

OpenAI'dan Gerçekçi Videolar Oluşturabilen Yapay Zekâ Uygulaması:

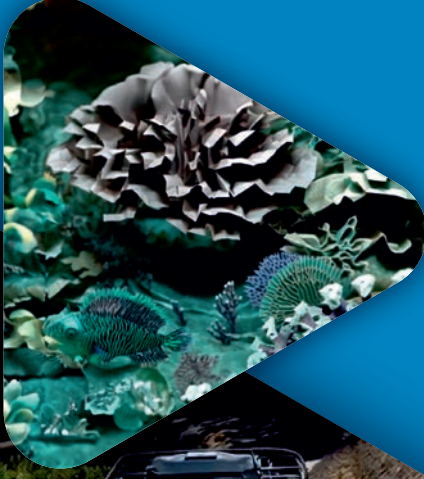
Sora

Dr. Tuba Sarıgöl [TÜBİTAK Bilim ve Teknik Dergisi



ChatGPT'yi geliřtiren OpenAI řirketi, metin formatındaki "prompt"lara yani istemlere gre gereki videolar oluřturabilen yapay zekâ uygulaması Sora'yı tanıttı.

Sora, kullanıcılar tarafından verilen yönergeleri takip ederek, bir dakika uzunluęunda videolar oluřturabiliyor. Uygulama gereki grntlerden animasyona farklı formatlarda videolar retebiliyor. Ayrıca farklı kamera aılarının bir arada kullanıldıęı karmařık grntler oluřturabiliyor. Sora, metinden video oluřturmanın yanı sıra sabit bir grntdeki geleri hareketlendirerek videoya dnřtrebiliyor.



dil modellerinde kullanılan bu yapı, kelimelerin anlamlı cümleler oluşturacak şekilde sıralanmasını sağlıyor. Sora'da ise araştırmacılar yama adı verdikleri küçük görüntü parçalarından bağlama uygun videolar oluşturmak için dönüşüm mimarisini kullandı. OpenAI uygulamayı tanıtırken,

verilen istemlere göre Sora tarafından üretilen çok sayıda video örneğini de paylaştı. Bu örneklerden birinde araştırmacılar uygulamaya şöyle bir istem verdi:

“Kamera, dik bir dağ yamacı boyunca etrafta çam ağaçları ile çevrili toprak bir yolda ilerleyen ve üzerinde siyah tavan bagajı bulunan beyaz renkli eski bir arazi aracını arkasından takip ediyor. Güneş, aracın üzerinden parlıyor. Araç ilerlerken arkasından toz kalkıyor. Kıvrılarak ilerleyen yolda başka bir araç bulunmuyor.”

Bu isteme göre Sora tarafından oluşturulan videoyu ve Sora tarafından oluşturulan diğer videoları kare kodu telefonunuza okutarak izleyebilirsiniz.



Sora'da yayılım modeli ve dönüşüm mimarisi olarak isimlendirilen yapay zekâ teknolojileri kullanılıyor. Yayılım modeli, OpenAI'nın metinden görsel üreten DALL-E uygulamasının da dayandığı model. Yayılım modelleri, rastgele piksellerin adım adım iyileştirilerek anlamlı görüntülere dönüştürülmesini sağlayacak şekilde eğitiliyor. Dönüşüm mimarisi ise sıralı verilerin işlenmesinde kullanılan bir model. OpenAI tarafından geliştirilen ChatGPT ya da Google DeepMind tarafından geliştirilen Gemini gibi büyük

Daha önce de metinden video oluşturabilen başka yapay zekâ uygulamaları tanıtılmıştı. Sora tarafından üretilen videolar ise yüksek çözünürlük, detayların netliği ve bir dakikaya ulaşabilen süresi ile önceki uygulamalardan daha üstün özelliklere sahip.

Sora, kullanıcılarının istemlerini anlamanın yanı sıra videoları oluştururken bu istemlerin fiziksel dünyada nasıl var olacağını ve aralarındaki etkileşimleri değerlendirebiliyor. Ancak uygulamanın şu an bazı zayıf yönleri var. Örneğin bazı olayların sonuçlarını doğru şekilde (bir kişi kurabiyeyi ısırdıktan sonra kurabiyedeki ısırk izi görünmemesi gibi) tahmin edemeyebiliyor. Ya da birden fazla varlığın olduğu sahnelerde insanların ve hayvanların hareketleri birbirine karışabiliyor. Ayrıca uygulama, kameranın izlediği yol boyunca gerçekleşen olayları tahmin etmekte zorlanabiliyor. Örneğin, karlı bir günde kameranın sokaktaki insanların hareketini takip ettiği örnekte videoda caddede ilerleyen araçlar yolun kenarındaki ağaç dallarının ardında kaybolduktan sonra tekrar ortaya çıkmayabiliyor.

OpenAI şirketi uygulamayı genel kullanıma açmadan önce uygulamanın neden olabileceği olası riskleri ve kötüye



kullanımları belirleyebilmek için “kırmızı takım” adı verilen ve alan uzmanlarından oluşan bir grubun değerlendirmesine sundu. Uzmanlar; yanlış bilgi, nefret içeriği ve ön yargılı içerik oluşturma gibi kritik konularda Sora’yı test ediyor. Ayrıca uygulamanın profesyoneller tarafından daha verimli bir şekilde kullanılmasını sağlayacak

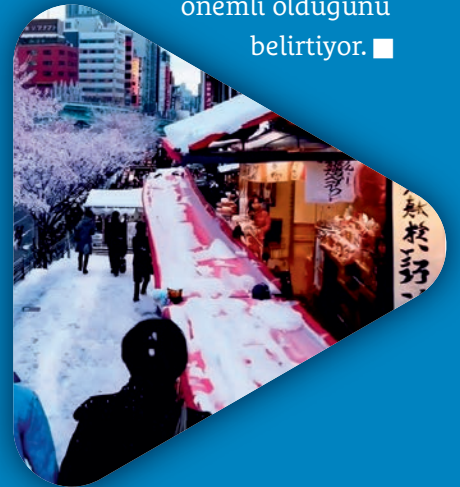
şekilde geliştirilebilmesi için bazı görsel tasarımcılara ve film yapımcılarına erişim izni verildi.

OpenAI şirketi uygulamanın olası kötüye kullanımlarının önüne geçebilmek için yanlıtıcı içeriklerin tespit edilmesini sağlayacak, örneğin bir videonun Sora tarafından ne zaman üretildiğini belirleyebilen, güvenlik araçları geliştiriyor. Ayrıca geliştirilen uygulamalar sayesinde kullanıcılar tarafından verilen istemlerin kullanım koşullarına uyup uymadığını



OpenAI kontrol edip gerektiğinde reddedebilecek. Örneğin, bir kişinin kimliğini açığa vuran, tanınmış kişilere benzeyen, aşırı şiddet içeren içeriklere yönelik istemler oluşturulamayacak.

OpenAI araştırmacıları geliştirdikleri uygulamanın olumlu ve olumsuz kullanımlarının sınırlarının belirlenebilmesi için gerçek dünyada deneyimlenmesinin önemli olduğunu belirtiyor. ■



Kaynaklar

<https://openai.com/sora>
<https://www.newscientist.com/article/2417639-realism-of-openais-sora-video-generator-raises-security-concerns/>
<https://www.technologyreview.com/2024/02/15/1088401/openai-amazing-new-generative-ai-video-model-sora/>