

Bilim insanları yapay zekayı yanılttı

Yapay zekanın bir okul otobüsünü devekuşu, kaplumbağayı tüfek, kediye köpek olarak algılayabileceği ortaya çıktı.



Deneyler, Japon bilim insanlarının yapay zekaya dayanan görüntü tanımlama sistemini yanıltması ile başladı.

UFAK DEĞİŞİKLİKLER, YAPAY ZEKAYI YANILTTI

Kyushu Üniversitesi'nden Su Jiawei ve meslektaşları, birçok fotoğrafta ufak değişiklikler yaptılar. Bu fotoğraflar daha sonra görüntü tanımlama sistemi tarafından analiz edildi.

Test edilen bütün sistemler "**derin sinir ağları**" denilen bir yapay zeka çeşidi etrafında kuruluydu.

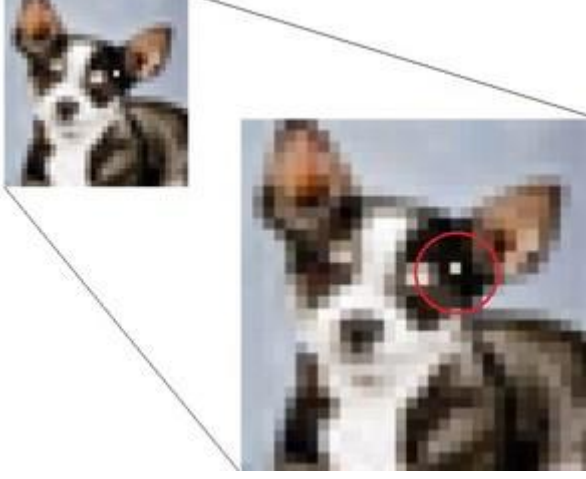
Sinir ağları temelli öğrenme sistemi, tipik olarak çok sayıda devre arasında bağ kurulmasını içeriyor-beyindeki sinir hücreleri gibi.

Bu sistemler, objelerin tipik olarak birbirinden nasıl farklılaştığını, pek çok farklı örneği karşılaştırarak öğreniyorlar.



FOTOĞRAFIN YÜZDE 74'ÜNDE BİR PİKSEL YETİYOR

Araştırmacılar, bir fotoğrafın yüzde 74'ünde bir pikseli değiştirmenin, yapay zekanın gördüğünü yanlış algılamasına neden olduğunu buldular.



HAYALET UÇAĞI KÖPEK SANAN YAPAY ZEKA

Bazı hatalar çok da büyük değildi. Örneğin kedi, köpekle karıştırıldı. Ancak bazıları oldukça şaşırtıcıydı. Örneğin hayalet uçak, köpek sanıldı.

3 BOYUTLU KAPLUMBAĞAYI TÜFEK OLARAK ALGILADI

Massachusetts Teknoloji Enstitüsü'nden (MIT) araştırmacılar deneyleri bir adım ileriye götürerek üç boyutlu baskılar kullandı.

Bu ekipte yer alan Massachusetts Teknoloji Enstitüsü'nden (MIT) Anish Athalye, dünya çapında araştırmacıların yapay zeka temelli görüntü tanımlama sisteminin zayıflığını ortaya koymak için deneyler yaptıklarını söylüyor.

Athalye ve iş arkadaşlarının yaptığı deneylerde üç boyutlu bir kaplumbağa, tüfek olarak algıladı.



"YAPAY ZEKAYA YANILTICI SALDIRILAR MÜMKÜN"

Athalye [BBC](#)'ye yaptığı açıklamada *"Giderek artan sayıda gerçek dünyada kullanılan sistemler sinir ağlarını kullanmaya başlıyorlar ve bu sistemlere yanıltıcı örneklerle saldırı düzenlemek mümkün"* diyor.

Şimdiye kadar böyle bir olay gerçekleşme de, Athalye bu akıllı sistemlerin çok kolay bir şekilde kandırılabilmelerinin endişe verici olduğunu söylüyor.

İNTERNET DEVLERİ, SORUNU ÇÖZMEYE ÇALIŞIYOR

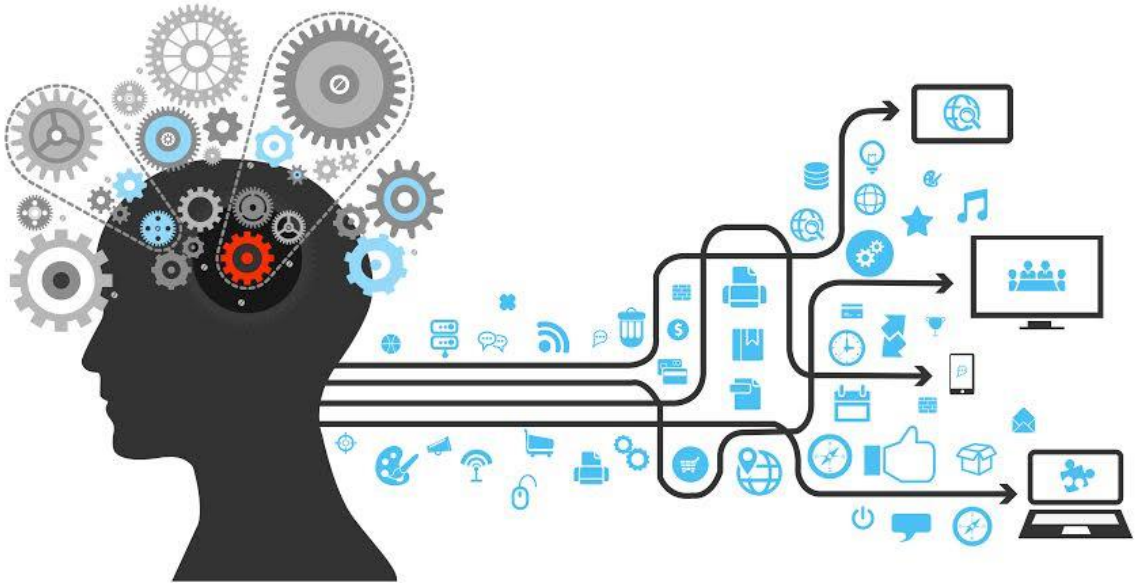
Facebook, Amazon ve [Google](#) gibi internet devlerinin bu tür çarpıtmalarla baş etme yöntemleri üzerinde çalıştıkları biliniyor.



"ÖNERİLEN TEKNİKLERİN HİÇBİRİ İŞE YARAMIYOR"

Athalye, derin sinir ağlarını bu tür yanıltmalardan korumanın zor olduğunu söylüyor ve devam ediyor:

"Bu çözülememiş bir sorun. Önerilen çok fazla sayıda teknik oldu ve hemen hepsi işe yaramaz çıktı."



"GARİP VE İLGİNÇ BİR ŞEYLER DÖNÜYOR"

Araştırmacı, öğrenme süreci sırasında bu yanıltıcı örneklerin gösterilmesinin işe yarayabileceğini düşünüyor, ancak bunun bile bu araştırmayla ortaya konan bütün sorunlara nihai bir çözüm getirmeyeceğini kaydediyor.

"Burada kesinlikle garip ve ilginç bir şeyler dönüyor. Biz sadece henüz bunun ne olduğunu bilmiyoruz."

