Derin Öğrenme ile Görüntü Kümeleme

Ömer Faruk Akmeşe[1*], Hasan Erbay[1], Hakan Kör[1]

1 Hitit University, **2** Ankara Yıldırım Beyazıt University, * Corresponding author, ofarukakmese@hitit.edu.tr

Özet

Elde ettiğimiz dijital veriler her zaman tablo biçiminde değildir. Veriler, video ve ses kayıtları, düz metinler ve resimlerden oluşabilir. Farklı formlardaki veriler üzerinde veri madenciliği işlemleri gerçekleştirmek için veriler dönüştürülür. Makine öğrenme yöntemleri kullanılarak görüntüleri tanımada, sınıflandırmada ve kümelemede başarılı sonuçlar elde edilmektedir. Bunun için bilgisayara çok sayıda nesnenin görüntüsü verilmeli ve öğrenme sağlanmalıdır. Böylece bilgisayar kendi bilişini kurmaya başlayacaktır. Buna denetimli öğrenme denmektedir ve çok sayıda veri gerekmektedir. Günümüzde, paylaşım veya depolama için bulut tabanlı sistemlere çok sayıda görüntü yüklenmiştir. Bu çalışmada eldeki görüntülerin kümelenmesi için derin öğrenme yöntemi kullanılmıştır. Teknik olarak, görüntü Google sunucularına gönderilir ve burada önceden eğitilmiş derin bir sinir ağından geçirilir. Derin öğrenmede yapay sinir ağları kullanılır ve görüntü kümelemede doğru sonuçlar elde etmek açısından mükemmel sonuçlar vermektedir. İnsan beyninin çalışma prensibini taklit eden yapay sinir ağları ile görüntü tanıma, görüntüleri sınıflandırma ve kümeleme işlemleri bilgisayar tarafından kolayca yapılabilmektedir.

Anahtar kelimeler: Derin öğrenme, Görüntü kümeleme.

Title

Image Clustering by Deep Learning

Abstract

The digital data we obtain is not always in the form of a table. The data may be composed of video and audio recordings, plain texts and pictures. Data is converted to perform data mining operations on data in different forms. Successful results are obtained in recognition, classification and clustering of images using machine learning methods. For this, a large number of object images should be given to the computer and learning should be provided. Thus, the computer will begin to establish its own cognition. This is called supervised learning and requires a lot of data. Today, a large number of images have been uploaded to cloud-based systems for sharing or storage. In this study, deep learning method was used for the clustering of the available images. Technically, the image is sent to Google servers and passed through a deep neural network earlier trained here. In deep learning, artificial neural networks are used and give excellent results in terms of accurate results in image aggregation. Image recognition by artificial neural networks mimicking the working principle of human brain, image classification and clustering can be done easily with computer.

Keywords: Deep learning, Image clustering.

Citation: Akmeşe, Ö. F., Erbay, H., Kör, H. (2018, October) *Derin Öğrenme ile Görüntü Kümeleme*. Paper presented at the Fifth International Management Information Systems Conference.

Editor: H. Kemal İlter, Ankara Yıldırım Beyazıt University, Turkey

Received: August 19, 2018, **Accepted:** October 18, 2018, **Published:** November 10, 2018

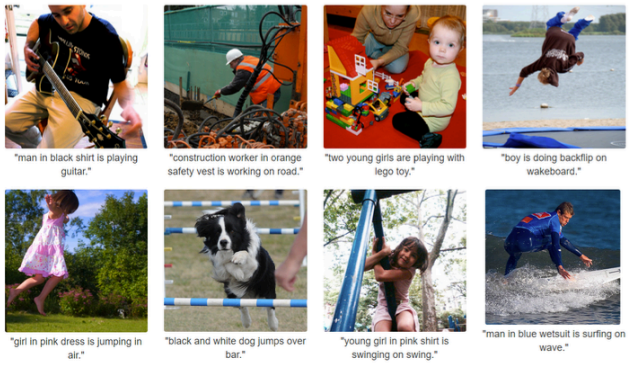
Copyright: © 2018 IMISC Akmeşe et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Giriş

Bir makine öğrenmesi sınıfı olan derin öğrenme, özellik çıkarma ve dönüştürme için birçok doğrusal olmayan işlem birimi katmanını kullanmaktadır. Her ardışık katman, önceki katmandaki çıktıyı girdi olarak alır (Deng & Yu, 2013). Önemli derecede yatırım ve araştırmanın yapıldığı bir alan haline gelen derin öğrenme, büyük miktarlardaki eğitim verilerinden özellik saptama yapabilen sistemler oluşturmak için ileri bir teknoloji olan derin sinir ağlarını kullanmaktadır (Kurt, 2015). Yapay zekâ alanlarından biri olan derin öğrenme, aynı zamanda makine öğrenme tekniklerinden birisidir. Günümüzde kamera sayılarının artmasıyla beraber dijital veride çok büyük miktarda bir birikim olduğu söylenebilir. Nesnelerin görüntülerinin bilgisayar tarafından anlamlı hale getirilmesi büyük kolaylıklar sağlayacaktır.

Video akışındaki her kare veya herhangi bir resim karesi için çeşitli filtrelemeler ve alt sınıflara ayırma gibi işlemler neticesinde resim karesi üzerinde tespit edilen nesneler önceden eğitilmiş ağa sokularak sınıflandırılmaktadır. Ayrıca derin öğrenme, yüz tanıma,

ses tanıma, parmak izi okuyucular, plaka okuyucular, sürücüsüz arabalar ve güvenlik kameraları gibi alanlarda da kullanılmaktadır.



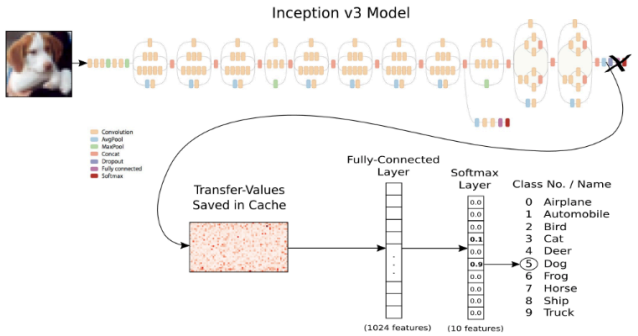
Şekil 1. Görüntüler için derin öğrenme ile elde edilen açıklamalar (Karpathy & Fei-Fei, 2015).

İnsanlar bir hayvan görüntüsü tanımada, işaretlerin okumasında veya bir insanın yüzünün tanımasında bir çaba göstermezler. Beynimiz görüntüleri kolay bir şekilde anlayabilirken bilgisayarlar için çözülmesi zor problemlerdir. Son birkaç yılda, bu zor problemleri ele alma konusunda muazzam bir ilerleme kaydedilmiştir. Özellikle, derin bir konvolüsyonel nöral ağ olarak adlandırılan bir tür model ile görsel tanıma görevleri için makul seviyede bir performans sağlanmıştır. Konvolüsyonel bir sinir ağı olarak adlandırılan özel bir tür nöral ağ, temelde, aynı nöronun birçok özdeş kopyasını kullanan bir tür sinir ağı olarak düşünülebilir. Aynı nöronun birden fazla kopyasını almak fonksiyonların soyutlanmasına kabaca benzemektedir. Programlama yaparken, bir fonksiyon bir kez yazılıp birçok yerde kullanılır ve aynı kod farklı yerlerde defalarca yazılmadan programlama hızlandırılır ve daha az hatayla sonuçlanır. Benzer şekilde, bir konvolüsyonel nöral ağ bir nöronu bir kez öğrenebilir ve birçok yerde kullanabilir, bu da modeli öğrenmeyi ve hatayı azaltmayı kolaylaştırır. Bu çalışmada derin öğrenme yöntemiyle görüntülerin kümelenmesi amaçlanmaktadır. Teknik olarak, sunucuya gönderilen görüntü, Google'ın Inception v3'ü gibi önceden eğitilmiş derin bir sinir ağından geçirilir. Çoğunlukla bazı özel amaçlarla eğitilen derin ağlar görüntüleri 1000 adet resim sınıfından herhangi birine sınıflandırabilmektedir. Bu çalışma kapsamında elimizdeki görüntülerin derin sinir ağlarından geçirilerek kümelenmesi üzerinde durulmuştur.

Yöntem

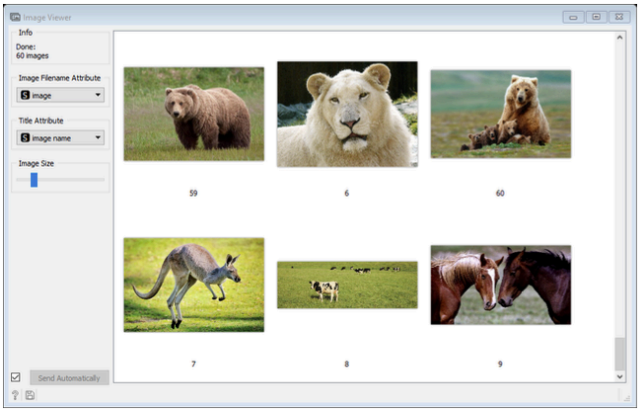
Günümüzde, paylaşım veya depolama için bulut tabanlı sistemlere çok sayıda görüntü yüklenmiştir. Bu büyük ölçekli görüntü verilerinin nasıl verimli bir şekilde düzenleneceği konusu da ortaya çıkan ve zorlayıcı olan diğer bir konudur (Hsu & Lin, 2017). Kümeleme, makine seçiminde özellik seçimi açısından kapsamlı bir şekilde çalışılmıştır (Boutsidis 2009).

Konvolüsyonel sinir ağları, biyolojik görüntü işleme sistemlerinden esinlenen bir tür sinir ağı mimarisidir. "Öğrenme" adı verilen bir süreçte ayarlanan birçok parametreli çok uzun operatör zincirlerinden oluşurlar. Tipik olarak tüm bir konvolüsyonel nöral ağın eğitimi haftalar sürebilir ve çok fazla hesaplama gücü gerektirir. Fakat farklı bir yol olan görüntüleri tanımak için önceden eğitilmiş bir ağ, farklı görüntü kümelerini tanımak üzere değiştirilebilir. Eğitilmiş olan InceptionV3 ağını kullanabiliriz. ImageNet, Inception-v3, 2012 verileri kullanılarak görsellerin bilgisayarlar tarafından tanınması için eğitilmiştir. ImageNet, hiyerarşinin her bir düğümünün yüzlerce ve binlerce resim tarafından tasvir edildiği düzenlenmiş bir görüntü veritabanıdır.



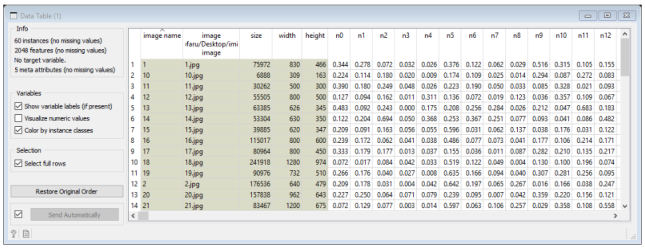
Şekil 2. Yeni bir sınıflandırıcıyı eğitmek için Inception-v3 ağı.

Bu çalışmada eldeki görüntülerin kümelenmesi için derin öğrenme yöntemi kullanılmıştır. Teknik olarak, görüntü Google sunucularına gönderilir ve burada önceden özel amaçlarla eğitilmiş derin bir sinir ağından geçirilir. Çalışmada, farklı hayvanlara ait 60 adet fotoğrafın bulunduğu bir görüntü verisi ile Orange programı kullanılarak fotoğraflarda aynı türden hayvanların hiyerarşik olarak kümelenmesi sağlanmıştır.



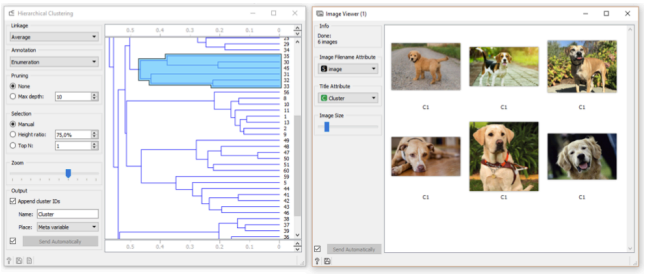
Şekil 3. Çalışmada kullanılan fotoğraflardan kesit.

Görüntülerin okunması ve program tarafından Google sunucularına yüklenmesiyle her görüntü için bir özellik vektörü hesaplanır ve bir derin öğrenme modeli kullanılır. Görüntüler, sayısal gösterimi elde etmek için vektörlerin temsillerinin hesaplandığı bir sunucuya gönderilir. Burada her bir görsel için ek sütunlarla 2048 ek özellik çıkarımı yapılarak gelişmiş veri tablosu haline döndürülür.



Şekil 4. Sunucudan alınan görüntü vektörleri.

Gönderilen resimler hiçbir yerde saklanmaz. Görüntü verilerinde en iyi sonucu veren kosinüs benzerliği yöntemiyle veri kümesindeki satır/sütun arasındaki mesafeler hesaplanır. Basitçe bu benzerlik, trigonometrideki kosinüs fonksiyonu üzerinden formülize edilmesi ve iki vektörün birbirine göre olan ilişkisi bir açı ile ifade edilmesine dayanır. Örneğin aynı yönü gösteren, yani birbiri ile aynı olan iki vektör için kosinüs değeri 1 olurken birbirine dik olan vektörler için kosinüs değeri 0, birbirinin zıddı olan vektörler için ise bu değer -1 olacaktır. Diğer bütün değerler, +1 ile -1 arasında yer alacaktır. Dolayısıyla kosinüs benzerliği ile benzerlik bu yöntemle sayısallaştırılmaktadır. Daha sonra görüntüler, hiyerarşik kümeleme analizi yapılarak gruplara ayrılır. Hiyerarşik kümeleme küme oluşturma için esnek bir yapıdır. Bu yöntemde küme sayısının önceden belirlenmesine ihtiyaç yoktur. Kümeleme işlemi sonrasında oluşturulan dendrogram adı verilen bir ağaç yapısı şekil 5'te görülmektedir.



Şekil 5. Hiyerarşik kümeleme analizi.

Sonuç

Son yıllarda büyük çıkış yapan derin öğrenme, sürücüsüz araçlardan cep telefonlarına, tıbbi uygulamalardan duygu analizlerine kadar çok geniş bir yelpazede kendisine kullanım alanı bulmuştur. Günümüzde depolanan veri miktarının sürekli artması görüntü işleme alanında özellikle görüntü sınıflandırma, kümeleme nesne tanıma gibi problemlerin ortaya çıkmasına sebep olmaktadır. Dijital verilerde çok büyük miktarlarda birikim olduğu, bu verileri analiz etmek için uzman sayısının yetersiz olduğu ve yapılan analizlerin insan yeteneklerini aştığı göz önüne alındığında dijital verilerin bilgisayar tarafından anlamlı hale getirilmesi büyük kolaylıklar sağlamaktadır. Bu konuda kendisini ispatlayan derin öğrenme metotları sıklıkla kullanılmakta ve en başarılı sonuçları vermektedir. Sonuç olarak; Bu çalışmada derin öğrenme ile görüntülerin hiyerarşik olarak kümelenmesi sağlanmıştır. Hiyerarşinin her bir düğümünün yüzlerce ve binlerce resim tarafından tasvir edildiği düzenlenmiş bir görüntü veritabanı olan ImageNet kullanılmıştır. Yapılan çalışma sonucunda veri setindeki 60 adet hayvanlara ait görselin her birinin başarılı bir şekilde kümelendiği görülmektedir. İşlem sonrasında oluşturulan dendrograma bakıldığında ayrı tür hayvanların görsellerinin ve bunlara ait alt türlerin doğru bir şekilde kümelendiği görülmektedir.

Kaynakça

Boutsidis, C., Drineas, P., & Mahoney, M. W. (2009). Unsupervised feature selection for the $ k $-means clustering problem. In Advances in Neural Information Processing Systems (pp. 153-161).

Deng, L., & Yu, D. Deep learning: Methods and applications. Found. Trends Signal Process. 7 (34), 197–387 (2013).

Karpathy, A., & Fei-Fei, L. (2015). Deep visual-semantic alignments for generating image descriptions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3128-3137).

Hsu, C. C., & Lin, C. W. (2017, September). Unsupervised convolutional neural networks for large-scale image clustering. In Image Processing (ICIP), 2017 IEEE International Conference on (pp. 390-394). IEEE.

Kurt, F. (21.07.2015). Derin öğrenme (deep learning) nedir?. Erişim tarihi: 29.07.2018, http://www.derinogrenme.com/2015/07/21/derin-ogrenme-deep-learning-nedir/