



## چالش بزرگی به نام «تشخیص تصاویر جعلی هوش مصنوعی»

شرکت اوپن‌ای‌آی (OpenAI) می‌گوید می‌تواند تصاویری را که توسط نرم‌افزار خودش ساخته شده‌اند، تشخیص دهد ...

شرکت اوپن‌ای‌آی (OpenAI) می‌گوید می‌تواند تصاویری را که توسط نرم‌افزار خودش ساخته شده‌اند، تشخیص دهد و در حال آزمایش ابزار جدیدی برای این منظور است، اما این ابزار برای تصاویر سایر مدل‌ها یا تصاویری که دستکاری شده‌اند، کارآمد نیست و این تهدید بزرگی است.

به گزارش ایسنا، همه ما فکر می‌کنیم که در شناسایی تصاویر ساخته شده توسط هوش مصنوعی بسیار خوب هستیم، چرا که تاکنون عادت داشته ایم تصاویری که توسط هوش مصنوعی ساخته می‌شوند را همراه با نقص‌های گاه‌فاحش و گاه‌ظریف ببینیم که تشخیص ساختگی بودن آنها را برایمان آسان می‌کرد.

به نقل از انگجت، با این حال فناوری به طور دائم در حال پیشرفت است و زمان زیادی نخواهد گذشت تا ببینیم دیگر نمی‌توانیم بگوییم چه چیزی واقعی است یا توسط هوش مصنوعی تولید شده است.

اکنون رهبر صنعت توسعه هوش مصنوعی یعنی شرکت اوپن‌ای‌آی در تلاش است تا با ایجاد مجموعه ابزاری که تصاویر ایجاد شده توسط هوش مصنوعی مولد DALL-E ۲ خود را شناسایی می‌کند، بر این مشکل غلبه کند.

این شرکت می‌گوید که می‌تواند تصاویری را که توسط DALL-۳ ساخته شده‌اند، در ۹۸ درصد مواقع به دقت تشخیص دهد که دقتی بسیار عالی است. با این حال، برخی نواقص وجود دارند. اول از همه اینکه یک تصویر باید حتماً توسط DALL-E ایجاد شده باشد و عیب اینجاست که این هوش مصنوعی تنها هوش مصنوعی در دسترس برای تولید تصاویر نیست.

در حالی که فضای اینترنت از تصاویر تولید شده توسط هوش مصنوعی پر شده است، بر اساس داده‌های ارائه شده توسط اوپن‌ای‌آی، سیستم آن تنها موفق شد ۵ تا ۱۰ درصد از تصاویر ساخته شده توسط سایر مدل‌های هوش مصنوعی را با موفقیت تشخیص دهد.

همچنین اگر تصویر به هر نحوی اصلاح شده باشد، این ابزار تشخیصی با مشکل مواجه می‌شود. البته در مورد تغییرات جزئی مانند برش و فشرده‌سازی به نظر نمی‌رسد که این موضوع چندان مهم باشد. در این موارد، میزان موفقیت کمتر بود، اما همچنان در محدوده قابل قبولی در حدود ۹۵ تا ۹۷ درصد بود. با این حال اصلاح رنگ، میزان موفقیت را به ۸۲ درصد کاهش داد.

اینجا جایی است که همه چیز پیچیده می‌شود، چرا که این ابزار زمانی که برای تشخیص و طبقه‌بندی تصاویری که دستخوش تغییرات گسترده‌تری شده‌اند، استفاده شد، با مشکل مواجه شد. اوپن‌ای‌آی حتی میزان موفقیت در این موارد را نیز منتشر نکرد و به گفتن اینکه سایر تغییرات می‌توانند عملکرد این ابزار را پایین بیاورند، بسنده کرد.

این یک اتفاق ناگوار است، چرا که تصاویر تولید شده توسط هوش مصنوعی پس از دستکاری ایجاد و تغییرات و اصلاح می‌توانند آثار سوئی بر مردم داشته باشند.

حداقل اوپن‌ای‌آی در مورد محدودیت‌های فناوری تشخیص خود شفاف است. این شرکت همچنین به آزمایش‌کنندگان خارجی دسترسی به ابزارهای ذکر شده را برای کمک به رفع این مشکلات می‌دهد.

اوپن‌ای‌آی همراه با مایکروسافت، ۲ میلیون دلار برای چیزی به نام صندوق تاب‌آوری اجتماعی (Societal Resilience Fund) سرمایه‌گذاری کرده است که امیدوار است آموزش و سواد در حوزه هوش مصنوعی را افزایش دهد.