

# Derin Öğrenme ile Tüketilen Besin İçeriklerinin Yaklaşık Kestirimi

## Approximate Estimation of the Nutritions of Consumed Food by Deep Learning

İbrahim Berkan AYDİLEK

Harran Üniversitesi, Mühendislik Fakültesi

Bilgisayar Mühendisliği

Şanlıurfa, Türkiye

berkanaydilek@harran.edu.tr

**Özetçe**—Derin öğrenme son yıllarda bilgisayar donanım sistemlerinin gelişmesi ve büyük veri kavramı sonucu makine öğrenmesinin popüler bir alt çalışma alanı haline gelmiştir. Derin öğrenme görüntü sınıflama, konuşma tanıma, doğal dil işleme gibi çalışmalarda başarılı bir şekilde kullanılmaktadır. Kontrollü beslenme, sağlıklı bir birey olabilmenin şartlarından biri olarak uzmanlar tarafından tavsiye edilmektedir. Gün içinde ne kadar besin tüketildiği ve bu besinlerin içerik ve kalori miktarlarının bilinmesi ve takip edilmesi sağlıklı beslenme kontrolüne yardımcı olmaktadır. Bu çalışmada, besin görüntüleri içeren Food-pics veri kümesi kullanılarak, görüntü düzeyinde bir besinin, besin değerleri yaklaşık olarak tahmin edilmeye çalışılmıştır. Görüntü tanıma ve sınıflama görevlerinde başarılı bir şekilde kullanılan derin öğrenme yaklaşımı olan Konvolüsyonel Sinir Ağları (Convolutional Neural Networks (CNN)) besin görüntü eğitim verisi ile eğitilmiş ve yüksek sınıflama başarıları elde edilmiştir. Ağ sonucu elde edilen en yüksek beş tahmin değerli besine ait değerlerin aritmetik ortalama besin değerleri, yaklaşık besin içerik kestirimi olarak kabul edilmiştir.

**Anahtar Sözcükler**—Derin öğrenme; Besin içeriği; Konvolüsyonel Sinir Ağları

**Abstract**— Deep learning has become a popular sub-study in machine learning in the recent years due to the conclusions of development in computer hardware systems and big data. Deep learning is used successfully in image classification, speech recognition, natural language processing, and so on. Controlled intake of nutrition is recommended as a condition for being a healthy individual. Knowing and monitoring how much food is consumed during the day, following the calorie and nutrition of these foods helps to control healthy nutrition. In this study, an attempt was made to approximate the nutritions of the food at the image level using the Food-pics dataset that contain nutrient images. Convolutional Neural Networks (CNN), a deep learning approach that has been used successfully in image recognition and classification tasks, has been trained with nutrition image training data, and a high classification success value has been achieved. The arithmetic average

nutritional values of the five highest predicted nutritional results obtained from the net were taken as the approximate nutrient content.

**Keywords**—Deep learning; Food Nutrition; Convolutional Neural Networks

### I. GİRİŞ

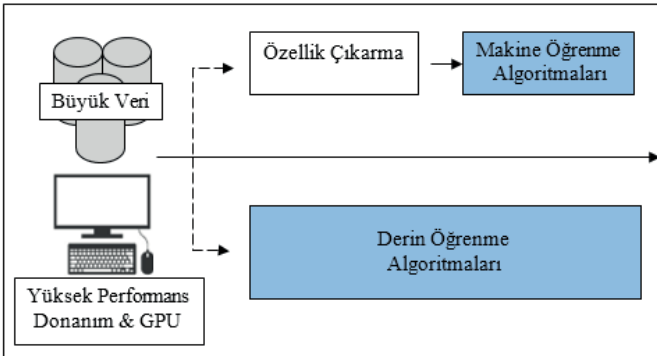
Obezite, aşırı, ihtiyaç fazlası gıda tüketiminin sonucu olarak gelişmekte böylece kişide diyabet tip 2, kalp hastalıkları ve kansere sebebiyet verebilmektedir. Başarılı ve sağlıklı bir diyet için tüketilen besinlerin ölçülmesi ve takip edilmesi önem arz etmektedir. Gün içinde tüketilen besinlerin kalori ve içeriklerinin takip edilmesi kolay olmamaktadır [1]. Obezitenin en büyük nedeni bireyler tarafından tüketilen enerji ile alınan gıdalar arasındaki orantısızlıktır. Bu nedenle, obez kişilerin normal, sağlıklı bir birey haline dönüşebilmeleri için günlük tüketilen besinlerin ölçülmesi, takip edilmesi gerekmektedir [2]. Obezite tedavisi için hastaların günlük aldığı gıda miktarlarını kayıt altına almaları gerekmektedir. Fakat bu kayıt altına alma işlemi, çoğu zaman, hastaların beslenme konusundaki bilgisizliği, ilgisizliği veya yorucu olması, unutma gibi sebeplerden dolayı yapılamamaktadır [3]. Bunun yanı sıra tüketilen besinler manuel olarak kayıt altına alınsa bile içerdiği besin değerlerini hesaplayacak yarı otomatik veya tam otomatik bir sistem olmadığı zamanlarda ne yazık ki hastalar çoğu zaman günlük tüketilen besinlerin içeriklerini ölçme ve hesaplama konusunda zorluklarla karşılaşmaktadırlar. Bu gibi otomatik sistemler, yani besin tanıma ve kayıt altına alma sürecine yardımcı olan yaklaşımlar, obeziteden etkilenen hastalara ve diyetisyenlere yardım sağlayacağı gibi bunun yanı sıra normal, sağlıklı bireylere de faydalar sağlayacağı düşünülmektedir [4].

Bugüne kadar görüntüden besin tanıma, besin içeriği ve kalorisi ile ilgili değişik ve farklı kategorilerde çalışmalar yapılmıştır [3], [5]–[12]. Bu çalışmalara göre bizim çalışmamızın özgün yönleri şu şekilde sıralanabilir. Derin öğrenme algoritmalarından konvolüsyonel sinir ağları ile 15

katmanlı bir ağ yapısı oluşturulmuştur. Bu sinir ağı ile birlikte Food-pics besin görüntü veri kümesi üzerinde çalışmalar yapılmış ve tahmin sonucu elde edilen en iyi beş besine ait, besin ağırlığı, içerdiği protein, yağ, karbonhidrat ve kalori değerleri ile ilgili aritmetik ortalama ile yaklaşık kestirimlerde bulunulmuştur.

## II. DERİN ÖĞRENME

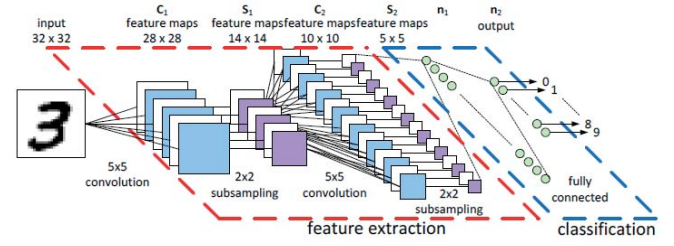
Büyük boyutlu görüntü ve ses veri kümelerinin artması ve bilgisayar donanım sistemlerinin gelişmesi özellikle grafik işleme birimlerinin (GPU) kapasitelerinin artması ile yapay sinir ağlarına olan ilgi tekrar artmış ve çok katmanlı farklı mimaride sinir ağları tasarlanmıştır. Bu sayede makine öğrenmesinin bir alt çalışma kolu olarak derin öğrenme kavramı son yıllarda özellikle Facebook, Google, Microsoft gibi büyük teknoloji firmaları tarafından yüksek ilgi görmüş ve başarılı sonuçlar elde edilen çalışmalar yapılmasına olanak sağlamıştır [13]. Derin öğrenme günümüzde daha çok görüntü sınıflama, ses tanıma, doğal dil işleme çalışmalarında kullanılmaktadır. Büyük boyutlu veri ve yüksek performanslı donanıma sahip olma durumuna göre derin öğrenme veya makine öğrenmesi algoritmaları tercih edilebilmektedir. Şekil 1’de görüldüğü gibi, görüntü veya ses gibi büyük boyutlu veri üzerinde normal makine öğrenme algoritmaları ile çalışma yapılırken genellikle özellik çıkarma (feature extraction) süreçleri gerekmektedir. Bu süreçler genellikle zor ve zahmetli çalışmalar olabilmektedir [14]. Özellik çıkarma sürecini daha fazla otomatik hale dönüştürebilen derin öğrenme algoritmaları tercih edilmektedir.



Şekil 1. Derin öğrenme ve makine öğrenme süreci

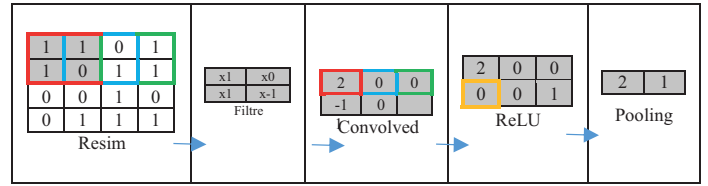
### A. Konvolüsyonel Sinir Ağları

Konvolüsyonel Sinir Ağları (Convolutional Neural Networks (CNN)), görüntüde nesne tanıma ve sınıflama için özelleştirilmiş Şekil 2’de görüldüğü gibi çok katmanlı bir yapay sinir ağı modelidir. İçerisinde her birinin değişik fonksiyonları olan farklı katmanlar bulunmaktadır. Bu mimaride ağı ilk katmanları özellik çıkarma sürecini sona doğru katmanları ise sınıflama görevini yapmaktadır.



Şekil 2. Basit CNN çalışma prensibi akış şeması [15]

Konvolüsyon, Pooling ve Tam bağlı (fully connected) katmanlar birbirini takip edecek şekilde arka arkaya gelerek konvolüsyonel sinir ağlarını oluşturmaktadırlar. Şekil 3’de görüldüğü gibi CNN’de temel olarak bir katmandan çıkan veriler diğer katman tarafından giriş olarak kabul edilmekte ve bu şekilde en son tam bağlı katmana kadar işlemler devam edilerek sınıflama gerçekleştirilmektedir.



Şekil 3. Basit CNN çalışma prensibi akış şeması

Özetle, şekilde Konvolüsyon katmanı, kayan bir filtrenin resim üzerinde gezdirilmesi ile filtre ve resim değerleri arasında nokta çarpım (dot product) işlemini yapmaktadır ve bu şekilde konvüle özellik değerleri (Convolved) elde edilmektedir. Konvüle özellik değerleri elde edildikten sonra Rectified Linear Unit (ReLU) aktivasyonu ( $y=\max(x,0)$ ) fonksiyonu uygulanmaktadır. Daha sonra ise Pooling katmanına ait değerler, örneğin filtredeki sayısal olarak en yüksek değerlerin seçilme işlemi yapılmaktadır. Bu aşamadan sonra çok katmanlı yapay sinir ağına benzer nöronların birbirine bağlı olduğu tam bağlı katman vasıtasıyla sınıflama yapılmaktadır.

## III. BESİN VERİ KÜMESİ: FOOD-PICS

Food-pics, 568 adet besin 315 adet besin olmayan toplam 883 resim ögesi içeren bir besin görüntü veri kümesidir [16]. Şekil 4’de bahsedilen veri kümesinden bir kesit gösterilmektedir. Bu veri kümesi aynı zamanda bir görüntünün besin ögesi olup olmadığını da başarılı bir şekilde sınıflayabilmektedir. Food-pics veri kümesi, yemek yeme alışkanlıkları üzerinde araştırmalar yapılması ve bazı standartlar oluşturarak yapılan çalışmalarda karşılaştırmaları kolaylaştırmak için ücretsiz olarak sunulmaktadır.



Şekil 4. Food-pics besin görüntü veri kümesine ait bir kesit

Bunun yanı sıra, TABLO I. 'de görüldüğü gibi bu veri kümesi içerdiği her bir görüntüye ait besin değerleri ile ilgili bilgiler içeren bir tablo sunmaktadır. Toplam ağırlık (gr), protein, yağ, karbonhidrat, kalori (Kcal) gibi sağlıklı beslenme için takip edilmesi gereken bu besin değerleri ayrı bir dosya (excel) olarak araştırmacıların kullanımına sunulmaktadır.

TABLO I. FOOD-PICS BESİN İÇERİKLERİ VERİ KÜMESİNE AİT BİR KESİT

Besin	Ağırlık	Protein	Yağ	Karbonhidrat	Kalori
cheesecake with cherry	150	9	25.71	49.335	466.5
Hamburger with bacon	360	54	54	50.4	907.2
cheese burger, french fries and cola	720	38.448	64.8	146.16	1332
chocolate cookie	19	1.178	4.94	11.97	96.9
cookie mix	18	1.4004	4.6152	10.2258	88.02
cheese cake with strawberries and crumbles	150	13.545	19.005	42.045	396
croissants	300	21.24	100.71	134.58	1524
snack mix	20	1.8	0.1	15.06	69.4
waffles	42	3.8178	10.08	23.94	197.4
Spaghetti with tomato sauce	250	8.35	16.825	40.125	347.5

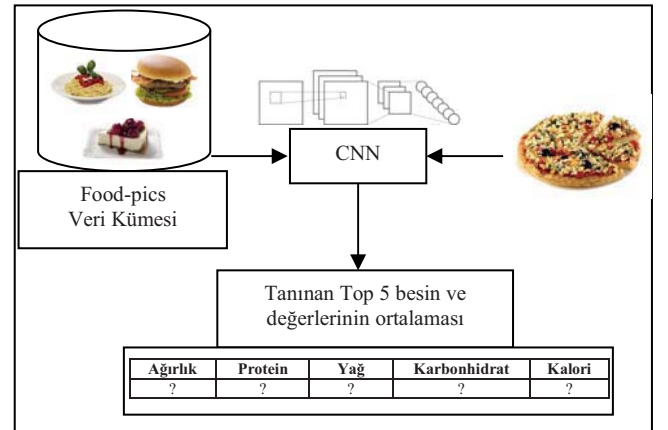
#### IV. ÖNERİLEN YÖNTEM

Önerilen besin değerlerinin ve kalorisinin yaklaşık hesaplanabilmesi için 3 Konvolüsyon, 3 Pooling, 4 ReLU ve 2 Tam bağlı katman olmak üzere toplam 15 katmana sahip bir konvolüsyonel sinir ağı tasarlanmıştır. Bu ağ, Food-pics besin görüntü veri kümesine ait 543 görüntü ile eğitilmiştir. 25 görüntü ise daha sonra modeli test için kullanılmak üzere ayrılmıştır. Deney için kullanılan sinir ağının mimarisi TABLO II. 'de verilmiştir.

TABLO II. KULLANILAN KONVOLÜSYONEL SINIR AĞI MİMARİSİ

Katman No	Tanım
1	Resim Giriş
2	Konvolüsyon
3	Max Pooling
4	ReLU
5	Konvolüsyon
6	ReLU
7	Average Pooling
8	Konvolüsyon
9	ReLU
10	Average Pooling
11	Tam Bağlı
12	ReLU
13	Tam Bağlı
14	Softmax
15	Sınıflama Çıkış

Ağ tasarım ve eğitim program kodları Matlab yazılımının 2017a sürümü kullanılarak yazılmış ve bu şekilde sonuçlar elde edilmiştir. Bu çalışmada, Windows 10 işletim sistemine sahip 8 GB bellekli, Intel Xeon 3.20 GHz işlemcili (CPU) ve Nvidia GeForce GTX TITAN 2880 CUDA core 888 MHz grafik işlemcili (GPU) bir masaüstü bilgisayar kullanılmıştır. Ağ eğitimi maksimum 20 epochs sonlanma kriteri ile yaklaşık 407 sn. çalışma süresine sahiptir. Eğitim veri kümesi üzerinde sınıflama doğruluk yüzdesi %92.45 olarak elde edilmiştir. Şekil 5'de önerilen derin öğrenme, konvüsyonel sinir ağı (CNN) kullanılarak yaklaşık besin içerikleri kestirim yaklaşımının akış şeması verilmiştir.









Şekil 5. Önerilen derin öğrenme besin içerikleri yaklaşık kestirim yöntemi

Buna göre, Food-pics görüntü veri kümesi ile eğitilmiş CNN ağına o anda tüketilen bir besin veya daha önce tüketilmiş besin görüntüleri verilerek çıkış değerleri elde

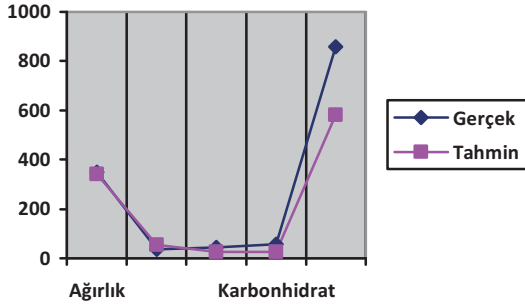
edilmektedir. Uluslararası görüntü sınıflama yarışmalarında kullanılan Top 5 başarı ölçütü, yani sınıflama yapılmak istenen görüntüye ait en başarılı beş tahmin çıkışın içinde yer alıp almama yüzdesine benzer bir yaklaşım benimsenmiştir. Önerilen yöntemde, sinir ağının son çıkış katmanında elde edilen en yüksek aktivasyon fonksiyonu değerlerine sahip olan beş adet besin içerik değerlerinin ortalaması alınarak yaklaşık bir sonuç üretilmektedir.

Şekil 6'da önerilen besin tahmin sistemi ile elde edilmiş temsili bir örnek çıktı görülmektedir. Test görüntü 'Pizza' olarak ağ girişine verilmiştir. Ağ çıkışında en yüksek değerli 5 adet besin görüntü çıktısının, besin içerikleri veri tablosundan ağırlık(gr), protein, yağ, karbonhidrat ve kalori(Kcal) değerleri sorgulanmıştır.

Test		1.Tahmin	2.Tahmin	3.Tahmin	4.Tahmin	5.Tahmin
						
Ağırlık	350	350	900	200	50	210
Protein	35	35	155.61	42	3.5	39.9
Yağ	42	42	73.26	8	0.1	6.09
Karbonhidrat	56	56	33.3	0	39	2.73
Kalori	857.5	857.5	1410.03	242	174.5	225.75

Şekil 6. Temsili bir tahmin çıktısı

Test besin görüntü verisinin değerlerine karşı elde edilen değerlerin aritmetik ortalaması alınarak karşılaştırma yapılmıştır. Buna göre, Şekil 7'deki gibi 350 gr olması gereken ağırlık değeri yaklaşık 342 gr olarak bulunmuştur. Diğer besin içerikleri protein, yağ, karbonhidrat ve kalori değerleri ise gerçekte, 35, 43, 56, 857.5 olması gerekirken, yaklaşık olarak ortalama 55, 25.89, 26, 582 şeklinde bulunmuştur.



Şekil 7. Yaklaşık besin içerikleri

Elde edilen besin içerikleri ve kalori sonuçları ile ilgili iyileştirmeler yapılabilmesi için daha fazla sayıda farklı açılardan çekilmiş besin görüntüsü içeren veri kümeleri ile sinir ağı eğitilebilir. Bunun yanı sıra görüntüdeki arka planların Food-pics veri kümesi için beyaz alanların ağ eğitimindeki etkisinin azaltılabilmesi için çıkarılması, ayrıştırılması (segmentation) fayda sağlayabilir. Aynı zamanda görüntüden farklı teknikler ile besin hacimleri bulunarak besin ağırlığı ile ilgili daha net hesaplamalar

yapılabilir. Böylece daha az hata payına sahip bir yaklaşım elde edileceği düşünülmektedir.

## V. SONUÇLAR

Bu çalışmada, bir derin öğrenme yaklaşımı, çoğu makine öğrenme görüntü sınıflama algoritmasından farklı olarak özellik çıkarım aşaması olmadan kullanılmıştır. Bu amaçla görüntüde nesne tanıma ve sınıflama çalışmaları için popüler olarak kullanılan bir konvülsiyonel sinir ağı (CNN) tasarlanmıştır. 15 katmana sahip bu ağ, herkese açık ve ücretsiz Food-pics besin görüntü veri kümesi ile eğitilmiştir. Eğitilen ağa daha önce bilinmeyen bir besin görüntüsü verilmiş ağ çıkışında en iyi beş adet besin tahminine ait besin içerik ve kalori değerlerinin aritmetik ortalaması alınmıştır. Bu sayede daha sağlıklı beslenme için gerekli olan tüketilen besinlerin kayıt altına alınması ve takibi derin öğrenme ile gerçekleştirilmiştir. Sonraki çalışmalarda görüntüler akıllı telefonların kameralarını kullanan mobil uygulamalar ile elde edilebilir ayrıca gün içinde alınan besinlerin yanında harcanan kalori ve yapılan aktiviteleri takip etmek için giyilebilir akıllı teknolojilerden faydalanılabilir.

## KAYNAKÇA

- [1] K. Ambhore and N. A. Dawande, "Measuring Calories and Nutrition from Food Image," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 5, no. 6, 2016.
- [2] Wenyan Jia, Ruizhen Zhao, N. Yao, J. D. Fernstrom, M. H. Fernstrom, R. J. Scabassi, and M. Sun, "A food portion size measurement system for image-based dietary assessment," in *2009 IEEE 35th Annual Northeast Bioengineering Conference*, 2009, pp. 1–2.
- [3] S. Anushadevi, "Calorie Measurement Of Food From Food Image," *Int. J. Appl. Inf. Commun. Eng.* *www.aetsjournal.com ISSN*, vol. 1, no. 7, pp. 14–17, 2015.
- [4] P. Pouladzadeh, S. Shirmohammadi, and R. Al-Maghrabi, "Measuring Calorie and Nutrition From Food Image," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 8, pp. 1947–1956, Aug. 2014.
- [5] A. A. Podutwar, A. V. Shinde, D. Yavatmal, and A. Professor, "Calorie and Nutrition Measurement Based on Food Image Processing."
- [6] A. S. Geeta Shroff, "Neural Network Based Food Recognition and Calorie Calculation for Diabetes Patients," 2009.
- [7] K. Kitamura, C. de Silva, T. Yamasaki, and K. Aizawa, "Image processing based approach to food balance analysis for personal food logging," in *2010 IEEE International Conference on Multimedia and Expo*, 2010, pp. 625–630.
- [8] P. Pouladzadeh, G. Villalobos, R. Almaghrabi, and S. Shirmohammadi, "A Novel SVM Based Food Recognition Method for Calorie Measurement Applications," in *2012 IEEE International*

- Conference on Multimedia and Expo Workshops*, 2012, pp. 495–498.
- [9] J. Shang, E. Pepin, E. Johnson, D. Hazel, A. Teredesai, A. Kristal, and A. Mamishev, “Dietary intake assessment using integrated sensors and software,” 2012, p. 830403.
  - [10] N. Dharman, S. PS, S. CM, S. Shanmughan, and N. Johnson, “IMG2 Calories,” *Int. J. Comput. Trends Technol.*, vol. 28, no. 3, pp. 144–148, Oct. 2015.
  - [11] Kavitha S and Pavithra S, “Performance Analysis of Nutritional Contents in Food Images Using Saran,” vol. 10, no. 12, 2015.
  - [12] R. Surya and S. Saru Priya, “Food-Image-Recognition-Using-Svm-Classifer-for-Measuring-Calorie-and-Nutrition-Values.docx,” *Int. J. Sci. Eng. Res.*, vol. 6, no. 4, 2015.
  - [13] G. Isik and H. Artuner, “Recognition of radio signals with deep learning Neural Networks,” in *2016 24th Signal Processing and Communication Application Conference (SIU)*, 2016, pp. 837–840.
  - [14] A. Moujahid, “A Practical Introduction to Deep Learning with Caffe and Python,” 2016. [Online]. Available: <http://adilmoujahid.com/posts/2016/06/introduction-deep-learning-python-caffe/>. [Accessed: 29-Apr-2017].
  - [15] M. Peemen, B. Corporaal, and Henk, “Speed Sign Detection and Recognition by Convolutional Neural Networks.”
  - [16] J. Blechert, A. Meule, N. A. Busch, and K. Ohla, “Food-pics: An image database for experimental research on eating and appetite,” *Front. Psychol.*, vol. 5, no. JUN, 2014.