



«دیپ فیک»

تیغی در کف زنگیان مست

سیدغلامرضا فلسفی
ارتباط پژوه

DEEP FAKE

آن است که تصور می‌شود. لذا بر همین اساس نسبت به آغاز موج جدید پویش‌های اطلاعات نادرست^۲ در انتخابات ۲۰۲۰ زنه‌ار داده‌اند؛ روندی نامیمون که تأثیر «اخبار جعلی»^۴ را بر افکار عمومی و نقش گروه‌های فریبکار فعال در «فیس‌بوک» را در تغییر نتیجه انتخابات ریاست جمهوری ۲۰۱۶ یادآور می‌شود.

کلیدواژه‌ها: دیپ فیک، هوش مصنوعی، اخبار جعلی، اطلاعات نادرست

الیس ساموئلز^۵، روزنامه‌نگار و سردبیر بخش راستی‌آزمایی فیلم‌های آنلاین در روزنامه واشنگتن پست، در این باره می‌گوید: دیپ فیک این ایده را که باورپذیری منوط به دیدن است، نقض می‌کند و این مهم در آینده مبارزات سیاسی و نتایج حاصل از آن بسیار تعیین‌کننده است. هانی فرید^۶، کارشناس حوزه قوانین دیجیتال در «دانشگاه

اندیشمندان حوزه «هوش مصنوعی»^۱ در عین حال که برای نابودی یک سلاح بسیار خطرناک سیاسی، یعنی فیلم‌های جعلی تولید شده توسط رایانه که بیشترین تأثیر را در تضعیف جایگاه نامزدهای انتخاباتی و گمراه‌سازی رأی‌دهندگان در انتخابات ریاست جمهوری ۲۰۲۰ ایالات متحده دارد، با یکدیگر در رقابت‌اند، هم‌صدا با هم پیامی مشترک را صلا می‌دهند: «جهان برای مواجهه با دیپ فیک^۲ آماده نیست.»

این پژوهشگران با مدنظر قرار دادن ملاک‌های کشف فیلم‌های جعلی، سامانه‌هایی خودکار طراحی کرده‌اند تا فیلم‌ها را تجزیه و تحلیل کنند. برخی از این ملاک‌ها عبارت‌اند از: ارزیابی نور، سایه‌ها، الگوهای پلک زدن، چگونگی حرکات سر، ابعاد واقعی چهره، مانند نسبت زاویه‌های سر و اجزای صورت نسبت به یکدیگر، یا نحوه تکان دادن سر و لبخند زدن. به‌رغم تمام این پیشرفت‌ها، محققان بر این باورند که گستردگی و تأثیرات عمیق این روش جاعلان بر افکار عمومی، بسی‌شدید و ماندگارتر از

مشخص شد این ویدیو با سرعتی ۲۵ درصد کمتر از حالت عادی بازتدوین و در نتیجه دست‌کاری شده بود.^۹

دیپ فیک که نخستین بار طی سال ۲۰۱۷ در یک پست «تارنمای ردیت»^{۱۰} از آن نام برده شد و در تولیدش از فن هوش مصنوعی «دیپ لرنینگ»^{۱۱} استفاده می‌شود، در ایالات متحده به یک اهرم فشار علیه سیاست‌یون، از جمله ساکنان منطقه «کپیتال هیل»^{۱۲} بدل شده است، تا جایی که قانون‌گذاران رجای واثق یافتند که دیپ فیک ضمن زیر سؤال بردن اعتبارشان، محل سلامت انتخابات و در ابعادی وسیع‌تر چالشی جدی برای امنیت ملی و حاکمیت مردم‌سالارانه محسوب می‌شود. با عنایت به این مهم، «کمیته اطلاعات مجلس نمایندگان» طی ژوئن ۲۰۱۹ برای استماع نظر متخصصان هوش مصنوعی جلسه‌ای تشکیل دادند و پیرامون اینکه دیپ فیک با سلب مشروعیت از رسانه، چگونه موجب انکار رسانه از سوی مخاطب می‌شود و امکان تشخیص صحیح را از وی می‌گیرد، و با این شرایط، مخاطب چگونه می‌تواند از تأثیر روان‌شناختی ماندگار دیپ فیک و هوش مصنوعی بر هوش و شخصیت خود رهایی یابد، اظهاراتی مطرح شد.

آدام شیف^{۱۳}، نماینده دموکرات حوزه انتخابیه کالیفرنیا و رئیس کمیته اطلاعات مجلس نمایندگان، پیرامون میزان آمادگی نظام سیاسی ایالات‌متحده در برابر دیپ فیک اظهار تردید کرده و افزوده است که گمان نمی‌کند مردم از آنچه انتظارشان را می‌کشد آگاه باشند.

راشل توماس^{۱۴}، بنیان‌گذار یک آزمایشگاه یادگیری ماشین^{۱۵} در سانفرانسیسکو، می‌گوید: یک پوشش اطلاعات نادرست با استفاده از فیلم‌های دیپ فیک بر بستر پلتفرم وب مدرن، با توجه به ساختار این پلتفرم، انبوهی از مخاطبان را جذب می‌کند و بیشتر و سریع‌تر از حقیقت اشاعه می‌یابد. توماس معتقد است محتوای جعلی به‌طور تضمین‌شده بر افکار عمومی مؤثر نیست، چراکه انسان‌ها در مقام موجوداتی اجتماعی اموری را باور می‌کنند که کثیری از هم‌نوعانشان به آن مؤمن باشند. لذا آنانی که در پی تأثیرگذاری بر افکار عمومی از طریق تولید و نشر مجعولات هستند، کاری ساده در پیش ندارند. در این بین، نبود قوانین مصرح ناظر بر دیپ فیک موضوعی قابل تأمل است. البته برخی کارشناسان حقوقی و فنی معتقدند قوانین مصوب که بر اعمالی مجرمانه چون افترا، سرقت هویت یا جعل عنوان‌های دولتی جاری‌اند، بر این مهم نیز دلالت می‌کنند. با این حال به نظر می‌رسد گسترده شدن وضع قوانین نیز نگرانی‌هایی در پی دارد، چرا که دیپ فیک بر مرز باریک میان محتوای طنزی که متمم اول قانون اساسی ایالات متحده از آن حمایت می‌کند و تبلیغات سیاسی مخرب در حرکت است؛ مرزی که ممکن است وجوه تمایز آن همواره واضح نباشد.

برخی نگران‌اند امکان تخدیر یا هیجانان ناشی از تماشای فیلم‌های جعلی، موجب شود مردم به استنادات و شواهد جعلی تبهکاران

برکلی» کالیفرنیا، با توجه به آمارهایی چون نسبت کپی تولیدکنندگان به تحلیل‌کنندگان دیپ فیک که به تعبیر وی صد به یک است، پیرامون امکان غلبه بر این فناوری اظهار تردید می‌کند. این اسناد علوم رایانه در دانشکده «دارتموث» که قریب به دو دهه در حوزه تشخیص فیلم‌های دستکاری شده فعالیت دارد، می‌گوید که با ریزینی در مواردی چون پیکسل‌های تصویر و سعی در یافتن ایرادهای جزئی آن، یا موشکافی در روند گردش خون در صورت شخصیت‌ها، گاه می‌توان به واقعی یا جعلی بودن فیلم واقف شد. به‌رغم این موارد فرید معتقد است هوش مصنوعی در کمترین زمان متصور، مرز این علائم و نشانه‌ها را نیز درمی‌نورد و روند شناسایی دیپ فیک‌ها، همچون رابطه موش و گربه در کارتون «تام و جری»، بین الگوریتم‌های هوش مصنوعی و خبرگان تشخیص فیلم‌های جعلی، به تسلسلی نافرجام دچار می‌شود.^۶

گرچه در ایالات‌متحده، دیپ فیک‌ها تبعات گسترده سیاسی در پی داشته‌اند، اما رسوایی‌های سیاسی فیلم محور لزوماً به دیپ فیک محدود نیست، بلکه گاه دست‌کاری‌های ساده در فیلم نیز در سپهر سیاسی غائله برپا می‌کند که شاهد مثال آن، ویدیوی پربیننده **نانسی پلوسی**^۸، رئیس مجلس نمایندگان آمریکا و رهبر حزب دموکرات این مجلس در می ۲۰۱۹ است. بیش از سه میلیون بازدیدکننده در این فیلم پلوسی را در حالی دیدند که دچار مشکل تکلم و عدم تعادل است. فیلم مذکور این گمان را در بینندگان ایجاد کرد که این عضو ارشد حزب دموکرات دارای مشکلات ذهنی است و از سلامت جسمی برخوردار نیست. علاوه بر این، برخی مخالفان وی، به‌ویژه جمهوری‌خواهان، در آن هنگام مدعی شدند، مشکل تکلم پلوسی ناشی از اعتیاد وی به مسکرات است که پس از چندی





ویدیوهای دیپ فیک نمونه‌ای از انقلابی است که هوش مصنوعی در عرصه اطلاعات نادرست ایجاد کرده؛ کما اینکه سامانه‌های جدید هوش مصنوعی زبانی، مانند «نسخه دوم مبدل پیش آموزش داده‌شده مولد»^{۲۳} که توسط آزمایشگاه تحقیقاتی اپن‌آی‌آی^{۲۴} طراحی شده، متون نوشتاری تولید می‌کنند. این الگوریتم مدل‌سازی زبان از ویژگی مولد بودن برخوردار است؛ در حالی که این ویژگی منحصر به زبان انسان است. مولد بودن زبان به معنای توانایی تولید صورت‌های جدید زبانی براساس قواعد موجود در زبان است که یکی از ویژگی‌های مهم آن به شمار می‌رود.^{۲۵}

در تولید ویدیوهای دیپ فیک به یاد داریم که تصویرهای هدف به سامانه معرفی می‌شدند و سامانه براساس بانک داده‌ای از تصویرها و صداهای موجود پیشین و کمک روش‌هایی نظیر «شبکه مولد متخاصم»^{۲۶} که شاخه‌ای مهم از یادگیری ژرف است، تصویرهای هدف را بر مبنای تصاویر موجود می‌ساخت و در نهایت شخصیتهای دیده می‌شد که در حال انجام کار یا ایراد سخن است؛ در حالی که روح شخص واقعی از آن بی‌خبر است. (فلسفی، ۱۳۹۸: ۸۲ و ۸۳). در تولید نوشتار نیز، با استفاده از آن دسته متون نوشتاری که در قالب داده‌های خام اولیه به سامانه وارد می‌شود، بندهایی در ادامه متن ورودی با همان لحن، مضمون و سبک نگارش تولید می‌شود. اگر به سامانه تیتیر جعلی بدهید، بقیه مطلب را با نقل‌قول‌ها و نمودارهای جعلی تولید می‌کند یا می‌تواند با در اختیار داشتن خط اول یک داستان کوتاه، شخصیت‌پردازی کند و به سرعت داستان کوتاه پرطرفداری بنویسد.^{۲۷} به‌رغم اینکه ممکن است متن به سمت مغلق‌گویی گرایش یابد، لیک این امکان مستسکی ارزشمند برای ربایتهای هرزه نگار و سازندگان اخبار جعلی محسوب می‌شود.

نسخه دوم مبدل پیش آموزش داده‌شده مولد، یک مدل‌ساز زبان طبیعی با ۱/۵ میلیارد مؤلفه در مجموعه‌ای مرکب از هشت میلیون صفحه وب است و هدف از طراحی آن، پیش‌بینی درون‌متنی واژه و پسین براساس واژگان پیشین است. سامانه مذکور نسخه‌ای ارتقا

ایمان بیاورند. محققان این مشکل را بی‌تفاوتی نسبت به امر واقع نامیده‌اند؛^{۱۶} بدین معنی که فرد برای فهم آنچه واقعیت دارد و آنچه واقعیت ندارد، بسیار تلاش می‌کند، لیکن در نهایت آنچه را که با وابستگی‌های پیشینش پیوند دارد، می‌پذیرد. نتایج «مرکز تحقیقات پیو»^{۱۷} که در ژوئن ۲۰۱۹ منتشر شد، مؤید این امر است، چرا که در آن تحقیق حدود دوسوم از پرسش‌شوندگان آمریکایی اظهار کردند، فیلم‌ها و تصویرهای جرح و تعدیل شده به مشکلی اساسی در تشخیص امر واقع در زندگی روزمره‌شان تبدیل شده است. بیش از یک‌سوم از پرسش‌شوندگان نیز معتقد بودند، اخبار ساختگی موجبات بی‌اعتمادی و کاهش کلی میزان دریافت اخبار توسط آنان را رقم زده است.

این نگرانی که دیپ فیک‌ها مردم را در پذیرش فیلم‌های واقعی دچار تردید و انکار کند، پدیده‌ای است که **رابرت چسینی**^{۱۸} و **دنیل سترون**^{۱۹} استادان حقوق، آن را سود حاصل از دروغ نام نهاده‌اند؛ به این معنی که قیج‌زدایی از جعلیات و دست‌کاری محتوا سبب شود، حتی پس از افشای دروغ، اعتماد به هرگونه اطلاعات پیرامون آن موضوع خاص برای افکار عمومی دشوار شود.^{۲۰}

فناوری دیپ فیک به سرعت در حال پیشرفت است، کما اینکه طی می ۲۰۱۹ پژوهشگران هوش مصنوعی در مؤسسه علم و فناوری «اسکولکوو»^{۲۱} در مسکو از سامانه «عکس معدود»^{۲۲} رونمایی کردند که می‌تواند با تعدادی محدود تصویر چهره، تصویرهای متعدد متقاعدکننده جعلی از آن چهره تولید و سپس آن تصویرها را به فیلم تبدیل کند. **ایگور زاخاروف**، پژوهشگر ارشد این گروه تحقیقاتی، با عنایت به بازنگری مداوم همکارانش در فناوری مورد بحث، از هرگونه اظهارنظر شخصی پیرامون این موضوع استنکاف می‌ورزد، اما در فرازی از مقاله منتشره از سوی گروه تحقیقاتی مذکور آمده که دستاورد حاصل از اثر شبکه‌ای تولید فناوری‌های جلوه‌های ویژه ویدیویی، فراتر از حد انتظار بوده و همچنین تأکید شده که این فناوری نوین با فناوری «آواتار عصبی» هم‌خوان است.



عبارت «گرفتگی عضلات هنگام شنا خطری بزرگ است» را با صدای **بیل گیتس، جین گودال**^{۳۵} و بسیاری دیگر بسازد. مدل طراحی شده، امکان تولید صداهای جعلی رایانه مینا را در قالب سخنرانی، موسیقی و سنتز متن به گفتار داراست که فراتر از تصورها، برآوردها و قضاوت‌های انجام گرفته در این زمینه است.^{۳۶} از مصداق‌های دیپ فیک‌های صوتی می‌توان به حادثه‌ای در اوت ۲۰۱۹ اشاره کرد که طی آن تبهکاران دیجیتالی به کمک دیپ فیک و جعل صدای مدیرعامل یک شرکت بریتانیایی، موفق شدند مبلغ ۲۴۳ هزار دلار از شرکت مذکور سرفت کنند!^{۳۷}

بخش عمده سرمایه‌گذاری به منظور پژوهش پیرامون یافتن راه‌های شناسایی دیپ فیک‌ها، توسط «دفتر نمایندگی پروژه‌های پژوهشی پیشرفته دفاعی»، موسوم به «دارپا»^{۳۸} که وابسته به وزارت دفاع آمریکا است، تأمین می‌شود. دارپا از سال ۲۰۱۶ حامی مالی برنامه طراحی زیرساخت تشخیص، تجزیه و تحلیل، تفسیر، ارائه شواهد و بازیابی خودکار فیلم‌ها و تصویرهای دیجیتال «دست‌کاری شده»^{۳۹} است که توسط جمعی از پژوهشگران طراز اول و گروه‌های تحقیقاتی بسیار معتبر این حوزه در حال انجام است. **مت توریگ**^{۴۰}، متخصص «بینایی ماشین»^{۴۱} که هدایت این برنامه را بر عهده دارد، می‌گوید: امکانی که برای شناسایی رسانه‌های جعلی و اخبار تولیدی آن‌ها با نام «رسانه‌های ترکیبی»^{۴۲} طراحی شده، یک فناوری تدافعی است که نه تنها علیه دشمنان خارجی، بلکه دشمنان سیاسی داخلی و وباش اینترنتی قابل به‌کارگیری است.

توریک اظهار می‌دارد: از زمان اولین رسانه‌ها تاکنون، حکومت‌ها و دولت‌ها امکان اعمال نفوذ و دست‌کاری در محتوای رسانه را داشته‌اند، اما طراحی سامانه‌های قوی کشف جعلیات، مانند رسانه‌های ترکیبی، سبب می‌شود گروه‌های تبهکار دارای منابع و امکانات محدود که در حوزه تولید و نشر اخبار جعلی فعال‌اند، با انجام محاسبه هزینه - فایده، از ارتکاب این عمل مجرمانه خودداری کنند. پژوهشگران بر این اعتقادند که فیلم‌های جعلی دارای کیفیت تصویری بالا، با سهولتی بیشتر قابل شناسایی‌اند؛ چراکه امکان مشاهده هرچه بیشتر جزئیات، ضریب تفکیک جعلی از واقعی بودن را به میزانی فراوان افزایش می‌دهد. اما قرار گرفتن چنین فیلم‌هایی بر «پلتفرم وب مدرن» این امکان تشخیص را سلب می‌کند. زیرا بیشتر قریب به اتفاق رسانه‌های اجتماعی و تارنماهای پیام‌رسان، به منظور سهولت و سرعت هرچه بیشتر به اشتراک گذاشته شدن فیلم‌های خود، آن‌ها را در قالب‌هایی فشرده بر این بستر بارگذاری می‌کنند که این امر موجبات حذف جزئیات فیلم‌ها را که سرخ‌هایی مهم برای کشف دیپ فیک محسوب می‌شوند، فراهم می‌کند.

موارد مذکور سبب شده‌اند، برخی افراد تشخیص فیلم‌های جعلی را ناممکن ببینند. لذا محققان بر آن شدند تا با تمرکز بر سامانه اصلت سنجی، بکوشند صحنه‌ها درست همان‌طور که فیلم‌برداری

یافته است که نسبت به نمونه پیشین^{۳۸} در حوزه مؤلفه و مقدار داده، ۱۰ برابر ارتقا یافته و به شکلی طراحی شده که به سهولت می‌تواند واژه پسین درون‌متنی را در چارچوب بانک داده متنی اینترنتی که معادل ۴۰ گیگابایت برآورد می‌شود، گمانه‌زنی و تولید کند.^{۳۹}

لازم به ذکر است که این روش روی متن سخنرانی رهبران سیاسی جهان در مجمع عمومی سازمان ملل متحد بین سال‌های ۱۹۷۰ تا ۲۰۱۵ آزمایش شده است.^{۴۰} در مقابل و برای مواجهه با این فناوری، پژوهشگران دانشگاه واشنگتن و مؤسسه هوش مصنوعی آلن یک سامانه کاشف متن جعلی به نام «گورو»^{۴۱} را در می ۲۰۱۹ معرفی کردند که ظرفیت آشکارسازی اخبار جعلی عصبی تولیدشده به وسیله ماشین را دارد.

مضاف بر پژوهش در حوزه دیپ فیک‌های ویدیویی و نوشتاری جعلی، موضوع تشخیص صداهای جعلی متقاعدکننده نیز در افق فکری کارشناسان این حوزه قرار دارد. پژوهشگران دانشگاه آلاباما در بیرمنگام که در حوزه صدا پژوهش می‌کنند، با استفاده از سه تا پنج دقیقه صدای فرد هدف که زنده ضبط شده یا از یوتیوب یا رادیو به دست آمده است، توانسته‌اند صدایی جعلی تولید کنند که انسان و سامانه‌های ایمنی بیومتریک در تشخیص جعلی بودن آن ناتوان‌اند. **نیتش ساکسنا**^{۴۲}، استاد و پژوهشگر این دانشگاه، پیرامون میزان تأثیر اقناع‌گرایانه این صداهای جعلی اظهار می‌دارد که با این فناوری می‌توان با صدای مادر برای فرزندش با کیفیتی پیام گذاشت که فرزند نتواند جعلی بودن صدا را تشخیص دهد یا شخصی را با انتشار یک فایل صوتی جعل شده از صدای خودش بدنام کرد [سلطانی‌فر، سلیمی و فلسفی، ۱۳۹۶: ۵۹].

محققان هوش مصنوعی شرکت رسانه‌ای «فیس‌بوک»، در ادامه مساعی پژوهشگران دانشگاه آلاباما، توانسته‌اند با شبیه‌سازی صدای افراد، گفتاری ماشینی تولید کنند که به شدت فریبنده است. برای انجام این مهم، سامانه «مل نت»^{۴۳} براساس داده‌های ورودی شامل صدها ساعت سخنرانی در همایش «تد» و کتاب‌های صوتی، در چارچوب مدل‌های صوتی طراحی شده به جعل هویت افراد اقدام می‌کند. برای مثال، این سامانه می‌تواند

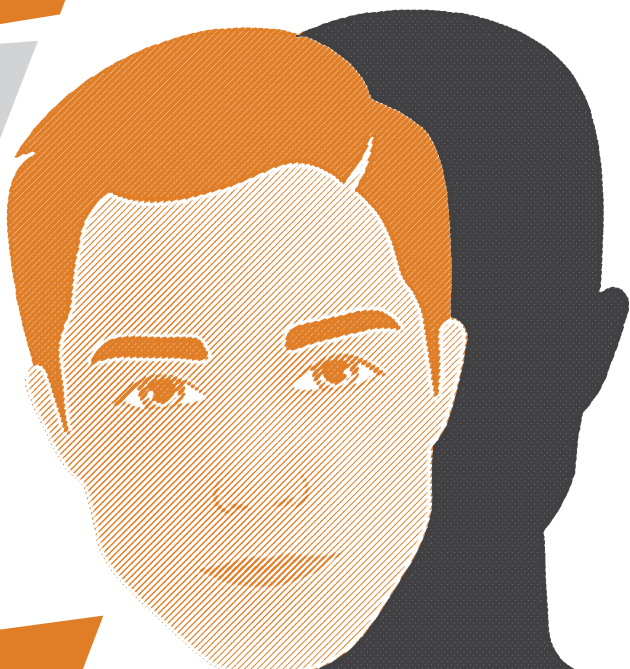
می‌شوند، به نمایش درآیند که از این روند در اصطلاح با عنوان «اثر انگشت» یاد می‌شود.^{۴۳} این امر می‌تواند به سهولت کشف جعلیات کمک کند، لیکن تحقیقش در گرو حصول توافق سازندگان گوشی‌های همراه هوشمند، دوربین‌ها و طراحان تارنماهاست؛ پیشنهادی با دورنمایی که تحقیقش سال‌ها به طول می‌انجامد.

اما شرکت‌های مذکور پیرامون حذف، نشانه‌گذاری، تنزل رتبه و حفظ جعلیات از موشی یکسان پیروی نمی‌کنند. برای مثال، «یوتیوب» فیلم تحریف‌شده پلوسی را که شرح آن به اختصار رفت، به سرعت از تارنمای خود حذف و اعلام کرد که فیلم مذکور سیاست این سایت شبکه اجتماعی را با عملی گمراه‌کننده نقض کرده است. اما «فیس‌بوک»، ضمن خودداری از حذف ویدیوی مذکور، به نشریه «پست» گفت که پلتفرم مذکور فاقد سیاستی است که کاربران را ملزم به بارگذاری اطلاعات درست در این سایت شبکه اجتماعی کند. **نیک کلگ**^{۴۴}، سخنگوی ارشد فیس‌بوک، تصریح کرد: گفتار و مباحث سیاسی سیاستمداران از بررسی واقعیت معاف می‌شوند. حتی اگر مطالبشان قوانین این سایت شبکه اجتماعی را نقض کنند، آن محتوا همچنان روی این پلتفرم باقی می‌ماند؛ چراکه به گفته وی، سانسور با آنچه فیس‌بوک به آن اعتقاد دارد، مغایر است.^{۴۵} یوتیوب اعلام کرد پیگیری و سرمایه‌گذاری در حوزه رسانه‌های ترکیبی را در دستور کار خود قرار داده است و این را با مساعی پیشین خود، مانند مبارزه با هرزه‌نگاری و یافتن فیلم‌هایی که قوانین حق تکثیر را نقض کرده‌اند - و به کمک توأمان نرم‌افزار و نظارت انسانی انجام می‌شود - قیاس کرده است.

فیس‌بوک نیز در توضیح اقدامش در تأمین منابع مالی پژوهش پیرامون دست‌کاری رسانه‌ای که در برخی دانشگاه‌ها انجام می‌شود، به نشریه پست گفت که مبارزه با «اطلاعات نادرست»^{۴۶} از مهم‌ترین کارهایی است که این شرکت انجام می‌دهد. البته فیس‌بوک پس از اینکه در می ۲۰۱۹ از حذف ویدیوی تحریف‌شده پلوسی از پلتفرم خود امتناع کرد، یک ماه بعد، یعنی در ژوئن همان سال، با ویدیوی دیپ فیک **مارک زاکربرگ** روبه‌رو شد که در سایت شبکه اجتماعی تحت مالکیتش، «اینستاگرام» منتشر شد. فیلمی جعلی براساس سخنرانی سال ۲۰۱۷ زاکربرگ که او را در حالی رو به دوربین نشان می‌دهد که اظهار می‌کند: کنترل کامل اطلاعات به سرقت رفته میلیاردها انسان، زندگی، اسرار و آینده آن‌ها را تهدید می‌کند. در ادامه این مساعی، برای مبارزه با اطلاعات نادرست، شرکت رسانه‌ای مذکور در اکتبر ۲۰۱۹ برنامه‌های خود را برای انتخابات ۲۰۲۰ ایالات متحده اعلام کرد که جلوگیری از گسترش اطلاعات نادرست و سرمایه‌گذاری در پروژه‌های مربوط به سواد رسانه‌ای بخشی از آن محسوب می‌شوند. فیس‌بوک در توضیح این برنامه بیان کرد که تلاش هم‌زمان خود را در جهت کاستن از اطلاعات نادرست در سایت شبکه‌های اجتماعی فیس‌بوک و اینستاگرام به کار می‌بندد

و برای مثال، از گسترش اطلاعات مذکور در دو شبکه مزبور به منظور ممانعت از پربیننده شدن آن‌ها اقدام می‌کند. طی این برنامه، اطلاعات نادرست را در اینستاگرام از جست‌وجو و هشتک‌ها حذف می‌کند و انتشار آن‌ها را در خبرخوان فیس‌بوک کاهش می‌دهد. به علاوه اگر صفحه‌ها، دامنه‌ها و گروه‌ها به تکرار اطلاعات نادرست دست بزنند، ضمن ممانعت از اشاعه کلی محتوایشان در فیس‌بوک، در تبلیغات و کسب درآمدشان محدودیت ایجاد می‌کند. این شرکت رسانه‌ای مدعی شد، طی نوامبر ۲۰۱۹ یک گروه مستقل راستی‌آزما، نسبت به علامت‌گذاری تمامی مطالب بارگذاری شده در فیس‌بوک و اینستاگرام اقدام می‌کند تا کاربران نسبت به اینکه کدامین مطلب را بخوانند، به اشتراک بگذارند و در مجموع بدان اعتماد کنند و تصمیماتی بهتر بگیرند.^{۴۷}

«تویتر» اعلام کرد، هر هفته با موضوع بیش از هشت میلیون حساب کاربری که در تلاش‌اند، ضمن دست‌کاری محتوا به انتشار آن بپردازند، درگیر است، با این حال راستی‌آزمایی هر توییت را نیز امکان‌پذیر نمی‌داند. این شرکت می‌افزاید: بر این باور نیست که بخواهد پیش از انتشار توییت در فضای آنلاین پیرامون چیستی و صحت محتوای آن تصمیم بگیرد و مداخله کند. آشکار شدن توییت دروغ در کسری از ثانیه اتفاق می‌افتد، چراکه دیگر کاربران، بلادرنگ به بررسی توییت‌ها مبادرت می‌کنند. توییت مدعی است محتوای نادرست تا هنگامی که کذب بودنشان اثبات شود، به میزانی ناچیز در توییت‌بازنشر می‌شوند. لازم به ذکر است که این شرکت هیچ‌گونه آمار و ارقامی برای اثبات ادعایش ارائه نکرده است.



DEEP FAKE

فیلم، و همچنین پردازش سیگنال، همراه با تحلیل کلیدواژه‌های محتوای صوتی و تصویری است.

توجه: این نوشتار ترجمه‌ای آزاد از بخشی از مقاله‌ای است به نام: **Top AI researchers race to detect 'deepfake' videos: We are outgunned**

که توسط **درو هارول (Drew Harwell)** در ژوئن ۲۰۱۹ به رشتهٔ تحریر درآمده است.

با توجه به در پیش بودن انتخابات ۲۰۲۰ ایالات متحده، ضرورت وجود سامانه‌های تشخیص دیپ فیک بیشتر احساس می‌شود، اما شرکت‌های بزرگ آمریکایی همچنان به صیانت از انتخابات در چارچوب تخلفات سایبری تمایلی فزاینده دارند. **شامیر علی‌بی**^{۴۸} که مؤسس «استارت آپ آمبر»^{۴۹} در حوزهٔ تشخیص جعلیات است، می‌گوید که شرکتش با همکاری بخشی از مشتریان در حال آزمایش یک پوشش محافظ در برابر دیپ فیک با استفاده از فناوری «بلاکچین» برای احراز هویت، و روشی جدید در تفکیک قطعات

پی‌نوشت‌ها

صاحب‌نظران دیپ فیک‌ها را ذیل رسانه‌های ترکیبی تعریف می‌کنند.

43. <https://hackernoon.com/detecting-fake-video-needs-to-start-with-video-authentication-224a988996ce>

44. Nick Clegg

45. <http://yon.ir/KguFO>

۴۶. Misinformation: اطلاعاتی نادرست بدون

تعمد در گمراه‌سازی مخاطب است که رشد آن به

واسطهٔ توسعهٔ بنیان‌های فناوریانه در حوزهٔ تبادل

اطلاعات، به‌ویژه در فضای آنلاین، فزونی یافته

است. تولید و نشر این اطلاعات به‌طورمعمول

از سوی خبرنگاران و سازمان‌های خبری که

صادقانه به حرفهٔ خود پایبندند، اما به دلایلی

چون خطای انسانی یا سوء عملکرد سازمانی

ناخواسته به ورطهٔ تولید و نشر اطلاعات ناراست

فرو می‌غلطند، انجام می‌پذیرد. البته فارغ از

خبرنگاران و سازمان‌های خبری، برخی مخاطبان

نیز این اطلاعات را به اشتراک می‌گذارند که به

زعم خود اطلاعاتی موقتی را با دیگران سهیم

می‌شوند [فلسفی، ۱۳۹۸: ۸۱].

47. <https://newsroom.fb.com/news/2019/10/update-on-election-integrity-efforts>

48. Shamir Allibhai

49. <https://ambervideo.com>

17. Pew Research Center

18. Robert M. Chesney

19. Danielle Citron

20. www.poynter.org/ethics-trust/2019/the-liars-dividend-is-dangerous-for-journalists-heres-how-to-fight-it

21. Skolkovo

22. Few Shot

23. Generative Pretrained Transformer-2 (GPT-2)

24. <https://openai.com>

25. www.zoomit.ir/2019/5/28/336405/gpt2-language-model

26. Generative Adversarial Network (GAN)

27. www.theverge.com/2019/2/14/18224704/ai-machine-learning-language-models-read-write-openai-gpt2

28. GPT

29. <https://openai.com/blog/better-language-models>

30. <https://arxiv.org/pdf/1906.01946.pdf>

31. <https://alleni.ai>

32. Grover

33. Nitesh Saxena

34. MelNet

35. Jane Morris Goodall

36. <https://arxiv.org/pdf/1906.01083.pdf>

37. www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

38. Defense Advanced Research Projects Agency (DARPA)

39. MediFor

40. Matt Turek

۴۱. Computer Vision: یا بینایی رایانه، شامل

روش‌های مربوط به پردازش، تجزیه و تحلیل و

درک محتوای تصویرها یا فیلم‌های دیجیتال است.

۴۲. Synthetic Media: رسانه‌هایی دیجیتال یا

در نگاهی کلان‌تر، رسانه‌هایی تغییر یافته‌اند که

غالباً توسط الگوریتم‌ها هدایت می‌شوند. برخی

Artificial Intelligence: به سامانه‌هایی گفته

می‌شود که واکنش‌هایی مشابه رفتارهای هوشمند

انسانی، از جمله درک شرایط پیچیده، شبیه‌سازی

فرایندهای تفکری و شیوه‌های استدلالی انسانی و

پاسخ موفق به آن‌ها، یادگیری و توانایی کسب دانش

و استدلال برای حل مسائل دارند.

۲. Deepfake: که با معادل‌هایی چون جعل عمیق،

جعل عمقی، جعل ژرف و یادگیری عمیق جعل

کردن از آن یاد می‌شود.

۳. Disinformation: به «کژ آگاه‌سازی عامدانه»

و نشر هدفمند و گستردهٔ اطلاعات نادرست گفته

می‌شود که به‌طور معمول خاستگاهی حاکمیتی

دارد. در چارچوب «اطلاعات نادرست» که غالباً

هم‌سنگ اخبار جعلی فرض می‌شود، مخاطب

بر بستر اطلاعاتی نادرست زندگی می‌کند و طی

فرایندی طولانی فریفته می‌شود. آن‌گاه پیرامون

امری به باوری کذب می‌رسد (فلسفی، ۱۳۹۸: ۸۲)

۴. Fake News عبارت است از تولید و نشر عامدانهٔ

اطلاعات نادرست و / یا شایعه به منظور جهت دادن

به افکار عمومی برای کسب منافع خاص.

5. Elyse Samuels

6. Hany Farid

7. www.latimes.com/business/technology/la-fi-tn-fake-videos-20180219-story.html

8. Nancy Pelosi

9. <http://yon.ir/dwCO>

10. www.reddit.com

۱۱. Deep learning: زیرشاخه‌ای از یادگیری ماشین

که از مجموعه‌ای از الگوریتم‌ها برای پردازش

سیگنال‌های حسی، مانند صدا و تصویر استفاده

می‌کند.

۱۲. Capitol Hill: مقر حکومتی ایالات متحده در

شهر واشنگتن دی سی که ساختمان‌های سن،

مجلس نمایندگان و دیوان عدالت عالی در آن واقع

است.

13. Adam Schiff

14. Rachel Thomas

15. www.fast.ai

16. <http://aviv.me>

منابع

۱. سلطانی‌فر، محمد؛ سلیمی، مریم؛ و فلسفی،

سیدغلامرضا (۱۳۹۶). **اخبار جعلی و**

مهارت‌های مقابله با آن. فصلنامهٔ رسانه. دورهٔ

۲۸. شمارهٔ سه. شمارهٔ پیاپی ۱۰۸.

http://qjmn.farhang.gov.ir/article_51733.html

۲. فلسفی، سیدغلامرضا (۱۳۹۸). «جهان

در خطر بازگشت به عصر تاریکی». *مجلهٔ*

بازخورد. سال دوم، شمارهٔ هشتم.

magiran.com/p1998506