

## 3. Veri Bilimi Yaz Okulu (VBYO) 2019

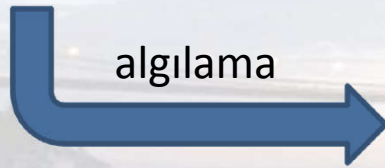
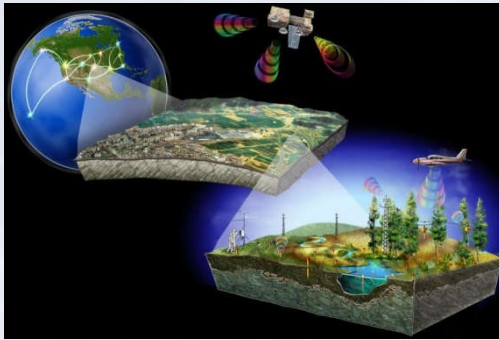
Veri Analitiği ve Araştırma Merkezi  
Sabancı Üniversitesi, Tuzla, İstanbul

*Derin öğrenme çağında  
uzaktan algılanmış görüntülerin sınıflandırılması*

Erchan Aptoula  
Bilişim Teknolojileri Enstitüsü  
Gebze Teknik Üniversitesi

7 Eylül 2019

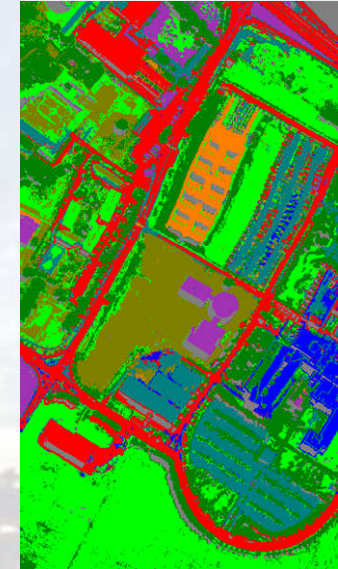
## Konu: piksel sınıflandırma



girdimiz

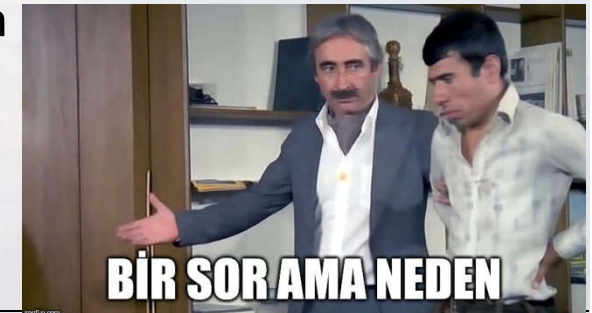


piksel  
sınıflandırma



sınıflandırma  
haritası

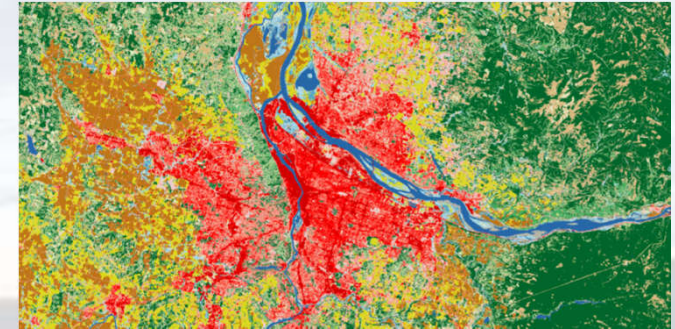
Görüntüdeki her piksele önceden tanımlı etiket kümesinden (örn. “asfalt”, “kestane ağacı ormanı”, “kiremit çatı”, vb) bir atama yapmak istiyoruz.



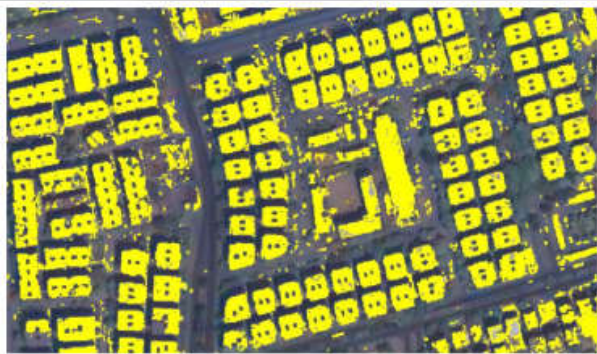
Çünkü sınıflandırma sonucu genelde birçok uygulamanın ilk ayağını oluşturuyor; örn. akıllı tarım, kirlilik takibi, sınır güvenliği, afet yönetimi, şehirleşme takibi, vb.

Dolayısıyla da:

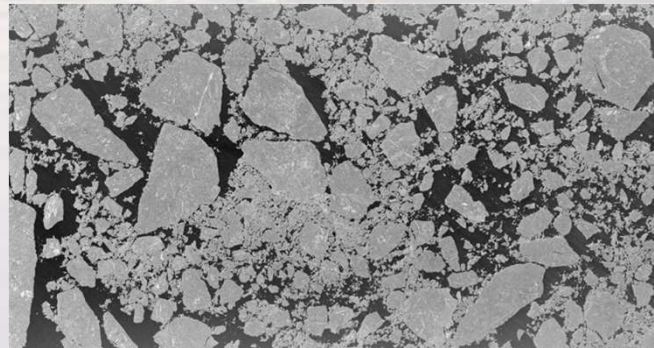
- Azami derecede isabetli,
- hızlı,
- ve otomatik olarak hesaplanması gerekmekte.



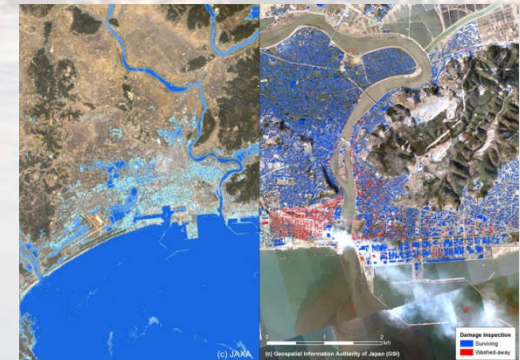
Arazi kullanım haritası



Bina tespiti



Çevre takibi



After yönetimi



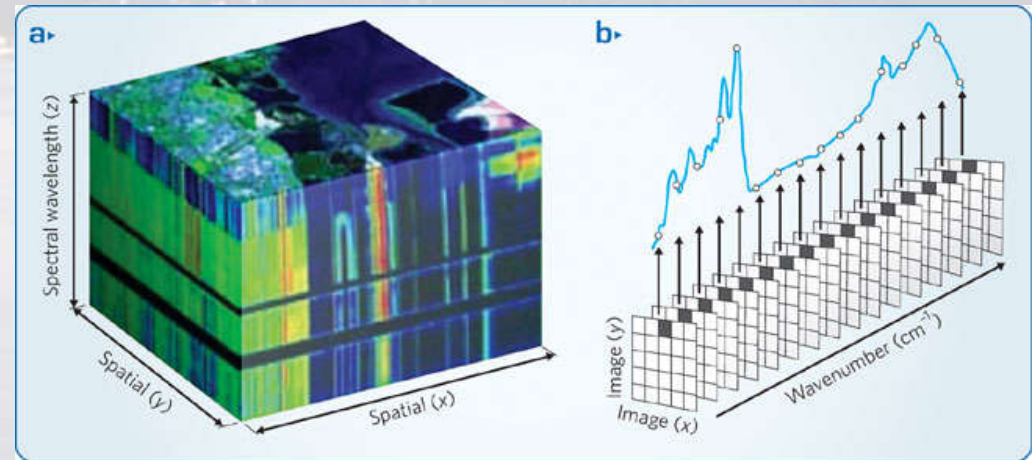
## Giriş

“Aman canım, ver işte birkaç bin etiketli eğitim pikselinin çevresindeki yamaları bir derin evrimsel sinir ağına olsun bitsin..”



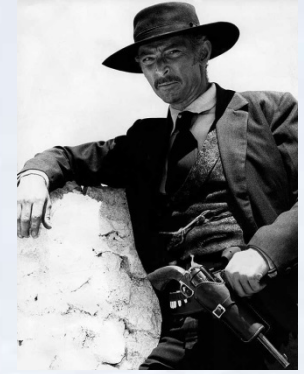
- **İşin “iyi” tarafı**

- Görüntüler 3 değil (RGB), yüzlerce farklı dalga boyu içerebilir (hiperspektral)
- Aynı sahne için hem optik, hem lidar görüntüsü olabilir – nasıl değerlendirilsin?
- Aynı sahnenin farklı tarihlerde çekilmiş görüntüleri olabilir (mevsimler boyunca tarım alanı görüntüleri gibi) – zaman boyutu nasıl değerlendirilecek?



## Giriş

- “kötü” tarafı
  - Görüntülerin çözünürlükleri ciddi değişim sergileyebiliyor;  
10 piksel bir tarla da olabilir, bir bina da, veya bir araba da...



Indian Pines, 20m/piksel

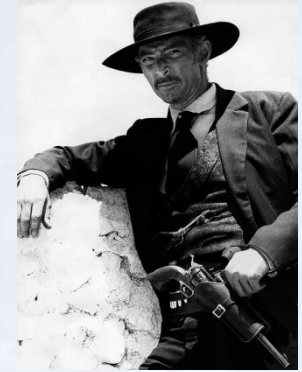


Pavia, 1.3m/piksel



Vaihingen, 0.23m/piksel

- “kötü” tarafı
  - Görüntülerin çözünürlükleri ciddi değişim sergileyebiliyor
  - Uzamsal verinin izgesel verinin yanı sıra kullanılması şart (nesnelerin boyutu belirsiz, piksel çevresinde küçük yamalar ile bir yere kadar...)



rastgele piksel  
düzeyinde permütasyon



izgesel sınıflandırma



aynı sonuç



izgesel sınıflandırma





- **ve “çirkin” tarafı**

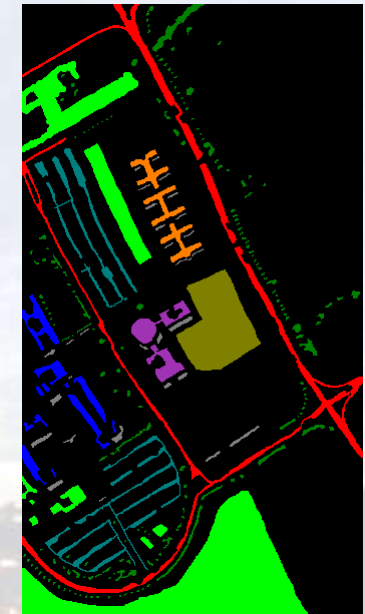
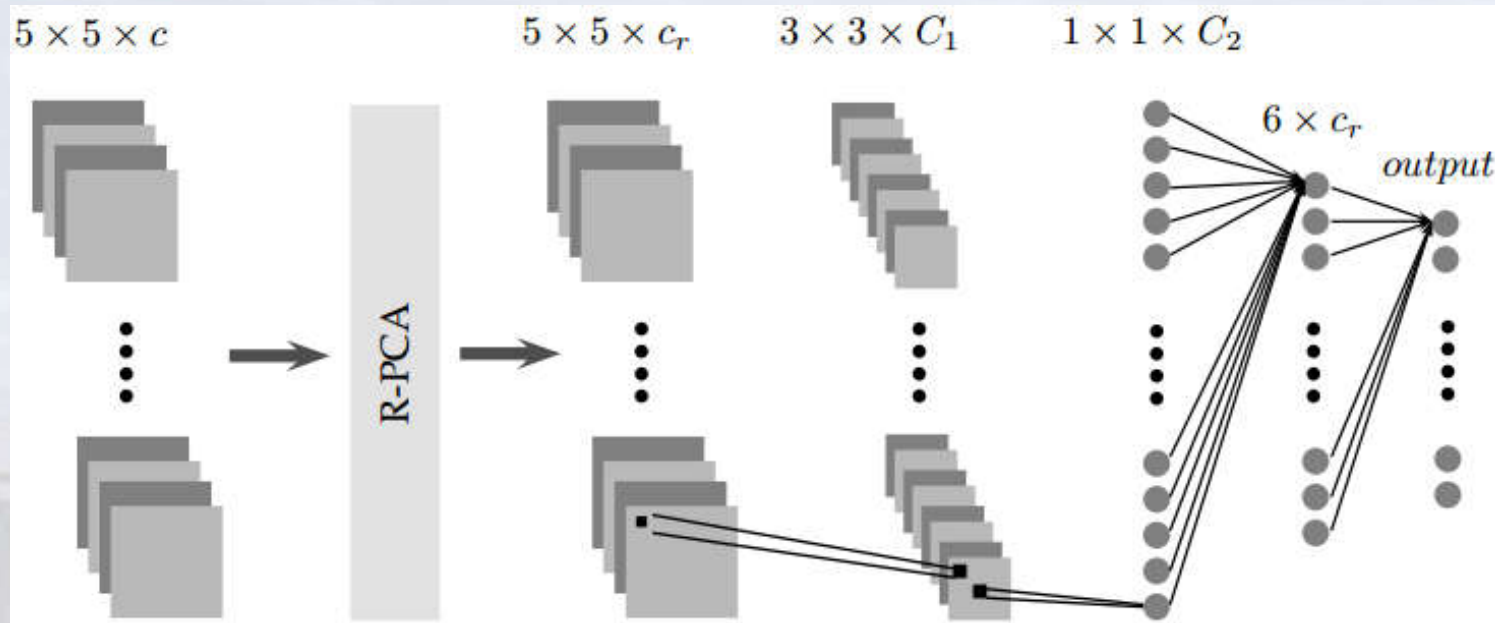
- Görüntü bantlarının yüksek sayıda olması hem karmaşıklığı artırıyor hem de bantlar arasında ciddi artıklık (*redundancy*) mevcut  
Ne yapalım? PCA, ICA, LDA, KPCA?



- Eğitim verisi genelde çok çok az..sınıf başına birkaç yüz veya bilemedin birkaç bin piksel örnek (sınıf sayısı 5 - 50 arası); zira görüntüler çok, uzmanlar az.
- Sınıflar genelde çok çok dengesiz; 10 bin piksel birinden, diğerinden ise 7.

## İlk teşebbüsler

### Makantasis vd. (2015): yamalar ile CNN eğitimi



Ağın girdisi piksel başına  $5 \times 5 \times c$  boyutlu tensörler. Yama boyutu çok küçük olduğundan «pooling» katmanları yer almamaktadır.

Makantasis vd, Deep supervised learning for hyperspectral data classification through convolutional neural networks, IGARSS, 2015



Yer doğrusunun piksellerinden rastgele seçilmiş %80 ile eğitiyorlar.

**Table 2.** Quantitative Evaluation Results. Comparison to SVM-based methods.

Dataset	Method (classification accuracy (%))				
	Our Approach	R-PCA RBF-SVM	R-PCA Linear-SVM	RBF-SVM	Linear-SVM
Pavia centre	<b>99.91</b>	98.87	97.63	99.01	97.86
Pavia university	<b>99.62</b>	93.82	84.39	93.94	84.67
Salinas	<b>99.53</b>	93.73	90.52	93.97	90.68
Indian Pines	<b>98.88</b>	82.71	79.47	82.79	79.56

Çok yüksek! İnanalım mı?

**Hayır!**

- Eğitim pikselleri rastgele seçilmiş. Komşuları sınama kümesi ile örtüşecek!
- Tek bir deney. Ne malum çok kolay bir %80-%20 ayrımı olmadığı?



Makantasis vd, Deep supervised learning for hyperspectral data classification through convolutional neural networks, IGARSS, 2015

**Yang vd. (2017)** ise uzamsal ve izgesel veriyi kullanarak iki bağımsız evrimsel sinir ağı eğitmiştir, tam bağlantılı katman aşamasında da onları kaynaştırmışlardır.

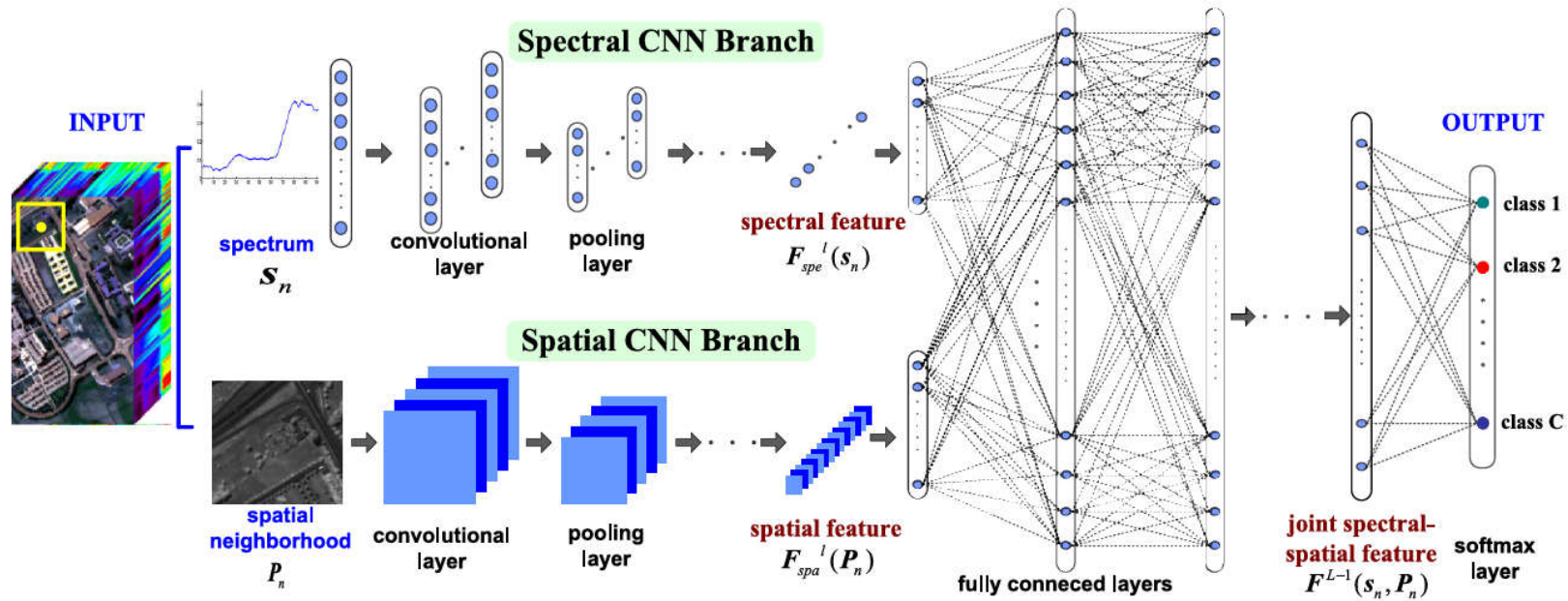


Fig. 3. Architecture of the proposed two branches of CNN for hyperspectral classification. (a) Pavia University data. (b) Salinas Valley data. (c) Spectral curves of different land-covers. (d) Spectral curves of same land-covers.

Yang vd, Learning and transferring deep joint spectral-spatial features for hyperspectral classification, TGRS, 2017

## İlk teşebbüsler

**Liu vd. (2018)** ise siyam ikizi evrimsel sinir ağlarını incelemiştir.

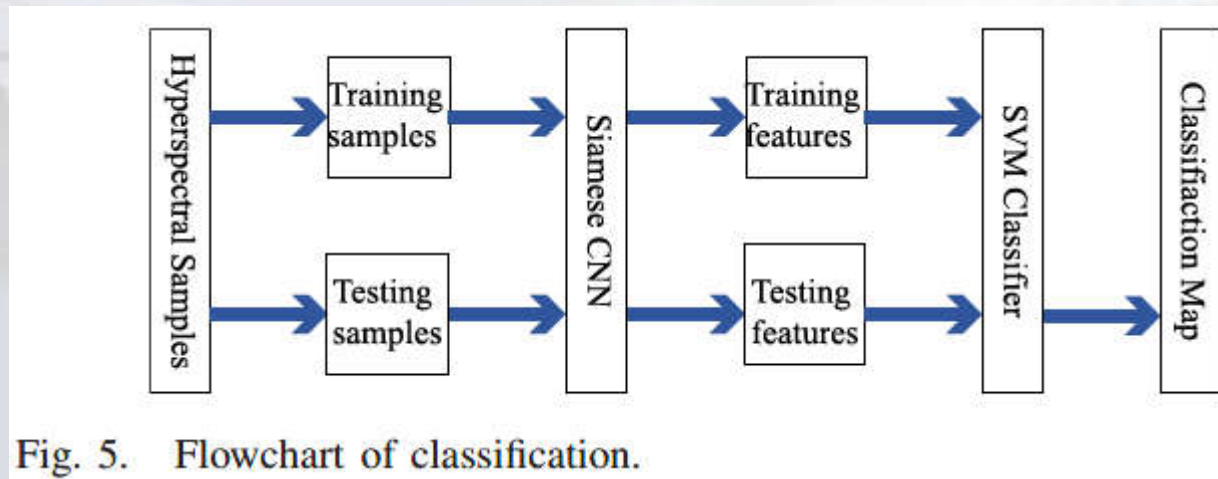
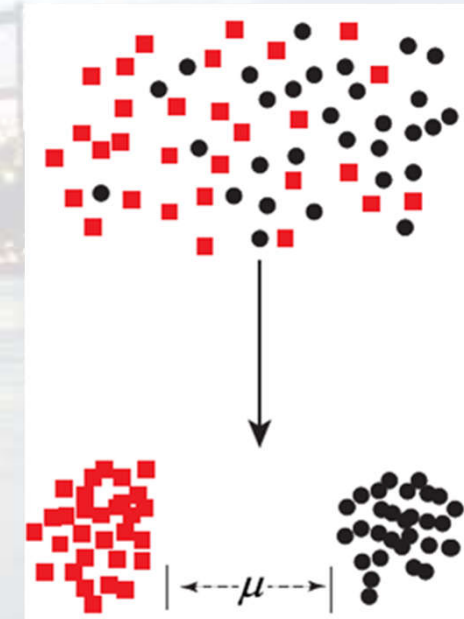


Fig. 5. Flowchart of classification.



Liu vd, Supervised deep feature extraction for hyperspectral image classification, TGRS, 2018



## Güncel yaklaşımlar

Ama bunun gibi yaklaşımlar yeterli olmadı.

- CNN'e her pikseli ayrı bir yama olarak sunmak çok yavaş
- Metrik öğrenme katkı sağlıyor, ama yetersiz.
- Otokodlayıcılar, derin inanç ağları..daha neler neler...



**Anlamsal bölütleme ağları** (*semantic segmentation networks*).

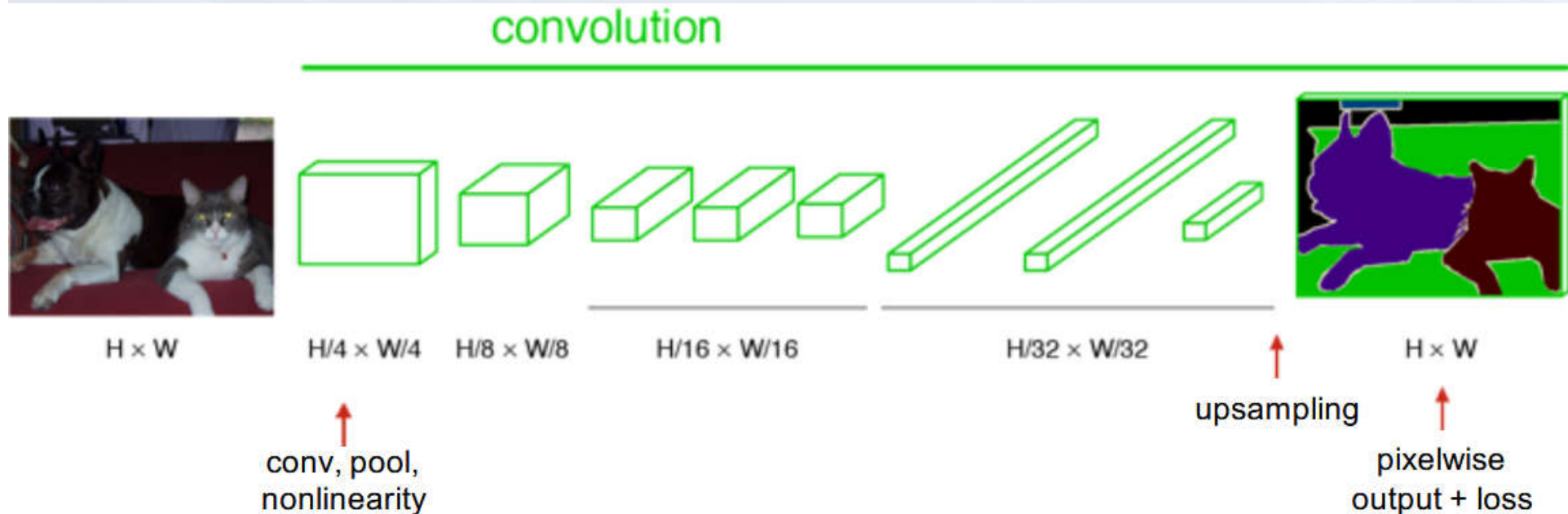
- Tüm görüntüyü **tek geçişte** sınıflandırıyor.
- Pratik olmanın ötesinde, **hızlı!**
- **Uzamsal** veriyi işlemeye uygun.

Anlamsal bölütleme dünyasında ufak bir gezintiye çıkalım.



## Anlamsal bölütleme: FCN

FCN: tam evrişimsel sinir ağları (*fully convolutional network*)

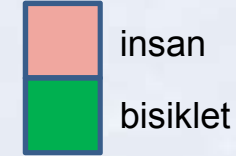
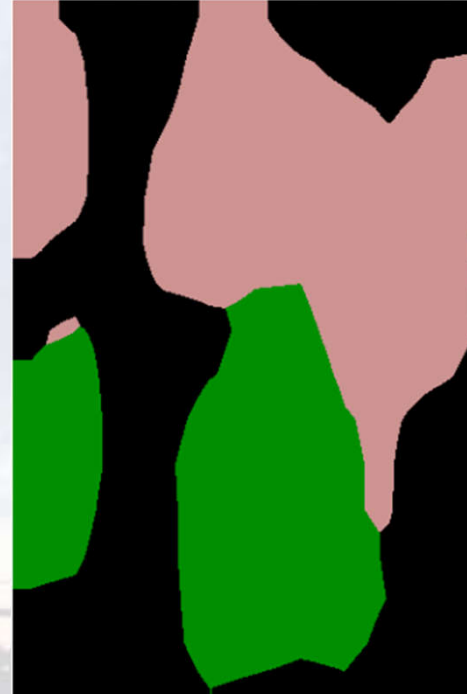


Öğrenilen nihai öznitelik vektörlerine komşuların katkısı var.

Ancak büyütürken? Dekonvolüsyon, çift doğrusal ara değer kestirimi, vb.

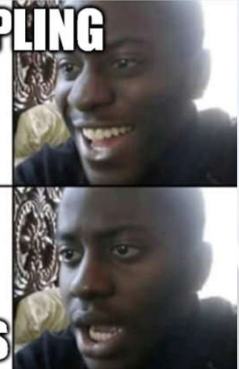
Long vd, Fully Convolutional Networks for Semantic Segmentation, CVPR, 2015

## Anlamsal bölütleme: FCN



SUBSAMPLING

AWFUL  
CONTOURS



Bir yanda ağın derinliği: ne kadar derin o kadar çok alt-örnekleme katmanı, ama o kadar çok da ağın zor/üst düzey sınıfları **tanıma yetkinliği** artıyor.

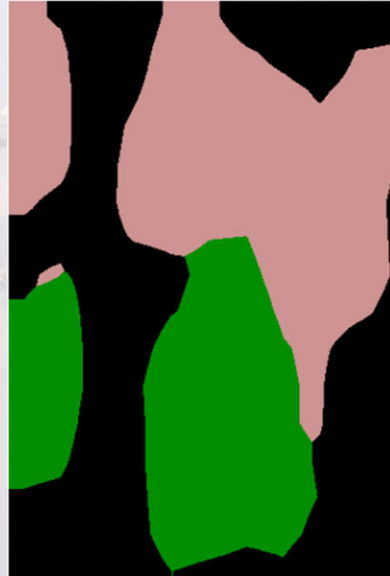
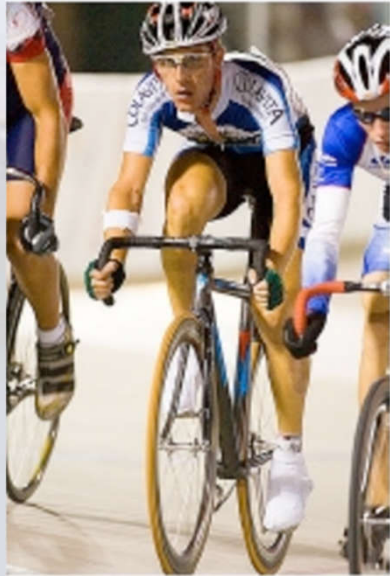
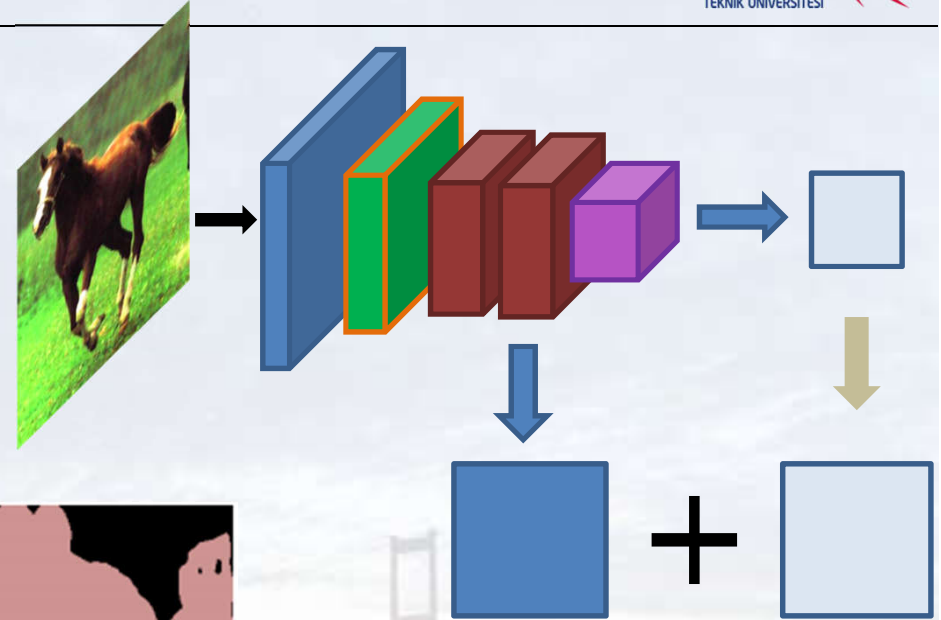
Diğer yandan konum isabetli: ne kadar çok alt-örnekleme katmanı, görüntü o kadar küçülüyor, haliyle tanınan nesnelerin **kenarlar isabeti** azalıyor.



## Anlamsal bölütleme: FCN

Atlamalı bağlantılar (*skip connections*)

Farklı çözünürlükte,  
farklı katmanların,  
ara değer kestirimi sonrası toplanıp,  
sınıfı topluca belirlemesi.



atlamasız



atlamalı



Long et al., Fully Convolutional Networks for Semantic Segmentation, CVPR, 2015

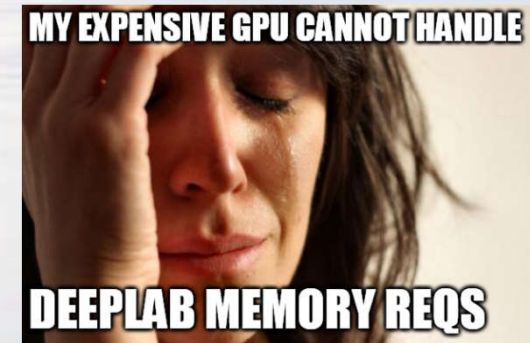
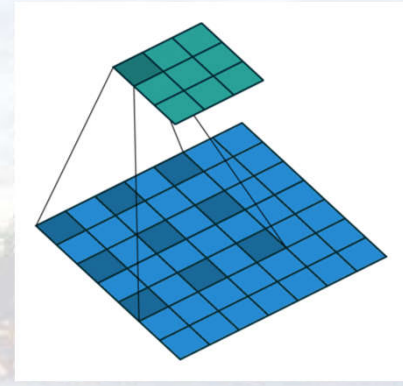
## Anlamsal bölütleme: DeepLab v1,v2

Sorunlar	Çözümler
Uzamsal hassasiyet, daha isabetli kenarlar	Koşullu rastgele alanlar (CRF)
Pooling görüntüleri küçültüp filtrelerin görüş açısını büyütür, ancak nesnelerin konum bilgisinin yozlaşmasına neden olur...	Genleşmiş ( <i>dilated</i> ) evrişimler

Genleşmiş evrişimler parametre sayısını artırmadan filtrelerin görüş açısını büyötmektedir.

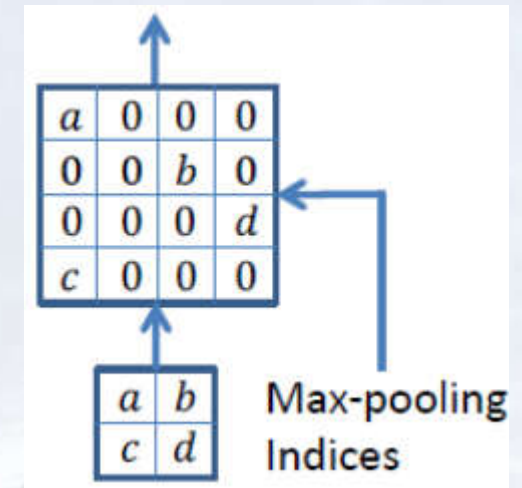
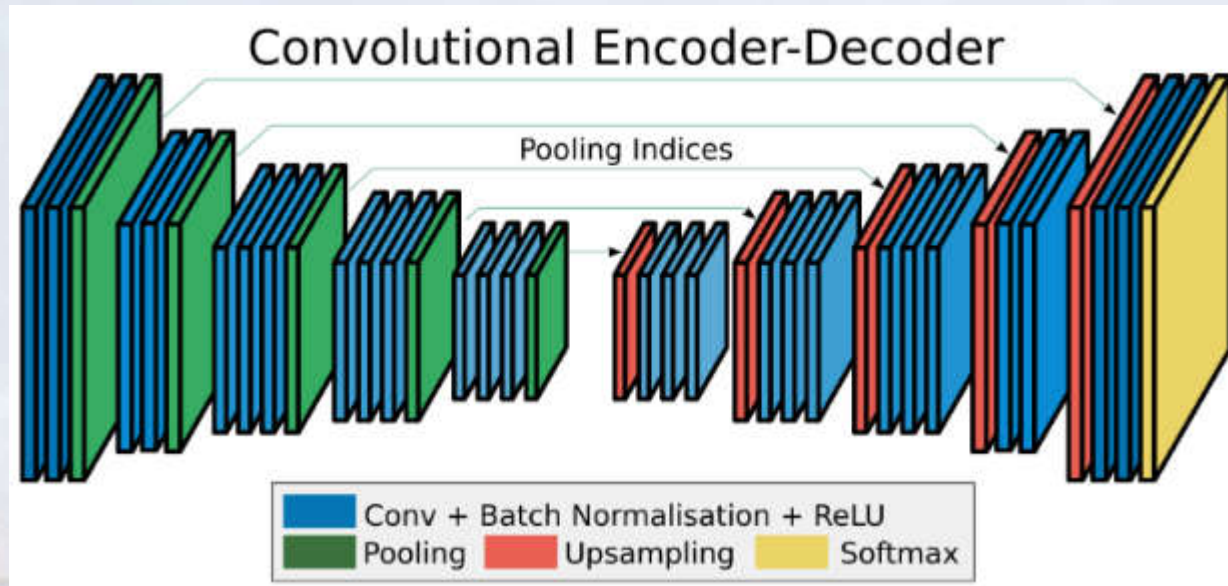
3x3'lük filtre delikler sayesinde 5x5'liğin görüş açısına sahip olabilmektedir.

Ancak artık alt-örnekleme sayısı azaldığından, **öznitelik haritaları çok daha büyük...**gelsin bellek yetersizliği!



Chen vd, DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, TPAMI, 2016

## Anlamsal bölütleme: SegNet



- + Düşük karmaşık model (hem eğitim süresi hem bellek tüketimi bakımından).
- + Özel ağırlıklı tekniklere gerek kalmadan uçtan uca eğitim.
- + İndis tabanlı ters ölçeklendirme sayesinde komşu piksellerin tahmine katkısı var.
- Büyük nesneler ile karşılaşınca sonuçlar kaba olabiliyor
- Çok nesneli kalabalık sahnelerde sonuçlar parlak olamayabiliyor

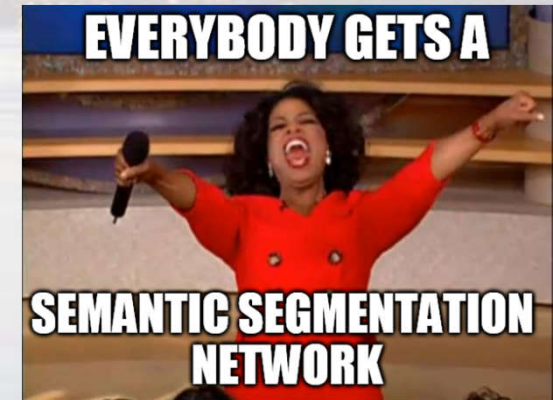
Badrinarayanan vd, SegNet: A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling, CVPR, 2015





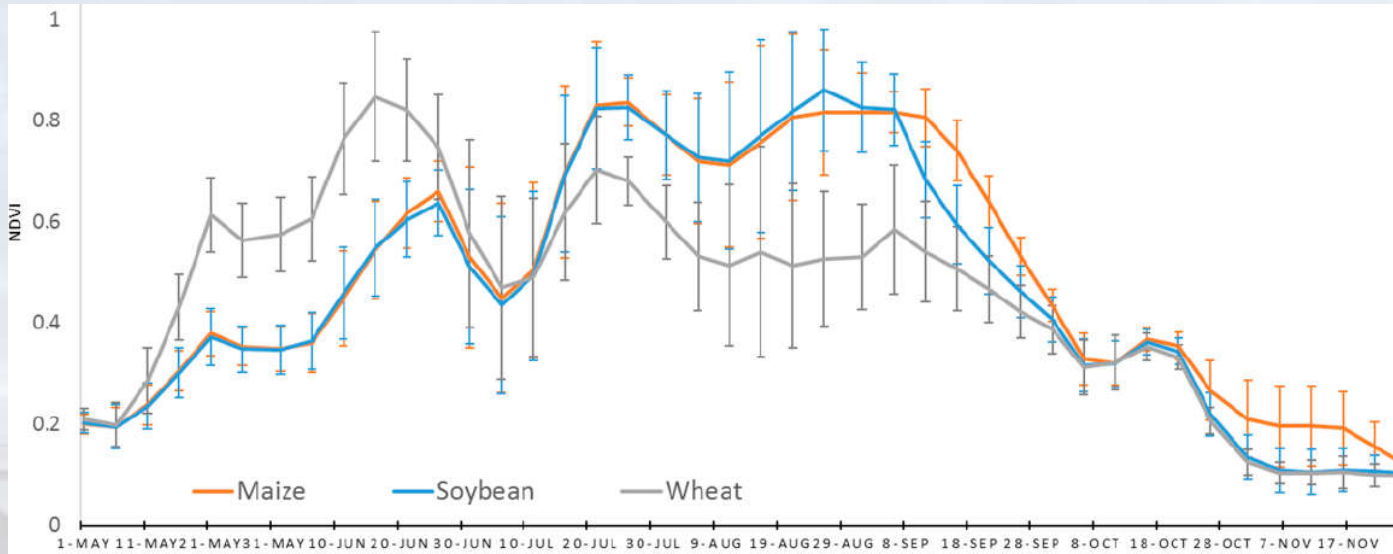
Figure 1: **DeepGlobe Challenges:** Example road extraction, building detection, and land cover classification training images superimposed on corresponding satellite images.

- Dereceye giren ilk 10 ekibin tamamı anlamsal bölütleme yaklaşımları kullanmış.
- Uzamsal veri işleme ağı yıkıldı..diğer mevzular ne olacak?



Demir vd, DeepGlobe 2018: A Challenge to Parse the Earth through Satellite Images, CVPR, 2018

Sentinel gibi uydular 10 günde bir tüm yeryüzünün görüntülerini almaktadır.



- Örneğin böylece farklı büyüme süreçlerine sahip bitkilerin sınıflarının ayırt edilmesi kolaylaşır.
- Aynı etiketleri farklı tarihlere de atayabilir ve ağı eğitebilirsiniz
- Aynı sahnenin farklı tarihli verilerini yığarak tensör olarak sunabilirsiniz.
- LSTM, RNN gibi zaman serileri işlemeye uygun özel mimarilere yönelebilirsiniz.



## Sınıf dengesizliği

Eğitim kümesindeki bazı sınıflar doğal olarak baskın olacaktır. Sınıflandırıcının yanlış her piksele o baskın sınıfı atamasını engellemek için:

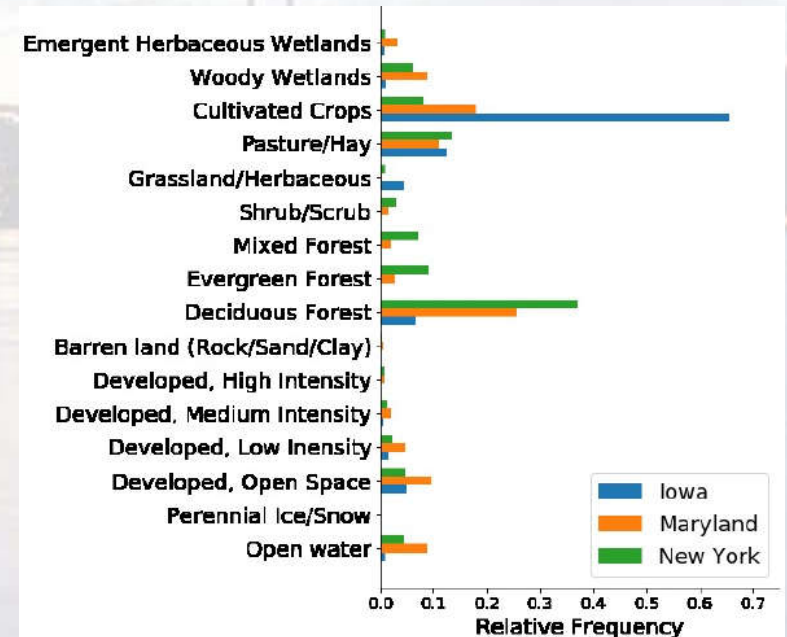
1) Ağırlıklar kullanıp, nadir sınıf piksellerinin hataları daha yüksek kayba sebep olabilir.

2) Focal loss gibi kayıp fonksiyonları kullanılabilir. Kolay sınıfların kayba etkisini azaltır (örnek sayısı) çok olsa da.

3) Eksik sınıfları sayısı yapay olarak artırılabilir

4) Baskın sınıf tek ise, önce onu tanıyıp, sonra diğerlerine odaklanabilirsiniz.

Not: başarıyı ölçümü için kappa istatistiği önerilir.





## Uygulama örneği

Görevimizi çiftçi bildirimlerinin doğrulanması amacıyla sınıflandırma haritası üretimi.

Gediz havzası, 10 sınıf:

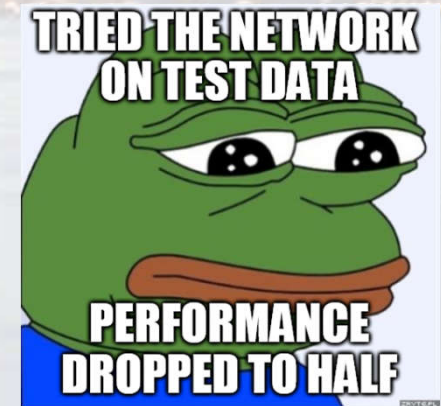
Pamuk, buğday, mısır, yonca, tütün, susam, bakla, papates, soğan ve fasülye.

Menemen bölgesi tüm sınıflara sahip.

Verisi de **bol** ve sınıflar **dengeli**.

Eğitim ve sinama olarak sınıf temelinde 80/20 ayırdık; doğrulama düzeyi: kappa 0.9

Gediz havzasının geri kalanına aynı sınıflandırıcıyı aynı veriler ile uygulayınca, başarıım 0.4

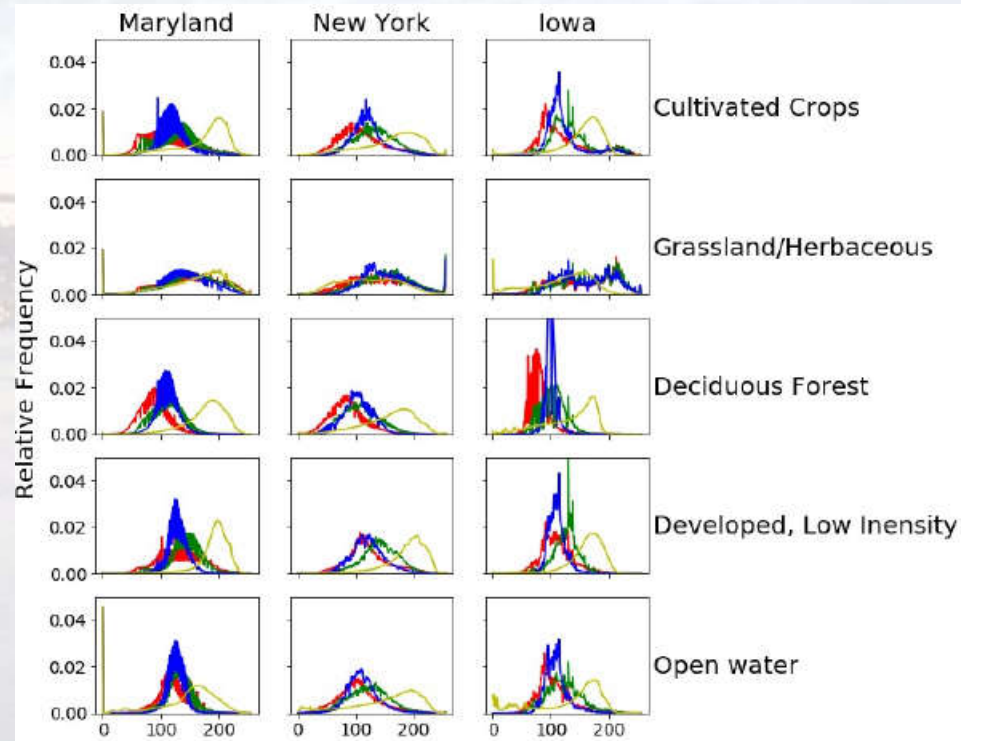
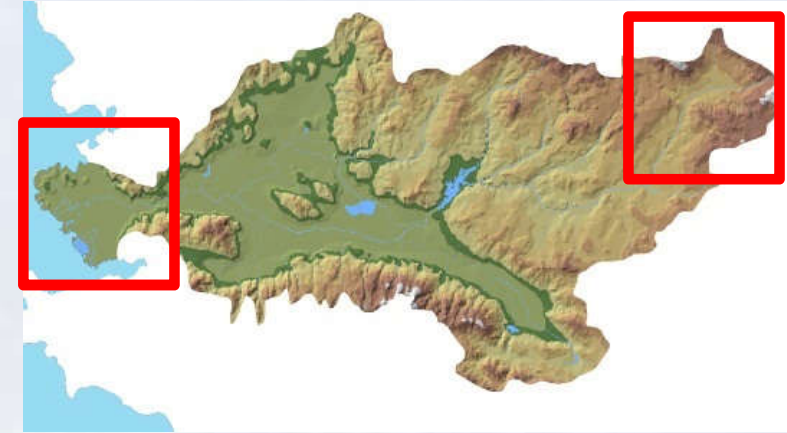
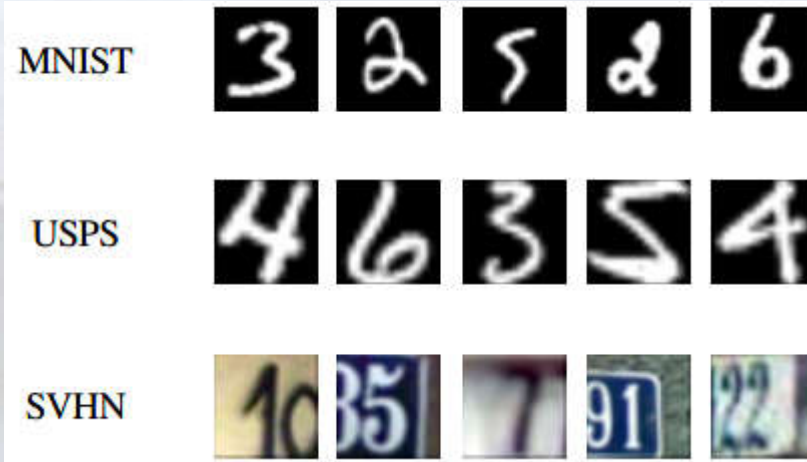


## Uygulama örneği

### Sorun bağlam!

Sınıflar aynı olsa da bağlam değişince en kral  
sinir ağı bile çaresiz

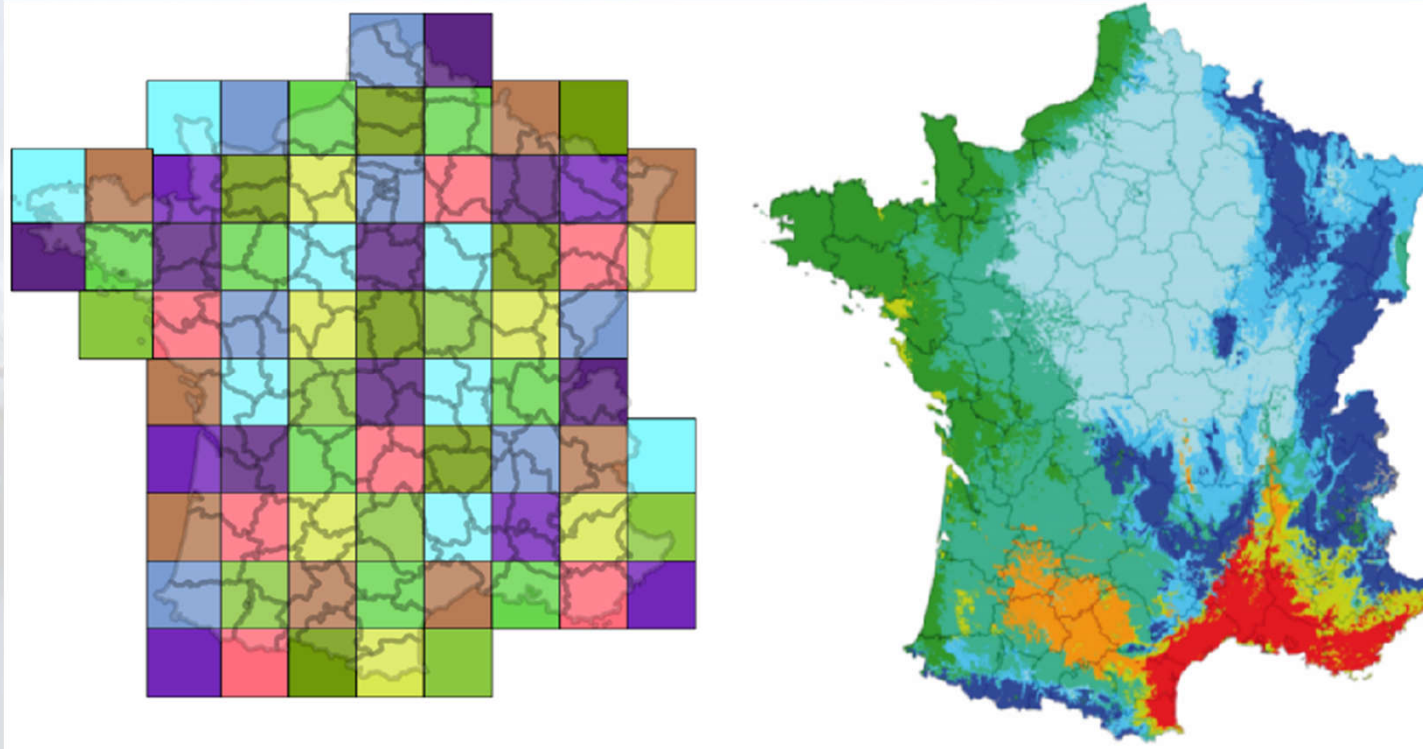
### MNIST örneği



## Uygulama örneği

Peki bağlamdan bağımsız olmak adına ne yapabiliriz?

- Kolay yolu, **eğer varsa** diğer bağlamlardan da eğitim verisi kullanımı

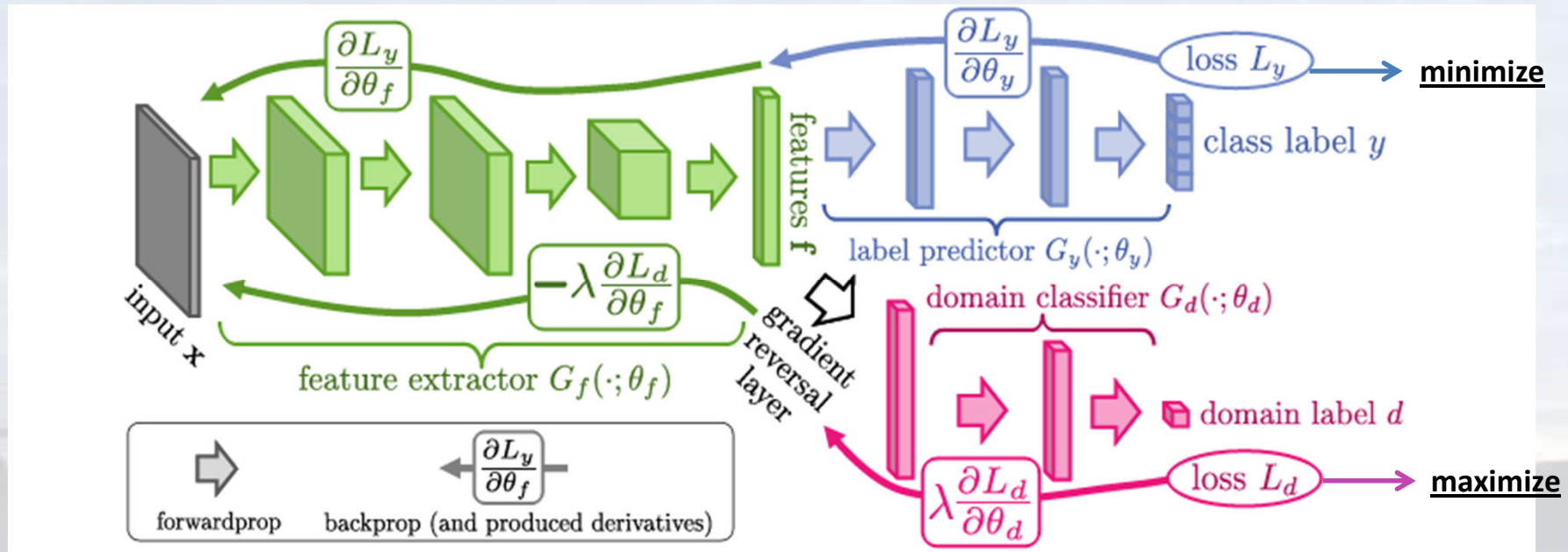


- Fakat yoksa?



## Uygulama örneği

Eğer hedef bağlamın eğitim verisi yoksa: **bağlam/alan çekişmeli ağlar**



Method	MNIST → USPS 173 → 105	USPS → MNIST 105 → 173	SVHN → MNIST 1435 → 173
Source only	0.752 ± 0.016	0.571 ± 0.017	0.601 ± 0.011
Gradient reversal	0.771 ± 0.018	0.730 ± 0.020	0.739 [16]
Domain confusion	0.791 ± 0.005	0.665 ± 0.033	0.681 ± 0.003
CoGAN	0.912 ± 0.008	0.891 ± 0.008	did not converge
ADDA (Ours)	0.894 ± 0.002	0.901 ± 0.008	0.760 ± 0.018

Table 2: Experimental results on unsupervised adaptation among MNIST, USPS, and SVHN.



Ganin vd, Domain adversarial training of neural networks, JMLR, 2016

- **Uzaktan algılama derin öğrenmenin en iyi müşterilerinden.**
  - Veri çok, durmadan geliyor.
  - Sorun çok: sınıflandırma, nesne/hedef/değişim tespiti, geri erişim, vb
- **Sınıflandırma bağlamında anlamsal bölütleme ağları gerçekten etkili**
  - Bütün görüntüyü tek geçişte sınıflandırabilirler, eğitimi ve sınaması hızlı
  - Uzamsal verinin işlenmesinde başarılılar.
- **Açık sorular**
  - Farklı kipleri (çözünürlük, algılayıcı türü, bantları) nasıl değerlendirmeli?
  - Çok büyük ölçekte (ülke, kıta, dünya) ölçeklenebilir yaklaşım şartı

**Dikkatiniz için teşekkür ederim.**

**eaptoula@gtu.edu.tr**  
**<http://bte.gtu.edu.tr/~eaptoula/>**