

Deepfake detected: Microsoft hakkab tehisintellekti abiga võitlema võltsvideote ja võltspiltide vastu

5 aastat tagasi - 09.09.2020 Autor: [AM](#)

Seoses viirusepuhangute, terroriaktide, valimiste ja alternatiivsete ravivõimalustega on Internetis igaühel võimalik otsa sattuda võltsvideotele ja -meediale, mida on manipuleeritud nii, et need näivad justkui ehtsad. Microsoft töötas välja selle avastamise vahendi, mis tulevikus võib meie arvutites aidata võltsitud sisu kindlaks teha samamoodi, nagu praegu turvatarkvara leiab õngitsus- ja petukirju.

Võltsmeediast oleme me [varem kirjutanud](#), sellest rääkis põhjalikumalt Helnese lahenduste arhitekt Markus Karileet. Nüüd on see probleem veel suuremaks paisunud. Microsofti asepresident Tom Burt [kirjutab](#), et Microsoft sai 1. septembriks valmis lahenduse, mis suudab eristada võltsitud videosisu päris sisust masinõppe ja tehisintellekti abiga. Microsoft Video Authenticator on tarkvara, mis analüüsib pilte või videot, andes sellele usaldatavuse protsendi. Video puhul jookseb usaldatavuse protsent video mängimise ajal jooksvalt iga kaadri kohta, mida vaadatakse. Lahendus on loodud avaliku andmekogu [Face Forensic++](#) põhjal ning testitud [DeepFake Detection Challenge](#)´i andmekogu peal.

Lisaks pakub Microsoft võimalust oma originaalsisu sertifikaadiga kaitsta, mida kontrollides saab kasutaja kindel olla, et meedia on pärit otseallikast.

Video Authenticatorit pakutakse organisatsioonidele AI Foundation´i vahendusel [Reality Defenderi](#) projekti kaudu. Üks osa sellest projektist on ka valimisvõitluses võltsingute varajane paljastamine, et need ei mõjutaks valijaid, poliitikuid ja meediat. Lisaks Microsoftile on ühinenud projektiga teiste seas ka näiteks Google ja Twitter.

Vaata siit üht MIT-is loodud näidet tehisintellektiga loodud Deepfake videost, kus USA president annab teada ebaõnnestunud kuundumisest:

Kui tahad end proovile panna, kas suudad ise eristada võltsingut originaalist, siis seda [saab testida siin](#).

- [Uudised](#)
- [Lahendused](#)
- [Turvalisus](#)

Pilt

Stanford University

Results

Expression-only Reenactment



Source Sequence



Unmodified
Target Sequence



Result