مرکز رصد و آینده پژوهی

خرداد 1401

تکنولوژی Dall-E؛ هیاهوی این روزهای شبکه‌های اجتماعی

فهرست

تکنولوژی Dall-E چیست؟ 3
Dall-E Mini سرگرمی جدید کاربران شبکه‌های اجتماعی 7
Dall-E؛ جرقه‌ی انقلابی بزرگ در خلق آثار هنری؟ 11
با تکنولوژی Dall-E نیازی به داشتن مهارت‌های پیشرفته فوتوشاپ نیست..... 11
OpenAI برای جلوگیری از سواستفاده از Dall-E سیاست‌های سختگیرانه‌ای دارد... 14
پلتفرم‌های هوش مصنوعی مشابه DALL-E 17
DALL-E؛ ترس‌ها و لبخندها..... 19

اگر اخیراً با تصاویر دیجیتالی عجیبی مثل «زرافه‌ای در فضا» یا «دارث ویدر گیتار الکتریک به‌دست» روبه‌رو شده‌اید، احتمالاً کارِ هوش مصنوعی Dall-E است که این‌روزها در توییتر غوغا به‌پا کرده است.

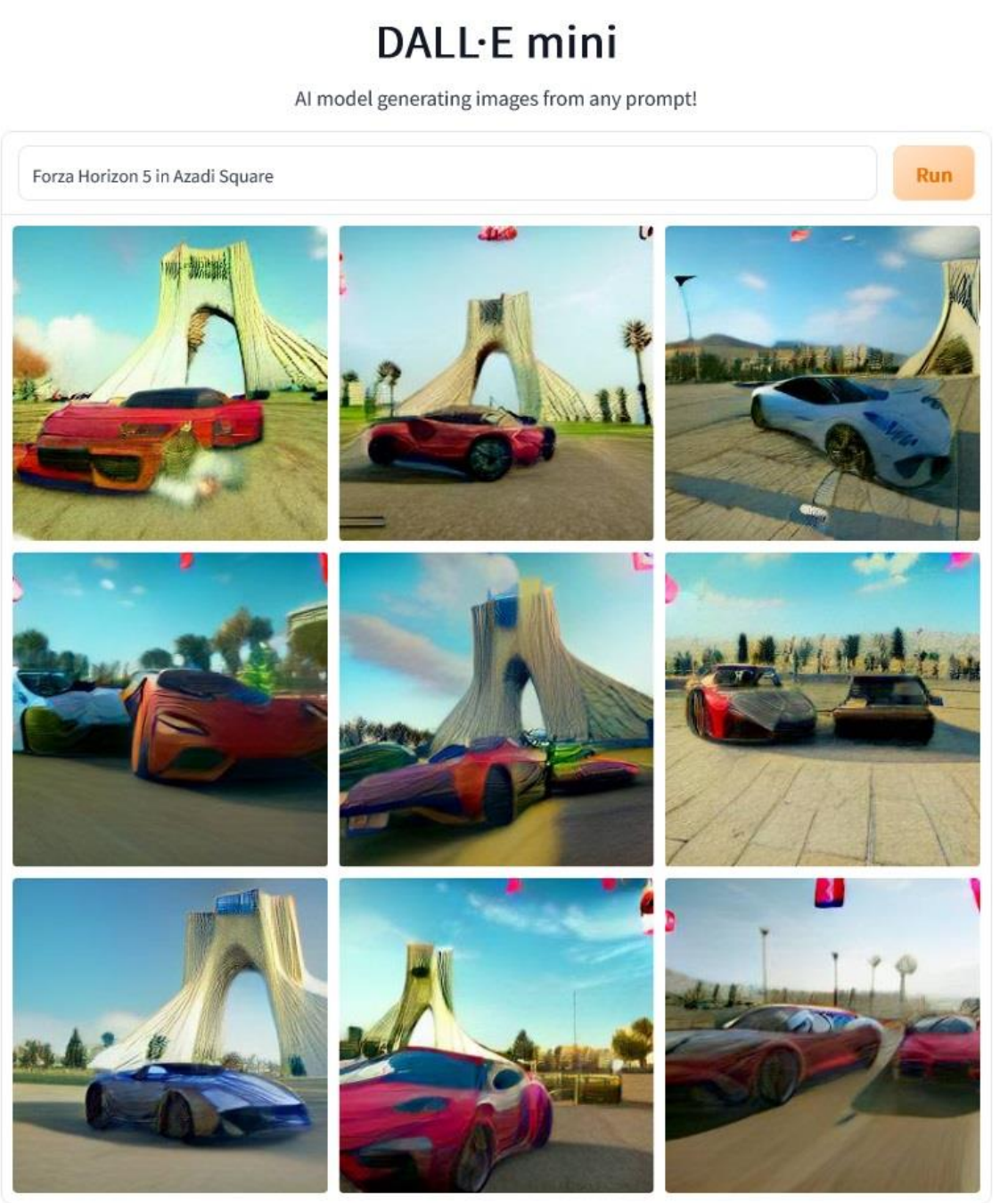
شاید شما هم مثل من چند روزی است که تایم‌لاین توییترتان پر از تصاویر دیجیتالی عجیب‌وغریبی شده است که توسط هوش مصنوعی Dall-E Mini ایجاد شده‌اند؛ تصاویری مثل کارل مارکس ساخته شده با برفک تلویزیون، کاراکتر والتر وایت با کنسول گیم‌کیوب در دست؛ گوردون رمزی در حال خوردن بیگ مک، یا کرمیت قورباغه در نقاشی جیغ ادوارد مونک.

سرویس Dall-E Mini که در وب‌سایت Hugging Face از آن میزبانی می‌شود، با استفاده از هوش مصنوعی، تراش داده و میلیاردها تصویر از گوشه‌وکنار اینترنت قادر است از متنی که کاربر به آن می‌دهد، تصاویر نسبتاً مرتبط ایجاد کند؛ حتی اگر متن تایپ شده شبیه مثال‌های بالا، عجیب و سورئال باشد و نمونه‌ای از آن‌ها در دنیای واقعی یافت نشود. مثلا من عبارت «پونیو آیفون به دست» را در این پلتفرم امتحان کردم و تصاویر زیر به دست آمد:

دلیل اینکه «اثر هنری» ایجادشده با Dall-E Mini اینقدر محبوب است و همه دارند درباره‌ی آن صحبت می‌کنند، به توانایی شگفت‌انگیز هوش مصنوعی به خلق تصاویر از ایده‌هایی برمی‌گردد که قبلا کسی آن‌ها را مطرح نکرده است. مثلا اگر در گوگل عبارت «گندالف در فضاپیما» را جستجو کنید، نتیجه‌ی موردنظر را دریافت نخواهید کرد؛ اما همین عبارت در پلتفرم‌های مولد تصویر از جمله Dall-E Mini نتایج نسبتاً مرتبطی را نشان می‌دهد.

هر چند سال یک بار، نوعی تکنولوژی پا به میدان می‌گذارد که دنیا را به پیش و پس از خودش تقسیم می‌کند. مثلا من اولین بار که آهنگی را «شزم» کردم یا اولین تماس ویدئویی که با وایبر گرفتم یا اولین عکسی را که با یک دوربین دومگاپیکسلی انداختم و بعد با یک فیلتر ساده در اینستاگرام پست کردم، خوب به خاطر دارم.

تصویر درخواست شده: فورتزا هورایزن ۵ در میدان آزادی



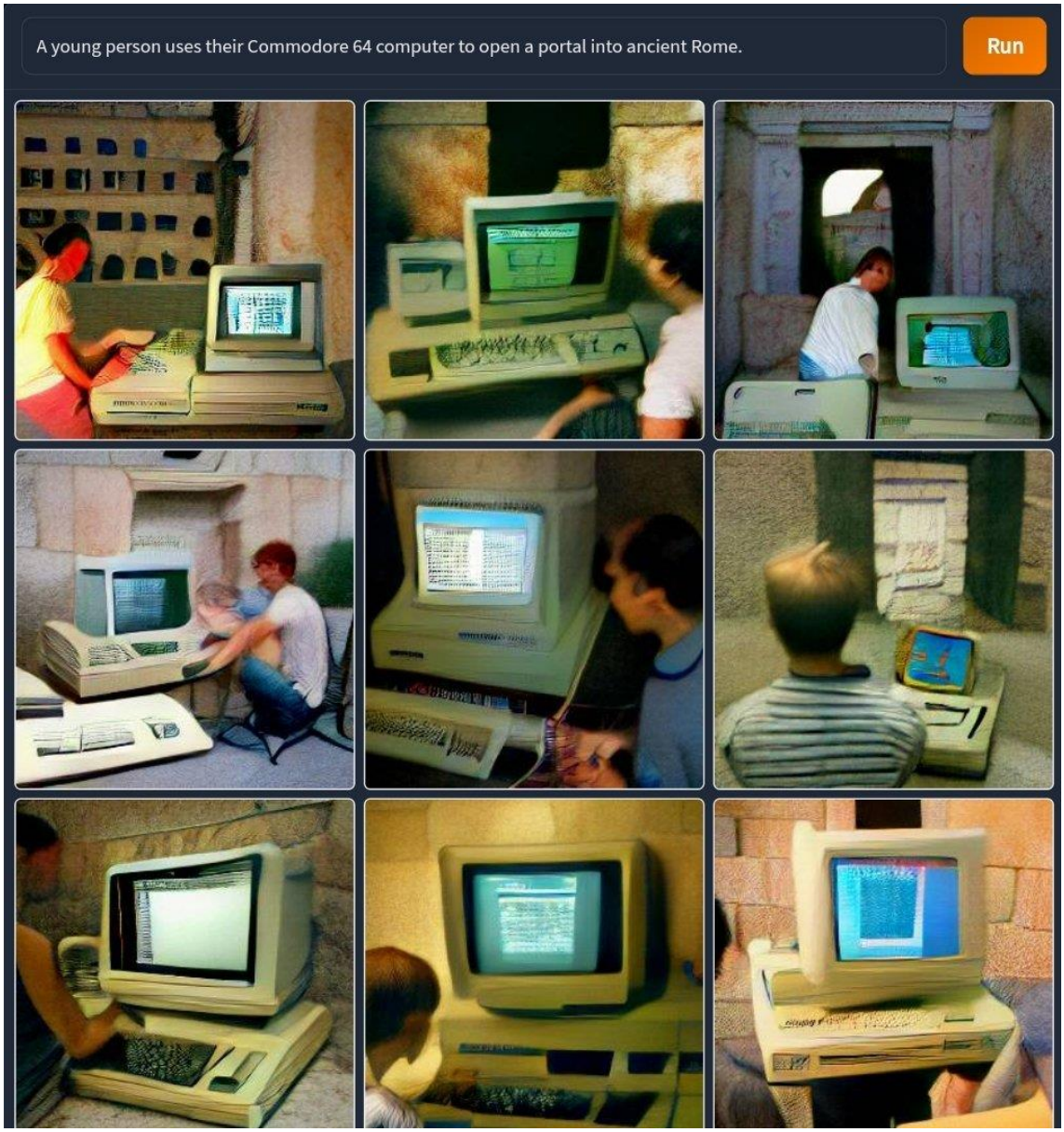
چیزی که این لحظات را به‌یادماندنی می‌کند، تصور اتفاقات غیر قابل پیش‌بینی و شگفت‌انگیزی است که با ظهور این تکنولوژی‌ها ممکن است دست‌یافتنی شوند. حالا که می‌شود تماس ویدیویی برقرار کرد، از فضای ابری هر فایلی را به هر دستگاهی فراخواند یا با یک گوشی و وای‌فای به‌صورت استریم زنده با هزاران نفر در سراسر دنیا ارتباط گرفت، دنیا چه شکلی خواهد شد؟

چند سالی می‌شد که خبری از این نوع تکنولوژی‌ها نبود؛ از همین‌هایی که دلمان بخواهد به دوستانمان نشان دهیم و بگوییم این را حتماً باید ببینید! البته Dall-E Mini همان‌طور که از اسمش مشخص است، تنها نمونه‌ی بسیار کوچکی از تکنولوژی بزرگ و آینده‌داری است که می‌تواند آینده را شکل دهد؛ دنیای الگوریتم‌ها و هوش مصنوعی.

Dall-E Mini سرگرمی جدید کاربران شبکه‌های اجتماعی

پروژه‌ی برنامه‌نویس اهل تگزاس به نام بوریس دیما (Boris Dayma) است که آن را در جولای ۲۰۲۱ برای شرکت در مسابقه‌ی محاسبات هوش مصنوعی زیر نظر گوگل و شرکت ارائه‌دهنده‌ی تکنولوژی یادگیری ماشین Hugging Face توسعه داد. این پروژه که در حال حاضر در وب‌سایت Hugging Face میزبانی می‌شود، این‌روزها به‌خاطر هیاهوی توییتر آنقدر محبوب شده که ترافیک روی سایت بسیار بالا است و شاید لازم به چند بار تلاش باشد تا سرانجام تصویری را که به دنبالش هستیم، نشان داده شود؛ اما ازآنجاکه استفاده از این سرویس کاملاً رایگان است، ارزش کمی صبر و تلاش را برای رفع حس کنجکاوی‌مان دارد.

تصویر درخواست شده: جوانی در حال باز کردن پرتالی به روم باستان به‌کمک کومودور ۶۴



در Dall-E Mini می‌توان هر چیزی را متصور شد؛ اتفاقا یکی از دلایلی که کاربران اینترنتی تا این حد شیفته‌ی این سرویس شده‌اند، به‌خاطر توانایی آن در به‌تصویر کشیدن عجیب‌ترین و بی‌ربط‌ترین درخواست‌ها و مسخره‌ترین موقعیت‌ها است که منجر به نتایج خنده‌دار و گاها غافلگیرکننده‌ای می‌شود. وقتی عبارت مدنظرتان را در کادر متنی Dall-E Mini وارد می‌کنید و دکمه‌ی Run را فشار می‌دهید، با ۹ تصویر به ابعاد ۲۵۶ در ۲۵۶ پیکسل روبه‌رو می‌شوید که از نظر الگوریتم به‌کار رفته در این پلتفرم، با چیزی که دنبالش هستید، مطابقت دارد. البته وقتی به این تصاویر دقیق می‌شوید، متوجه خواهید شد که چقدر ایراد دارند؛ به‌ویژه اگر به جای حیوان یا کاراکترهای خیالی،

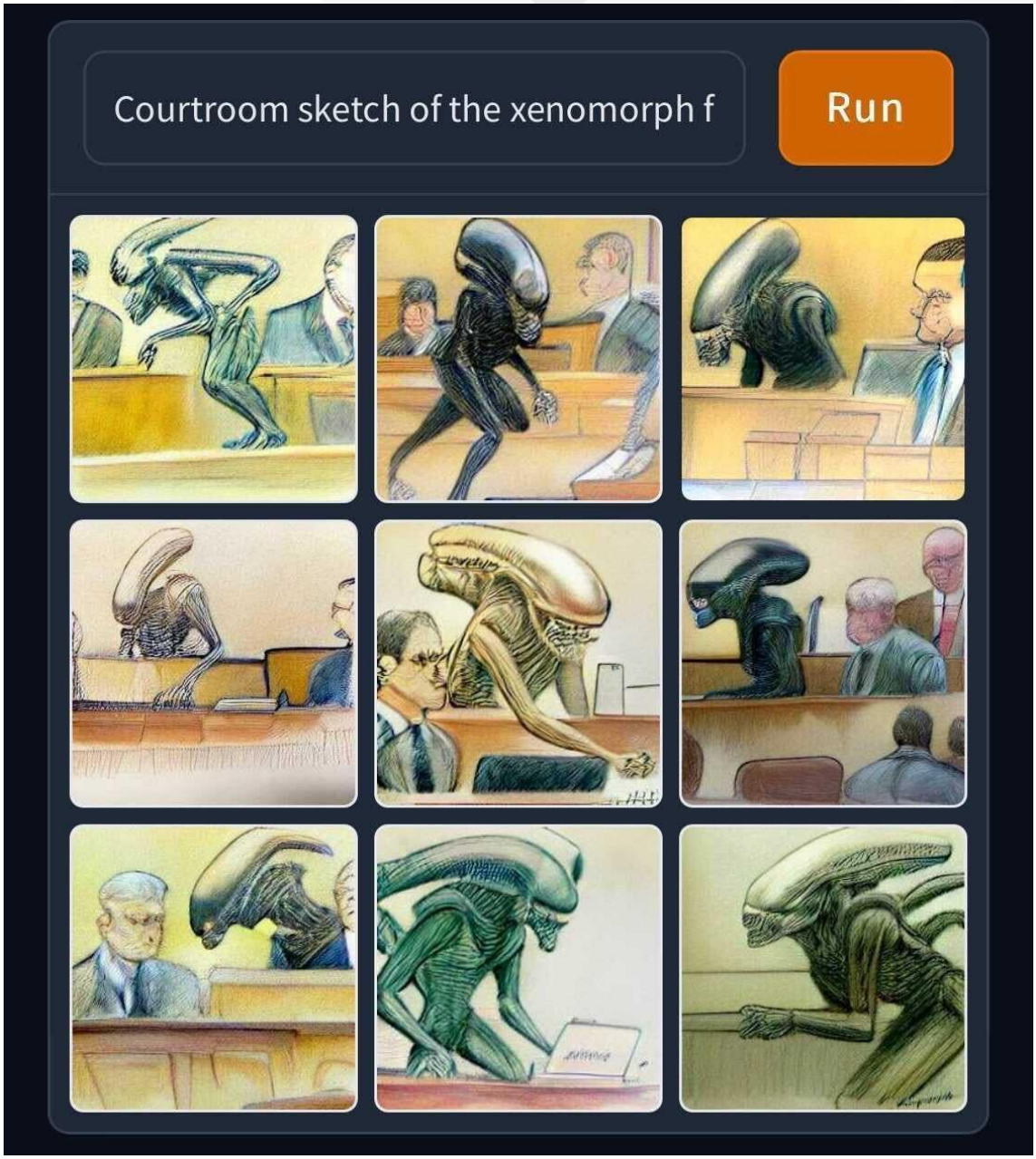
نام افراد واقعی را وارد کنید. اما از دور و در نگاهی گذارا، تصاویر خلق شده اغلب به چیزی که انتظار داریم، شباهت زیادی دارند.

دیما اذعان دارد که این پلتفرم نتایج بهتری را در نقاشی‌های انتزاعی و آبستره نشان می‌دهد، اما در مورد تصاویر واقعی با جزییات بیشتر، با مشکل روبه‌رو می‌شود.

سخت‌ترین بخش بدون شک تصاویر افراد است. اگر از Dall-E یک منظره بخواهید، نتیجه عالی خواهد بود؛ چون اگر درختی ایراد داشته باشد، کسی متوجه آن نخواهد شد. اما اگر بخشی از صورت آدم، مثلا چشم دچار مشکل باشد، ما بلافاصله متوجه آن خواهیم شد.

اگرچه Dall-E Mini می‌تواند تصاویر زیبا و «هنری» ایجاد کند، پای ریاضیات و الگوریتم بدون دخالت هرگونه ذوق هنری در میان است. کاری که هوش مصنوعی این پلتفرم انجام می‌دهد، جاری کردن ذوق و قریحه‌ی هنری‌اش نیست. متأسفانه یا خوشبختانه، هوش مصنوعی هنوز آنقدر پیشرفت نکرده که بخواهد از خود خلاقیت نشان دهد؛ بلکه الگوریتم Dall-E Mini فقط به تصاویر بی‌شماری در اینترنت که متن با درخواست کاربر مطابقت دارند، نگاه می‌کند و بعد الگوهایی را که در اکثر آن‌ها تکرار شده است، نظیر اَشکال، رنگ‌ها و توضیحات را پیدا می‌کند. Dall-E Mini سپس به کمک این الگوها یاد می‌گیرد چگونه تصویر متناسب با درخواست متنی کاربر را ایجاد کند.

تصویر درخواست شده: موجود فضایی فیلم Alien به سبک پیش‌طراحی اتاق دادگاه



دیما Dall-E Mini را تقلیدی از پروژه‌ی DALL-E شرکت OpenAI اما در مقیاس بسیار کوچک‌تر با معماری ساده‌تر از نسخه‌ی اصلی توصیف می‌کند که اگرچه نتایج آن کیفیت بسیار پایین‌تری از DALL-E دارد، به‌طور کاملاً رایگان برای همه دردسترس است و برای اجرایش نیازی به سخت‌افزارهای آن‌چنانی نیست؛ حتی می‌توانید Dall-E Mini را در گوشی هوشمند خود امتحان کنید، هرچند به گفته‌ی سازنده، روی وب نتیجه‌ی بهتری می‌دهد.

جالب است بدانید مدل‌های یادگیری ماشین که برای تبدیل متن به تصویر به‌کار می‌روند، تنها چند سالی است که به این حد از توانایی Dall-E Mini رسیده‌اند. مثلا در این مطلب که در سال ۲۰۱۸ منتشر شده بود، می‌توانید چالش‌ها و ضعف‌های مدل‌های قدیمی‌تر را مشاهده کنید. وقتی به این مدل، کپشن «گله‌ای زرافه در کشتی» داده شد، فقط توانسته بود چند شکل زرافه‌مانند ایستاده روی آب ایجاد کند. این مدل حتی از پس پردازش درخواست بسیار ساده‌ی «یک گوسفند» هم برنیامده بود. اینکه حالا می‌توانیم از یک پروژه‌ی شخصی کوچک که صرفاً برای شرکت در مسابقه طراحی شده بود، چنین نتایج نزدیک به واقعیتی دریافت کنیم، پیشرفت چشمگیری را در «فهم و درک» الگوریتم‌ها نشان می‌دهد.

Dall-E؛ جرقه‌ی انقلابی بزرگ در خلق آثار هنری؟

Dall-E Mini در مقایسه با نسخه اصلی آن، یعنی DALL-E ، یا دقیق‌تر بگوییم DALL-E 2، کودک نوپایی بیش نیست؛ اما متأسفانه پروژه‌ی اصلی در حال حاضر در فاز بتای خصوصی است و کمتر از ۵ هزار نفر به آن دسترسی دارند.

سرویس DALL-E که نامش از ترکیب سالوادور دالی، نقاش سورئالیست و انیمیشن WALL-E پیکسار گرفته شده است، در ژانویه ۲۰۲۱ در شرکت OpenAI مستقر در سانفرانسیسکو متولد شد. OpenAI در سال ۲۰۱۵ به‌دست ایلان ماسک، سم آلتمن، ایلیا سوتسکور و سه نفر دیگر تأسیس شد، اما ماسک در سال ۲۰۱۸ از هیئت مدیره خارج شد. در سال ۲۰۱۹، مایکروسافت یک میلیارد دلار در این شرکت سرمایه‌گذاری کرد. OpenAI به خاطر توسعه‌ی GPT-3، ابزاری برای تولید متن‌های پیچیده و سنگین از عبارت‌های ساده و همین طور Copilot، ابزاری برای خودکار کردن فرایند کدنویسی برای مهندسان نرم‌افزار، معروف است و کد متن‌باز آن‌ها برای استفاده‌ی عموم در گیت‌هاب موجود است.

با تکنولوژی Dall-E دیگر نیازی به داشتن مهارت‌های پیشرفته فوتوشاپ نیست

نسخه‌ی اول DALL-E نیز براساس مدل GPT-3 توسعه یافت و تنها به ایجاد تصاویری در ابعاد ۲۵۶ در ۲۵۶ پیکسل محدود بود. اما نسخه‌ی دوم که در آوریل ۲۰۲۲ وارد فاز بتای خصوصی شد،

جهش بزرگی در حوزه‌ی مولدهای تصویر مبتنی بر هوش مصنوعی محسوب می‌شود. تصاویری که DALL-E 2 قادر به ایجاد آن‌ها است حالا ۱۰۲۴ در ۱۰۲۴ پیکسل هستند و از تکنیک‌های جدیدی چون «inpainting» استفاده می‌کنند که در آن بخش‌هایی از تصویر به انتخاب کاربر با تصویر دیگری جایگزین می‌شوند؛ مثلا فرض کنید از پرتقالی در ظرف عکس گرفته‌اید و بعد به DALL-E می‌گویید آن را با تصویر سیب جایگزین کند و DALL-E این کار را به تمیزترین شکل ممکن انجام می‌دهد، طوری که دیگر نیازی به مهارت‌های پیشرفته فوتوشاپ نیست!

نسخه دوم DALL-E علاوه‌بر قابلیت ویرایش و رتوش عکس، می‌تواند توضیح متنی ساده‌ای مانند «مهمانی چای فیل‌ها روی چمن» را که پیش از این وجود خارجی نداشته‌اند، به چنان تصاویر هنری یا واقعی تبدیل کند که از دیدن آن‌ها شگفت‌زده خواهید شد. درواقع، جادوی DALL-E نه صرفاً به شناخت اشیا به‌صورت جداگانه بلکه در درک فوق‌العاده‌ی آن از روابط بین اشیا است، به‌طوری که وقتی از آن می‌خواهید «فضانوردی سوار بر اسب» را ایجاد کند، خوب می‌داند منظور شما از این خواسته دقیقاً چیست. در این رشته توییت می‌توانید برخی از تصاویری را که با DALL-E ایجاد شده‌اند، تماشا کنید.

تصویر درخواست شده: دایناسور رباتی درمقابل هیولاهای کامیونی در کولوسئوم

OpenAI پروژه DALL-E را نمونه‌ای از همکاری بین انسان‌های خلاق با سیستم‌های هوشمند برای به تصویر درآوردن ایده‌های جدید و تقویت خلاقیت انسان‌ها توصیف می‌کند. این شرکت همچنین اضافه می‌کند که تصاویر ایجاد شده با DALL-E می‌توانند به ما بگویند که آیا سیستم واقعا حرف ما انسان‌ها را می‌فهمد یا اینکه فقط چیزهایی را که یاد گرفته تکرار می‌کند. علاوه‌براین، DALL-Eبا ما نشان می‌دهد سیستم‌های هوش مصنوعی، دنیای ما را چگونه می‌بینند و درک می‌کنند که این مورد به گفته OpenAI ، برای توسعه‌ی هوش مصنوعی مفید و امن بسیار مهم است.

نکته‌ی مهمی که باید درباره‌ی DALL-E بدانید این است که شرکت توسعه دهنده‌ی آن به‌شدت مراقب است که از آن برای مصارف نادرست استفاده نشود. کاربرانی که به استفاده از این پلتفرم دعوت شده‌اند، بعد از ایجاد اکانت لازم است با سیاست محتوای این شرکت موافقت کنند. برای مثال، DALL-E اجازه‌ی به‌کار بردن الفاظ حاوی نفرت، خشونت، برهنگی و موارد غیراخلاقی و هرگونه محتوای سیاسی را به کاربر نمی‌دهد. این پلتفرم همچنین از روش‌هایی برای جلوگیری از تولید تصاویر واقع‌گرایانه از چهره‌ی افراد حقیقی از جمله افراد مشهور استفاده می‌کند.

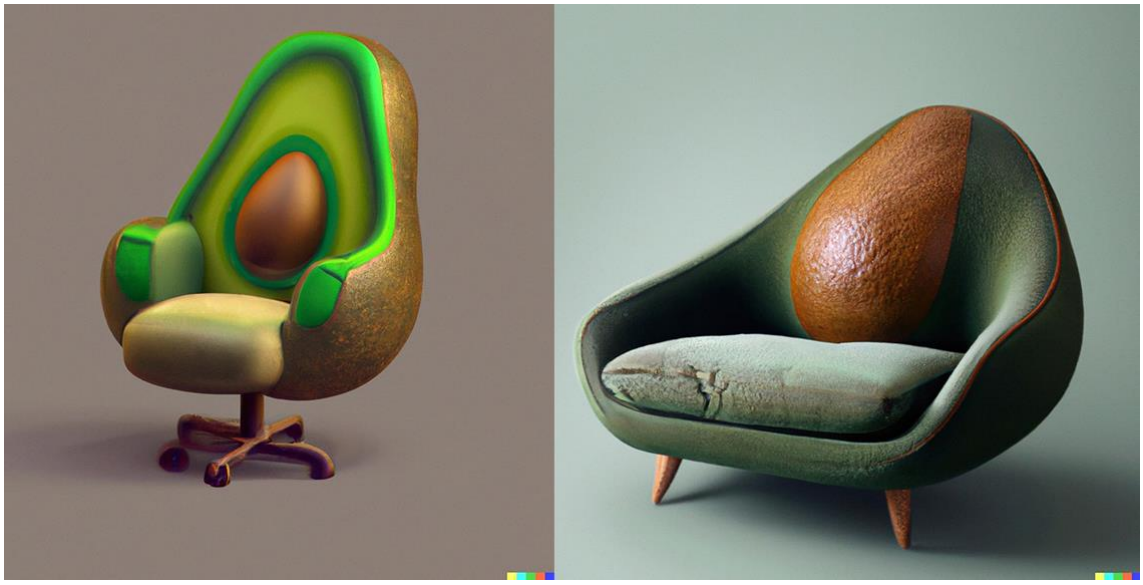
OpenAI برای جلوگیری از سواستفاده از Dall-E سیاست‌های سختگیرانه‌ای دارد

DALL-Eهمچنین به فهرستی از کلمات ممنوع مانند «تیراندازی» دسترسی دارد تا جلوی ایجاد تصاویر حاوی محتوای حساس گرفته شود. کاربران همچنین اجازه‌ی استفاده از این پلتفرم برای ایجاد تصاویری با هدف فریب، مثل دیپ‌فیک را ندارند. اگرچه می‌توان در DALL-E تصاویری براساس چهره‌ی افراد مشهور ایجاد کرد، امکان آپلود عکس بدون اجازه‌ی آن‌ها ممکن نیست و این پلتفرم از روش‌هایی مثل محو کردن چهره برای جلوگیری از تولید تصاویر واقع‌گرایانه استفاده می‌کند تا کاملاً مشخص شود تصاویر دستکاری شده‌اند و واقعی نیستند.

کار با DALL-E به‌سادگی تایپ کردن عبارت مدنظر در کادر متنی است؛ انگار که نوار جستجوی گوگل به فوتوشاپ تبدیل شده باشد. این پلتفرم حتی با الهام گرفتن از گوگل، دکمه‌ای به نام «surprise me»(غافلگیرم کن) دارد که به انتخاب خودش عبارت‌هایی را براساس تصاویر قبلی ایجاد شده، در کادر متنی وارد می‌کند. این دکمه به‌ویژه برای شرایطی کاربردی است که کاربر در جستجوی ایده‌ی جدیدی برای خلق یک اثر هنری است، اما چیزی به ذهنش نمی‌رسد.

مدت زمانی که طول می‌کشد DALL-E متناسب با عبارت تایپ شده، ۱۰ تصویر مرتبط ایجاد کند، چیزی حدود ۱۵ ثانیه است؛ البته به‌تازگی تعداد تصاویر برای هر درخواست به ۶ کاهش یافته است تا افراد بیشتری بتوانند از این پلتفرم استفاده کنند.

تصور DALL-E از عبارت «صندلی آواکادو»



یکی از تکنولوژی‌های اصلی به‌کار رفته در این پلتفرم، «diffusion» است که سال گذشته واحد هوش مصنوعی گوگل درباره‌ی آن توضیحاتی ارائه داد. به‌طور کلی، مدل‌های مبتنی بر diffusion، داده‌های وارد شده به شبکه برای آموزش هوش مصنوعی را با اضافه کردن نویز گاوسی (Gaussian noise) تخریب می‌کنند و به‌آرامی جزییات داده‌ها را پاک می‌کنند تا درنهایت تنها چیزی که از آن باقی می‌ماند، نویز خالص باشد. سپس، شبکه‌ی عصبی دیگری این فرایند تخریب داده را در جهت معکوس انجام دهد تا تمام نویزها به‌تدریج حذف شده و نمونه‌ای کاملاً عاری از نویز ایجاد شود.

تمام این توضیحات به‌کنار، چیزی که آدم را واقعا متحیر می‌کند، خلاقیت فوق‌العاده‌ی این تکنولوژی در تولید تصویر است. مثلا تصاویر زیر را در نظر بگیرید که از این عبارت‌ها ایجاد شده‌اند:

یک خرس اقتصاددان درمقابل نمودار نزولی بازار سهام، هنر دیجیتال

یک گاو نر اقتصاددان درمقابل نمودار صعودی بازار سهام، هنر دیجیتال

قدرت DALL-E در ثبت احساسات در این دو مورد واقعا شگفت‌انگیز است؛ ترس و درماندگی خرس درمقابل خشم و عصبانیت گاو. اگرچه استفاده از لفظ «خلاقانه» برای توصیف این فرایند اشتباه است، چراکه اتفاقی که واقعا می‌افتد برپایه‌ی حدس و احتمال است نه ذوق هنری. بااین‌حال، احساسی که این تصاویر در بیننده ایجاد می‌کند مثل این است که واقعا داریم به اثری هنری و خلاقانه نگاه می‌کنیم.

یکی دیگر از ویژگی‌های جالب‌توجه DALL-E توانایی‌اش به حل مسئله به روش‌های گوناگون است. برای مثال، وقتی از آن خواسته شد «یک شیرینی دارچینی خوشمزه با چشمان متحرک عروسکی» را نشان دهد، برای تصویر کردن این مدل چشم‌ها راه‌های مختلفی را امتحان کرد که یکی از آن‌ها، به‌طرز خنده‌داری شیرینی‌های دارچینی مینیاتوری بود.

شرکت OpenAI هنوز تصمیم نگرفته که روزی DALL-E را دردسترس عموم قرار دهد یا خیر. به گفته‌ی این شرکت، هدف پروژه‌ی کنونی این است که به افراد معدودی نحوه‌ی کار با این تکنولوژی را نشان دهند و در صورت نیاز، هم پلتفرم و هم سیاست‌های محتوا را براساس بازخوردی که دریافت می‌کنند، به‌روزرسانی کنند.

اگرچه DALL-E بی‌شک پیشرفته‌ترین ابزار مولد تصویری است که تا به امروز توسعه یافته، تا زمانی که هنوز به‌طور گسترده دردسترس قرار نگرفته است می‌توان خود را با نمونه‌های مشابه از جمله Dall-E Mini و چند مورد دیگر که در ادامه معرفی شده‌اند، سرگرم کرد.

پلتفرم‌های هوش مصنوعی مشابه DALL-E

علاوه‌بر Dall-E Mini که این‌روزها کاربران اینترنتی را به‌شدت به خود معتاد کرده است، نمونه‌های دیگری نیز دردسترس است که کار کم‌وبیش مشابه‌ای را انجام می‌دهند. پلتفرم‌هایی نظیر StarryAI و NightCafe بسیار شبیه DALL-E عمل می‌کنند، با این تفاوت که تمام تصاویر ایجاد شده کاملاً غیرواقعی و هنری هستند و امکان ندارد کسی آن‌ها را با عکس اشتباه بگیرد. مثلا NightCafe براساس عبارت‌های تایپ شده، تصاویر رویاگونه‌ای تولیدمی‌کند و سبک‌های مختلفی نظیر «سایبرپانک» یا «فانتزی» را دراختیار کاربر قرار می‌دهد.

AI Art Maker که به گفته‌ی خود سایت، «تخیل را به هنر تبدیل می‌کند»، شامل گزینه‌های مختلفی ازجمله انیمه، آبرنگ و همچنین تصویر واقع‌گرایانه است که برای هر درخواست، تنها یک تصویر ۲۵۶ در ۲۵۶ پیکسل را به‌طور رایگان نمایش می‌دهد، اما برای دریافت ابعاد بزرگ‌تر باید پول پرداخت کرد. علاوه‌براین، امکان تبدیل اثر هنری ایجاد شده به NFT در این پلتفرم وجود دارد.

تصاویر ایجاد شده در AI Art Maker براساس درخواست متنی «گربه بستنی به‌دست» در 4 سبک متفاوت

اپلیکیشن Dream by WOMBO که نسخه‌ی تحت وب آن نیز موجود است، شبیه NightCafe سبک‌های هنری مختلفی را دراختیار کاربر قرار می‌دهد تا نتیجه شبیه یک اثر هنری به نظر برسد. این اپلیکیشن مبتنی بر دو شبکه‌ی عصبی به‌نام‌های VQGAN و CLIP است که وظیفه‌ی اولی تولید تصاویر مشابه تصاویر دیگر است و CLIP طوری آموزش دیده تا میزان تناسب توضیح متنی با عکس را مشخص کند.

تصاویر ایجاد شده توسط Imagen گوگل

پروژه GauGAN2 انویدیا نیز که سال گذشته درباره‌ی آن صحبت کردیم، به‌کمک یادگیری عمیق می‌تواند از کلمات و عبارت‌های تایپ‌شده، تصاویری خلق کند که گاه بسیار شبیه به عبارت تایپ‌شده و گاه هنری و گاه بسیار دلهره‌آور هستند. الگوریتم GauGAN2 با ۱۰ میلیون تصویر منظره به‌کمک ابررایانه سلین انویدیا، یکی از ده اَبَرکامپیوتر پرقدرت جهان، آموزش دیده است و می‌تواند در بهترین حالت تصاویری کاملاً واقع‌گرایانه خلق کند. این پروژه اگرچه همچنان در فاز بتا قرار دارد، می‌توانید آن را به‌طور رایگان امتحان کنید.

گوگل نیز در حال کار روی نسخه‌ی مشابه Dall-E به نام Imagen است که برای خروجی تصاویر باکیفیت‌تر از مدل‌های OpenAI استفاده می‌کند. متأسفانه گوگل نسخه‌ای نمایشی از Imagen ندارد تا بتوان مانند Dall-E Mini با آن کار کرد؛ اما می‌توان چند نمونه تصویری را که با موتور Imagenساخته شده‌اند در بالا و همچنین در خود وب‌سایت رسمی این پروژه مشاهده کرد.

DALL-E؛ ترس‌ها و لبخندها

در دنیای تکنولوژی، ظهور پدیده‌ای چون DALL-E که نمایش فوق‌العاده‌ای از قدرت و پیشرفت هوش مصنوعی است، می‌تواند نقطه‌ی شروع انقلابی به بزرگی و تاثیرگذاری اینترنت و گوشی‌های هوشمند باشد. اگرچه OpenAI هنوز کاربردهای احتمالی این تکنولوژی را مشخص نکرده است، افرادی که تاکنون فرصت تجربه‌ی آن را داشته‌اند، کاربردهای جالبی برایش کشف کرده‌اند. برای مثال، هنرمندی با استفاده ازDALL-E ، فیلترهای واقعیت افزوده برای اپلیکیشن‌های شبکه‌های اجتماعی طراحی کرده است؛ یا آشپزی برای تزیین غذاهایش از DALL-E ایده می‌گیرد. بن تامپسون، تحلیلگر دنیای فناوری نیز در مقاله‌ای درباره‌ی قابلیت‌های احتمالیDALL-E ، به خلق محیط‌ها و اشیای دیجیتال در متاورس به‌روشی بی‌نهایت ارزان اشاره کرده است.

ابزارهایی چون DALL-E می‌توانند برای طراحان گرافیست مفید باشند؛ مثلا پیش از آنکه خودشان دست به کار شوند، می‌توانند از DALL-E بخواهند چند ایده‌ی مفهومی را برایشان به تصویر بکشد. این پلتفرم همچنین برای افرادی که قدرت مالی استخدام طراح ندارند، می‌تواند راهگشا باشد. شاید شما هم در کودکی دلتان می‌خواسته کتاب کمیک خودتان را بکشید، اما چون مهارت نقاشی‌تان خوب نبوده، این ایده هیچ وقت به واقعیت تبدیل نشده است.

برخی از عاشقان هوش مصنوعی نیز برای DALL-E کاربرد جالب دیگری کشف کرده‌اند؛ این افراد سراغ آثار هنری کلاسیک رفته‌اند و از هوش مصنوعی خواسته‌اند تا ادامه‌ی این آثار را نقاشی کند یا آن‌ها را در سبک‌های کاملاً متفاوتی از نو متصور شود. یکی از کاربران ردیت با DALL-E تلاش کرد تا نقاشی نیمه‌کاره جورج واشنگتن را که توسط گیلبرت استوارت در سال ۱۹۹۶ کشیده شده بود، کامل کند که نتیجه زیر به دست آمد:

SOURCE RESULT

فعلا به نظر نمی‌رسد DALL-E از آن دسته ابزاری باشد که اکثر افراد بخواهند از آن به‌طور روزمره استفاده کنند؛ اما می‌توان تصور کرد در ماه‌ها و سال‌های آینده، کاربردهای خلاقانه‌ی دیگری برای این تکنولوژی در تجارت الکترونیک، شبکه‌های اجتماعی، خانه و محیط کار کشف شود.

اغلب این چنین است که با ظهور تکنولوژی جدید، تمام تمرکز ما به جنبه‌ها و کاربردهای مثبت آن معطوف می‌شود و سواستفاده‌های احتمالی از آن در آینده را نادیده می‌گیریم؛ اما به همان اندازه که از ظهور تکنولوژی DALL-E هیجان‌زده‌ایم، جا دارد بابت استفاده‌ی نادرست از چنین ابزاری در دستان افراد و شرکت‌هایی با قوانین و خط قرمزهای کمتر از OpenAI نگران بود.

شاید شرکتی مثل OpenAI علیه سواستفاده از DALL-E سیاست‌های سختگیرانه‌ای داشته باشد، اما با ظهور ابزارهای جدید و مشابه، مانند Dall-E Mini که نظارت جدی بر محتوا نداشته باشند،

می‌توان انتظار کاربردهای مخربی داشت. همین حالا نیز برخی افراد با استفاده از تکنولوژی دیپ‌فیک به آزار و اذیت دیگران مشغولند؛ بعید نیست افرادی باشند که بخواهند از پلتفرمی مانند DALL-E برای اهدافی مخرب استفاده کنند.

استفاده از Dall-E برای بسط آثار هنری کلاسیک

از سوی دیگر، فرایند اتوماسیون همواره با خود نگرانی از بین رفتن شغل‌ها را به همراه داشته است. حالا که هوش مصنوعی می‌تواند هر چیزی را که بتوان متصور شد، برایمان نقاشی کند، دیگر چه نیازی به تصویرگران حرفه‌ای است؟ یکی از همین هنرمندان از نگرانی خود در این باره در توییتر نوشت:

من این حس بد را دارم که هنر مبتنی بر هوش مصنوعی ثبات اقتصادی حرفه‌ی تصویرگری را خواهد بلعید. نه به این خاطر که هنر قرار است به‌طور کامل با هوش مصنوعی جایگزین شود؛ بلکه به این خاطر که این مدل هنر بسیار ارزان‌تر و برای اکثر افراد و سازمان‌ها مناسب‌تر خواهد بود.

گفتن اینکه «من برای هنر فقط پیش هنرمندان واقعی می‌روم» آسان است. اما صبر کنید تا مجبور شوید بین پرداخت کارمزد ۵۰۰ دلار و صفر دلار برای سیستمی که می‌تواند تا ۹۵ درصد کارتان را راه بیندازد، انتخاب کنید.

مشکل دیگر مدل‌های هوش مصنوعی که شبکه‌های عصبی آن‌ها براساس داده‌های استخراج شده از اینترنت آموزش دیده است، بحث تبعیض و محتوای توهین‌آمیز است. چند سال پیش، گروهی از پژوهشگران MIT مجبور شدند مجموعه عظیمی از داده متشکل از ۸۰ میلیون تصویر را که برای آموزش الگوریتم‌شان استفاده می‌کردند، به‌دلیل دربرداشتن «اصطلاحات و تصاویر توهین‌آمیز» حذف کنند. در اکثر این مدل‌ها نیز اگر از کلمات مربوط به کسب‌وکار استفاده کنید، اغلب تصاویر تولید شده مردان را نشان می‌دهد که این حاکی از تبعیض علیه زنان است.

از طرف دیگر باید به جنبه‌های مثبت این تکنولوژی در مقیاس گسترده‌تر نیز فکر کرد. وقتی اکثر تصاویری که در اینترنت با آن روبه‌رو می‌شویم به‌دست هوش مصنوعی تولید شده است، چه بلایی سر درک ما از واقعیت می‌آید؟ از کجا می‌توان واقعیت را از هوش مصنوعی تمیز داد؟

در حال حاضر، DALL-E ابزار سرنوشت‌سازی در دنیای تکنولوژی مخصوص مصرف‌کننده به نظر می‌رسد. سؤال اینجا است که آیا تا چند سال دیگر نگاه ما به DALL-E همچنان به چشم انقلابی شگفت‌انگیز در حوزه‌ی هنر و خلاقیت خواهد بود یا آغازگر ماجرایی با ابعاد نگران‌کننده‌تر؟

منبع: زومیت