

Bilgisayarların G z nden

Yapay zekâ alanında çalışan bilim insanlarının temel amacı insan beyninin kapasitesine ulaşabilen bilgisayarlar geliştirmek. Bunu başarabilmek için bilgisayarların algılayabilmesi, öğrenebilmesi ve doğru kararlar verebilmesi gerekiyor. İnsan beyninin çok hızlı ve doğru kullanabildiği bu yeteneklerden biri de nesneleri algılayıp tanımlayabilmek.

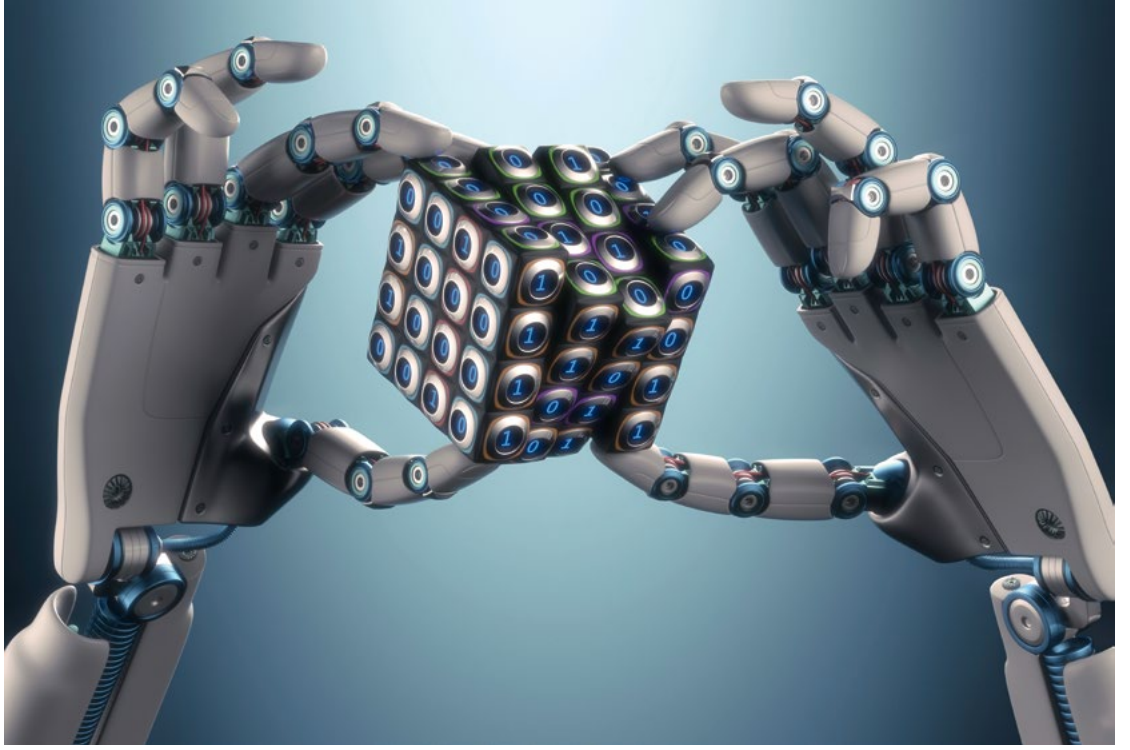
Fotoğraflarımızı sınıflandıran, sorularımızı cevaplayan, bize film öneren yapay zekâ uygulamalarını akıllı telefonlardan biliyoruz. Bu yazılımların başarısı bilgisayarların öğrenme yeteneklerine bağlı.

Nesnelerin görsel olarak algılanabilmesi ve tanımlanabilmesi hayli karmaşık bir problem. Çünkü bir nesnenin konumuna göre sonsuz sayıda iki boyutlu görüntüsü olabilir. Bunun yanı sıra başka birçok faktör, örneğin ışık ve arka plandaki görüntüler de nesnenin nasıl algılandığını etkiler. Beynimiz bu problemi kolayca çözebiliyor. Ayrıca insanlar bir görüntüyü algıladıklarında sadece o görüntüdeki nesnelerin ne olduğunu fark etmekle kalmıyorlar. Örneğin aile üyelerinizin bir arada yer aldığı bir fotoğrafı ele alalım. O fotoğrafa baktığınızda fotoğrafta sadece insanlar olduğunu görmez aynı zamanda kaç kişi olduğunu belirleyebilir ve insanların yüz ifadelerinden o anki duygularını tahmin edebilirsiniz.

Nesnelerin bilgisayarlar tarafından tanınması iki aşamada gerçekleşiyor. Birinci aşamada görüntüdeki nesneler algılanıyor. İkinci aşamada ise bu nesneler sınıflandırılıyor. Yani algılanan nesnenin bir otomobil mi, kedi mi ya da sandalye mi olduğu belirleniyor.

Bilgisayarlar görmeyi yani bir görüntüdeki nesnelerin bilgisayarlar tarafından algılanması ve tanınmasının önündeki en önemli engel nesnelerin konumunun, nesnenin şeklini belirleyen dış hatların belirgin olup olmamasının, aydınlatma etkisiyle oluşan yansımaların ve gölgelerin bir nesnenin farklı görüntülerinin elde edilmesine neden olması. Ayrıca çevrelerindeki nesneler ve arka plandaki görüntüler de nesnelerin algılanmasını ve tanınmasını zorlaştırıyor. Bilim insanları bu problemi nesneleri bilgisayarlara tanıtarak çözebileceklerini düşünüyor. Yani bilgisayarların insan beyni gibi öğrenebilmesi gerekiyor.

İnsan beyninde yaklaşık 100 milyar sinir hücresi var ve bu hücrelerin birbirleri ile olan etkileşimleri bilişsel etkinliklerin gerçekleştirilmesini sağlıyor. Bilgisayarların yüzleri tanıyabilmesini, konuşulanları anlayabilmesini sağlayan yaklaşımın temelinde de yapay sinir ağlarının oluşturulması fikri var. Yapay sinir ağlarının toplanan verileri insan beyni gibi işlemesi amaçlanıyor. İnsan beyni retinadan alınan görsel verileri beynin farklı bölümlerinde değerlendirilerek algılıyor. Yapay sinir ağları yönteminde de görsel veriler binlerce bilgisayar tarafından değerlendirilerek nesnelerin en doğru şekilde tanınmasına çalışılıyor.



Bilgisayarla görme yöntemlerinde temel olarak, analiz edilecek görüntüden elde edilen veriler değerlendirilerek her nesneyi tanımlayan farklı sembollerin üretilmesi amaçlanıyor. Örneğin bir köpek fotoğrafı “hayvan”, “köpek”, “evcil köpek” şeklinde etiketlenebilir. Bu süreci sosyal medya platformlarında fotoğrafların etiketlenmesine benzetebiliriz. Nesnelerle ilişkili olan bu sembollerin üretilmesi ve sınıflandırılması bilgisayarlar tarafından tanınmalarındaki en önemli aşamalardan biri. Çünkü nesnenin ham görüntüsünden elde edilen verilere göre belirlenen sembollerin nesneyi ne kadar doğru tanımladığı geliştirilen yöntemin başarısını yansıtır. Örneğin yapay sinir ağı sistemleri başlangıçta nesneleri tanımakta yeterince başarılı değildi. Ancak sisteme girilen veri sayısı arttıkça bilgisayarların nesneleri birbirinden ayırma ve doğru bir şekilde tanıma oranı yükseldi.

Nesneleri tanımlayan sembollerin sınıflandırılması bilgisayarlar tarafından gerçekleştirilse de bu sembollerin belirlenme işlemi genellikle insanlar tarafından gerçekleştiriliyor. Bu sürecin her nesne için gerçekleştirilmesi gerektiği düşünüldüğünde bunun uygulanabilir bir yöntem olmadığı anlaşılabılır. Brigham Young Üniversitesi'nden bilim insanları *Pattern Recognition* dergisinde yayımlanan araştırmalarında nesnelerin etiketlenme basamağının bilgisayarlar tarafından başarılı bir şekilde gerçekleştirilmesini sağlayan bir algoritma geliştirdi.

Araştırmacılar Dah-Jye Lee bilgisayarların nesneler arasındaki farkları belirlemesini sağlamak-tansa, her nesnenin kendine has özelliklerini bilgisayarlara öğretmenin yararlı olabileceğini söylüyor. Bu, bir çocuğa kedi ve köpek arasındaki farkı öğretmek için kedilere ve köpeklere has özellikleri kavratmaya benzetilebilir. Araştırmacılar geliştirdikleri algoritmayı dört farklı veri grubu (motosiklet, yüz, uçak ve otomobil) kullanarak sınavı ve bu yöntemle her veri grubu için %100 doğru tanımlama yapılabildiği anlaşıldı.

Aslında bir görüntüdeki nesnelerin algılanması, nesnenin ne olduğunun belirlenmesinden daha zor bir süreç. Özellikle de arka planda karmaşık görüntüler olduğu ve nesnelerin arka plan görüntüsünden zor ayırt edilebildiği durumlarda. Bilim insanları bu sorunun üstesinden gelmek için genellikle görüntüyü çok küçük bölümlere ayırıyor. Piksel gruplarından oluşan bu bölümlerin renk, doku, parlaklık gibi özellikleri karşılaştırılarak görüntüde yer alan nesnelerin hatları belirlenebiliyor. Ancak günümüzde kullanılan algoritmaların görüntüdeki nesneleri algılama hızı çok yavaş.

Görüntülerdeki nesnelerin analiz edilmesi için kullanılan geleneksel yöntemlerde, özgün boyutu ne olursa olsun, görüntünün analiz edilen bölümünün boyutunun sabit olması gerekiyor. Bu amaçla görüntü kesiliyor ya da genişletilerek ve uzatılarak görüntünün özgün şekli bozuluyor. Bu durum nesnelerin

doğru şekilde algılanmasını engelleyebiliyor. Microsoft araştırmacılarının nesnelerin doğru ve hızlı bir şekilde algılanabilmesini sağlamak için geliştirdiği yazılım, boyutu ne olursa olsun her görüntüyü eşit sayıda bölüme ayırarak analiz ediyor. Örneğin bir fotoğrafı kare şeklinde küçük kâğıt parçalarıyla kapladığınızı düşünün. Farklı boyuttaki başka bir fotoğrafı bir önceki fotoğrafla aynı sayıda kâğıt parçasıyla kaplayabilmek için kullandığınız kâğıt parçalarının boyutunun farklı olması gerekir. Araştırmacılar geliştirdikleri yöntemde analiz edilen görüntü boyutu ile ilgili bir sınırlama olmamasının, nesnelerin bilgisayarlar tarafından algılanma hızını önceki yöntemlere göre 20-100 kat arasında artırdığını belirledi.

Bilgisayarla görme günümüzde birbirinden çok farklı alanlarda yararlı olabilecek bir teknoloji. Temel amacı daha akıllı bilgisayarlar yani ileri yapay zekâ uygulamaları geliştirmek olan bu teknoloji sürücüsüz otomobillerden robot teknolojisine, tıbbi görüntüleme alanındaki araştırmalara kadar pek çok alanda önemli katkılar sağlıyor. Örneğin genetik, sinirbilim, ekoloji gibi birçok alanda bilim insanları, uzaktan görüntüleme yöntemleri ile elde edilen görüntüleri analiz edilerek önemli bilgilere ulaşabiliyor. Bu yöntem aynı zamanda kendi yaşam alanlarında fark edilmeleri zor olan canlı türlerinin tespit edilmesine ve uzun süreyle izlenmelerine imkân sağlıyor.

Gelecekte otomobillerin sürücüyü ihtiyaç duymayan otonom araçlar olacağı düşünülüyor. Henüz yollarda olmasalar da sürücüsüz otomobillerin ilk örneklerinin yakın gelecekte kullanılmaya başlanacağı düşünülüyor. Sürücüsüz otomobillerde sensörler ve kameralar sayesinde elde edilen veriler analiz edilerek aracın bu verilere göre hareket etmesi sağlanıyor. Dolayısıyla aracın çevresindeki diğer otomobilleri, yayaları, engelleri, yol çizgilerini, trafik ışıklarını ve trafik işaretlerini algılaması ve tanımlaması bu teknolojinin temelini oluşturuyor.



Robot teknolojilerinin karşı karşıya olduğu önemli sorunlardan biri robotların karmaşık ortamlarda da çalışabilmesi olduğu için, nesneleri algılayan sistemler robot teknolojilerinin de vazgeçilmez bir parçası.

Doğadan ilham alan teknolojilerden biri olan bilgisayarla görme teknolojisi, sağlanan ilerlemelere rağmen henüz emekleme aşamasında olsa da gelecekte hayatımızın önemli bir parçası olacak.



Kaynaklar

- Jones, N., "Computer science: The learning machines", *Nature*, Cilt 505, Sayı 7482, s. 146-148, 2014.
- Pinto, N., Cox, D. D., DiCarlo, J. J., "Why is Real-World Visual Object Recognition Hard?", *PLOS Computational Biology*, Cilt 4, Sayı 1, s. 151-156, 2008.
- Lillywhite, K. ve ark., "A feature construction method for general object recognition", *Pattern Recognition*, Cilt 46, Sayı 12, s. 3300-3314, 2013.
- <http://news.office.mit.edu/2014/computer-neural-networks-identify-objects-like-primate-brain-1218>
- <http://research.microsoft.com/en-us/news/features/spp-102914.aspx>
- He, K. ve ark., "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition", *Computing in the 21st Century Conference*, 2014. <http://arxiv.org/pdf/1406.4729v1.pdf>

