DERİN ÖĞRENME YÖNTEMLERİ İLE YÜZ GÖRÜNTÜLERİNDEN YAŞ KESTİRİMİ

Yusuf Çağrı DAŞKIN¹, Füsun BALIK ŞANLI², Taşkın ÖZKAN³, Hakan KARTAL⁴

¹Müh., AKSA Enerji, 34810, İstanbul, cagri.daskin@aksa.com.tr

²Doç. Dr., Yıldız Teknik Üniversitesi, Harita Mühendisliği Bölümü, 34220, İstanbul, fbalik@yildiz.edu.tr

³Yük. Müh, NetcadYazılım A.Ş., 34742, İstanbul, taskin.ozkan@netcad.com.tr

⁴Yük. Müh, UHUZAM, 34396, İstanbul, hakan@cscrs.itu.edu.tr

ÖZET

Yapay zekâ çalışmalarında kaydedilen ilerlemeler ses, görüntü ve yazı işleme konularında başarılı sonuçlar elde edilebilmesine olanak sağlamıştır. Yapay sinir ağlarının son hâli olan derin yapay sinir ağları görüntü işlemede robot görmesi, görüntü etiketleme, görüntü analiz etme ve sınıflandırma konularında kullanılmaya başlanmıştır. Askerî amaçlar ve güvenlik konularında görüntü ve video işleme tekniklerinde doğru kurgulanmış derin yapay sinir ağları mimarileri ile desteklenen sistemlerde doğruluk ve güvenilirliğin arttığı gözlenmiştir.

Bu çalışmada, yaş kestirimi için yüz görüntülerinden oluşan bir veri seti ile derin yapay sinir ağı eğitilerek farklı iterasyonlardaki sonuçlara göre doğruluk analizi yapılmış ve değişen parametrelere göre sonuç doğruluğunun ne yönde etkilendiği gösterilmiştir.

Anahtar Sözcükler: derin yapay sinir ağları, yaş ve cinsiyet kestirimi

ABSTRACT

AGE ESTIMATION FROM FACIAL IMAGES USING DEEP NEURAL NETWORKS

Recent improvements in artificial intelligence studies have enabled people to get successful results in audio, image and text processing. Deep artificial neural networks, which is the latest form of artificial neural networks, are used in all image processing, robot vision, image labeling, image analysis and image classification. It is observed that the accuracy and the reliability are enhanced in systems supported by deep artificial neural network architectures that have been correctly designed for image and video processing techniques in military purposes and security issues.

In this study, for the age estimation a data set consisting of face images was used to train the deep artificial neural network with different parameters, and accuracy analysis was performed and it was shown how the result accuracy is affected according to the changing parameters.

Keywords: deepneural networks, age and gender estimation

1. GİRİŞ

1.1. Derin Yapay Sinir Ağı

Görüntü ve video işleme teknikleri, askerî amaçlardan robot görmesine, güvenlik konularından uzaktan algılamaya pek çok alanda kullanılmaktadır. Güvenlik ihtiyaçları, üretim bantlarının otomasyonu, robot ve makine görmesi gibi konularda artan doğruluk ihtiyacı piksel bazlı klasik yöntemlerden, yapay zekâ bazlı yöntemlere geçilmesi ihtiyacını doğurmuştur.

Derin yapay sinir ağları son yıllarda görüntü ve video işleme alanlarında en sık kullanılan modellerdir. Sistemin örnek veri seti ile eğitilmesi ve kararları bu eğitimin sonucuna göre vermesi esasına dayanır. Çalışma prensibi olarak klasik sinir ağlarından bir farkları yoktur. Derinlik kavramı, ağı oluşturan tam bağlantılı gizli katman sayısının fazlalığından gelir (Şekil 1). Gizli katman sayısı ne kadar fazlaysa, ağ o kadar derin olarak nitelendirilir.

Şekil 1. a) Klasik yapay sinir ağı b) Derin yapay sinir ağı.

Bir ağı eğitilmesi, örnek veri setindeki her bir elemanın ağı iteratif olarak sokulması ile yapılır. Her bir iterasyonda ağı sinir uçlarındaki ağırlıklar geri beslemeli olarak tekrar düzenlenerek ağı çıktısının belli bir eşik değerin altında veya üstünde kalması amaçlanır.

Probleme yönelik olarak tasarlanan birkaç farklı derin yapay sinir ağı mimarisi vardır. Girdi katmanının 2 boyutlu matris olduğu durumlarda konvolüsyonel derin yapay sinir ağları tercih edilmektedir. Bu çalışmada da konvolüsyonel derin yapay sinir ağları ile çalışılmıştır.

1.2. Derin Yapay Sinir Ağı Katmanları

Konvolüsyonel derin yapay sinir ağları, öğrenme işlemini bir dizi sıralı işlemi tekrarlı şekilde yaparak gerçekleştirir. Her bir aşamında farklı girdilere farklı işlemler uygulanan bu işlemlere ağı katmanları denir. Konvolüsyonel derin yapay sinir ağlarının katmanları girdi katmanı, konvolüsyon katmanı, havuzlama katmanı, gizli katmanlar ve çıktı katmanı olarak sayılabilir (Şekil 2). Derin yapay sinir ağının yapısı, ağı tasarlayan kişi tarafından bu katmanların sayısı ve bazı katmanların sıralamadaki yerinin belirlenmesiyle belirlenir. Görüntü işleme konularında başarısı kabul edilmiş ve yaygın olarak kullanılan bazı derin yapay sinir ağı modelleri vardır. Bunlar AlexNet, ZF Net, GoogleNet, ResNet ve VGG Net olarak sıralanabilir. Bu çalışmada VGG-16 modeli kullanılmıştır.

Şekil 2. Derin yapay sinir ağı katmanları.

1.2.1. Girdi Katman

Bu katmanda derin yapay sinir ağına, üzerinde eğitim yapılacak veri setinden örnekler girdi olarak verilir. Bu çalışmadaki girdi verileri, kişilerin yüz görüntüleridir.

1.2.2. Konvolüsyon Katmanı

Konvolüsyonel derin yapay sinir ağını klasik yapay sinir ağından ayıran bir diğer özellik, ağı denetimsiz olmasıdır. Klasik yapay sinir ağları ile görüntü sınıflandırma problemlerinde girdi görüntü, ağı tasarımcısı tarafından belirlenen bazı filtrelerden geçirilerek görüntü üzerinden bazı özellikler çıkarılır. Sinir ağına eğitimi ham görüntü üzerinde değil, özellik verileri üzerinde yapılır. Konvolüsyonel derin yapay sinir ağlarında ise girdi katmana uygulanacak özellik çıkarımı matrisleri bir kişi tarafından verilmez. Bu matrisler eğitim sonucunda ağı kendisi tarafından oluşturulur. Konvolüsyon katmanı, girdi veriden özellik çıkarımının yapıldığı katmandır (Şekil 3). Konvolüsyon işleminden sonra çıkarılan özellik haritalarının boyutları, girdi katmanın boyutlarından küçük olur (Goodfellow et al., 2016).

Şekil 3. Konvolüsyon katmanında özellik çıkarımı yapılıır.

1.2.3. Havuzlama Katmanı

Havuzlama katmanında özellik haritaları üzerinde çekirdek bir matris gezdirilir. Gezdirilen çekirdek matris ile özellik haritasını temsil eden matrisin çakıştığı bölgedeki en büyük, en küçük veya ortalama değer alınarak yeni bir matris oluşturulur (Şekil 4). Havuzlama işlemi en büyük havuzlama veya ortalama havuzlama olarak iki farklı şekilde yapılabilir (URL 1, 2018). Konvolüsyonel derin yapay sinir ağlarında çoğunlukla en büyük havuzlama tekniği kullanılır. Gezdirilen çekirdek matris, boyut ve kaydırma adımı olarak iki değişken ile ifade edilir (Scherer et al., 2010). Boyut değişkeni çekirdek matrisin $N \times N$ ifadesi ile 2 boyutunu belirler. Kaydırma adımı parametresi ise çekirdek matrisin, uygulanacağı matriste her adımda kaç hücre yer değiştireceğini belirler (Goodfellow et al., 2016).

Şekil 4. En büyük havuzlama işlemi.

1.2.4. Gizli Katmanlar

Derin yapay sinir ağlarını klasik yapay sinir ağlarından ayıran en önemli özellik, gizli katman sayısının 1'den fazla olmasıdır. Gizli katman sayısı ne kadar fazlaysa ağ o kadar derin sayılır. Gizli katmanlardaki her bir değer bir sonraki katmanın bütün girdi değerlerine doğrudan etki eder. Bu yapıdaki gizli katmanlara tam bağlantılı gizli katmanlar denir (Şekil 5).

Şekil 5. Tam bağlantılı gizli katmanlar.

1.2.5. Çıktı Katmanı

Çıktı katmanı, ağa verilen girdi katmanı değerlerinin bütün katmanlardaki işlemlerden geçtikten sonra aldığı değerler katmanıdır. Derin yapay sinir ağının son katmanıdır. Öğrenme aşamasında her bir iterasyonun sonucunda çıktı katmanının değerleri ile etiket değerleri karşılaştırılarak hata hesaplanır. Hatalar geriye dönük olarak dağıtılarak, ağıdaki her katman arasındaki bağlantıların ağırlıklarına ve bias değerlerine düzeltme getirilir. Bir sonraki iterasyonda güncellenmiş ağırlık ve bias değerleri ile işlemler yapılır. Yapılan çalışmada yaş grupları 0 – 79 olarak belirlendiğinden, çıktı katman düğüm sayısı 80 olarak belirlenmiştir (Şekil 6).

Şekil 6. Çıktı katman.1.3. Tensorflow Çatısı.

Derin öğrenme uygulamalarının geliştirilebilmesi için hazırlanmış bazı yazılım çatıları vardır. Bunların en yaygın olarak kullanılanları Caffee, Torch, Keras ve Tensorflow olarak sıralanabilir. Bu çalışmada Tensorflow çatısı kullanılmıştır. Tensorflow, Google Brain ekibi tarafından geliştirilmiştir. Apache 2.0 Açık Kaynak Kodu Lisansı ile lisanslanmıştır. Python kodlama dili ile tam uyumludur (URL 1, 2018)). Windows ve Linux işletim sistemlerinde kullanılabilir. x86 ve x64 mimarili işlemcilerde kullanılabildiği gibi ARM mimarili işlemcilerde de kullanılabilmektedir. Grafik işlemci ile çalışmaya olanak vermesi sayesinde büyük bir işlem gücü desteği sağlamaktadır.

2. UYGULAMA

2.1. Veri Seti

Derin öğrenme uygulamalarındaki büyük problemlerden biri, derin yapay sinir ağlarının eğitimi için gerekli olan örneklem veri setinin temin edilmesidir. Eğitim için aynı özellikleri barındıran binlerce örneklem bulmak gerekebilir. Bazı üniversite veya vakıflarca yaş – cinsiyet, ruh hâli kestirimi gibi problemler için hazırlanmış veri setleri mevcuttur. Bunlardan en yaygın olarak kullanılanları Imagenet veri seti ve IMDB-WIKI veri setidir (URL 2, 2018), Çalışmada IMDB-WIKI veri seti kullanılmıştır. Veri seti imdb.com sitesinden toplanmış 460723 ve wikipedia.org sitesinden toplanmış 62328 yüz görüntüsünü içermektedir. Görüntüler JPEG formatında kaydedilmiştir. Her bir görüntü, ait olduğu kişinin doğum tarihi ve fotoğrafın çekildiği yıl bilgileri ile etiketlenmiştir (Şekil 7).

Şekil 7. Veri setine ait görüntüler, kişilerin doğum tarihi ve fotoğrafın çekildiği tarih ile etiketlenmiştir.

2.2. Görüntülerin Tasnif Edilmesi

Veri setine ait görüntüler, C# dili ile yazılan yardımcı bir program ile otomatik olarak yaşa göre tasnif edilmiştir. Her bir görüntüdeki kişinin yaşı, görüntünün adında barındırdığı doğum tarihi ve fotoğrafın çekildiği yıl bilgilerine

göre hesaplanmıştır. Yaşı hesaplanan görüntü, aynı yaş ile adlandırılmış klasöre kopyalanmıştır. Bu şekilde bütün görüntüler tasnif edilmiştir. İşlem sonucunda 0 – 79 arasında yaş grupları oluşturulmuş ve her bir görüntü ait olduğu klasöre kaydedilmiştir (Şekil 8).

Şekil 8. Görüntüler yaşlara göre tasnif edilip klasörleştirilmiştir.

2.3. Derin Yapay Sinir Ağı'nın Eğitilmesi

Tasnif edilen görüntüler ile VGG-16 mimarisi ile kurgulanmış derin yapay sinir ağı eğitilmiştir. Eğitim işlemi Tensorflow çatısı kullanılarak Python dilinde yazılmış bir programla yapılmıştır. Program Nvidia GTX 1080 işlemcili ekran kartı üzerinde çalıştırılmıştır. Eğitim yaklaşık 10 saatte ve 207000 iterasyonda tamamlanmıştır. Her 1000 iterasyondaki eğitim sonucu vektör dosya olarak kaydedilmiştir.

Teoride ağı son katmanındaki loss ve çapraz entropi fonksiyonu değerinin 0 olması beklenir (URL 1, 2018). Loss fonksiyonunun 0 değeri vermesi, örnek verinin tam doğrulukla tahmin edildiği anlamına taşır. Çapraz entropi değerinin 0 olması ise tahmin edilen değerin, ait olduğu sınıfa tam korelasyonla ait olduğu anlamına gelir (de Boer et al., 2005).

2.4. Farklı İterasyonlarda Kaydedilen Eğitim Sonuçlarının Yaş Kestirimi Doğruluk Analizi

Farklı iterasyonlardaki eğitim sonuçları ile veri seti üzerinde doğruluk analizi yapılmıştır. Analiz, Python kodlama dili ile yazılmış bir yardımcı program vasıtası ile yapılmıştır. Yaş doğruluğu, eğitim sonucundan elde edilen görüntülere ait yaş bilgilerinin, görüntülerin bulunduğu klasörlerdeki yaş etiketleriyle mukayese edilmesi ile hesaplanmıştır. Her tahminin, etiket değerlerinden farklarının mutlak değerleri toplamının analize giren tahmin sayısına bölünerek ortalama yaş hataları bulunmuştur (Şekil 9). İterasyon sayısının loss ve çapraz entropi fonksiyonu değerine etkisi ve loss ve çapraz entropi fonksiyonu değerindeki değişimin doğruluk üzerindeki etkisi gözlemlenmiştir.

Şekil 9. 207000. iterasyonun tahmin sonuçları. Soldan sağa: TD: 79 GD: 82, TD: 5 GD: 3, TD: 22 GD: 33 (TD: Tahmin değeri, GD: Gerçek değer).

Çizelge 1. Belirli iterasyonlardaki eğitim sonuçlarının loss, çapraz entropi ve ortalama yaş hatası değerleri.

İterasyon	Loss	ÇaprazEntropi	OrtalamaYaşHatası
1	9602,99	4,25	16,66
1001	5,04	3,52	12,40
2001	4,86	4,2	11,81
10001	4,21	3,76	8,67
23001	4,39	2,89	8,47
115001	9,01	0,32	8,09
130001	8,93	0,43	7,72
140001	8,97	0,21	8,00
150001	8,91	0,24	7,95
160001	9,00	0,23	7,93
201001	8,98	0,09	8,01

3. SONUÇLAR

Doğruluk analizi sonuçlarına göre iterasyon sayısı arttıkça çapraz entropi değerinin 0'a yaklaşma eğiliminde olduğu söylenebilir. Loss değerinin eğitimin ilk aşamalarında büyük değerle başlayıp, küçülme eğilimine girse de belirli bir sayıdaki iterasyondan sonra aynı değerlerde seyrettiği gözlemlenmiştir. Ortalama yaş hatası değerindeki değişimin loss değerindeki değişimle benzerlik gösterdiği görülmüştür. Ancak iterasyon sayısının artması, çapraz entropi değerini 0'a yaklaştırdığı için derin yapay sinir ağının yaş kestirimi sonucunun daha güvenilir yapılmasını sağladığı söylenebilir.

KAYNAKLAR

Goodfellow et. al. (2016), Deep Learning

Scherer, D., Müller, A., Benkhe, S., 2010, Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition

Fukai, H., Takimoto, H., Mitsukura, Y., Fukumi, M., 2009, Age and Gender Estimation System Based on Human Perception (2009) 1143-1148

Eran Eidinger, Roe Enbar, Tal Hassner, 2014, Age and Gender Estimation of Unfiltered Faces (2014) 2170 – 2179

Rasmus Rothe, Radu Timofte, Luc Van Gool, 2015, DEX: Deep Expectation of Apparent Age From a Single Image

Gil Levi and Tal Hassner, 2015, Age and Gender Classification using Convolutional Neural Networks

de Boer, P-T., Kroese, D., Mannor, S., & Rubinstein, R. Y. 2005, A Tutorial on the Cross- Entropy Method. *Annals of operations research*, 134(1), 19-67. DOI: 10.1007/s10479-005-5724- z

URL 1, 2018, https://www.tensorflow.org/api_docs/, (Tensorflow API Dokümanı)

URL 2, 2018, <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki>