

علی بابا از هوش مصنوعی جدیدی به نام AtomoVideo رونمایی کرده که می تواند تصاویر را به ویدئو تبدیل کند. این فناوری در مقایسه با مدل های مشابه مانند Pika و Gen-2، نتایج بهتری ارائه می دهد. عملکرد AtomoVideo در مقایسه با رقبای خود، کاهش گذارهای عجیب و غریب بین فریم ها؛ در نمونه ای که یک فضا نورد در حال حرکت در فضا است، کلاه ایمنی او در Gen-2 ناپدید می شود، اما AtomoVideo حرکت را به صورت طبیعی تر نشان می دهد. جلوگیری از ناپدید شدن اشیاء؛ در نمونه ای دیگر، اسکی بازان در Gen-2 ناپدید می شوند، اما AtomoVideo این مشکل را ندارد.

## ۶ سایت هوش مصنوعی تبدیل متن به صدا

**cloudtts**

<https://cloudtts.com>

بر اساس منابعی که پیدا کردم، به نظر می رسد یک سرویس رایگان تبدیل متن به گفتار است که روی همه دستگاه ها از جمله موبایل کار می کند. این وبسایت امکانات زیر را ارائه می دهد: تبدیل متن به گفتار رایگان، شما می توانید متن خود را تایپ یا کپی و پیست کنید تا CloudTTS آن را با صدای طبیعی برای شما بخواند.

**pi.ai**

<https://pi.ai/onboarding>

یک هوش مصنوعی (AI) شخصی است که توسط استارت آپ Inflection AI ساخته شده است. این چت بات قابلیت های زیر را دارد: گفتگوی متنی و صوتی؛ شما می توانید هم به صورت نوشتاری و هم با صدا با Pi.ai صحبت کنید. شخصی سازی؛ Pi.ai ادعا می کند که با صحبت کردن بیشتر با آن، می تواند شخصیت شما را یاد بگیرد و بر اساس علایق و اطلاعاتی که از شما به دست می آورد، گفتگو را پیش ببرد.

**elevenlabs.io**

<https://elevenlabs.io>

تولید صدای با کیفیت بالا؛ ElevenLabs ادعا می کند که PrimeVoice AI صدایی با کیفیت بالا و بسیار طبیعی تولید می کند که از سایر نرم افزارهای تبدیل متن به گفتار (TTS) متمایز است. تنوع صدا و لهجه؛ Prime Voice AI به شما امکان دسترسی به طیف گسترده ای از صداها و لهجه ها را در زبان های مختلف می دهد تا بتوانید صدای مناسب برای پروژه خود را پیدا کنید.

**play.ht**

<https://play.ht>

یک ابزار تبدیل متن به گفتار با هوش مصنوعی است که صداهای بسیار طبیعی تولید می کند. این وبسایت به شما امکان می دهد تا متن خود را به گفتار با صدای طبیعی در ۱۴۲ زبان و لهجه مختلف تبدیل کنید، ۹۰۷ صدا برای انتخاب وجود دارد! برخی از ویژگی های کلیدی Play.ht عبارتند از: کیفیت صدای بالا؛ Play.ht ادعا می کند که با استفاده از هوش مصنوعی، گفتاری با صدای بسیار طبیعی تولید می کند.

**pollyspeak**

<https://www.pollyspeak.com>

tts از طیف وسیعی از مدل های TTS از جمله Bark، Tortoise، MusicGen، MAGNet و RVC، Vocos، Demucs، SeamlessM4T می کند. تنظیمات مختلف؛ tts-generation-webui به شما امکان می دهد تا تنظیمات مختلفی را برای هر مدل TTS تنظیم کنید تا صدای مورد نظر خود را بدست آورید.

**github.com**

<https://github.com/rsxdalv/tts-generation-webui>

PollySpeak یک ابزار مفید برای کاربران iOS به نظر می رسد که می خواهند متن را برای نیازهای مختلف به گفتار تبدیل کنند. این برنامه قابلیت هایی مانند اسکن اسناد و کارکرد بدون اینترنت را ارائه می دهد که ممکن است در همه برنامه های تبدیل متن به گفتار وجود نداشته باشد.

حرکت طبیعی تر؛ در مقایسه با Pika ۱۰۰ که حرکات غیرطبیعی و غیرقابل توضیح را نشان می دهد، AtomoVideo حرکات را به صورت طبیعی تر شبیه سازی می کند.

ویژگی های کلیدی AtomoVideo:

**حفظ دقت تصویر:** AtomoVideo دقت تصویر ورودی را در ویدئو حفظ می کند.

**انتقال روان حرکت:** این فناوری حرکت را به صورت روان و طبیعی در ویدئو شبیه سازی می کند.

پیش بینی فریم های بعدی؛ AtomoVideo می تواند فریم های بعدی ویدئو را با دقت قابل قبولی پیش بینی کند.

**سازگاری با مدل های AtomoVideo T2I:** با مدل های مختلف متن به تصویر (T2I) سازگار است.

کنترل پذیری معنایی؛ کاربران می توانند محتوای ویدئو را مطابق با سلیقه خود تغییر دهند.

**نحوه عملکرد AtomoVideo:** مدل های پیش فرض T2I به عنوان پایه استفاده می کند و آنها را با مدل های فضایی-زمانی یک بعدی و ماژول های توجه تقویت می کند. این لایه های اضافی به AtomoVideo کمک می کنند تا جزئیات و سبک های پیچیده را ثبت کند و در عین حال ثبات زمانی را در طول ویدئو حفظ کند.

**محدودیت های فعلی:** در حال حاضر، علی بابا پلتفرمی آنلاین برای کاربران ارائه نکرده است تا AtomoVideo را به صورت مستقیم تجربه کنند.

AtomoVideo علی بابا علی رغم محدودیت های فعلی، پیشرفتی قابل توجه در زمینه تبدیل تصویر به ویدئو محسوب می شود. این فناوری پتانسیل بالایی برای کاربردهای مختلف مانند انیمیشن، تولید فیلم، و بازی های ویدئویی دارد.

CREATE IMAGES FROM TEXT WITH AI

