Bilgisayarlı Görü

Yapay Zekâ ve Veri Dersleri Rehberi · Kısım III: Sinyal, Görüntü ve Ses

Görüntüden anlam | Öznitelikler | Tanıma ve nesne bulma | Kamera ve 3B | Video ve hareket | Komşu dersler

Sercan KÜLCÜ · Giresun Üniversitesi · Bilgisayar Mühendisliği

© Sercan Külcü. Tüm hakları saklıdır. · Sunumun etkileşimli HTML sürümü: https://sercankulcu.github.io/files/ai_courses/sunumlar/Bolum_18_Bilgisayarli_Goru.html

GİRİŞ

Ders Hangi Soruyu Sorar?

> **Bilgisayar bir görüntüde ya da videoda ne olduğunu nasıl anlar?**

Bilgisayarlı görü, bir görüntüden ya da videodan **anlam çıkarmayı** inceler: hangi nesneler var, nerede, sahne nasıl bir üç boyutlu yapıya sahip. Girdi görüntüdür ama çıktı görüntü değil; etiket, kutu, maske ya da derinliktir. Görüntüyü dönüştürmek görüntü işlemenin işidir.

Bir bakışta

YAKLAŞIM görüntüden anlam çıkarma	**GİRDİ** görüntü ya da video	**ÇIKTI** etiket, kutu, derinlik, 3B yapı
BAŞARI ÖLÇÜTÜ doğruluk, mAP, IoU	**MATEMATİK YÜKÜ** yüksek: geometri, doğrusal cebir	**VERİ İHTİYACI** çok etiketli görüntü

GİRİŞ

Kısa Tarihçe

1963
Larry Roberts'ın doktora tezi, çizgi çizimlerinden üç boyutlu blokları çıkarır: alanın ilk çalışması.

1966
MIT'de "Summer Vision Project": görünün bir yaz projesiyle çözülebileceği düşünülür; çözülmez.

1982
Marr'ın Vision kitabı yayımlanır: görüyü katmanlı bir hesaplama süreci olarak tanımlar.

1999
Lowe SIFT özniteliklerini tanıtır: ölçek ve dönmeden etkilenmeyen yerel nokta eşleme.

2001
Viola ve Jones gerçek zamanlı yüz bulucuyu yayımlar; dijital fotoğraf makinelerine girer.

2012
AlexNet ImageNet yarışmasını büyük farkla kazanır; görüde derin öğrenme dönemi başlar.

DERS

Temel Kavramlar

Yerel öznitelikler
Köşe ve anahtar noktalar: Harris, SIFT, ORB.

Sınıflandırma
Görüntünün bütününe etiket vermek: "kedi".

Nesne bulma
Her nesneye kutu ve etiket: YOLO, Faster R-CNN.

Anlamsal bölütleme
Her piksele bir sınıf: yol, yaya, araç.

Kamera modeli
İğne deliği modeli, kalibrasyon, iç ve dış parametre.

Çoklu görüş
Stereo, epipolar geometri, hareketten yapı.

Hareket ve izleme
Optik akış; videoda nesneyi kareden kareye izleme.

Evrişimli ağlar
CNN ve görüntü dönüştürücüleri (ViT).

DERS

Tipik İş Akışı



Örnek: Sürücü destek sistemi ön kameradan kare alır, parlaklığı dengeler, evrişimli bir ağla öznitelik çıkarır ve yayaları kutularla işaretler; stereo kameradan uzaklıklarını da kestirir. Çıktı bir görüntü değil, "8 metrede yaya var" bilgisidir; sistem buna göre fren uyarısı verir.

DERS

Derste Neler İşlenir?

1. Görüntü oluşumu ve kamera modeli
2. Süzgeçler ve kenar bulma (görüntü işleme tekrarı)
3. Yerel öznitelikler: köşe, SIFT, ORB
4. Öznitelik eşleme ve RANSAC
5. Homografi ve panorama birleştirme
6. Kamera kalibrasyonu ve stereo görü
7. Hareketten yapı ve çoklu görüş
8. Optik akış ve nesne izleme
9. Evrişimli ağlarla sınıflandırma
10. Nesne bulma: R-CNN ailesi, YOLO
11. Anlamsal ve örnek bölütleme
12. Görü-dil modelleri ve güncel yönelimler

Kazandırdığı beceriler

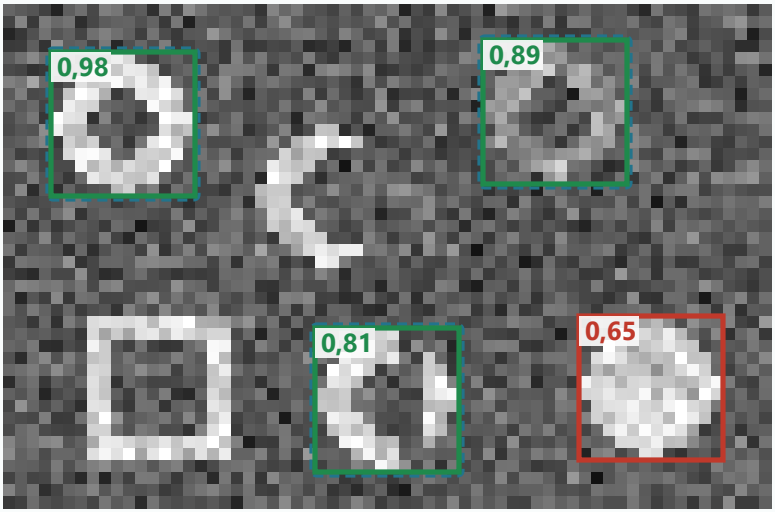
OpenCV | PyTorch ile hazır modelleri kullanma ve ince ayar | kamera kalibrasyonu | etiketli veri hazırlama | IoU ve mAP ile değerlendirme

Tipik bir dönemin akışıdır; sıra ve kapsam üniversiteden üniversiteye değişir.

DENE

Görüntüde Halkaları Bul

Bir montaj hattında kamera, bantta duran contaları (halka biçimli) saymalı. Şablon her konuma kaydırılıp benzerliği ölçülüyor. Güven eşiğini ve bastırmayı değiştirip kutuların doğruluğuna bakın.





12×12 şablon

Güven eşiği = **0,62**

Maksimum olmayanı bastır | Skor haritası

1643 pencere tarandı, eşiği geçen **15**
Bastırmadan sonra: 4 kutu
Doğru: **3** Yanlış: **1** Kaçan: **0**
Ortalama IoU: 1,00

Kesik mavi: gerçek kutu. Yeşil: doğru tespit, kırmızı: yanlış.

Görüntü işleme görüntüyü dönüştürür; bilgisayarlı görü ondan **karar** çıkarır: kutu, etiket, güven. Ders, tespiti IoU, doğru/yanlış tespit ve eşik seçimiyle ölçmeyi öğretir.

KONUM

Ön Koşullar

Bu dersten önce bilinmesi yararlı olan dersler ve neden gerektikleri.

Ders	Neden gerekir
Görüntü İşleme	kenar, filtre, bölütleme
Doğrusal cebir	matris, öz değer ve izdüşüm: kamera geometrisinin dili
Programlama	Python ile dizi işlemleri ve kütüphane kullanımı
Algoritmalar	arama, eşleme ve karmaşıklık: öznitelik eşleme ve izleme

KONUM

Bu Ders Sana Göre mi?

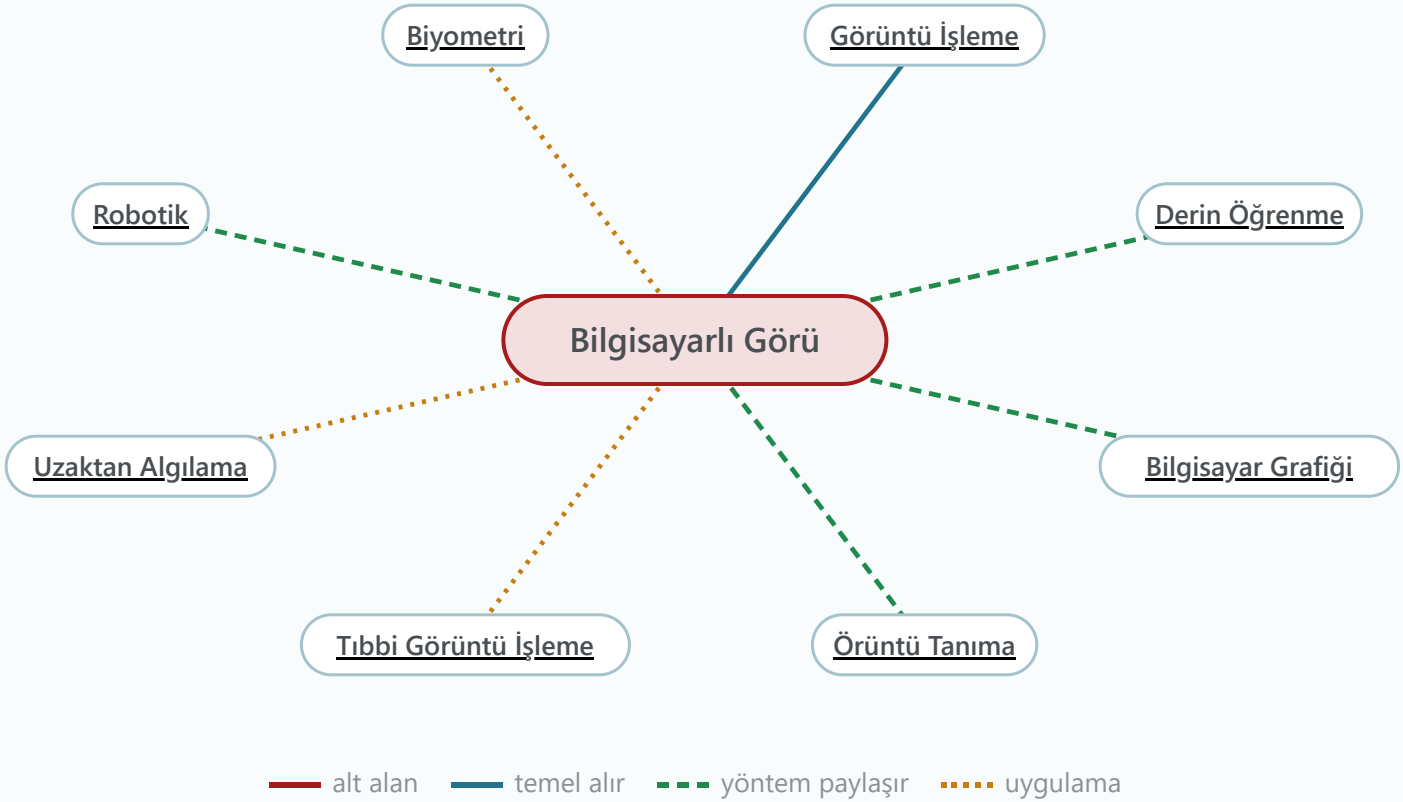
Kendine sor; "hayır" dediklerin için sağdaki sütun nereden başlayacağını gösterir.

Kendine sor	Hayırsa önce
Matris çarpımı, ters matris ve öz değeri açıklayabiliyor musun?	Doğrusal Cebir
Görüntünün bir piksel matrisi olduğunu ve evrişimin ne yaptığını biliyor musun?	Görüntü İşleme
Bir sinir ağının kayıp ve gradyan inişiyle nasıl eğitildiğini anlatabiliyor musun?	Yapay Sinir Ağları
Python'da NumPy dizileriyle çalışabiliyor musun?	Programlamaya giriş, ardından NumPy'a kısa bir giriş

Çoğuna "evet" diyebiliyorsan derse hazırsın; bir ya da iki "hayır" dönem içinde kapanabilir.

KONUM

Dersin Komşuları



ÖRTÜŞME

Örtüştüğü Dersler

İlişki	Ders	Ortak konu
temel alır	Bu ders Görüntü İşleme dersini temel alır	kenar, filtre, bölütleme
uygulama	Tıbbi Görüntü İşleme bu dersin yöntemlerini uygular	lezyon bulma, sınıflandırma
uygulama	Uzaktan Algılama bu dersin yöntemlerini uygular	arazi sınıflandırma, değişim tespiti
uygulama	Biyometrik Sistemler bu dersin yöntemlerini uygular	yüz tanıma
yöntem paylaşır	Derin Öğrenme ile yöntem paylaşır	evrişimli sinir ağları, görüntü dönüştürücüleri
yöntem paylaşır	Örüntü Tanıma ile yöntem paylaşır	nesne sınıflandırma
yöntem paylaşır	Bilgisayar Grafiği ile yöntem paylaşır	kamera modeli; grafik sahneden görüntü, görü görüntüden sahne
yöntem paylaşır	Robotik ve Robot Görüsü ile yöntem paylaşır	robot görüsü, SLAM

AYRIŞMA

Ayrıştığı Noktalar (1/2)

Ölçüt	Bilgisayarlı Görü	Görüntü İşleme	Derin Öğrenme
Temel soru	Bilgisayar bir görüntüde ya da videoda ne olduğunu nasıl anlar?	Bir görüntü nasıl iyileştirilir, dönüştürülür ve parçalara ayrılır?	Çok katmanlı ağlar ham veriden kullanışlı temsilleri nasıl kendisi öğrenir?
Yaklaşım	görüntüden anlam çıkarma	piksel düzeyinde dönüşüm	çok katmanlı temsil öğrenme
Girdi	görüntü ya da video	görüntü	ham veri: görüntü, metin, ses
Çıktı	etiket, kutu, derinlik, 3B yapı	işlenmiş görüntü ya da bölgeler	öğrenilmiş temsil ve tahmin
Başarı ölçütü	doğruluk, mAP, IoU	görsel kalite, PSNR	doğruluk, kayıp, ölçeklenme
Matematik yükü	yüksek: geometri, doğrusal cebir	orta: evrişim, Fourier	yüksek: türev, matris, olasılık
Veri ihtiyacı	çok etiketli görüntü	az; öğrenme şart değil	çok; GPU gerekir

- **Görüntü İşleme:** Görüntü işlemede girdi de çıktı da görüntüdür; bilgisayarlı görüde girdi görüntü, çıktı anlamdır: etiket, kutu, derinlik.
- **Derin Öğrenme:** Derin öğrenme genel bir araçtır; görü onu kamera geometrisi, çoklu görüş ve hareket bilgisiyle birlikte kullanır.

Ayrıştığı Noktalar (2/2)

Ölçüt	Bilgisayarlı Görü	Bilgisayar Grafiği	Örüntü Tanıma
Temel soru	Bilgisayar bir görüntüde ya da videoda ne olduğunu nasıl anlar?	Bir sahnenin tanımından gerçekçi bir görüntü nasıl üretilir?	Ölçümlerden bir nesnenin hangi sınıfa ait olduğuna nasıl karar verilir?
Yaklaşım	görüntüden anlam çıkarma	geometri ve ışık benzetimi	istatistiksel karar kuramı
Girdi	görüntü ya da video	sahne: model, ışık, kamera	ölçüm ya da öznitelik vektörü
Çıktı	etiket, kutu, derinlik, 3B yapı	görüntü ya da animasyon	sınıf etiketi
Başarı ölçütü	doğruluk, mAP, IoU	gerçekçilik, kare hızı	sınıflandırma hatası
Matematik yükü	yüksek: geometri, doğrusal cebir	yüksek: doğrusal cebir, geometri	orta-yüksek: olasılık, doğrusal cebir
Veri ihtiyacı	çok etiketli görüntü	gerekmez	etiketli örnekler

- **Bilgisayar Grafiği:** Grafik sahne tanımından görüntü üretir; görü tersini yapar, görüntüden sahneyi kurmaya çalışır.
- **Örüntü Tanıma:** Örüntü tanıma her tür ölçümü sınıflandırır; görü sınıflandırmayı uzamsal yapı, konum ve 3B ile birlikte ele alır.

UYGULAMA

Uygulama Alanları

Otonom sürüş
Şerit, yaya ve trafik işareti algılama.

Yüz tanıma
Telefon kilidi, geçiş kontrolü.

Perakende
Kasasız mağaza, raf takibi.

Artırılmış gerçeklik
Kameranın konumunu bulup sahneye nesne eklemek.

Endüstri
Hatalı ürünü tanıyıp hattan ayırmak.

Spor analizi
Oyuncu ve top izleme, istatistik.

UYGULAMA

Riskler ve Etik

Yüz tanıma ve gözetim

Kalabalıkta kişi tanıma, rıza olmadan izlemeye dönüşebilir; birçok ülke kamusal alanda kullanımını sınırlıyor.

Eşit olmayan başarım

Eğitim verisinde az temsil edilen ten renklerinde ve yaş gruplarında hata oranı belirgin biçimde yüksek çıkabilir.

Saldırgan örnekler

İnsan gözünün fark etmediği küçük bir çıkartma, bir dur levhasını model için başka bir levhaya çevirebilir.

Güvenlik açısından kritik hatalar

Otonom sürüşte kaçırılan tek bir yaya, ortalama doğruluğun anlatmadığı bir bedeldir.

Ayrıntılı tartışma: Açıklanabilir Yapay Zekâ ve Yapay Zekâ Etiği.

UYGULAMA

Araçlar ve Örnek Veriler

Araçlar ve kütüphaneler

OpenCV | PyTorch / torchvision | Ultralytics YOLO | Detectron2 | COLMAP | Open3D

Tanınmış veri setleri

ImageNet | COCO | PASCAL VOC | KITTI | Cityscapes | ADE20K

Görü, görüntü işlemenin aksine **çok sayıda etiketli görüntüye** ihtiyaç duyar; COCO'da her nesne için kutu ve maske elle çizilmiştir.

UYGULAMA

Sık Karıştırılanlar

Yaygın düşünce	Doğrusu
Bilgisayarlı görü, görüntü işlemenin başka adıdır.	Görüntü işleme görüntüyü dönüştürür; görü görüntüden anlam çıkarır. Görü, görüntü işlemeyi ön adım olarak kullanır.
Görü artık çözülmüş bir problemdir.	Ağlar, eğitim verisinden farklı ışık, hava ya da kamerada hâlâ belirgin biçimde yanılabilir.
Görü yalnızca derin öğrenmedir.	Kalibrasyon, stereo, SLAM ve 3B yeniden yapılandırma büyük ölçüde geometriye dayanır.
Model nesneyi insan gibi "görür".	Ağlar dokuya ve bağlama aşırı güvenebilir; küçük bir piksel değişikliği kararı çevirebilir.

SONRASI

Bu Dersten Sonra

Bu dersi temel alan ya da genişleten dersler

Tıbbi Görüntü İşleme | Uzaktan Algılama | Biyometrik Sistemler

İlgili meslekler

Bilgisayarlı görü mühendisi | Otonom sistemler mühendisi | AR/VR geliştiricisi | Robotik algı mühendisi | Araştırmacı

Bilgisayarlı görü üçüncü kısmın merkezidir: robotik, biyometri, tıbbi görüntü işleme ve uzaktan algılama onun yöntemlerini kendi alanlarına uygular. Derinleşmek için derin öğrenme ve üretken modeller, geometri tarafında bilgisayar grafiği ve robotik ile devam edilebilir.

SONRASI

Kaynaklar

- Szeliski, *Computer Vision: Algorithms and Applications*, 2. baskı, Springer, 2022.
- Hartley ve Zisserman, *Multiple View Geometry in Computer Vision*, 2. baskı, Cambridge University Press, 2004.
- Prince, *Computer Vision: Models, Learning, and Inference*, Cambridge University Press, 2012.
- Marr, *Vision*, W. H. Freeman, 1982.

Son

Bölüm 18: Bilgisayarlı Görü

← Önceki bölüm · Alanın haritası · Sonraki bölüm →