을 찾아내는 기술도 발전하였다. 그러나 문제는 딥페이크 동영상을 찾아내는 “딥페이크 판별기(deepfake detectors)”로도 동영상의 진위 여부를 완벽하게 찾아내지는 못한다는 데 있다.⁵⁵⁾ 미국의 국방첨단과학기술연구

50) 보도에 의하면, 미국의 컴퓨터 설계회사인 엔비디아(NVIDIA)는 생산적 적대 신경망 기술을 “유명한 가짜 이미지 만들기에 적용했다. 엔비디아 테슬라 P100 GPU·엔비디아 딥 러닝 라이브러리 CUDA·개발 도구 cuDNN 기술로 구성된 신경망 네트워크는 유명한 사진으로부터 이목구비, 배경, 색상 등 정보를 취득했다. 이어 신경망 네트워크는 유명한 사진에 스스로 학습한 배경 흐림 효과를 넣고 이미지 구도를 변경, 새로운 이미지를 만들어냈다. 필요한 경우 화소 보간(화소를 채워넣어 해상도나 색상을 보완하는 기술)도 가한다. 실제로는 찍힌 적이 없는, 유명한 사진을 임의로 만들어낸 것. 생산적 적대 신경망이 이 기술을 습득하기까지 약 20일간 소요됐다”고 한다. 차주경 (2017, 10, 31). 엔비디아, 인공지능으로 ‘가짜 이미지 생성’ 기술 논문 발표. <ITChosun>, URL: http://it.chosun.com/site/data/html_dir/2017/10/31/2017103185046.html.

51) Porup, J. M. (2021, March 18). How and why deepfake videos work—and what is at risk, *CSCQ*, URL: <https://www.csoonline.com/article/3293002/deepfake-video-s-how-and-why-they-work.html>.

52) 생산적 적대 신경망은 2014년 몬트리올대학교의 이안 굿펠로우(Ian Goodfellow)와 동료 연구원들이 작성한 논문에서 처음 소개된 것으로 알려져 있다. Nicholson, Chris, A Beginner’s Guide to Generative Adversarial Networks (GANs), URL: <https://wiki.pathmind.com/generative-adversarial-network-gan>.

53) Porup, J. M. (2021, March 18). 앞의 자료.

54) Peng, Tony. (2019, April 4). Father of GANs Ian Goodfellow Splits Google For Apple, *Synced*, URL: <https://medium.com/syncedreview/father-of-gans-ian-good-fellow-splits-google-for-apple-279fcc54b328>.

55) Brown, Nina I. (2020). Deepfakes and the Weaponization of Disinformation, *Virginia Journal of Law & Technology* 23, 2; 캘리포니아대학교와 서던캘리포니아대학교의 연구원들이 개발한 딥페이크 판별 기술은 최소 92% 정도의 정확성을

소(Defense Advanced Research Projects Agency)는 수백만 불을 들여 딥페이크를 판별하는 도구를 개발하려 하고 있지만 생산적 적대 신경망을 이용하여 딥페이크를 만든 경우에는 대부분 딥페이크 탐지 소프트웨어에 감지되지 않는다고 한다.⁵⁶⁾

첨단기술은 급속히 발전하고 있으며, 예전과 달리 이제는 누구나 손쉽게 딥페이크 동영상을 만들 수 있게 되었다. 동영상 합성 프로그램인 페이크앱(FakeApp)은 인터넷에서 무료로 다운로드받을 수 있으며,⁵⁷⁾ 유튜브에는 페이크앱을 설치하고 사용할 수 있는 방법을 알려주는 동영상이 게시되어 있다.⁵⁸⁾

Ⅲ. 미국에서의 유튜브 딥페이크와 허위사실공표

1. 서설

미국의 경우 연방 차원에서 딥페이크를 규제하기 위해 여러 가지 법안들이 연방의회에 제출되었으나 모두 통과되지는 못하였다. 예를 들어, 2018년 12월에 연방의회에 제출된 「악의적 딥페이크 금지법」(Malicious

갖고 있다고 한다. Knight, Will. (2019, June 21). A new deepfake detection tool should keep world leaders safe—for now, *MIT Technology Review*, URL: <https://www.technologyreview.com/2019/06/21/134815/a-new-deepfake-detection-tool-should-keep-world-leaders-safefor-now/>.

56) Knight, Will. (2018, May 23). The US Military is Funding an Effort to Catch Deepfakes and Other AI Trickery, *MIT Technology Review*, URL: <https://www.technologyreview.com/2018/05/23/142770/the-us-military-is-funding-an-effort-to-catch-deepfakes-and-other-ai-trickery/>.

57) 예컨대, Filehorse, FakeApp. (2021, December 3). URL: <https://www.filehorse.com/download-fakeapp/>.

58) YouTube, tech4tress, Deepfakes Guide: FakeApp 2.2 Tutorial Installation (totally simplified, model folder included), URL: <https://www.youtube.com/watch?v=Lsv38PkLsGU>.

Deep Fake Prohibition Act of 2018)은 “연방법, 주법, 자치법규(local law) 또는 인디언법(Tribal law)상 범죄 또는 불법행위를 야기하는 (facilitate) 딥페이크를 배포할 의도로 제작”하는 경우 벌금이나 2년 이하의 징역 또는 양자를 병과하는 내용의 법안이었다.⁵⁹⁾ 특히, 동 법안은 선거의 실시나 외교관계에 관한 행위 또는 폭력을 조장하는 행위를 비롯하여 국가기관의 작용에 영향을 미칠 수 있는 딥페이크를 가중처벌하여 벌금이나 10년 이하의 징역 또는 양자를 병과하는 것으로 규정하였다.⁶⁰⁾ 또한 동 법안은 연방 수정헌법 제1조에 의해 보호되는 행위는 책임을 면제하는 규정도 두고 있었다.⁶¹⁾

한편 주 차원에서는 캘리포니아와 텍사스가 선거에 영향을 미칠 목적으로 가짜 동영상을 만들어 배포하는 등의 행위를 규제하는 법률을 제정하였다. 이하에서는 캘리포니아주의 선거법상 딥페이크금지법을 중심으로 법적인 문제점들을 살펴보기로 한다.

2. 유튜브 딥페이크와 허위사실공표

가. 캘리포니아주 딥페이크금지법

(1) 법률 규정

캘리포니아주는 2019년 9월에 미국에서 처음으로 선거법을 개정하여 딥페이크를 규제하고 있으며,⁶²⁾ 동 법은 2023년 1월 1일까지 한시적으

⁵⁹⁾ Malicious Deep Fake Prohibition Act(2018), GovTrack, URL: <https://www.govtrack.us/congress/bills/115/s3805>.

⁶⁰⁾ 위 법률안 참조.

⁶¹⁾ 위 법률안 참조.

⁶²⁾ Cal. Elec. Code § 20010(a). “(a) Except as provided in subdivision (b), a person, committee, as defined in Section 82013 of the Government Code, or other entity shall not, within 60 days of an election at which a candidate for elective office will appear on the ballot, distribute, with actual malice, materially

로 시행된다.⁶³⁾ 동 법은 다음과 같이 규정하고 있다.

(b)항에 규정된 경우를 제외하고, 주정부 법령집 제82031조에 규정된 개인, 위원회 또는 그 밖의 권리주체(entity)는 공직후보를 선출하기 위한 선거일로부터 60일 이내에 현실적 악의로(with actual malice) 후보자의 명예를 훼손하거나 유권자가 후보자에게 찬성 또는 반대의 투표를 하도록 기망하기 위해 (e)항에 규정된 바와 같이 현저히 기망하는(materially deceptive) 내용의 후보자에 관한 음성 또는 동영상 매체를 배포하여서는 아니 된다.

동 법은 이미지, 비디오 또는 오디오가 조작된 것이라는 표시를 한 경우에는 예외적으로 동 법을 적용하지 않는 것으로 규정하고 있다.⁶⁴⁾ 한편 동 법은 딥페이크 대신에 “현저히 기망하는 오디오 매체나 비디오 매체”라는 용어를 사용하면서 그 의미를 다음과 같이 정의하고 있다.⁶⁵⁾

deceptive audio or visual media, as defined in subdivision (e), of the candidate with the intent to injure the candidate’s reputation or to deceive a voter into voting for or against the candidate.”

63) Cal. Elec. Code § 20010(g). “This section shall remain in effect only until January 1, 2023, and as of that date is repealed, unless a later enacted statute, that is enacted before January 1, 2023, deletes or extends that date.”

64) Cal. Elec. Code § 20010(b). “(b)(1) The prohibition in subdivision (a) does not apply if the audio or visual media includes a disclosure stating: ‘This ___ has been manipulated.’ (2) The blank in the disclosure required by paragraph (1) shall be filled with whichever of the following terms most accurately describes the media: (A) Image. (B) Video. (C) Audio.”

65) Cal. Elec. Code § 20010(e). “As used in this section, ‘materially deceptive audio or visual media’ means an image or an audio or video recording of a candidate’s appearance, speech, or conduct that has been intentionally manipulated in a manner such that both of the following conditions are met: (1) The image or audio or video recording would falsely appear to a reasonable person to be authentic. (2) The image or audio or video recording would cause a reasonable person to have a fundamentally different understanding or impression of the expressive content of the image or audio or video recording than that person would have if the person were hearing or seeing the unaltered, original version of the image or audio or video recording.”

“현저히 기망하는 오디오 매체나 비디오 매체(materially deceptive audio or visual media)”란 다음 2가지 요건 즉, (1) 합리적인 사람에게 진짜로 잘못 인식될 수 있는 이미지나 음성 또는 동영상 녹화물로서 (2) 그러한 이미지나 음성 또는 동영상 녹화물이 합리적인 사람으로 하여금 조작되지 않은 원본 이미지나 음성 또는 동영상 녹화물을 보았을 경우보다 그 이미지나 음성 또는 녹화물의 표현 내용에 대하여 본질적으로 다른 인식이나 인상을 갖게 만드는 것이라는 요건을 모두 충족하는 방법을 통해 고의로 조작된 후보자의 외모, 말 또는 행동을 기록한 이미지 또는 음성이나 동영상 녹화물을 의미한다.

동 법의 시행 후 2020년 9월에 대통령 선거를 앞두고 딥페이크 동영상 만들게 해 주는 스마트폰 애플리케이션인 임프레션스(Impressions)는 트럼프 대통령의 얼굴을 이용하여 딥페이크 비디오를 만들 수 있는 기능을 임프레션스에서 삭제하였다.⁶⁶⁾ 한편 딥페이크 동영상을 만들 수 있는 애플리케이션인 리페이스(Reface)는 정치인의 동영상에 이용자의 얼굴을 넣어 딥페이크 동영상을 만들 수는 있지만 정치인이 등장하는 가짜 동영상을 만들 수는 없도록 하였다.⁶⁷⁾ 이처럼 캘리포니아주가 미국에서 처음으로 선거와 관련한 딥페이크 동영상을 배포하는 행위를 규제하는 법률을 제정함에 따라 딥페이크 동영상을 만들 수 있는 애플리케이션들도 선거와 관련한 가짜 동영상을 만드는 기능을 삭제하는 등의 조치를 하고 있다.

66) Thalen, Mikael. (2020, September 2). Deepfake app takes Trump videos offline until after the election, *Daily Dot*, URL: <https://www.dailydot.com/debug/deep-fake-app-trump-2020-election/>.

67) Lomas, Natasha. (2020, August 18). Deepfake video app Reface is just getting started on shapeshifting selfie culture, *Techcrunch*, URL: <https://techcrunch.com/2020/08/17/deepfake-video-app-reface-is-just-getting-started-on-shapeshifting-selfie-culture/>.

(2) 딥페이크와 언론의 자유

미국 연방대법원은 *Ashcroft v. American Civil Liberties Union* 사건에서 일반적으로 언론의 자유를 보장하고 있는 연방 수정헌법 제1조는 “정부가 어떤 표현의 메시지, 아이디어, 대상(subject matter) 또는 그 내용을 이유로 표현을 제한할 수 있는 권한이 없다”는 것을 의미한다고 하였고,⁶⁸⁾ *Reed v. Town of Gilbert* 사건에서는 어떤 표현의 내용을 규제하는 법률은 “위헌이라는 추정을 받으며, 정부가 긴절한 국가의 이익(compelling state interests)을 달성하기 위해 필요한 최소한의 범위에서(narrowly tailored) 표현의 자유를 제한하는 것이라는 점을 입증한 경우에만 정당화될 수 있다”고 판시하였다.⁶⁹⁾ 그런데 캘리포니아주의 딥페이크금지법은 미국 연방대법원이 말하는 표현의 내용에 대한 규제에 해당한다. 딥페이크 동영상의 전달하려는 내용이 풍자나 패러디인 경우는 물론 특정한 후보자를 당선되게 하거나 당선되지 못하게 하는 경우에도 마찬가지이다. 특히 캘리포니아주의 딥페이크금지법은 “현저히 기망하는 오디오 매체나 비디오 매체”라고 함으로써 내용을 기준으로 딥페이크를 규제하고 있다. 그런데 미국의 경우 내용에 근거한 규제, 특히 정치적인 의사표현에 대한 규제의 위헌여부에 대한 심사는 “엄격심사(strict scrutiny)”가 적용되며, 그러한 표현행위를 규제해야 하는 국가의 긴절한 이익이 있고 필요한 최소한의 범위에서만 기본권을 제한하였다는 사실을 정부가 입증해야 한다.⁷⁰⁾

⁶⁸⁾ *Ashcroft v. American Civil Liberties Union*, 535 U.S. 564, 573 (2002).

⁶⁹⁾ *Reed v. Town of Gilbert*, 576 U.S. 155, 163 (2015).

⁷⁰⁾ *Reed v. Town of Gilbert*, 576 U.S. 155, 163-164 (2015). 한편 캘리포니아주 딥페이크금지법의 위헌 여부에 대해서는 연방대법원이 ‘*Citizens v. FEC*, 558 U.S. 310 (2010)’ 판결에서 적용한 “적정심사(exacting scrutiny)” 즉, “충분히 중요한 정부의 이익(sufficiently important governmental interest)”이 있는지의 여부를 심사하는 원칙이 적용될 것이므로, 패러디나 풍자가 아닌 정치적 딥페이크가 선거에 입박하여 공표된 경우에는 결국 합헌으로 결정될 것이라는 견해가 있다. Bodi, Matthew. (2021). The First Amendment Implications of Regulating Political

딥페이크규제법과 관련하여 공직선거의 후보자를 당선되게 하거나 당선되지 못하게 할 목적으로 딥페이크 동영상을 만들어 공표한 행위는 선거의 공정성을 해치는 행위로서 국가는 그러한 행위를 규제해야 할 ‘긴절한 이익’이 있다고 주장할 것으로 보인다. 그러나 미국 연방대법원은 *Brown v. Hartlage* 사건에서 “선거절차의 엄결성(integrity)을 보호하는 것”은 “적법한(legitimate)” 국가의 이익은 필지언정 긴절한 이익은 아니라고 하였다.⁷¹⁾ 이 사건에서 연방대법원은 특히 “유권자들이 잘못된 정보를 근거로 선택을 할 수도 있다는 우려는 국가가 언론을 제한하는 긴절한 정당화 사유가 되지 못한다”고 하면서, “어떤 이슈가 논의 또는 토론할 가치가 있는지를 선택하는 것은 정부의 역할이 아니”라고 판시하였다.⁷²⁾ 연방대법원의 이러한 논거는 만약 국가가 어떤 정치적 표현이 옳은 것인지의 여부를 판단하고 규제할 수 있게 되면 국가는 자신의 정책에 대해 비판적인 입장을 보인 정치적 의사표현을 규제하려고 할 것이며, 이는 표현의 자유에 대한 중대한 제약이 될 수 있다는 점을 고려한 것이라고 할 수 있다. 더 나아가서 어느 특정한 후보가 허위의 사실을 주장하는 내용의 딥페이크 동영상이 제작되어 공표된다고 하더라도 유권자들이 그러한 동영상을 진짜라고 믿지 않을 수도 있고, 그러한 동영상이 선거에서 유권자가 후보자를 선택하는데 영향을 미쳐서 선거의 공정성을 저해한다는 것을 입증하기 곤란한 경우도 있을 수 있다.⁷³⁾

한편 기본권 제한이 ‘필요한 최소한의 범위’에 머물러야 한다는 원칙

Deepfakes, *Rutgers Computer & Technology Law Journal* 47, 164. 따라서 규제의 필요성의 측면에서 보면 캘리포니아주 딥페이크금지법의 위헌여부는 국가가 정치적 딥페이크를 금지하는 데에 ‘긴절한 이익(compelling interest)’이 있는지의 여부를 심사하는 ‘엄격심사’ 기준을 적용할 것인지, 아니면 ‘충분히 중요한 이익(sufficiently important interest)’이 있는지의 여부를 심사하는 ‘적정심사’ 기준을 적용할 것인지에 따라 위헌인지의 여부가 달라진다고 할 수 있다.

71) *Brown v. Hartlage*, 456 U.S. 45, 53-54 (1982).

72) *Brown v. Hartlage*, 456 U.S. 45, 60 (1982).

73) Rowbottom, Jacob. (2012). Lies, Manipulation and Elections—Controlling False Campaign Statements, *Oxford Journal of Legal Studies* 32, 516.

과 관련하여, 미국 연방대법원은 *McCullen v. Coakley* 사건에서 언론의 자유에 대한 제한은 “국가의 정당한 이익을 달성하는데 필요한 그 이상 이어서는 아니 된다”고 하면서,⁷⁴⁾ 국가는 “실질적으로 언론을 덜 제한하는 다른 조치들이 국가의 이익을 달성하는데 실패할 것이라는 점을 입증하여야 하며, 단지 국가가 선택한 규제 방법이 용이하다는 점만을 입증하는 것으로는 부족하다”고 판시하였다.⁷⁵⁾ 캘리포니아주의 딥페이크금지법은 패러디나 풍자가 아닌 한 공직선거의 후보자에 관한 조작된 내용의 동영상을 전면적으로 금지하고 있으므로 그 규제의 범위가 넓다. 또한 동 법은 기록으로 남아 있지 않은 과거의 실제 사실을 재연하기 위해 딥페이크 동영상을 만들거나 후보자가 자신의 딥페이크 동영상을 만드는 것을 금지할 수도 있다는 문제점이 있다.⁷⁶⁾ 동 법의 입법 목적은 악의적인 허위사실공표로부터 후보자를 보호하고 선거의 공정성을 지키는 것이지만, 실제로는 그 적용범위가 넓어서 후보자에 관한 딥페이크 이미지나 음성 또는 동영상 녹화물이 광범위하게 규제를 받게 된다. 나아가 동 법이 적용 예외로 하고 있는 패러디나 풍자에는 해당하지 않지만 딥페이크를 이용한 사진이나 음성 또는 동영상 중에는 정치적인 표현의 자유로서 보호되어야 할 가치가 있는 표현도 있을 수 있다.

(3) 딥페이크와 명예훼손

캘리포니아주는 명예훼손을 범죄로 규정하여 처벌하지 않고, 단지 민사상 불법행위의 하나로 규율하고 있다.⁷⁷⁾ 캘리포니아주 민법은 명예훼손

⁷⁴⁾ *McCullen v. Coakley*, 573 U.S. 464, 486 (2014).

⁷⁵⁾ *McCullen v. Coakley*, 573 U.S. 464, 495 (2014).

⁷⁶⁾ Fink, David E. & Diamond, Sarah E. (2020, September 3). Deepfakes: 2020 and Beyond, *The Recorder*, URL: <https://www.law.com/therecorder/2020/09/03/deep-fakes-2020-and-beyond/?slreturn=20220211072801>.

⁷⁷⁾ Cal. Civ. Code § 45. “Libel is a false and unprivileged publication by writing, printing, picture, effigy, or other fixed representation to the eye, which exposes any person to hatred, contempt, ridicule, or obloquy, or which causes him to be

손을 “원고를 혐오(hatred), 경멸, 조롱(ridicule) 또는 오욕(obloquy)을 당하게 만들거나 따돌림을 당하게 하거나 소외되도록 만드는 허위사실의 공표로서 명예훼손의 예외로 인정되는 특권의 경우가 아닌(a false and unprivileged publication) 공표행위”라고 정의하고 있는데,⁷⁸⁾ 1964년 미국 연방대법원의 *New York Times Co. v. Sullivan* 판결에 따라, 명예훼손 소송에서 원고가 공적 인물(public figure)인 경우에는 피고가 ‘현실적 악의’를 가지고 명예를 훼손하는 사실을 적시하였다는 것을 원고가 입증해야 한다.⁷⁹⁾ 민사상 불법행위인 명예훼손과 마찬가지로 캘리포니아주 뱃페이크금지법에 따르면 ‘현실적 악의’ 즉, “사진이나 그림에 합성된(superimposed) 사람의 이미지가 허위의 적시(false representation)라는 인식 또는 허위의 적시를 위해 사진이나 그림에 어떤 사람의 이미지가 합성되었는지 여부를 확인하지 않고 미필적으로 무시(disregard)”하였다는 것을 원고가 입증해야 한다.⁸⁰⁾ 그런데 미국의 경우 명예훼손이 성립하기 위해서는 적시된 사실이 진실이라고 믿을 만한 것(believability)이어야 한다.⁸¹⁾ 캘리포니아주 뱃페이크금지법 역시 명예훼손과 유사하게 “합리적인 사람이 조작되지 않은 원본 이미지나 음성 또는 동영상 녹화물을 듣거나 보았을 경우와는 달리 이미지나 음성 또는 녹화물의 표현

shunned or avoided, or which has a tendency to injure him in his occupation.” Cal. Civ. Code § 45(a). “A libel which is defamatory of the plaintiff without the necessity of explanatory matter, such as an inducement, innuendo or other extrinsic fact, is said to be a libel on its face. Defamatory language not libelous on its face is not actionable unless the plaintiff alleges and proves that he has suffered special damage as a proximate result thereof. Special damage is defined in Section 48a of this code.”

78) *Brodeur v. Atlas Ent., Inc.*, 204 Cal. Rptr. 3d 483, 492 (Ct. App. 2016).

79) *Dickinson v. Cosby*, 250 Cal. Rptr. 3d 350, 364 (Ct. App. 2019).

80) Cal. Elec. Code § 20010(a). “... For purposes of this section, ‘actual malice’ means the knowledge that the image of a person has been superimposed on a picture or photograph to create a false representation, or a reckless disregard of whether or not the image of a person has been superimposed on a picture or photograph to create a false representation.”

81) *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46, 52-54 (1988).

내용에 대하여 본질적으로 다른 인식이나 인상을 갖게 만드는 것”을 딥페이크 규제 요건으로 규정하고 있다.⁸²⁾

한편 미국 캘리포니아주의 경우 공직선거에서 딥페이크 동영상을 이용한 허위사실공표가 있는 경우에 후보자는 민사상 명예훼손을 이유로 법원에 소송을 제기할 수도 있고 딥페이크금지법 위반을 이유로 게시 금지 가처분(injunctive relief) 신청이나 손해배상을 청구할 수도 있다.⁸³⁾ 그런데 딥페이크금지법은 “선거일로부터 60일 이내”⁸⁴⁾에 공표된 딥페이크 동영상만을 규율 대상으로 하고 있다는 점에서 명예훼손에 비해 그 보호범위가 좁다고 할 수 있다.

(4) 딥페이크와 프라이버시 침해

딥페이크 동영상이 명예훼손에 해당되지 않는 경우에는 딥페이크 동영상으로 인해 일반인들에게 “잘못된 인상(a false light)”을 심어주었다는 이유로 보통법상 프라이버시의 침해를 주장할 수도 있다.⁸⁵⁾ 명예훼손과 달리 프라이버시 침해의 경우에는 피고가 고의로 잘못된 인상을 심어 주었다는 사실을 원고가 입증할 필요가 없다. 이 점에서 명예훼손과 프라이버시의 침해는 구별된다.⁸⁶⁾ 캘리포니아주 딥페이크금지법이 규정하고 있는 “허위의 적시(false representation)”에 대응하는 개념이 프라이

82) Cal. Elec. Code § 20010(e). “... cause a reasonable person to have a fundamentally different understanding or impression of the expressive content of the image or audio or video recording than that person would have if the person were hearing or seeing the unaltered, original version of the image or audio or video recording.”

83) Tashman, 앞의 논문, 1404쪽.

84) Cal. Elec. Code § 20010(a). “... within 60 days of an election at which a candidate for elective office will appear on the ballot, ...”

85) Cal. Civ. Prac. Torts Common Law Action § 20:12; Tashman, 앞의 논문, 1405쪽.

86) Tashman, Alexandra. (2021). “Malicious Deepfakes”—How California’s A.B. 730 Tries (and Fails) to Address the Internet’s Burgeoning Political Crisis, *Loyola of Los Angeles Law Review* 54, 1405.

버시 침해의 요건인 “잘못된 인상”이라고 할 수 있는데, 딥페이크 동영상에 대한 법적인 대처방법으로는 딥페이크금지법 위반보다 프라이버시 침해를 주장하는 것이 더 실효적인 방법이 될 수 있다. 왜냐하면 프라이버시 침해를 주장하는 경우에는 원고가 동영상을 공표한 피고의 ‘현실적 악의’를 입증하지 않아도 되기 때문이다.

(5) 딥페이크와 퍼블리시티권

캘리포니아주 민법 제3344조는 사전 동의 없이 타인의 “성명, 음성, 서명, 사진 또는 초상을 고의로 이용하는” 행위를 퍼블리시티권(right of publicity)의 침해로 규정하고 있다.⁸⁷⁾ 따라서 딥페이크 동영상으로 피해를 입은 후보자는 퍼블리시티권의 침해를 주장할 수 있는지가 문제될 수 있으나, 퍼블리시티권의 침해는 경제적 또는 “상업적 이득(commercial gain)”을 목적으로 피해자인 원고의 초상 등을 이용할 것을 요건으로 하고 있으므로, 적어도 순수한 정치적인 표현의 의도로서 특정 후보자의 당선이나 낙선을 목적으로 하는 경우에는 ‘상업적 이득’이라고 보기 어려우므로 퍼블리시티권의 침해를 주장하기는 어려울 것으로 보인다.⁸⁸⁾

(6) 딥페이크와 저작권

딥페이크 동영상은 대부분 기존에 존재하던 동영상이나 저작물을 이용하여 만들어진다. 따라서 딥페이크에 저작권의 보호를 받는 동영상이

⁸⁷⁾ Cal. Civ. Code § 3344(a). “(a) Any person who knowingly uses another’s name, voice, signature, photograph, or likeness, in any manner, on or in products, merchandise, or goods, or for purposes of advertising or selling, or soliciting purchases of, products, merchandise, goods or services, without such person’s prior consent, ...”

⁸⁸⁾ Chesney, Bobby & Citron, Danielle. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, *California Law Review* 107, 1794.

이용되는 경우에는 「디지털 밀레니엄 저작권법」(Digital Millennium Copyright Act)에 따라 저작권 침해가 문제될 수 있으며, 저작권자가 아닌 사람이 저작권의 보호를 받는 저작물을 온라인상에서 배포하는 경우에는 ‘삭제(takedown)’를 요구할 수 있다.⁸⁹⁾ 「디지털 밀레니엄 저작권법」에 따라 저작권자뿐만 아니라 영상물에 등장하는 사람이면 누구나 영상물의 삭제를 요구할 수 있으므로, 허위사실을 공표하는 내용의 딥페이크 동영상에 등장하는 후보자는 영상물의 저작권자인지의 여부와 관계없이 딥페이크를 삭제하라고 요구할 수 있다.⁹⁰⁾ 따라서 캘리포니아주의 딥페이크금지법에 따라 영상물에 대한 게시 금지 가처분을 신청하는 것보다는 「디지털 밀레니엄 저작권법」에 의하여 신속한 딥페이크 영상물의 삭제를 요구하는 것이 더 효과적인 구제방법이라고 할 수 있다.

(7) 딥페이크와 법적용의 곤란성

유튜브에 게시된 딥페이크 사진, 음성 또는 동영상의 게시자를 찾아내는 것이 불가능한 경우가 많다. 앞서 살펴본 바와 같이 누구나 손쉽게 딥페이크 동영상을 만들어 유튜브에 올릴 수 있는 반면에 게시자는 국내가 아닌 국외에 있을 수도 있으며, 신원을 감추면 추적하는 것이 사실상 불가능할 수도 있다.⁹¹⁾ 한편 캘리포니아주 딥페이크금지법은 “현실적 악의(with actual malice)로 후보자의 명예를 훼손하거나 유권자가 후보자에게 찬성 또는 반대의 투표를 하도록 기망하기 위해” 딥페이크를 이용한 음성이나 동영상 매체를 배포하는 행위를 금지하고 있다. 그런데

⁸⁹⁾ 17 U.S.C. 512.

⁹⁰⁾ Delfino, Rebecca A. (2019). Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn’s Next Tragic Act, *Fordham Law Review* 88, 936.

⁹¹⁾ Reynolds, Matt. (2020, June 9). Courts and lawyers struggle with growing prevalence of deepfakes, *ABA Journal*, URL: <https://www.abajournal.com/web/article/courts-and-lawyers-struggle-with-growing-prevalence-of-deepfakes>.

문제는 공직선거의 후보자가 딥페이크를 배포한 자의 ‘현실적 악의’ 즉, “사진이나 그림에 합성된 사람의 이미지가 허위의 적시라는 인식 또는 허위의 적시를 위해 사진이나 그림에 어떤 사람의 이미지가 합성되었는지 여부를 확인하지 않고 미필적으로 무시”하였다는 것을 입증하는 것이 쉽지 않다는 점이다. 물론 딥페이크는 “현저히 기망하는 내용의 후보자에 관한 음성 또는 동영상 매체”이므로 이를 만든 사람은 딥페이크의 내용이 거짓이라는 것을 알고 있었다고 할 것이지만,⁹²⁾ 문제는 그것을 입증해야 하는 책임이 딥페이크 동영상으로 인해 피해를 입은 후보자에게 있어서 실제로는 딥페이크를 배포한 자의 현실적 악의를 입증하는 것이 어려울 수도 있다는 점이다.

(8) 딥페이크와 「통신품위법」

앞서 살펴본 바와 같이 미국은 「통신품위법」 제230조를 규정하여 “서비스 제공자나 이용자가 헌법적으로 보호를 받는지 여부와 관계없이 선의로 음란하거나, 외설적이거나, 선정적이거나, 저속하거나, 과도하게 폭력적이거나, 괴롭히거나 그 밖의 불쾌한 자료에 대한 접근이나 이용을 제한하는 조치를 하거나 또는 그러한 자료에 대한 접근을 제한하기 위하여 정보 제공자나 그 밖의 자가 취한 기술적 조치와 관련하여 상호 연결된 컴퓨터 서비스의 제공자나 이용자는 어떠한 책임도 지지 아니”하도록 해왔다.⁹³⁾ 따라서 유튜브와 같은 플랫폼들은 이용자가 게시한 공

⁹²⁾ Bodi, Matthew. (2021). The First Amendment Implications of Regulating Political Deepfakes, *Rutgers Computer & Technology Law Journal* 47, 162.

⁹³⁾ 47 U.S.C. 230(c)(2). “(2) Civil liability. (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or (B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1).”

직선거의 후보자에 관한 허위사실의 딥페이크 동영상에 대해 법적인 책임을 면제받기 때문에 그러한 동영상을 규제하는데 적극적인 노력을 기울이려고 하지 않을 수도 있다.

나. 텍사스주 딥페이크금지법

텍사스주의 딥페이크금지법은 “누구든지 후보자에게 해를 가하거나 선거 결과에 영향을 미치려는 의도로 (1) 딥페이크 비디오를 제작하고 (2) 선거일로부터 30일 이전에 그 딥페이크 비디오를 공표하거나 배포하는 행위는 범죄로 한다”고 규정하고,⁹⁴⁾ 여기서 말하는 딥페이크 비디오란 “속이려는 의도로 제작된 비디오로서 실존 인물(a real person)이 실제로는 하지 아니한 행위를 한 것처럼 보이게 하는 비디오를 의미”하는 것으로 규정하고 있다.⁹⁵⁾ 그런데 동 법은 캘리포니아주 딥페이크금지법과 달리 ‘현실적 악의’를 형사처벌의 구성요건으로 규정하고 있지 않다. 또한 동 법은 딥페이크가 패러디나 풍자인 경우에는 동 법을 적용하지 않는다는 예외규정을 두고 있지도 않다.

앞서 살펴본 바와 같이 미국 연방대법원이 *New York Times Co. v. Sullivan* 사건에서 제시한 ‘현실적 악의’의 법리에 비추어 볼 때 텍사스주 딥페이크금지법은 위헌으로 선언될 가능성이 높다는 견해가 유력하다.⁹⁶⁾ 또한 동 법은 “속이려는 의도로 제작된 비디오로서 실존 인물이 실제로는 하지 아니한 행위를 한 것처럼 보이게 하는” 표현물을 제작하

⁹⁴⁾ Tex. Elec. Code § 255.004(d). “(d) A person commits an offense if the person, with intent to injure a candidate or influence the result of an election: (1) creates a deep fake video; and (2) causes the deep fake video to be published or distributed within 30 days of an election.”

⁹⁵⁾ Tex. Elec. Code § 255.004(e). “In this section, ‘deep fake video’ means a video, created with the intent to deceive, that appears to depict a real person performing an action that did not occur in reality.”

⁹⁶⁾ Bodi, Matthew. (2021). The First Amendment Implications of Regulating Political Deepfakes, *Rutgers Computer & Technology Law Journal* 47, 171.

여 공표하는 등의 행위를 처벌하는 것이므로, 정치적 표현물의 ‘내용’을 기준으로 형사제재를 부과하는 것이어서 그 위헌 여부를 심사함에 있어서 ‘엄격심사’ 기준이 적용될 것이다. 그런데 캘리포니아주 딥페이크금지법과 달리, 텍사스주 딥페이크금지법은 풍자나 패러디를 적용예외로 규정하고 있지 않으므로, 언론의 자유에 대한 최소한의 제한으로 인정될 수 없어서 위헌으로 선언될 가능성이 높다.⁹⁷⁾

3. 소결

미국의 경우 현재까지 새로운 법률을 제정하여 딥페이크를 규제하는 주는 캘리포니아와 텍사스밖에 없다. 그런데 앞에서 살펴본 바와 같이 캘리포니아주의 딥페이크금지법은 ‘현실적 악의’의 원칙을 규정하고 있으므로, 실제에 있어서는 동 법을 적용하기 어려운 경우가 많을 것으로 예상된다. 따라서 딥페이크를 규제하는 새로운 법률을 만들기보다는 명예훼손법이나 저작권법과 같이 기존에 있던 법적 규제를 통해 딥페이크의 문제를 해결하는 것이 바람직하다는 주장이 제기되고 있다.⁹⁸⁾ 반면에 텍사스주의 딥페이크금지법은 캘리포니아주 딥페이크금지법과 달리 ‘현실적 악의’의 법리를 명시적으로 규정하지 않고 있고, 풍자나 패러디의 경우를 적용대상에서 제외하는 규정도 없으므로, 엄격심사 기준에 의할 경우 연방 헌법에 위반된다고 선언될 가능성이 높다고 할 수 있다.

97) Bodi. (2021). 위의 논문, 171쪽.

98) 캘리포니아주의 딥페이크금지법은 표현의 자유를 규정한 연방 수정헌법 제1조에 위반된다는 견해가 있다. Tashman, Alexandra. (2021). “Malicious Deepfakes”-How California’s A.B. 730 Tries (and Fails) to Address the Internet’s Burgeoning Political Crisis, *Loyola of Los Angeles Law Review* 54, 1408-1416. 이와는 달리 동 법이 ‘현실적 악의’의 법리를 규정하고 있고, 패러디나 풍자의 경우에는 적용되지 않으므로 합헌이라는 견해도 있다. Bodi. (2021). 앞의 논문, 164쪽.

IV. 우리나라에서의 유튜브 딥페이크와 허위사실공표

1. 서설

최근 우리나라에서도 딥페이크 선거운동을 금지하는 내용의 「공직선거법」 개정안이 발의되었다. 이하에서는 먼저 거짓말이 우리 헌법상 표현의 자유의 보호영역에 포함되는 것인지에 관하여 살펴보고, 이어서 유튜브에 게시된 딥페이크 동영상이 「공직선거법」상 허위사실공표죄에 해당하는 것인지를 검토해 보기로 한다. 아울러 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」에 따라 딥페이크 동영상의 삭제를 요구할 수 있는지의 여부도 간략하게 살펴보기로 한다.

2. 거짓말과 표현의 자유

우리 헌법상 의도적인 거짓말은 어떠한 법적 성격을 가지고 있는가? 2010년 공익을 해할 목적의 허위의 통신을 금지한 「전기통신기본법」 규정의 위헌여부가 문제된 이른바 미네르바 사건에서, 5명의 헌법재판관들은 “‘허위사실’이라는 것은 언제나 명백한 관념은 아니다. 어떠한 표현에서 ‘의견’과 ‘사실’을 구별해내는 것은 매우 어렵고, 객관적인 ‘진실’과 ‘거짓’을 구별하는 것 역시 어려우며, 현재는 거짓인 것으로 인식되지만 시간이 지난 후에 그 판단이 뒤바뀌는 경우도 있을 수 있다. 이에 따라 ‘허위사실의 표현’임을 판단하는 과정에는 여러 가지 난제가 뒤 따른다”고 하면서,⁹⁹⁾ “표현이 어떤 내용에 해당한다는 이유만으로 표현의 자유의 보호영역에서 애당초 배제된다고는 볼 수 없고, ‘허위사실의 표현’이 일정한 경우 사회윤리 등에 반한다고 하여 전체적으로 표현의 자유의

⁹⁹⁾ 현재 2010.12.28. 선고 2008헌바157등, 재판관 이강국, 재판관 이공현, 재판관 조대현, 재판관 김종대, 재판관 송두환의 과잉금지원칙 위반 여부에 관한 보충의견, 판례집 22-2하, 699.

보호영역에서 배제시킬 수는 없다. ‘허위사실의 표현’도 헌법 제21조가 규정하는 언론·출판의 자유의 보호영역에는 해당하되, 다만 헌법 제37조 제2항에 따라 국가안전보장·질서유지 또는 공공복리를 위하여 제한할 수 있는 것이라고 해석하여야 할 것”이라고 하였다.¹⁰⁰⁾ 이 사건에서 반대의 견을 제시한 2명의 헌법재판관 역시 “허위사실이라고 하여 반드시 타인의 명예나 권리 또는 공중도덕이나 사회윤리를 침해하는 것은 아니므로 허위사실의 표현도 표현의 자유의 보호영역에서 배제되는 것은 아니”라고 하였다.¹⁰¹⁾

이러한 헌법재판소의 판단과 관련하여, 다수의견이 “이상한 논리와 결론에 이르게 된 것은 ‘허위임을 알고도 이를 전파하는 행위’와 ‘허위임을 모르고 이를 전파하는 행위’를 구분하지 않기 때문”이라고 하면서,¹⁰²⁾ “허위임을 알면서 행하는 허위의 의도적 표현은 위축효과와 관계없기 때문에 ‘불가피’한 것이 아니다. 보호할 이유가 전혀 없다”고 주장하는 견해가 있다.¹⁰³⁾ 더 나아가 이 견해는 “허위사실의 표현은 표현의 자유의 보호영역에 속하는 것이 아니기 때문에 이 사건에서 제한받는 기본권은 신체의 자유이지, 표현의 자유가 아니다. 그럼에도 불구하고 헌법재판소는 이를 제대로 검토하지 않았다”고 주장한다.¹⁰⁴⁾ 다시 말해서 이 견해는 거짓말이 표현의 자유의 보호영역 내에 있는 것인지의 여부는 의도적인 거짓말이었는지 아니면 모르고 한 거짓말이었는지 여부에 따라 달라진다고 한다.

한편 미국 연방대법원은 2012년 *United States v. Alvarez* 사건에서,

100) 헌재 2010.12.28. 선고 2008헌바157등, 재판관 이강국, 재판관 이공현, 재판관 조대현, 재판관 김종대, 재판관 송두환의 과잉금지원칙 위반 여부에 관한 보충의견, 판례집 22-2하, 700.

101) 헌재 2010.12.28. 선고 2008헌바157등, 재판관 이동흡, 재판관 목영준의 반대의견, 판례집 22-2하, 705.

102) 문재완 (2011), 허위사실의 표현과 표현의 자유-한국과 미국의 판례 비교를 중심으로, <공법연구>, 39집 3호, 136.

103) 문재완 (2011), 위의 논문, 136-137쪽.

104) 문재완 (2011), 위의 논문, 137쪽.

자신이 “25년간 복무하고 전역한 해병”이며 “1987년에 무공훈장 (Congressional Medal of Honor)을 받았다”고 거짓말을 한 알바레즈 (Xavier Alvarez)에게 적용된 「군경력 사칭 금지법」(Stolen Valor Act)의 처벌 규정은 위헌이라고 판결하였다.¹⁰⁵⁾ 이 사건에서 연방대법원은 사회에 중대한 해악을 끼치지 않는 경우에는 의도적인 거짓말도 언론의 자유의 보호영역 내에 있다고 본 것이다.¹⁰⁶⁾ 우리 헌법은 제37조 제2항에서 “국민의 모든 자유와 권리는 국가안전보장·질서유지 또는 공공복리를 위하여 필요한 경우에 한하여 법률로써 제한할 수 있으며, 제한하는 경우에도 자유와 권리의 본질적인 내용을 침해할 수 없다”고 규정하고 있다. 따라서 비록 의도적인 거짓말이라고 하더라도 언론의 자유의 보호영역에 속하는 것이며, 다만 “필요한 경우에 한하여” 법률로써 제한할 수 있다고 해석하는 것이 우리 헌법 제37조 제2항의 해석에도 합치되는 것이라고 생각된다.

3. 유튜브 딥페이크와 허위사실공표

가. 「공직선거법」상 허위사실공표죄

우리나라의 「공직선거법」 제250조 제1항은 ‘당선목적 허위사실공표죄’에 관하여 다음과 같이 규정하고 있다.

당선되거나 되게 할 목적으로 연설·방송·신문·통신·잡지·벽보·선전문서 기타의 방법으로 후보자(후보자가 되고자 하는 자를 포함한다. 이하 이 조에서 같다)에게 유리하도록 후보자, 후보자의 배우자 또는 직계존비속이나 형제자매의 출생지·가족관계·신분·직업·경력등·재산·행위·소속단체, 특정인 또는 특정단체로부터의 지지여부 등에 관하여

¹⁰⁵⁾ *United States v. Alvarez*, 567 U.S. 709, 714-715 (2012).

¹⁰⁶⁾ *United States v. Alvarez*, 567 U.S. 709, 723 (2012).

허위의 사실[학력을 게재하는 경우 제64조제1항의 규정에 의한 방법으로 게재하지 아니한 경우를 포함한다]을 공표하거나 공표하게 한 자와 허위의 사실을 게재한 선전문서를 배포할 목적으로 소지한 자는 5년 이하의 징역 또는 3천 만원 이하의 벌금에 처한다.

한편 같은 법 같은 조 제2항은 ‘낙선목적 허위사실공표죄’에 관하여 다음과 같이 규정하고 있다.

당선되지 못하게 할 목적으로 연설·방송·신문·통신·잡지·벽보·선전문서 기타의 방법으로 후보자에게 불리하도록 후보자, 그의 배우자 또는 직계존·비속이나 형제자매에 관하여 허위의 사실을 공표하거나 공표하게 한 자와 허위의 사실을 게재한 선전문서를 배포할 목적으로 소지한 자는 7년 이하의 징역 또는 500만원 이상 3천만원 이하의 벌금에 처한다.

후보자 본인이 당선될 목적인 경우 또는 후보자 아닌 자가 특정한 후보자를 ‘당선되게 할 목적’인 경우에는 「공직선거법」 제250조 제1항의 적용을 받게 될 것이며, 유튜브는 “연설·방송·신문·통신·잡지·벽보·선전문서” 중 ‘통신’에 속한다고 볼 것이다. 딥페이크 동영상의 내용이 의견이나 주장이 아니라 ‘사실’에 관한 것인 때에는 동 조항이 규정하고 있는 ‘허위의 사실’에 해당될 수 있다. 또한 딥페이크 동영상을 유튜브에 게시하는 행위는 동조가 규정하고 있는 ‘공표’에 해당한다. 한편 딥페이크 동영상이 특정한 후보를 ‘당선되지 못하게 할 목적’으로 공표된 경우에는 「공직선거법」 제250조 제2항이 규정하고 있는 ‘낙선목적 허위사실공표죄’의 적용을 받게 될 것이다. 딥페이크 동영상을 유튜브에 게시하는 행위는 같은 법 제250조 제1항과 마찬가지로 ‘통신’의 방법으로, 허위의 사실을 공표하는 것으로 해석된다.

중앙선거관리위원회는 ‘더불어민주당 법률지원단장 송기현’이 “당선되

지 못하게 할 목적으로 실제 후보자가 하지 않은 말이나 행동을 직접한 것처럼 제작한 동영상, 음성메시지 등의 딥페이크 영상을 제작하여 공표함으로써 후보자를 비방하는 행위가 법 제250조 제2항 위반에 해당 하는지 여부” 등 총 7가지 사안에 관하여 2021년 12월 17일 서면질의를 한 것에 대해, “딥페이크 영상을 이용한 선거운동에 관한 질의회답”이라는 제목 아래 “귀문의 경우 딥페이크 영상은 「공직선거법」(이하 ‘법’이라 함) 상 동영상의 일종으로 선거운동을 할 수 있는 사람이 법상 허용되는 방법으로 딥페이크 영상을 이용하여 선거운동을 할 수 있으나, 딥페이크 영상(AI ○○○)임을 표시하여야 합니다. 딥페이크 영상(AI ○○○)임을 표시하지 않는 경우 선거인이 실제 후보자의 행위로 오인하게 되어 법 제250조(허위사실공표죄)에 위반될 것”이라는 답변내용을 홈페이지에 게시하였다.¹⁰⁷⁾

나. 중앙선거관리위원회의 ‘딥페이크 영상 관련 법규운용 기준’

중앙선거관리위원회는 ‘딥페이크 영상 관련 법규운용 기준’을 통해 “딥페이크 영상을 이용하여 법상 허용되는 방법으로 선거운동을 할 수 있으나, 딥페이크 영상(AI ○○○)임을 표시하지 않는 경우 「공직선거법」(이하 ‘법’이라 함) 제250조(허위사실공표죄)에 위반됨”이라는 기준을 공표하였다.¹⁰⁸⁾ 또한 동 위원회는 딥페이크 동영상을 게시하여 선거운동을 하는 경우 법적 책임의 주체는 ‘영상 게시자’라고 하였다.¹⁰⁹⁾ 그런데 딥페이크 동영상에 ‘딥페이크 영상(AI ○○○)’임을 표시하도록 하는 것은 캘리포니아주의 딥페이크금지법의 내용과 유사한 것으로서 미국에서는

107) 중앙선거관리위원회 (2022, 1, 12). 서면/인터넷 질의보기, 딥페이크 영상을 이용한 선거운동에 관한 질의회답. URL: <https://www.nec.go.kr/site/nec/law/lawView.do>.

108) 중앙선거관리위원회 (2022, 1, 11). 딥페이크 영상 관련 법규운용기준. URL: <http://www.nec.go.kr/site/nec/ex/bbs/View.do?cbIdx=1084&bcIdx=156972>.

109) 중앙선거관리위원회 (2022, 1, 11). 위 기준 참조.

그 실효성에 의문이 제기되고 있다.¹¹⁰⁾ 예컨대, 조작되지 아니한 원본 동영상에 게시하면서 ‘딥페이크 영상’이라고 표시하는 경우에는 유튜브 이용자들에게 혼란을 야기할 수 있다는 것이다.¹¹¹⁾ 더 나아가 유튜브 이용자에게 ‘딥페이크 영상’이라고 표시되지 않은 동영상만 신뢰할 수 있다는 잘못된 믿음을 심어주게 되는 결과가 초래될 수도 있다.

한편 앞서 살펴본 바와 같이 캘리포니아주의 딥페이크금지법은 이미지, 비디오 또는 오디오가 조작된 것이라는 표시를 한 경우에는 동 법을 적용하지 않는 것으로 규정하고 있다.¹¹²⁾ 그런데 이와 달리 우리나라의 「공직선거법」은 “딥페이크 영상(AI ○○○)임을 표시”하는 경우에는 ‘허위사실공표’에 해당하지 않는다는 규정을 두고 있지 않다. 그럼에도 불구하고 중앙선거관리위원회가 “딥페이크 영상(AI ○○○)임을 표시하지 않는 경우”에는 허위사실공표에 해당한다는 ‘딥페이크 영상 관련 법규운용 기준’을 제시한 것은 법률에 명시적인 근거가 없는 유권해석이라고 할 것이다. 비록 딥페이크 동영상임을 표시하지 않은 경우라고 할지라도 경우에 따라서는 허위사실공표에 해당되지 않는다고 보아야 할 경우가 있을 수 있다. 위에서 살펴본 캘리포니아주 딥페이크금지법과 같이 정치적 패러디나 풍자는 언론의 자유에 의해 보호되어야 하기 때문이다. 나아가 중앙선거관리위원회가 ‘딥페이크 영상 관련 법규운용 기준’을 통해 사실상 딥페이크에 “딥페이크 영상(AI ○○○)임을 표시”하도록 강제하는 것은 언론의 자유에 대한 사전적 제약이며 위축효과를 야기할 수 있다.

110) Tashman, Alexandra. (2021). “Malicious Deepfakes”—How California’s A.B. 730 Tries (and Fails) to Address the Internet’s Burgeoning Political Crisis, *Loyola of Los Angeles Law Review* 54, 1420.

111) Tashman, Alexandra. (2021). 위의 논문, 1420.

112) Cal. Elec. Code § 20010(b). “(b)(1) The prohibition in subdivision (a) does not apply if the audio or visual media includes a disclosure stating: ‘This ___ has been manipulated.’ (2) The blank in the disclosure required by paragraph (1) shall be filled with whichever of the following terms most accurately describes the media: (A) Image. (B) Video. (C) Audio.”

다. 정보통신망법상 명예훼손과 정보의 삭제요청 등

「정보통신망 이용촉진 및 정보보호 등에 관한 법률」(이하 ‘정보통신망법’이라고 함) 제70조 제2항은 「형법」상 명예훼손죄에 대한 특별법으로서, “사람을 비방할 목적으로 정보통신망을 통하여 공공연하게 거짓의 사실을 드러내어 다른 사람의 명예를 훼손한 자는 7년 이하의 징역, 10년 이하의 자격정지 또는 5천 만원 이하의 벌금에 처한다”고 규정하고 있다. 「형법」상 명예훼손죄와 달리, 동 조항은 ‘사람을 비방할 목적’이라는 초과주관적 구성요건을 규정하고 있고, ‘정보통신망을 통하여’ 명예를 훼손할 것을 요건으로 하고 있다. 따라서 유튜브에 딥페이크 동영상을 게시하여 공직선거 후보자의 명예를 훼손하는 경우에는 동 조항과 「공직선거법」상 허위사실공표죄와의 관계가 문제될 수 있으나, 「공직선거법」상 허위사실공표죄가 특별법에 해당하는 것으로 해석해야 할 것이다.

한편 ‘정보통신망법’은 “사생활 침해나 명예훼손 등 타인의 권리가 침해된 경우”에는 “정보통신서비스 제공자에게 침해사실을 소명하여 그 정보의 삭제 또는 반박내용의 게재”를 요청할 수 있는 것으로 규정하고 있다.¹¹³⁾ 유튜브에 게시된 딥페이크 동영상이 허위의 사실로서 후보자의 명예를 훼손하는 경우에 후보자는 동 법에 따라 정보의 삭제를 요청할 수 있으며, 요청을 받은 정보통신서비스 제공자는 “지체없이 삭제·임시조치 등의 필요한 조치”를 해야 한다.¹¹⁴⁾ 따라서 동 법에 의하면, 공직선거의 후보자가 당선되거나 되게 할 목적 또는 당선되지 못하게 할 목

113) 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」 제44조의2 제1항 “정보통신망을 통하여 일반에게 공개를 목적으로 제공된 정보로 사생활 침해나 명예훼손 등 타인의 권리가 침해된 경우 그 침해를 받은 자는 해당 정보를 처리한 정보통신서비스 제공자에게 침해사실을 소명하여 그 정보의 삭제 또는 반박내용의 게재(이하 “삭제등”이라 한다)를 요청할 수 있다.”

114) 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」 제44조의2 제2항 “정보통신서비스 제공자는 제1항에 따른 해당 정보의 삭제등을 요청받으면 지체 없이 삭제·임시조치 등의 필요한 조치를 하고 즉시 신청인 및 정보제공자에게 알려야 한다. 이 경우 정보통신서비스 제공자는 필요한 조치를 한 사실을 해당 게시판에 공시하는 등의 방법으로 이용자가 알 수 있도록 하여야 한다.”

적으로 딥페이크를 이용하여 만든 동영상을 유튜브에 게시하는 행위가 사생활 침해나 명예훼손에 해당하는 경우에 딥페이크 동영상에 등장하는 후보자는 유튜브에 침해사실을 소명하여 딥페이크 동영상의 삭제를 요구할 수 있다.

라. 「공직선거법」 개정 법률안

2021년 12월 이해식 의원을 비롯한 28명의 국회의원들은 “최근 원본 이미지나 동영상 위에 다른 영상을 중첩하거나 결합하여 원본과는 다른 가공 콘텐츠를 생성하는 기술인 딥페이크(Deepfake) 기술이 발달하여 이를 선거운동에 이용할 여지가 높아지고 있다”고 주장하면서, “선거운동 시 후보자 또는 예비후보자가 실제로 말한 것이나 행동한 것처럼 보이도록 만든 거짓의 음향·화상 또는 영상 등의 정보를 포함하지 않도록 하고 이를 위반할 경우 3년 이하의 징역 또는 600만원 이하의 벌금”에 처하는 내용의 「공직선거법」 개정안을 국회에 제출하였다.¹¹⁵⁾ 동 법률안이 신설할 것을 제안하고 있는 제82조의8의 내용은 다음과 같다.

제82조의8(인공지능 기술을 이용한 영상 등의 제작·사용 금지) 누구든지 선거운동을 위하여 후보자 또는 예비후보자가 실제로 말한 것이나 행동한 것처럼 보이도록 인공지능 기술을 이용하여 만든 거짓의 음향·화상 또는 영상 등의 정보를 제작하거나 이를 배포·방영 기타의 방법으로 사용하여서는 아니 된다.

동 법률안은 초과주관적 구성요건요소로서 ‘당선되거나 되게 할 목적’ 또는 ‘당선되지 못하게 할 목적’을 요건으로 규정하고 있지 않다. 또한 현행 「공직선거법」 제250조 제1항이 “당선되거나 되게 할 목적으로 ...

115) 공직선거법 일부 개정 법률안 (이해식, 2021, 12, 20), URL: http://likms.assembly.go.kr/bill/billDetail.do?billId=PRC_L2O1S1Y2C1V5B1R1X5Q0V4B1M6O8I1.

후보자, 후보자의 배우자 또는 직계존비속이나 형제자매의 출생지·가족관계·신분·직업·경력등·재산·행위·소속단체, 특정인 또는 특정단체로부터의 지지여부 등에 관하여 허위의 사실”을 공표하는 행위를 규율하고 있는 것과 달리, 동 법률안은 “거짓의 음향·화상 또는 영상 등의 정보”라고만 규정하고 있을 뿐 그 범위를 제한하고 있지 않다. 한편 현행 「공직선거법」 제250조 제2항이 “후보자, 그의 배우자 또는 직계존·비속이나 형제자매에 관하여 허위의 사실을 공표”하는 등의 행위를 처벌하고 있는 것과 비교해 보아도 동 법률안은 그 범위가 매우 넓다. 따라서 우리 헌법 제37조 제2항이 규정하고 있는 과잉금지의 원칙에 비추어 볼 때 동 법률안은 입법목적의 정당성과 수단의 적절성은 인정될 수 있겠으나, 언론의 자유에 대한 침해는 최소한에 머물러야 한다는 점에서 침해의 최소성 원칙을 충족하기는 어려울 것으로 보인다.

V. 맺음말

우리 생활에서 딥페이크 기술은 유용하게 이용될 수도 있다. 영화를 만들 때 딥페이크 기술을 이용하면 배우의 나이를 어리게 보이도록 하거나 나이들어 보이게 할 수도 있으므로, 한 명의 배우가 다양한 연령대의 등장인물을 더 쉽게 연기할 수 있게 된다.¹¹⁶⁾ 또한 딥페이크를 이용한 패러디나 풍자는 현실 정치를 더 실감나게 비판하는데 이용될 수도 있다. 그리고 딥페이크를 이용한 동영상은 익명의 정치적 표현일 수도 있다. 자신이 한 말이나 표현으로 인해 보복을 당하거나 처벌을 받을 수도 있다는 두려움은 표현의 자유를 위축시키기 때문에 익명표현을 보장

116) Miller, Matt. (2020, January 7). Some Deepfakes on YouTube Spent Seven Days Fixing the Shitty De-Aging in the Irishman, *Esquire*, URL: <https://www.esquire.com/entertainment/movies/a30432647/deepfake-youtube-video-fixes-the-irishman-de-aging/>.

하는 것이 중요하다는 점에서 딥페이크에 대한 규제는 신중해야 한다.

이와는 반대로 딥페이크 기법을 이용한 가짜 동영상은 보는 이로 하여금 진짜라고 믿게 할 가능성이 높기 때문에 기존에 언론의 자유에 적용되었던 법리와는 다른 새로운 법리가 적용되어야 한다는 견해도 있다.¹¹⁷⁾ 이에 캘리포니아주는 딥페이크금지법을 한시법으로 제정하여 시행해 보고 있는 것이라고 할 수 있다. 그러나 동 법은 ‘현실적 악의’를 요건으로 규정하고 있어서, 앞서 살펴본 바와 같이 실제의 경우에 있어서는 딥페이크 동영상을 게시한 자에게 법적 책임을 묻는 것이 쉽지 않을 것으로 보인다.

한편 우리나라의 「공직선거법」상 허위사실공표죄는 그 적용범위가 넓고 미국과 달리 ‘현실적 악의’의 법리를 규정하고 있지 아니하므로,¹¹⁸⁾ 선거에서 특정한 후보자를 당선 또는 낙선시킬 목적으로 딥페이크 동영상을 유튜브에 게시하는 행위는 별도의 법률을 제정하지 않고서도 동 법에 의해 충분히 규제될 수 있다고 본다. 따라서 최근에 이해식 의원이 대표 발의한 법률안과 같이 「공직선거법」에 딥페이크 규제에 관한 새로운 처벌규정을 신설할 필요는 없을 것으로 생각된다.

117) Chemerinsky, Erwin. (2018). False Speech and the First Amendment, *Oklahoma Law Review* 71, 2; Sunstein, Cass. (2020). Falsehoods and the First Amendment, *Harvard Journal of Law & Technology* 33, 425-426; Bodi, Matthew. (2021). The First Amendment Implications of Regulating Political Deepfakes, *Rutgers Computer & Technology Law Journal* 47, 169.

118) 우리나라는 미국과 달리 ‘현실적 악의’의 법리를 완전히 도입하고 있지는 않다. 왜냐하면 미국의 ‘현실적 악의’의 법리는 피고의 ‘현실적 악의’를 입증할 책임을 원고에게 부담시키고 있는 반면에, 우리 법원은 그러한 입증책임을 원고에게 부과하고 있지 않기 때문이다.

■ 참고 문헌

- 문재완 (2011). 허위사실의 표현과 표현의 자유-한국과 미국의 판례 비교를 중심으로. <공법연구>, 39집 3호, 113-142.
- 최종선 (2020). 허위조작정보 규제 시도에 대한 비판적 고찰. <홍익법학>, 21권 1호, 363-391.
- 강문규·유오상 (2021, 12, 9). ‘AI 윤석열’ 논란…美·유럽선 “유권자 판단 왜곡” 딥페이크 활용 금지. <헤럴드경제>, URL: <http://news.heraldcorp.com/view.php?ud=20211209000455>.
- 강주리 (2021, 6, 14). 文 사진 중앙에 “대한민국의 위상”…남아공 대통령 잘라냈다. <서울신문>, URL: https://www.seoul.co.kr/news/newsView.php?id=20210614500158&wlog_tag3=naver.
- 금준경 (2022, 1, 20). 극단적 유튜브 채널 방치 이유 있냐 질문에 구글의 답은. <미디어오늘>, URL: <http://www.mediatoday.co.kr/news/articleView.html?idxno=301900>.
- 김명일 (2022, 2, 6). 故노무현이 “정의로운 이재명 지지”… 與 영상에 친문도 발끈했다. <조선일보>, URL: <https://www.chosun.com/politics/election2022/2022/02/06/TOFLQ5BNG5EXTF6HEYCS243TIU/>.
- 김성민 (2020, 10, 17). 페이스북·유튜브 가짜뉴스 놔두면 법적 책임 묻는다. <조선일보>, URL: https://www.chosun.com/economy/tech_it/2020/10/17/FY7BQ4NUSZG67B5UPQ32623NFQ/.
- “딥페이크 범죄 기승, 타깃 안 되려면” 국정원의 경고. (2021, 5, 26). <JTBC 뉴스>, URL: https://news.jtbc.joins.com/article/article.aspx?news_id=NB12007096.
- ‘뉴 노멀’은 ‘새 기준, 새 일상’으로, ‘웨비나’는 ‘화상 토론회’로-국립국어원 새말모임(4.20.~22.) 대체어 마련. (2020, 4, 28). 문화체육관광부.
- 배진영 (2021, 6, 14). 지도자 우상화 위해 사진 조작하는 북한 답아가나?. <월간조선>, URL: http://monthly.chosun.com/client/mdaily/daily_view.asp?idx=12711&Newsnumb=20210612711.
- 양정우 (2020, 10, 26). ‘프롭테크’→‘부동산 정보기술’…쉬운 우리말 대체어 선정. <연합뉴스>, URL: <https://www.yna.co.kr/view/AKR20201026040200005>.

- 엄민재 (2018, 7, 5). 목소리도 행동도 똑같아...진짜 같은 가짜 ‘딤페이크’. <SBS 뉴스>, URL: https://news.sbs.co.kr/news/endPage.do?news_id=N1004834758&plink=SEARCH&cooper=SBSNEWSSEARCH.
- 이상현 (2020, 2, 4). 페이스북 이어 유튜브도 ‘선거관련 조작 동영상’ 삭제 방침. <연합뉴스>, URL: <https://www.yna.co.kr/view/AKR20200204051200009?input=1195m>.
- 이세영 (2020, 12, 26). AI 기술로 만들어낸 춤추는 ‘가짜 영국 여왕’. <조선일보>, URL: https://www.chosun.com/international/2020/12/26/BL2DA2RS5JGMJCTFGAU6R76AVE/?utm_source=naver&utm_medium=referral&utm_campaign=naver-news.
- 이정우 (2021, 12, 20). 이해식 의원, 유권자 기만하는 ‘딤페이크 선거운동 금지법’ 대표발의. <The Public>, URL: <https://www.thepublic.kr/news/newsview.php?ncode=1065576005911254>.
- 정성호 (2019, 5, 29). WP “펠로시 조작 동영상, 새 가짜뉴스 대응할 준비 부족 드러내”. <연합뉴스>, URL: <https://www.yna.co.kr/view/AKR20190529027900091?input=1195m>.
- 정형모 (2020, 6, 6). 비말 대신 칩방울, 슬로푸드 대신 정성 음식 어때요. <중앙일보>, URL: <https://www.joongang.co.kr/article/23795039>.
- 차주경 (2017, 10, 31). 엔비디아, 인공지능으로 ‘가짜 이미지 생성’ 기술 논문 발표. <ITChosun>, URL: http://it.chosun.com/site/data/html_dir/2017/10/31/2017103185046.html.
- 네이버영화, 포레스트 검프, URL: <https://movie.naver.com/movie/bi/mi/basic.naver?code=17159>.
- 네이버TV, 포레스트 검프 CG메이킹 영상, URL: <https://tv.naver.com/v/1076460>.
- 중앙선거관리위원회 (2022, 1, 11). 딤페이크 영상 관련 법규운용기준. URL: <https://www.nec.go.kr/site/nec/ex/bbs/View.do?cbIdx=1084&bcIdx=156972>.
- _____ (2022, 1, 12). 서면/인터넷 질의보기, 딤페이크 영상을 이용한 선거운동에 관한 질의회답. URL: <https://www.nec.go.kr/site/nec/law/lawView.do>.
- YouTube 고객센터, Misinformation policies, 잘못된 선거 정보 관련 정책, URL: <https://support.google.com/youtube/answer/10835034?hl=ko&r>

ef_topic=10833358#zippy=%2C%EC%84%A0%EA%B1%B0-%EA%B3%B5%EC%A0%95%EC%84%B1.

- Bodi, Matthew. (2021). The First Amendment Implications of Regulating Political Deepfakes, *Rutgers Computer & Technology Law Journal* 47, 143-172.
- Brown, Nina I. (2020). Deepfakes and the Weaponization of Disinformation, *Virginia Journal of Law & Technology* 23, 1-59.
- Chemerinsky, Erwin. (2018). False Speech and the First Amendment, *Oklahoma Law Review* 71, 1-15.
- Chesney, Bobby & Citron, Danielle. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, *California Law Review* 107, 1753-1819.
- Delfino, Rebecca A. (2019). Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act, *Fordham Law Review* 88, 887-938.
- Rowbottom, Jacob. (2012). Lies, Manipulation and Elections—Controlling False Campaign Statements, *Oxford Journal of Legal Studies* 32, 507-535.
- Sunstein, Cass. (2020). Falsehoods and the First Amendment, *Harvard Journal of Law & Technology* 33, 387-426.
- Tashman, Alexandra. (2021). “Malicious Deepfakes”—How California’s A.B. 730 Tries (and Fails) to Address the Internet’s Burgeoning Political Crisis, *Loyola of Los Angeles Law Review* 54, 1391-1422.
- Ajder, Henry et al. (2019, September). The State of Deepfakes: Landscape, Threats, and Impact, *Deeptrace*, URL: https://regmedia.co.uk/2019/10/08/deepfake_report.pdf.
- Binder, Matt. (2020, January 20). Deepfakes Are Getting Easier to Make and the Internet’s Just Not Ready. *Mashable*, URL: <https://in.mashable.com/tech/10633/deepfakes-are-getting-easier-to-make-and-the-internets-just-not-ready>.
- Dixon Jr., Herbert B. (2019, August 12). Deepfakes: More Frightening Than Photoshop on Steroids, *American Bar Association*, URL: <https://www>

- w.americanbar.org/groups/judicial/publications/judges_journal/2019/summer/deepfakes-more-frightening-photoshop-steroids/.
- Fink, David E. & Diamond, Sarah E. (2020, September 3). Deepfakes: 2020 and Beyond, *The Recorder*, URL: <https://www.law.com/therecorder/2020/09/03/deepfakes-2020-and-beyond/?slreturn=20220211072801>.
- Gessen, Masha. (2018, July 15). The Photo Book that Captured How the Soviet Regime Made the Truth Disappear, *The New Yorker*, URL: <https://www.newyorker.com/culture/photo-booth/the-photo-book-that-captured-how-the-soviet-regime-made-the-truth-disappear>.
- Hardesty, Larry. (2017, April 14). Explained: Neural networks, *MIT News*, URL: <https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>.
- Knight, Will. (2019, June 21). A new deepfake detection tool should keep world leaders safe-for now, *MIT Technology Review*, URL: <https://www.technologyreview.com/2019/06/21/134815/a-new-deepfake-detection-tool-should-keep-world-leaders-safe-for-now/>.
- _____. (2018, May 23). The US Military is Funding an Effort to Catch Deepfakes and Other AI Trickery, *MIT Technology Review*, URL: <https://www.technologyreview.com/2018/05/23/142770/the-us-military-is-funding-an-effort-to-catch-deepfakes-and-other-ai-trickery/>.
- Lecher, Colin. (2019, October 7). California has banned political deepfakes during election season, *The Verge*, URL: <https://www.theverge.com/2019/10/7/20902884/california-deepfake-political-ban-election-2020>.
- Lomas, Natasha. (2020, August 18). Deepfake video app Reface is just getting started on shapeshifting selfie culture, *Techcrunch* URL: <https://techcrunch.com/2020/08/17/deepfake-video-app-reface-is-just-getting-started-on-shapeshifting-selfie-culture/>.
- Marr, Bernard. (2018, October 1). What is deep Learning AI?, A Simple Guide with 8 Practical Examples, *Forbes*, URL: <https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/?sh=14e204a58d4b>.
- Metz, Cade. (2019, December 24). Internet Companies Prepare to Fight the

- 'Deepfake' Future, *The New York Times*, URL: <https://www.nytimes.com/2019/11/24/technology/tech-companies-deepfakes.html>.
- Miller, Matt. (2020, January 7). Some Deepfaker on YouTube Spent Seven Days Fixing the Shitty De-Aging in the Irishman, *Esquire*, URL: <https://www.esquire.com/entertainment/movies/a30432647/deepfake-youtube-video-fixes-the-irishman-de-aging/>.
- Morris, Allie. (2019, October 9). Texas is first state to ban political 'deepfake' videos, *San Antonio Express-News*, URL: <https://www.expressnews.com/news/local/politics/article/Texas-is-first-state-to-ban-political-14504294.php>.
- Nicholson, Chris. A Beginner's Guide to Generative Adversarial Networks (GANs), URL: <https://wiki.pathmind.com/generative-adversarial-network-gan>.
- Peng, Tony. (2019, April 4). Father of GANs Ian Goodfellow Splits Google For Apple, *Synce4*, URL: <https://medium.com/syncedreview/father-of-gans-ian-goodfellow-splits-google-for-apple-279fcc54b328>.
- Porup, J. M. (2021, March 18). How and why deepfake videos work—and what is at risk, *CSO*, URL: <https://www.csoonline.com/article/3293002/deepfake-videos-how-and-why-they-work.html>.
- Reynolds, Matt. (2020, June 9). Courts and lawyers struggle with growing prevalence of deepfakes, *ABA Journal*, URL: <https://www.abajournal.com/web/article/courts-and-lawyers-struggle-with-growing-prevalence-of-deepfakes>.
- Thalen, Mikael. (2020, September 2). Deepfake app takes Trump videos of fine until after the election, *Daily Dot*, URL: <https://www.dailydot.com/debug/deepfake-app-trump-2020-election/>.
- Vincent, James. (2018, May 22). Why we need a better definition of 'deep fake', *The Verge*, URL: <https://www.theverge.com/2018/5/22/17380306/deepfake-definition-ai-manipulation-fake-news>.
- Adobe, Photoshop, Britannica, URL: <https://www.britannica.com/technology/Adobe-Photoshop>.
- Filehorse, FakeApp. (2021, December 3). URL: <https://www.filehorse.com/download-fakeapp/>.

U.S. Copyright Office, Section 512 of Title 17: Resources on Online Service Provider Safe Harbors and Notice-and-Takedown System, URL: <https://www.copyright.gov/512/>.

YouTube, tech4tress, Deepfakes Guide: FakeApp 2.2 Tutorial Installation (totally simplified, model folder included), URL: <https://www.youtube.com/watch?v=Lsv38PkLsGU>.

Webwise.ie, Explained: What is YouTube?, URL: <https://www.webwise.ie/parents/what-is-youtube/>.

ABSTRACT

Deepfake Videos on YouTube and Publication of False Statement of Facts

Huh, Soon Chul
Professor of Law at Kyungnam University, Ph.D., S.J.D.

With the advent of Artificial Intelligence technology, fake videos that look completely real can be created, making new legal problems that have not existed before. At the end of 2020, a fake video of British Queen Elizabeth II greeting and dancing for Christmas was released. Also, a deepfake video that former U.S. President Obama called President Trump “a total and complete dipshit” was disclosed to alarm the danger of deepfakes in 2018. Recently, even in South Korea, with the presidential election just around the corner, the ruling party posted on its official YouTube channel a deepfake video to the effect that a former Korean President declared support for a presidential candidate. The controversy over the video, however, made the party take down it from the channel.

If a deepfake video about a candidate is posted on YouTube ahead of the election, which negatively or positively affects voters’ choice of candidates, it may pose a serious threat to democracy that is based on fair elections. Therefore, California and Texas in the United States have passed election laws regulating an act of distributing a deepfake video which may injure a candidate or influence the result of an election.

The term “deepfake”, which refers to a fake video, is the compound

word of “deep learning” and “fake,” and can be said to be “manipulated images, audio or video, designed to make the altered media seem authentic.” However, the California law regulating deepfakes immediately before an election is criticized on the ground that it is not a narrowly tailored restriction on freedom of speech to achieve the compelling government interest such as fair elections. It has also been criticized because it can be rarely applied to a case due to the actual malice rule of the U.S. Supreme Court. According to this criticism of the law, it is more desirable to regulate deepfakes with the defamation or copyright laws. Meanwhile, some critics argue that the Texas anti-deepfake law is unconstitutional not only because it does not prescribe for the actual malice rule, but also because it does not exempt its application to parody or satire.

Recently, the National Election Commission in South Korea promulgated the “Standard for the Application of the Law on Deepfake Videos” and issued an authoritative interpretation that an act of posting a deepfake video on the Internet without indicating that it is “a deepfake video (AI OOO)” violates Article 250 of the Public Official Election Act. However, there is no expressive legal ground under the law to display that it is a deepfake video. In addition, this requirement may cause a chilling effect on freedom of speech because it is a prior restraint. In the meantime, it seems reasonable to conclude that no additional criminal punishment law needs to be passed because the current law punishing an act of publishing false statement of facts during an election can adequately regulate deepfake videos although a bill to criminalize deepfakes was introduced in the National Assembly.

Keywords: deepfake, YouTube, publication of false statement of facts.
election, Public Official Election Act

[논문투고일 2022. 3. 13. 논문수정일 2022. 4. 6. 게재확정일 2022. 4. 7.]