딥페이크 벗겨내기

어쩌면 여러분은 "스티브 부세미가 민소매 옷을 입을 때는 대체로 빨간 옷은 안 입을텐데?"라고 생각하셨을지 모르겠습니다. 사실 이 사진은 진짜가 아닙니다. 이 사진은 딥페이크(deepfake)라고 알려진 컴퓨터로 만든 영상에서 나온 것입니다. 기계학습이 발달하고 계산 능력이 증가한 덕분에, 불행히도 요즈음 딥페이크 영상은 만들기는 쉽지만 가려내기는 어려워졌습니다. 하지만 모든 게 사라지지는 않습니다. 컴퓨터가 인간의 안내를 받아 딥페이크 영상을 만들어내듯, 반대로 찾아낼 수도 있습니다. (머리와 입술의 움직임 에 대한) 기하학을 포함하여, (한 얼굴을 다른 얼굴로 바꿀 때 생기는 불일치를 찾기 위해) 선형대수학과, (영상이 진짜가 아닐 확률을 재기 위해) 확률론을 이용하는 것이 현재의 접근법입니다. 모든 것을 액면가대로  받아들이지 않는 것이 어쩌면 위조와의 싸움에서 가장 중요한 무기일지 모릅니다.

현재 연구자들은 위조꾼들을 좌절시키기 위해 더 탄탄한 방법을 연구 중입니다. 영상 파일의 비트들을 이용하여 수학적으로 암호화한 숫자를 할당하여 디지털 서명으로 기능하 게끔 하는 방법입니다. 이러한 서명은 디지털 화폐에서 거래를 인증하고 조작을 가려내 기 위해 사용하는 것과 유사한 방식으로 블록체인의 일부가 됩니다. 비디오를 조작하면 원본 파일의 비트가 변하는데 원래의 서명은 변하지 않으므로, 새 파일과 원본의 서명이 일치하지 않게 됩니다. 이러한 인증 수단을 갖추면, 안전하지 않은 웹사이트에 접근하려할 때 팝업이 뜨듯 딥페이크 영상을 접할 때마다 경고를 내보내어, 여러분이 보는 것이 보기와는 다른 것일 수도 있음을 알려줍니다.

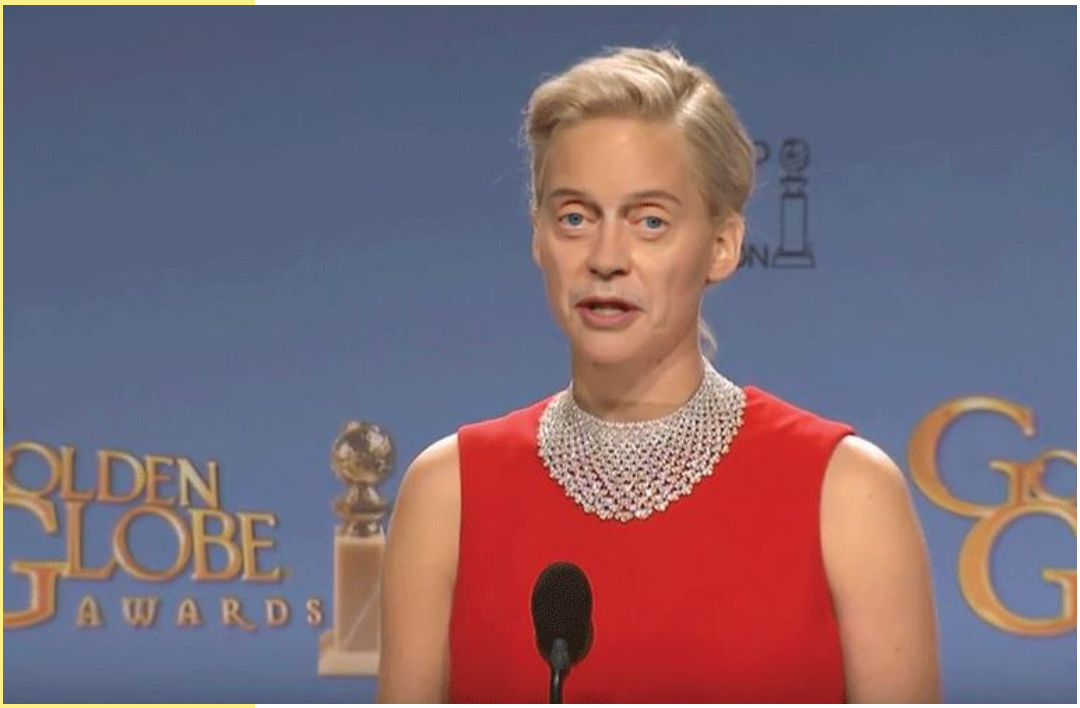

Image: Screen grab from video by VillainGuy

더 알아보기:
"Protecting World Leaders Against Deep Fakes," by Agarwal, Farid, Gu, He, Nagano, and Li, 2019.

Translation courtesy of the Korean Mathematical Society

Listen Up!

MM/146/KR

AMS
AMERICAN MATHEMATICAL SOCIETY

Mathematical Moments 프로그램은 과학, 자연, 기술, 그리고 인간의 문화에서 수학이 하는 역할에 대한 올바른 평가와 이해를 촉진합니다.

www.ams.org/mathmoments