

Makine Öğrenmesi-1

Akış

- **Makine Öğrenmesi nedir?**
- **Günlük Hayatımızdaki Uygulamaları**
- **Verilerin Sayısallaştırılması**
- **Özellik Belirleme**
 - Özellik Seçim Metotları
 - Bilgi Kazancı (Informaiton Gain-IG)
 - Sinyalin Gürültüye Oranı: (S2N ratio)
 - Alt küme seçiciler (Wrappers)
 - Yeni Özelliklerin Çıkarımı
 - Temel Bileşen Analizi (Principal Component Analysis)
 - Doğrusal Ayırte den Analizi (Linear Discriminant Analysis)
- **Sınıflandırma Metotları**
 - Doğrusal Regresyon
 - Karar Ağaçları (Decision Trees)
 - Yapay Sinir Ağları
 - En Yakın K Komşu Algoritması (k - Nearest Neighbor)
 - Öğrenmeli Vektör Kuantalama (Learning Vector Quantization)
- **Kümeleme Algoritmaları:**
 - Hiyerarşik Kümeleme
 - K-means
 - Kendi Kendini Düzenleyen Haritalar (Self Organizing Map -SOM)
 - DBscan
- **Regresyon Algoritmaları**
- **Çok Boyutlu Verilerle Çalışmak**
- **Veri Sızıntısı**
- **Pekiştirmeli Öğrenme**

Makine Öğrenmesi

Çok büyük miktardaki verilerin elle işlenmesi, analizinin yapılması mümkün değildir.

Bu tür problemlere çözüm bulmak amacıyla makine öğrenmesi metotları geliştirilmiştir.

Bu metotlar

geçmişteki verileri kullanarak

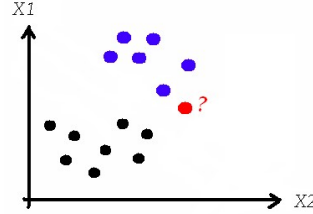
veriye en uygun **modeli** bulmaya çalışırlar.

Yeni gelen verileri de bu modele göre analiz edip sonuç üretirler.

Metot türleri

- Farklı uygulamaların analizlerden farklı beklentileri olmaktadır.
- Makine öğrenmesi metotlarını bu beklentilere göre sınıflandırmak mümkündür.

Sınıflandırma

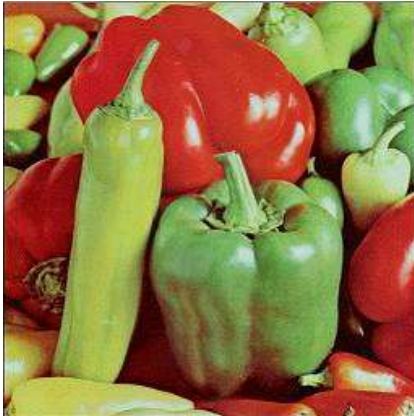


x1	x2	x3	x4	x5	Sınıf
1.2	3.4	2.3	6.7	7.8	A
1	2	2	4	5	A
9	8	6	5	3	B
6	7	8	9	1	B
8	8	9	9	9	B

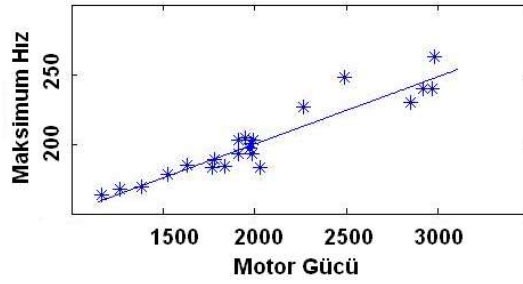
Kümeleme

- 256 rengi 16 renge nasıl indiririz?
- $n \text{ bit} = 2^n$ farklı değer

x1	x2	x3	x4	x5
1	3	2	6	7
1	2	2	4	5
9	8	6	5	3
6	7	8	9	1
8	8	9	9	9



Regresyon Eğri Uydurma



Motor gücü	Max hız
1500	230
2500	280
1600	210
3500	290
1200	210

Pekiştirmeli Öğrenme

- Modele çıkışının ne olması gerektiği söylenmez sadece doğru ya da yanlış yaptığı söylenir. ML3'te göreceğiz.

Birliktelik Kuralları Keşfi

- Sepet analizi
 - hangi ürünler birlikte satılıyor?
- Raf düzenlemesi
 - hangi ürünler yan yana konmalı?
- Promosyonlar
 - neyin yanında ne verilmeli?

Makine Öğrenmesinin

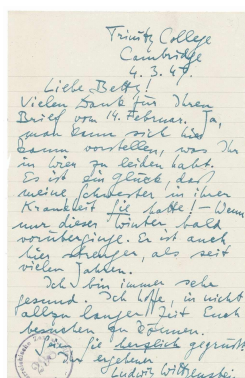
Günlük Hayatımızdaki Uygulamaları

Dolar kuru tahmini

- 6.1 6.2 6.3 6.3 6.3 6.2 6.4 6.1

D-3	D-2	D-1	D
6.1	6.2	6.3	6.3
6.2	6.3	6.3	6.3
6.3	6.3	6.3	6.2
6.3	6.3	6.2	6.4
6.3	6.2	6.4	6.1

El yazısı / Kitap Yazısı Tanıma HCR /OCR



The main purpose works that tried to address the problem of widely supervised multi-line recognition were [1], [2]. Besides these methods, other methods that work on full page recognition require the localization of ground truth of each line during training, a detailed comparison between the training data required by our proposed method vs. other methods is further presented in Table 1.

In this work, we present a simple and novel neural network architecture, named *OrganicNet*, that can be added to any existing convolutional neural network (CNN) text line recognition framework to a full page recognition. It can recognize full text pages in a widely supervised manner with the few given line localization ground truth (other usual in the image) is enough to be supervised during training, and without performing any explicit segmentation. In contrast to previous work, this is a very efficiently using feed-forward connections only (no recurrent connections), essentially as a single neural forward pass.

Our main intuition in this work is, instead of the traditional two-stage framework that first requires line segmentation followed by a separate neural network approach for learning to simultaneously implicitly segment and recognize. This works by learning a representation transformation that transforms the input into a representation where both segmentation and recognition is shared.

We explicitly define an input multi-line image into a single-line image (i.e. from a 2D representation of characters to a 1D representation). In the original image, we match together lines on long line, two-line representation is actually similar. Both representations are compared on their own in the same single-line image network forward pass instead of being carried out separately two each line, and thus all computations are shared between recognition and explicit segmentation, and the whole process is fast.

The main requirement to achieve this are: Using the idea of a spatial bottleneck followed by a simple, and using the CTC loss function [3] which strongly indicates a simple 1D input. We consider a simple neural network with multiple layers that apply these neural ideas, and demonstrate both in effectiveness and generality by applying it to a number of state-of-the-art text recognition neural network architectures. We show that our neural network can learn to recognize text lines without any localization information with exactly the same training procedure (i.e. without requiring to explicitly and highly training expert, this is a general training procedure or shared pre-training strategies).

On the challenge of KITD-2017 RTR [1] full page benchmark we achieve state-of-the-art Character Error Rate (CER) without any localization data. On full page recognition of the IAM dataset, we even gain an absolute state-of-the-art CER learning results that work on carefully pre-segmented text lines, without using any localization information during training or testing.

In summary, we address the problem of widely supervised full page text recognition. In particular, we make the following contributions:

- We conceptually propose a new approach for widely supervised, segmentation, other segmentation and recognition, and apply it to text.
- We propose a simple and generic neural network sub-module that can be added to any CNN-based text line recognition framework to a multi-line text recognition that address the same simple training procedure.
- We carry out extensive sets of experiments on a number of state-of-the-art text recognition benchmarks and datasets. The achieved performance demonstrates state-of-the-art performance on KITD-2017 RTR and the full paragraph IAM dataset.

2. Related Work

There is one such prior work in the literature regarding full page recognition. Segmentation-free multi-line recognition has been mainly considered in [1], [2]. The idea of both is to use selective attention to focus only on a specific part of the input image, either character in [1] or line in [2]. These works have two major drawbacks. First, both are difficult to train and deploy, and their neural network on single-line images before training on multi-line images, which reduce the efficiency of the work. Second, though [1] is much faster than [2], both are very slow compared to neural networks that learn on segmented text lines.

Besides these two segmentation-free methods, other methods that work on full-page recognition either require the localization of ground truth of each line (e.g. [4]), or part [1] of the training data to train either a separate network or a sub-module on all the provided neural ground truth line localization. Also, all these methods require line breaks to be annotated on all the provided neural ground truth line recognition (i.e. not lines must be segmented both results in the image and manually in the transcription [1]).

On the other hand, [5] presents the idea of a widely supervised manner without requiring line breaks in the transcription by using the alignment between the provided line transparency and the ground truth as a confidence optimization problem, and finally solving it. However [5] still requires the same pre-training as [1] and performs worse.

3. Methodology

Figure 1 presents the core idea of our proposed *OrganicNet* model, and how it can be attached to any fully convolutional neural network. Both before and after our work are shown for easy comparison.

The Convolutional Temporal Classification (CTC) loss function allows the training of neural text recognition on

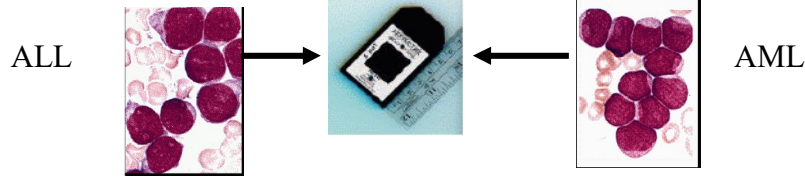
İşlem: Şekillerin hangi harf olduğunu tahmin etme

Kredi Taleplerini Değerlendirme

- Birisi bankadan borç ister.
- Banka borcu versin/vermesin.
- Nasıl?

e-ticaret

- Birisi Amazon.com dan bir kitap yada ürün alıyor.
Müşteriye alması muhtemel kitapları önerelim.
Ama nasıl?
Kitapları
konularına
yazarlarına
birlikte satışlarına
göre kümelemek?



Gen Mikrodizilimleri

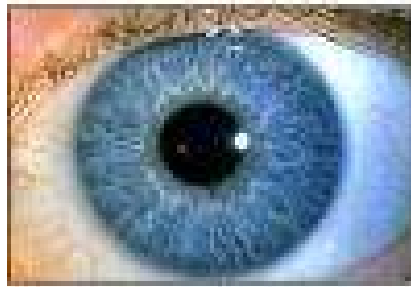
100 kişinin (hasta/sağlam) elimizde gen dizilimleri var.

Bu dizilimleri analiz ederek hasta olup olmadığı bilinmeyen birisinin hasta olup olmadığını yada hastalığının türünü öğrenebilir miyiz?

En iyi tedaviyi önerebilir miyiz?

Nasıl? Elimizde hangi bilgiler olmalı?

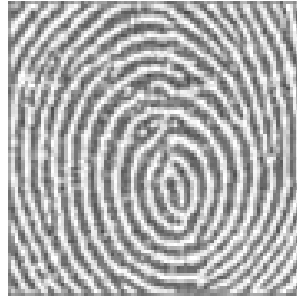
Bu adam kim? İçeri girsin mi?



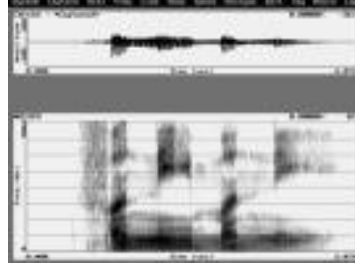
Bu adam kim?
Bu adam havaalanında mı?



Bu parmak izi kimin?
Bu adamı tutuklayalım mı?



Bu ses kimin?
Bu ses ne diyor?



Bu imza kimin? Yoksa taklit mi?

Taklit olup olmadığını nasıl anlarız?
Zaman bilgisi ?

Bu metnin konusu nedir? Bu mail spam mi?



Anti spam yazılımları nasıl çalışır?

Spamciler nasıl çalışıyor?

Yeni nesil spam mailler: Mesaj resimde,

metinde ise anti spamlardan kaçmak için gereken kelimeler var.

Makine öğrenmesi metotlarını hem spamciler hem anti spamciler kullanıyor.

Olağan dışı bir durum var mı? Güvenlik kamerası kayıtları



Kamera kaydındaki kiři ne anlatıyor? İřaret Dili Tanıma



Mehmet Fatih AMASYALI Yapay Zeka Ders Notları

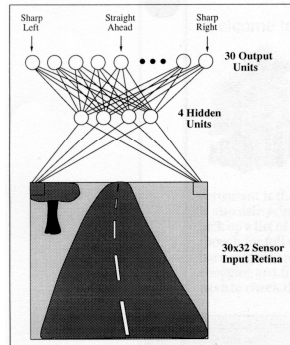
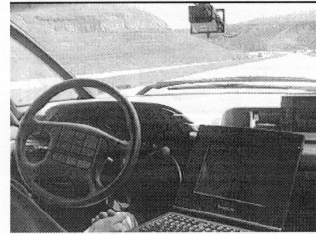
YILDIZ TEKNİK ÜNİVERSİTESİ BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

ALVIN*

Otobanda saatte 70 mil hızla **sürücüsüz** gidebilen bir otomobil

Bütün denemeler **trafiğe kapalı alanlarda** gerçekleştirilmiştir 😊

Neden şehir içi değil?
Neden otoban?
Neden diğer arabalar yok?
Araba birine çarparsa suçlu kim?



[*] <https://papers.nips.cc/paper/1988/file/812b4ba287f5ee0bc9d43bbf5bbe87fb-Paper.pdf>

Mehmet Fatih AMASYALI Yapay Zeka Ders Notları

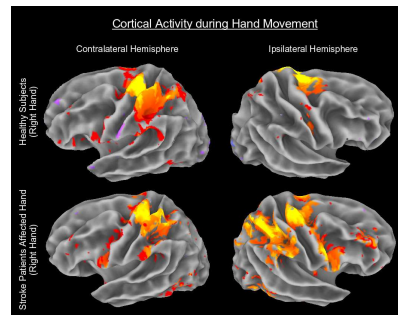
YILDIZ TEKNİK ÜNİVERSİTESİ BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Adalet

- Çin’de pilot uygulama:
 - bir şehrin mahkeme hakimleri bir bilgisayar programı
 - Amaç: Daha adil bir dünya
 - Aynı özelliklere sahip davalarda aynı kararların alınması
 - Sistemin eğitimi için neler gerekli?
 - Milyonlarca/Milyarlarca (orası Çin) davaya ait verilerin kategorilenmesi

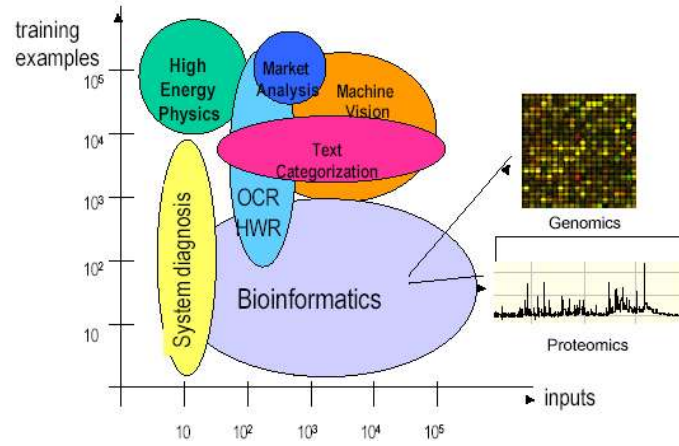
Beyin Aktiviteleri

- İnsanların
 - değişik şeyler düşünürkenki,
 - değişik duygulara sahipkenki,
 - problem çözerken ki beyin aktiviteleri kaydedilir.
- Görev?



[*] <https://www.nmr.mgh.harvard.edu/neurorecovery/technology.htm>

Uygulamalardaki boyut örnek sayıları



[*] Isabelle Guyon, <http://clopinet.com/isabelle/Projects/ETH/lecture1.pdf>

Mehmet Fatih AMASYALI Yapay Zeka Ders Notları

YILDIZ TEKNİK ÜNİVERSİTESİ BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Öğrenmenin geleceği

- Öğrenme modülleri birçok uygulamaya girmiş durumda.
- Ticari olarak satılan (gerçek dünya için tasarlanmış) birçok ürün var.
- Bu tür ürün ve modüllerin sayısı giderek artmakta.
- Çözüm bekleyen ve büyük potansiyelleri olan birçok problem (?) var.

Mehmet Fatih AMASYALI Yapay Zeka Ders Notları

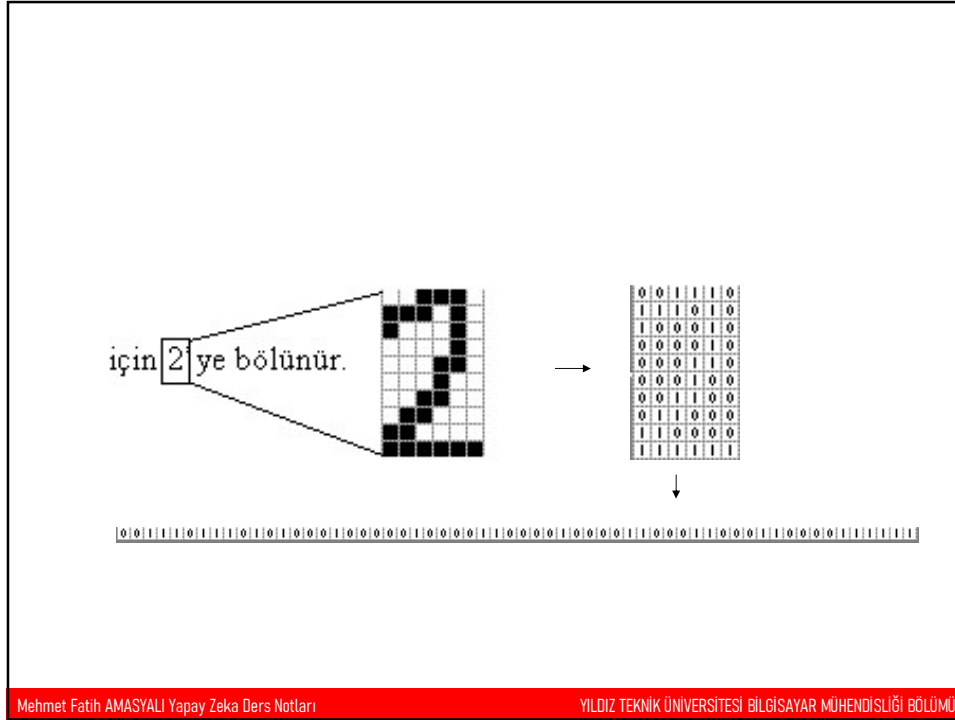
YILDIZ TEKNİK ÜNİVERSİTESİ BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Verilerin Sayısallaştırılması

Resim	Resmin her bir pikselinin renkli resimlerde R,G,B değerleri, siyah-beyaz resimlerde 1–255 arası gri seviyesi kullanılarak sayılara çevrilir. Renkli resimler 3 adet, siyah beyazlar 1 adet en*boy büyüklüğünde matrisle ifade edilir.
Metin	Metindeki harfler, heceler ve kelimeler genelde frekanslarına göre kodlanarak sayılara çevrilir.
Hareketli görüntü	Resim bilgisine ek olarak resmin hangi resimden sonra geldiğini gösteren zaman bilgisini de içerir. Bu ek bilgi haricinde yapılan işlem resim ile aynıdır.
Ses	Ses, genlik ve frekansın zaman içinde değişimiyle kodlanır.

Örnek sınıflandırma uygulaması

- Sistem: bir kitap fotokopisinin içindeki yazılarının metne dönüştürülmesi
- Öncelikle metindeki satırlar bulunur.
- Her bir satırdaki harfler bulunur. Her harfe ait onlarca örnek resimden etiketlenmiş bir veritabanı oluşturulur. Her bir resim için



- Bu şekilde tanınmak istenen harf için çeşitli fontlarla yazılmış birçok örneği temsil eden 60 boyutlu vektörler elde edilir.
- Bu uygulamamız için özellik sayımız 60'tır. Diğer bir deyişle örneklerimiz 60 boyutlu bir uzayda temsil edilmektedirler.
- Elimizde 10 rakama ait farklı fontlarla yazılmış 10'ar resim olursa veri kümemiz 100 örnek* 60 boyutluk bir matris olacaktır.
- Elimizde her örneğin hangi harf olduğunu gösteren sınıf bilgiside bulunmaktadır.
- Bu matris eğitim ve test kümesi oluşturmak için 2'ye bölünür.
- Eğitim kümesi bir sınıflandırıcıya verilir.
- Modellenir.
- Modelin başarısını ölçmek için sınıflandırıcının daha önce görmediği, modelini oluşturmakta kullanmadığı test kümesi için tahminde bulunması istenir.
- Bu tahminlerle gerçek sınıfların aynılığının ölçüsü sınıflandırıcının başarı ölçüsüdür.