

Türk Sofrası

Yemek Tanıma için Türk Yemekleri
Barındıran Bir Denektaşlı Veri Kümesi



Cem Güngör • Fatih Baltacı • Aykut Erdem • Erkut Erdem



Yaprak sarma: 0,564

Pekin ördeği: 0,384

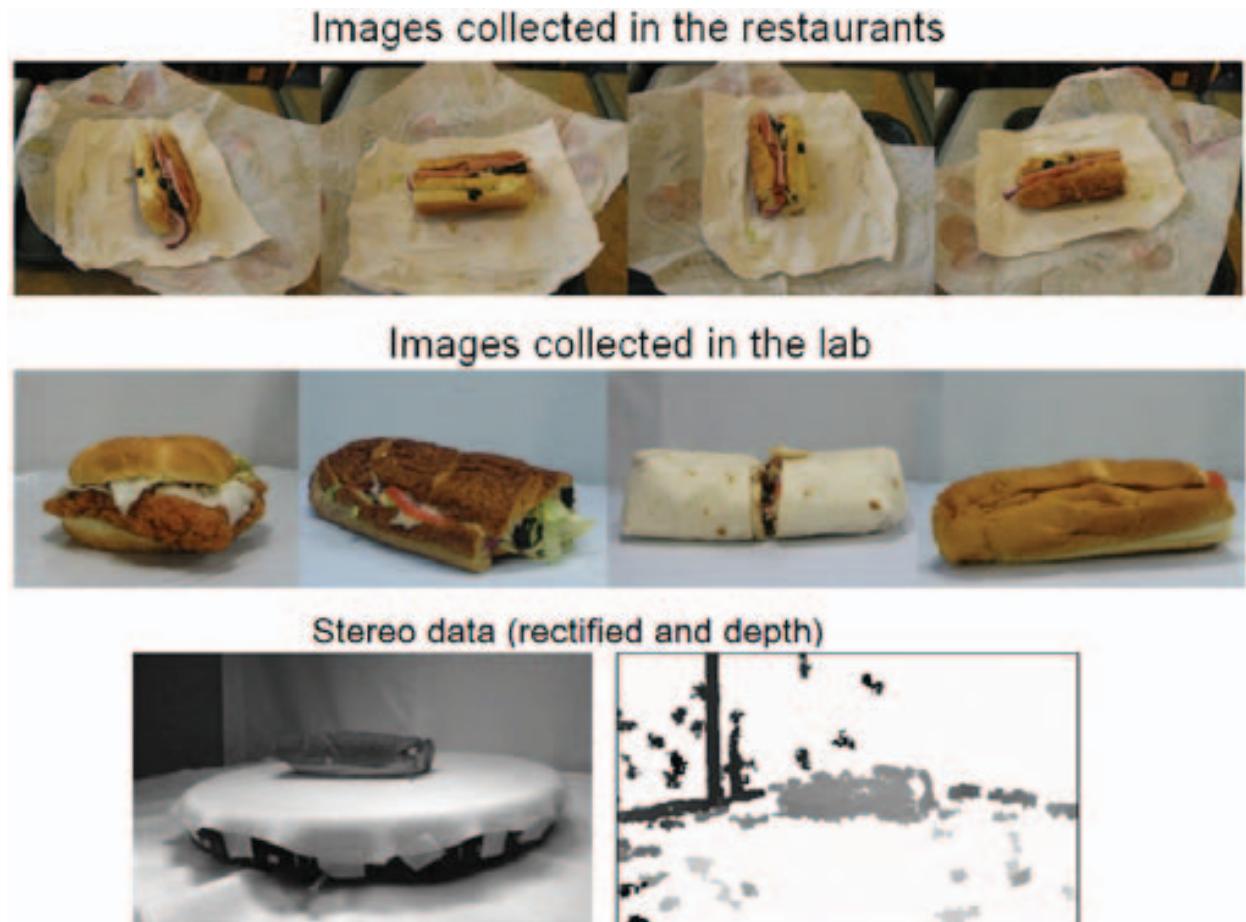
Biftek: 0,014

Baklava: 0,011



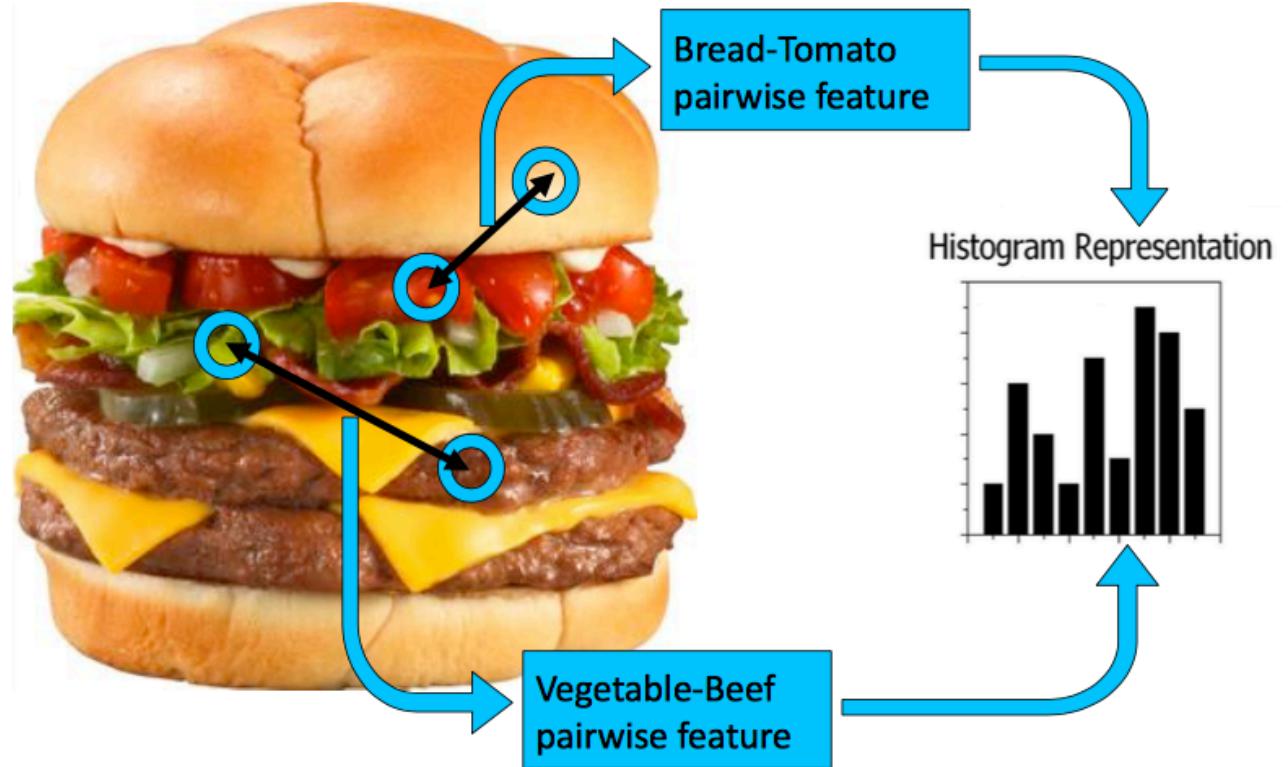
PFID: Pittsburgh Fast-Food Image Dataset (Chen vd., 2009)

- 4.545 görüntü
- 7 yemek sınıfı
- SIFT öznitelikleri
- Destek Vektör makinesi
(*Support vector machine – SVM*)



Food Recognition Using Statistics of Pairwise Local Features (Yang vd., 2010)

- Yerel öznitelikler üzerinden yemek bileşenlerinin ikili ilişkilerini kodlayan bir temsil
- PFID veri kümesi
- SVM



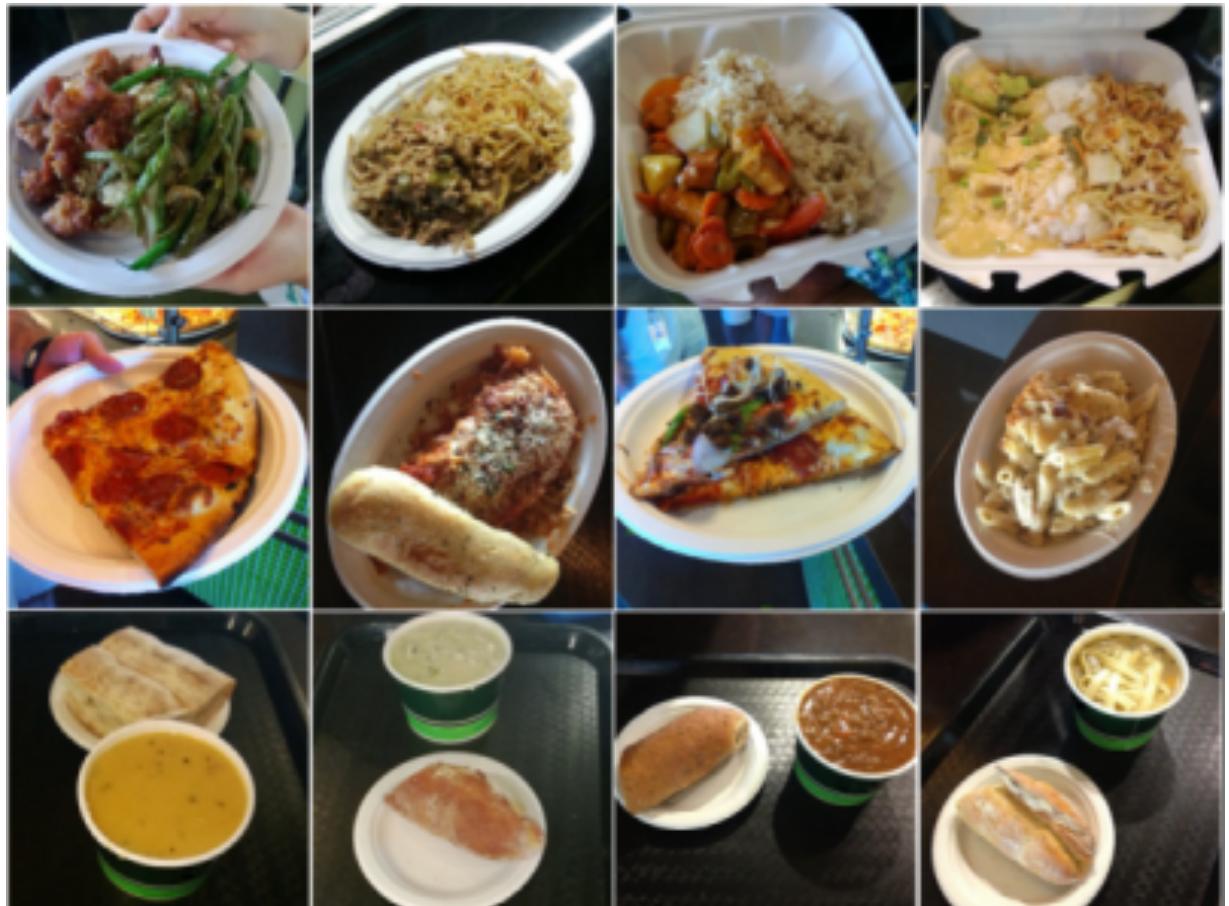
Food-101 -- Mining Discriminative Components with Random Forests (Bossard vd., 2014)



- 101K görüntü, 101 yemek sınıfı
- Rastgele Orman (*Random Forest*)
- Derin Evrişimsel Sinir Ağrı (*Deep Convolutional Network*)
 - AlexNet (Krizhevsky, 2012)

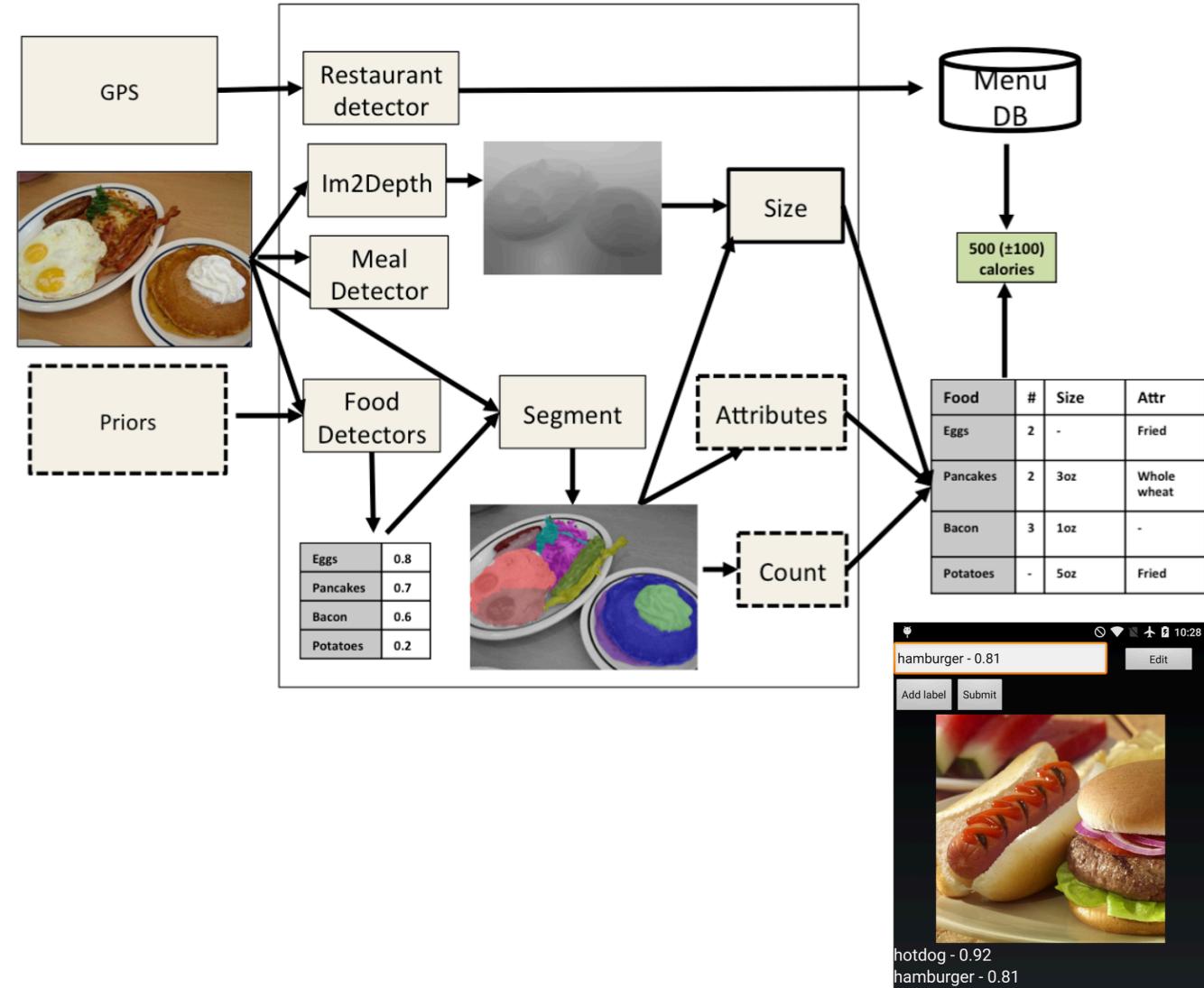
Menu-match: Restaurant-specific food logging from images (Beijbom vd., 2015)

- 646 görüntü, 41 kategoriden 1386 etiket
 - Çin lokantası
 - İtalyan restoranı
 - Çorbacı
- Kalori değerleri
- Yakın restoranların menüleri (GPS bilgisi)
- {HOG,SIFT,LBP}+SVM
- Kategori ve kalori tahmini



Im2calories: towards an automated mobile vision food diary (Myers vd., 2015)

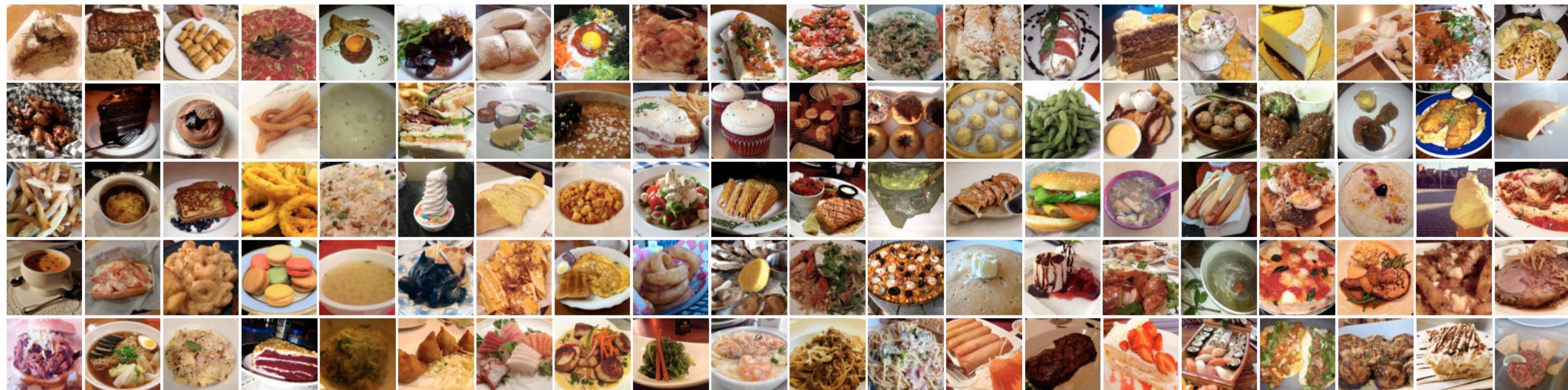
- Food201 verikümesi
 - 23 restoran
 - 2517 yemek bileşeni
 - 201 kategori
- Derin evrişimsel sinir ağları
- Anlamsal bölütleme
(Cordeiro vd., 2015)
- Derinlik kestirimi (Eigen ve Fergus, 2014)
- Kalori hesabı



Bir şey farkettiniz mi?

Food-101 veri kümesi (Bossard vd., 2014)

- 101 yemek sınıfı
- Her yemek sınıfı için 1000 görüntü
- Sadece 1 Türk yemeği: baklava



UNICT-FD889 veri kümesi (Farinella vd., 2014)

- 889 farklı yemek için 3583 yemek görüntüsü
- Türk yemeği bulunmamaktadır.



Food-5K veri kümesi (Singla vd., 2016)

- 5.000 görüntü (2.500 yemek ve 2.500 yemek olmayan)
- İkili sınıflandırma: yemek veya yemek değil



Türk Sofrası-15

- Türk yemeklerine özelleşmiş ilk veri kümesi



Google Images ile Veri Toplama

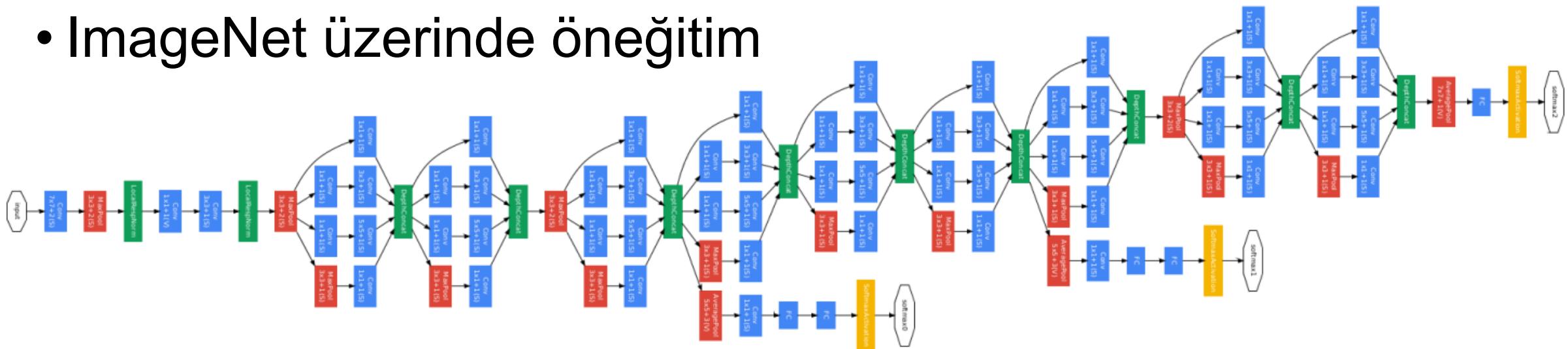
Yemek Tanıma için Bütünleşik Veri Kümesi

- 113 farklı yemek sınıfı
- Food-101 (97)
- Türk Sofrası (15)
- Food-11K (1)

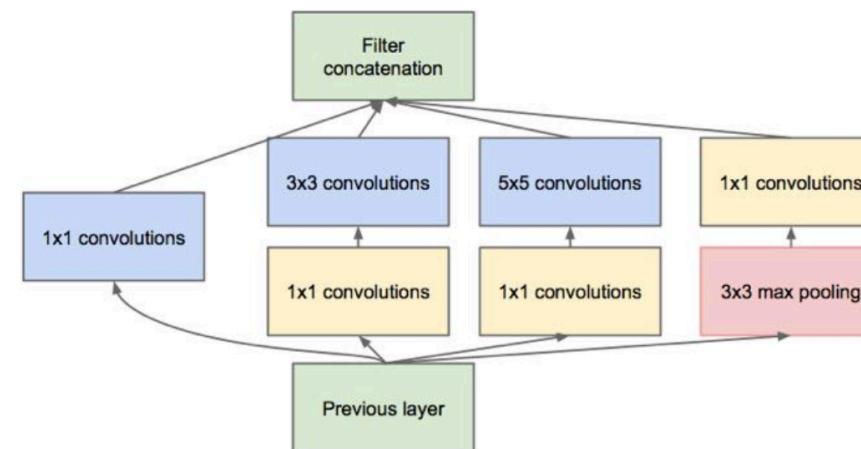


Derin Evrişimsel Ağlar ile Yemek Tanıma

- Google Inception v3 modeli (Szegedy vd., 2016)
- ImageNet üzerinde öneğitim



- ~5 milyon parametre
- Inception modülü



Veri Büütme



orjinal görüntü



parlaklık artırma



parlaklık azaltma



yatay çevirme



dikey çevirme



gürültü ekleme

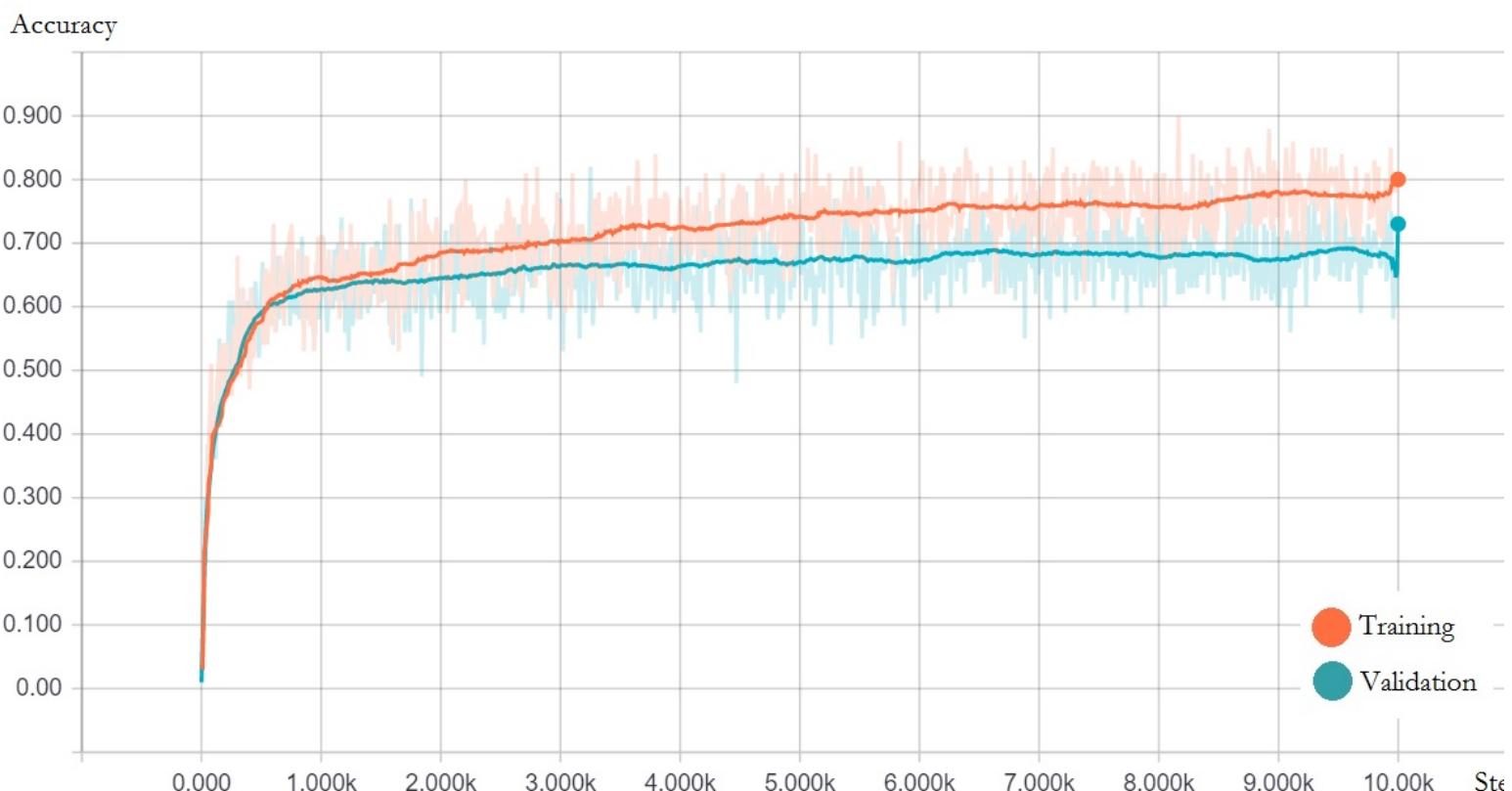
Gerçekleştirim Detayları

- Bütünleşik veri kümelerinin %10'luk kısmı test için, %10'luk bir kısmı doğrulama (*validation*) kullanılmıştır.
- Veri kümelerindeki tüm görüntüler 299×299 piksel boyutuna indirgenmiştir.
- Nvidia GTX960M GPU kartı üzerinde eğitim ve test
- Tensorflow kütüphanesi

- Derin öznitelik çıkarımı (son katmana kadar): 2 saat 24 dakika
- 1.000 adım kullanılarak gerçekleştirilen öneğitim: 3 dakika 34 saniye
- 10.000 adım kullanılarak gerçekleştirilen öneğitim: 25 dakika 24 saniye

Deneysel sonuçlar

- Öğrenme oranı: 0,3
- Parti büyüklüğü: 100
- Adım sayısı: 10.000
- Doğruluk oranı: %68,2



Örnek Sonuçlar



cig kofte: 0.87750

salad: 0.03749

falafel: 0.02075

tuna tartare: 0.01002

Örnek Sonuçlar



hamsi: 0.58653

baklava: 0.30801

carrot cake: 0.05741

humus: 0.01253

Örnek Sonuçlar



simit: 0.96864

donut: 0.00905

garlic bread: 0.00673

onion rings: 0.00340

Örnek Sonuçlar



Gnocchi

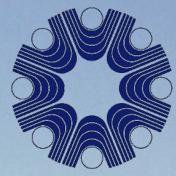
gnocchi: 0.31065
kuru fasulye: 0.22848
chop: 0.15065
grilled salmon: 0.15038

Sonuç ve Gelecek Çalışmalar

- Türk yemeklerine özelleşmiş yeni bir veri kümesi
<https://vision.cs.hacettepe.edu.tr/data.php>
- Derin evrişimsel ağlar (Google Inception v3) ile ilk sınıflandırma sonuçları
- Veri kümelerinin genişletilmesi
- Kalori hesaplama, yemek tarifi belirleme (Malmaud vd., 2015) gibi problemlere uyarlama

Teşekkürler

Ufak Bir Duyuru



BİLİM AKADEMİSİ

<http://byoyo2017.vision.cs.hacettepe.edu.tr>



HACETTEPE
ÜNİVERSİTESİ

Bozkırda Yapay Öğrenme Yaz Okulu 2017

Bilgisayarlı Görü ve Robotik

21-24 Ağustos 2017, Beytepe, Ankara

SON BAŞVURU TARİHİ: 15 HAZİRAN 2017



Gökberk
Cinbis
Bilkent
Üniversitesi



Hakan
Çevikalp
Eskişehir
Osmangazi
Üniversitesi



Hamdi
Dibeklioğlu
Delft Teknoloji
Üniversitesi



Bernard
Ghanem
Kral Abdullah
Bilim ve Teknoloji
Üniversitesi



Sinan
Kalkan
Orta Doğu Teknik
Üniversitesi



Melih
Kandemir
Özyegin
Üniversitesi



Pınar
Duygulu Şahin
Hacettepe
Üniversitesi



Aykut
Erdem
Hacettepe
Üniversitesi



Koray
Kavukçuoğlu
Google
DeepMind



Sanem
Sariel
İstanbul Teknik
Üniversitesi



Erol
Şahin
Orta Doğu Teknik
Üniversitesi



Ozan
Şener
Stanford
Üniversitesi



Emre
Uğur
Boğaziçi
Üniversitesi



Zeynep
Yücel
Okayama
Üniversitesi



Erkut
Erdem
Hacettepe
Üniversitesi



Nazlı
İkizler Cinbis
Hacettepe
Üniversitesi

DANIŞMA KURULU



Aydın Alatan
Orta Doğu Teknik
Üniversitesi



Bülent Sankur
Boğaziçi
Üniversitesi



HACETTEPE
ÜNİVERSİTESİ
BİLGİSAYARLI
GÖRÜ LABI

Kaynaklar

- S. L. Yang, M. Chen, D. Pomerleau, and R. Sukthankar, “Food recognition using statistics of pairwise local features,” in CVPR, 2010.
- L. Bossard, M. Guillaumin, and L. Van Gool, “Food-101 – mining discriminative components with random forests,” in European Conference on Computer Vision, 2014.
- D. M. S. S. Oscar Beijbom, Neel Joshi and S. Khullar, “Menu-match: Restaurant-specific food logging from images,” in WACV, 2015.
- A. Myers, N. Johnston, V. Rathod, A. Korattikara, A. Gorban, N. Silberman, S. Guadarrama, and et al., “Im2calories: towards an automated mobile vision food diary,” in ICCV, 2015.
- A. Singla, L. Yuan, and T. Ebrahimi, “Food/non-food image classification and food categorization using pre-trained googlenet model,” in Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management, pp. 3–11, ACM, 2016.
- Y. Kawano and K. Yanai, “Automatic expansion of a food image dataset leveraging existing categories with domain adaptation,” in Proc. Of ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV), 2014.
- J. Malmaud, J. Huang, V. Rathod, N. Johnston, A. Rabinovich, and K. Murphy, “What’s cookin’? interpreting cooking videos using text, speech and vision,” in NAACL, 2015.
- Giovanni Maria Farinella, Dario Allegra, Filippo Stanco, “A Benchmark Dataset to Study the Representation of Food Images”
- Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna, “Rethinking the Inception Architecture for Computer Vision”