

■ از تکنولوژی دیپ فیک صوتی که منجر به کلون کردن اصوات یا بازتولید تقریباً بی نقص صدای افراد می شود می توان هم برای مقاصد خوب و هم مقاصد بد استفاده کرد. برای مثال می توان آن را در سرویس های ترکیب صدا به کار گرفت و صدا را به افرادی که این توانایی را به خاطر عارضه هایی مانند اسکروز جانبی آمیوتروفیک، آپراکسی، ضربه های مغزی تروماتیک، سکنه یا هر چیز مشابهی از دست داده اند بازگرداند.

از این تکنولوژی همین حالا در صنایع فیلم و تلویزیون، گیمینگ و مراکز تماس استفاده می شود و بی شمار پتانسیل نیز برای استفاده از آن در امور رمزنگاری و تراپی نیز وجود ندارد. در عین حال، نمی توان این را زیر سوال برد که تکنولوژی دیپ فیک صوتی تهدیدی بزرگ برای پروسه های دموکراتیک است، خصوصاً پروسه هایی که به حریم شخصی مربوط می شوند. در صورت استفاده بدخواهانه، تکنولوژی دیپ فیک صوتی می تواند به کلاهبرداری، فریب و آزار منجر شود.

دقیقا به خاطر همین پتانسیل های خوب و بد، اکنون زمانش رسیده که توجهی مضاعف به تکنولوژی دیپ فیک صوتی نشان دهیم و به طریقی به بررسی به آن پردازیم که منجر به کاهش خطر فریب دادن مردم شود، خصوصاً این فریب که شخصی سرشناس، چیزی را گفته که در حقیقت بیانش نکرده است.

موارد استفاده برای کسب و کارها چه خواهد بود؟ ارقام نشان می دهند که فرصتی بزرگ برای سازمان های اقتصادی فراهم شده تا از تکنولوژی های صوتی استفاده کرده و وارد تجارت های تازه شوند. موسسه AppDynamics در سال ۲۰۱۸ گزارش کرد که تا سال ۲۰۲۰، بیش از نیمی از سرچ های انجام شده در سطح وب، با صدا صورت خواهند گرفت.

از سوی دیگر نیز، ۶۱ درصد از تصمیم گیران برجسته حوزه فناوری اطلاعات گام را فراتر گذاشته و انتظار دارند فرمان های صوتی به صورت کامل جایگزین فرمان های تایپ شده به صورت دستی برای یافتن اطلاعات در سراسر اینترنت شوند. و نسل جوان در این حوزه پیشرو است - ۸۴ درصد از متولدین قرن بیست و یکم همین حالا از دستیارهای صوتی برای پایش برنامه روزانه و مسئولیت های خود استفاده می کنند.

بیاید به مرور مزایا، معایب و آینده که در انتظار تکنولوژی دیپ فیک صوتی است پردازیم.

امکان صحبت به صورت طبیعی از سوی افرادی که از مشکلات پزشکی رنج می برند

امکان برقراری ارتباط و اشتراک گذاری افکار و احساسات با کلمات، از اهمیتی بسیار زیاد برخوردار بوده و در واقع این یکی از معدود ویژگی هایی است که ما انسان ها را منحصر به فرد می کند. اما تکنولوژی واگردانی صوتی سنت شکنانه کنونی، کارهای بسیار بیشتری می تواند برای افرادی انجام دهد که از توانایی صحبت برخوردار نیستند.

برای مثال گسترش تکنولوژی اتوماسیون خانگی را کنترل صوتی را در نظر بگیرید. کلون کردن صدا می تواند باعث شود افرادی که قادر به صحبت به صورت طبیعی نیستند استقلال بیشتری داشته باشند و بهتر به استفاده از دیوایس هایی پردازند که از با فرمان های صوتی کنترل می شوند.

دستیارهای صوتی

بنابر گزارش «پیش بینی وضعیت دستیار دیجیتال و دیوایس صوتی مبتنی بر هوش مصنوعی» شرکت Ovum، تا سال ۲۰۲۱ میلادی شمار دستیارهای صوتی از تعداد انسان های حاضر روی کره خاکی بیشتر خواهد بود. بیاید برای مثال به دستیار صوتی گوگل نگاه کنیم. صدای این دستیار با استفاده از سیستم نوشتار به گفتار Tacotron 2 تولید می شود که خود بر دو شبکه عصبی عمیق متکی است.

نخستین شبکه، متن را به تصویری بصری (یا طیف نگاره) از فرکانس های صدا به مرور زمان تبدیل می کند و سپس یک سیستم WaveNet به تحلیل طیف نگاره پرداخته و المان های صوتی را شکل می دهد. خروجی، گفتاری است که تفاوت چندانی با گفتار عادی انسانی ندارد و حتی می تواند کلمات بسیار دشوار را به خوبی تلفظ کند.

محتوای تعاملی برای کورس های یادگیری آنلاین

کلون کردن صدا با هوش مصنوعی باعث می شود که دیگر در هر کلاس درس نیازی به نوت برداری یا بازنگری در نوت ها برای یافتن اشتباهات بالقوه نباشد. از سوی دیگر، هزینه پولی و زمانی سخنرانی هایی که به صورت حرفه ای ضبط می شوند نیز کاهش می یابد و کورس های یادگیری آنلاین شکلی دسترس پذیرتر به خود می گیرند. این اتفاق کوچک نیست، خصوصاً در دوران بحران کنونی که به خاطر پاندمی کووید-۱۹، محدودیت های فراوان برای مردم به وجود آمده است.

اخادی

از ویدیوهای جعلی اما شدیداً واقع‌گرایانه با صحنه‌های جنسی یا خشونت‌آمیز که با ترکیب تکنولوژی دیپ فیک ویدیویی و صوتی ساخته شده‌اند می‌توان برای اخاذی از افراد استفاده کرد.

ایمیل‌های اسپم

اگر تا به حال ایمیل برایتان آمده باشد که در آن نوشته شده «با بانک X ارتباط برقرار کنید تا اطلاعات بیشتر راجع به روند حواله پول در اختیارتان قرار بگیرد» به احتمال زیاد نام ایمیل‌های اسپم به گوشتان خورده و هیچ توجهی هم به آن‌ها نشان نداده‌اید. اما یک تماس تلفنی از سوی کسی که صدایی بسیار شبیه به یکی از مخاطبان مورد اعتمادتان دارد و شما را به پاسخ دادن به یک ایمیل ترغیب می‌کند ممکن است نظرتان را عوض کرده و باعث شود دست به انجام کاری بزنید که بعداً آرزو می‌کردید هیچوقت انجامش نمی‌دادید.

رقابت غیر قانونی

یک نفر ممکن است خود را مدیر عامل فلان کمپانی جای بزند و در این جایگاه مهم، شروع به نشر اطلاعات دروغین در تماس‌های تلفنی مربوط به گزارش‌های مالی کند. سهام‌داران و سرمایه‌گذاران فریب خواهند خورد و قیمت سهام را از آنچه در واقعیت هست خواهند پنداشت. از همین تکنیک بدخواهانه می‌توان برای تخریب رقبای صنعتی نیز استفاده کرد.

استفاده در مدل بنگاه-بج-بنگاه

با در نظرگیری یافته‌های موجود در گزارش AppDynamics که بالاتر به آن‌ها اشاره شد، استفاده از تکنولوژی بازتولید صدا به صورت گسترده، آنقدرها دور نیست. ۶۹ درصد از تصمیم‌گیرندگان حوزه فناوری اطلاعات در شرکت‌هایی کار می‌کنند که یا همین حالا روی تکنولوژی سرمایه‌گذاری کرده‌اند یا می‌خواهند طی سه سال آتی روی آن سرمایه‌گذاری کنند.

چه به دنبال بازتولید صدا باشید، چه درمان برای مشکلات کلامی، چه گیمینگ، چه رمزنگاری و چه هرچیز دیگر، از تکنولوژی کلون صدا می‌توان در گستره وسیعی از صنایع نفع برد. برای مثال با توسعه یک سیستم بلادرنگ در دنیای گیمینگ می‌توان به گیمرها اجازه داد که از صداهای مختلف در چت‌های داخل بازی استفاده کنند. وقتی صحبت از تکنولوژی دیپ فیک صوتی باشد، جهان دارد با سرعت بسیار زیادی تغییر می‌کند.

کلون صدا برای مراکز تماس

کمپانی Reespecher اکنون مشغول کار روی تکنولوژی‌هایی سنت‌شکن است که به اپراتورها در آن سوی آب‌ها اجازه می‌دهد صدایی شبیه به افراد بومی داشته باشند. بنابراین هنگام تماس با یکی از این مراکز، احساس خواهید کرد که با یکی از شهروندان کشور خود صحبت می‌کنید. از سوی دیگر، اپراتورهای رباتیک هم به زودی صدایی انسانی‌تر خواهند داشت.

جمع‌بندی

تکنولوژی دیپ فیک صوتی بدون هیچ تردید ریسک‌های خود را به همراه می‌آورد، اما نخستین گام برای تمرکز مضاعف روی جنبه‌های مثبت تکنولوژی اینست که جنبه‌های منفی را شناسایی کرده و در صدد تقلیل دادن آن‌ها برآییم. صداهای کلون شده‌ای که هیچ تفاوتی با صدای سخنگوی اصلی ندارند می‌توانند بهترین ابزار ممکن برای فیلم‌سازان، بازی‌سازان و دیگر تولیدکننده مدیا باشند.

خیلی زود صداهای بازتولید شده به صورت دیجیتالی، احساسات و حالات گوناگون را نیز در بر می‌گیرند، همانطور که شواهدش را با دستیارهای صوتی نظیر بیکسبی سامسونگ، سیری اپل یا الکسای آمازون دیده‌ایم. و با از راه رسیدن صداهایی که قادر به انتقال احساسات باشند، مشاوره و مصاحبت شکلی کاملاً جدید به خود خواهند گرفت.

منبع: دیجیاتو