

Derin Öğrenme Nedir?

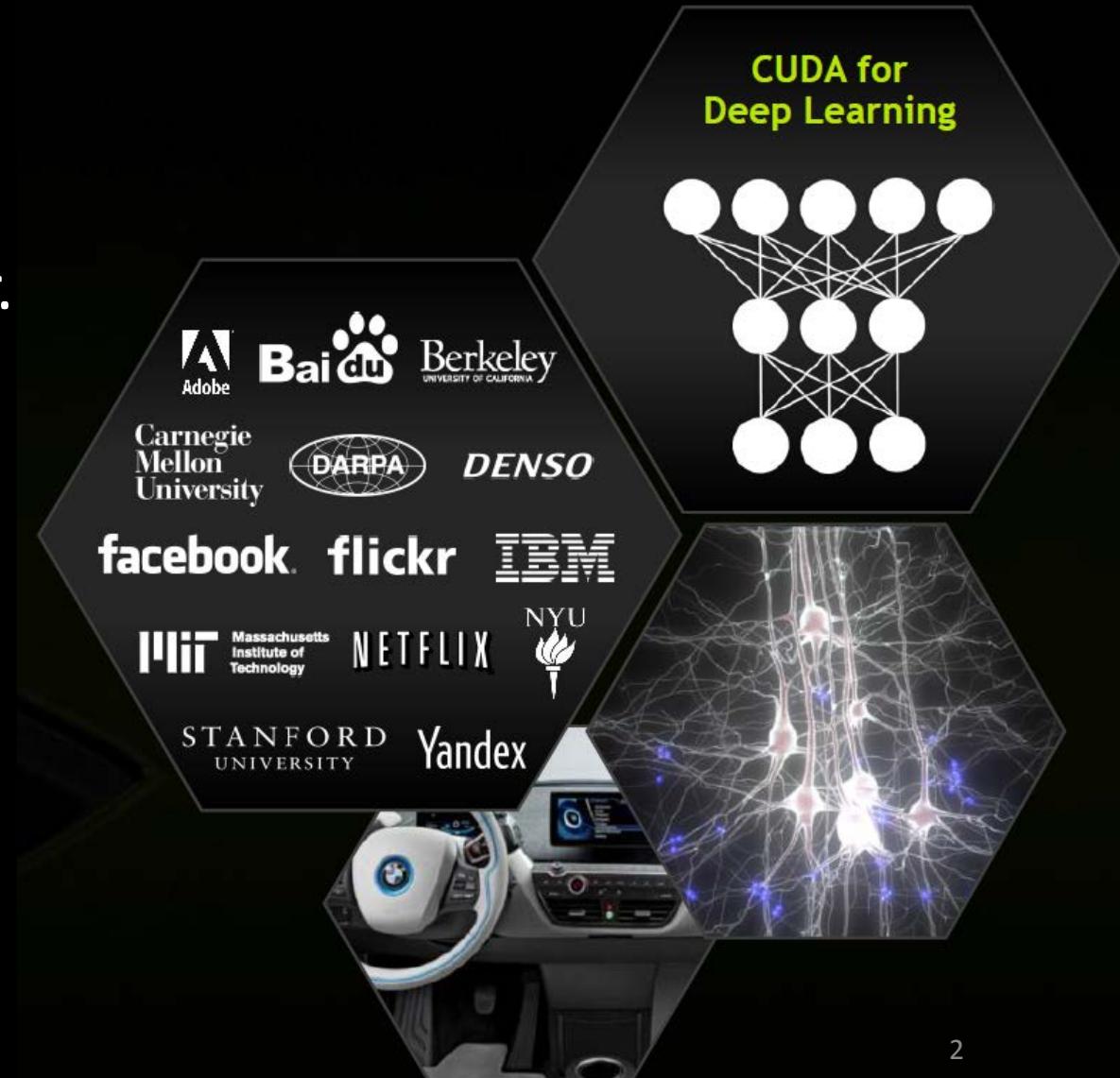
Ferhat Kurt

info@derinogrenme.com

Bu sunum 10 Aralık 2015 tarihinde ODTÜ Kongre ve Kültür Merkezi’nde Yapılmıştır.
Diğer sunumlar için <http://www.derinogrenme.com> sitesini ziyaret edebilirsiniz.

Derin Öğrenme & Yapay Zekâ

Derin öğrenme, makinelerin dünyayı algılama ve anlamasına yönelik yapay zekâ geliştirmede en popüler yaklaşımındır. Şu anda ağırlıklı olarak belirli anlamıyla ilgili görevlere odaklanılmış ve bu alanlarda birçok başarı elde edilmiştir. Bugün, en önemli araştırma kuruluşlarının yanı sıra dünyanın en büyük internet şirketlerinin bazıları araştırma ve ürünlerinde derin öğrenme kullanmak için GPU'ları kullanmaktadır.



Geniş-kapsamlı Tanıma

Flüt



Veri

Veri Çeşitliliği

Veri, gerçek dünyayı yansıtacak yeterlilikte çeşitliliğe sahip olmalıdır.

Keçi

Veri Açıklaması

Veri, algoritmanın öğrenebilmesi için yeterli bilgiye sahip olmalıdır.

Raket

Algoritma

Algoritma Kompleksliği

Algoritmalar veriden öğrenecek kadar yeterli gücü sahip olmalıdır.

Kibrıt

Algoritma Hızı

Algoritmalar, geniş veriyi işleyecek kadar yeterli hızda sahip olmalıdır.

Sırt Çantası

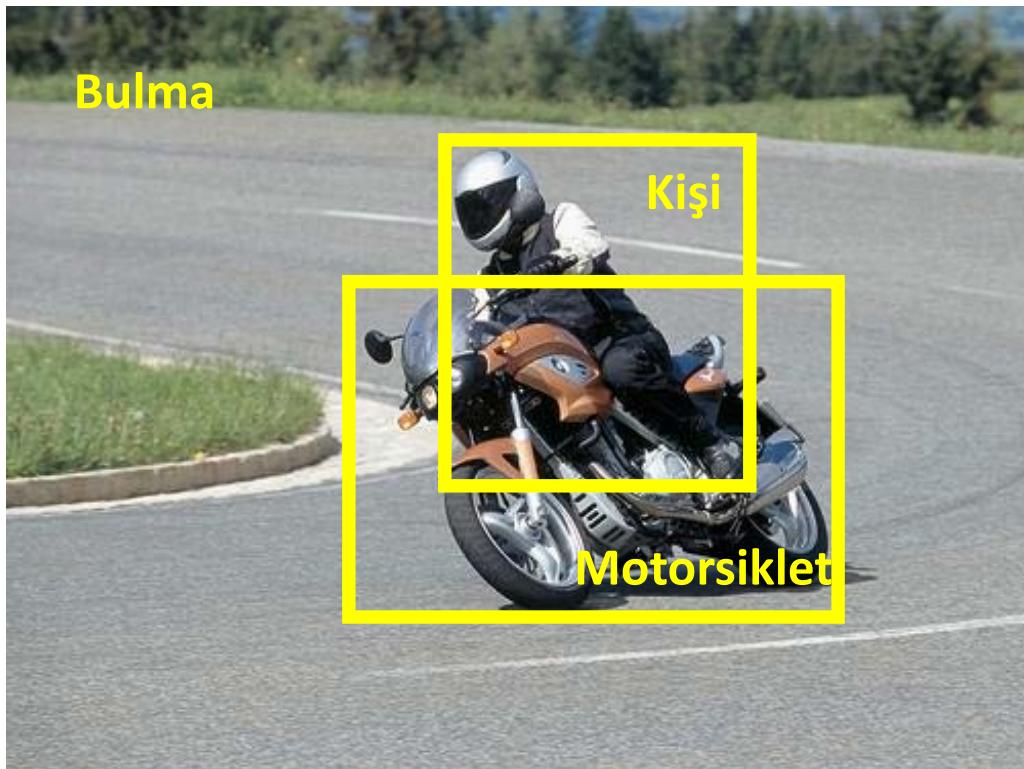


PASCAL VOC 2005-2012

20 Nesne Sınıfı

22,951 resim

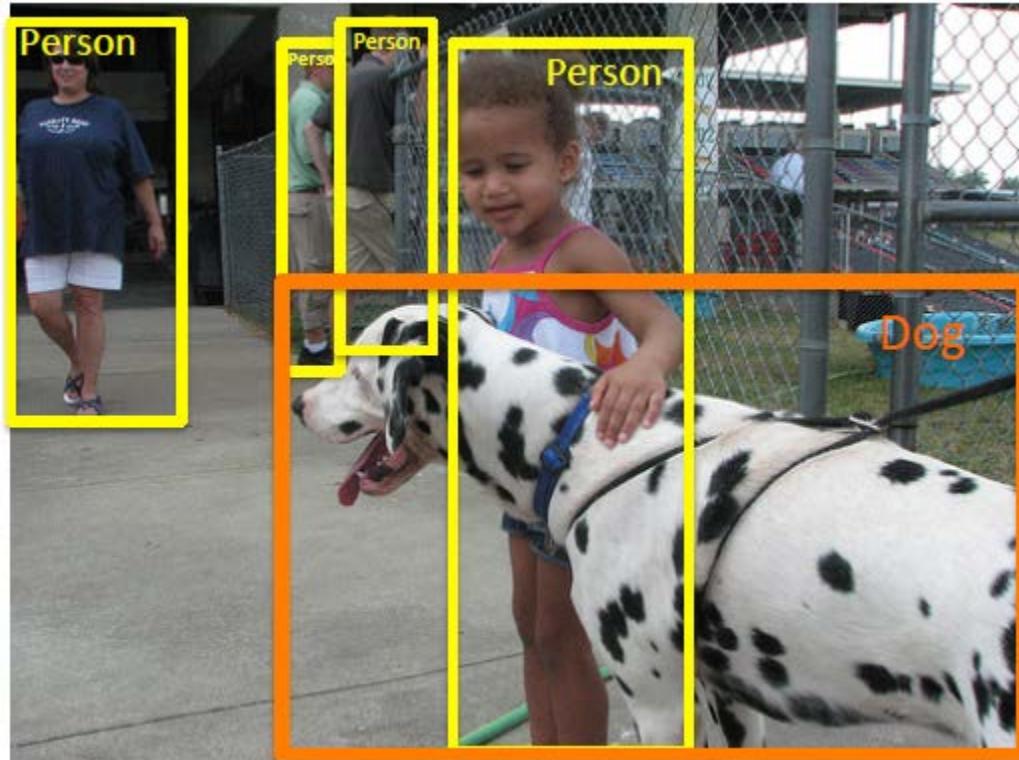
Sınıflandırma: kişi, motosiklet



Parçalara ayırma -Segmentasyon



IMAGENET BÜYÜK KAPSAMLI GÖRSEL TANIMA YARIŞMASI (ILSVRC) 2010-2015



14.197.122 resim, 21.841 synset indekslenmiştir.

<http://image-net.org/challenges/LSVRC/>

200 Nesne Sınıfında 517,840 resim

1000 Nesne Sınıfında 1,431,167 resim

Yarışma Kategorileri

İki Ana Yarışma Kategorisi

1. Nesne tespiti (200 farklı etiketlenmiş kategoride),
2. Nesne Yeri Bulma (1000 farklı kategoride)

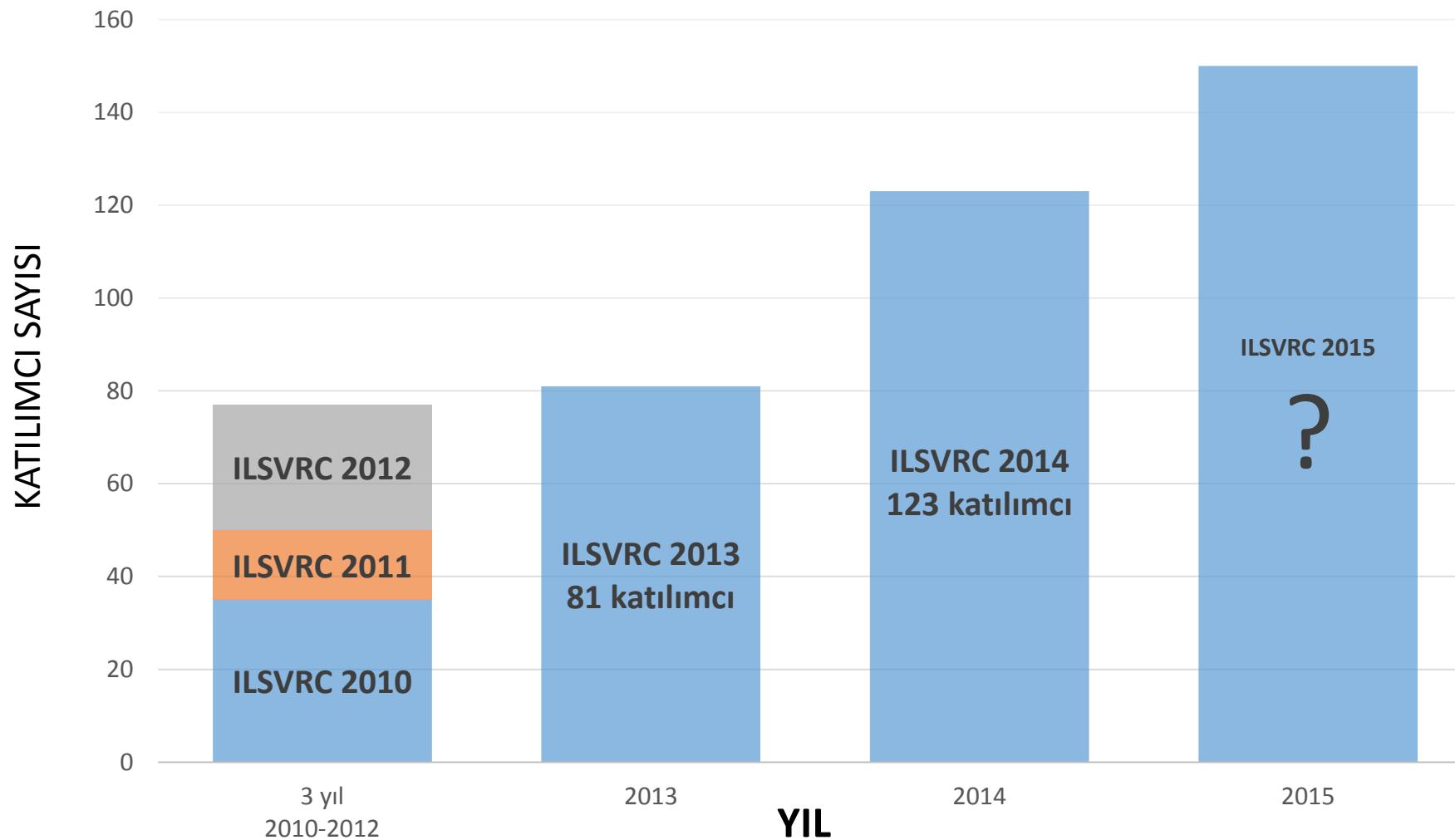
İki Taster Yarışma Kategorisi (**Yeni**)

1. Videoda nesne tespiti (30 farklı etiketlenmiş kategoride),
2. Sahne sınıflandırma (401 kategoride-MIT Places takımı)

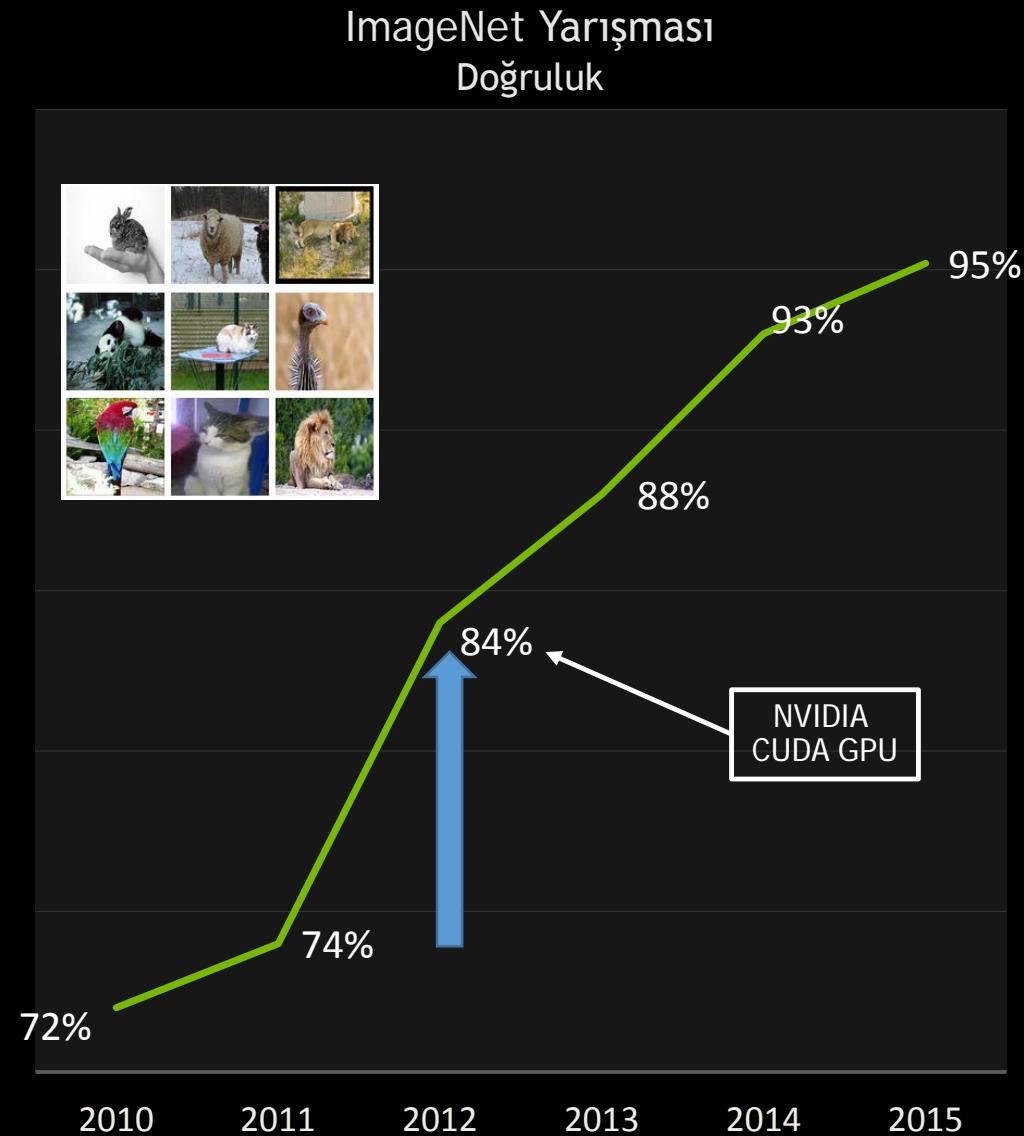
ILSVRC 2015 - YENİ EKLENEN VİDEO KATEGORİSİ

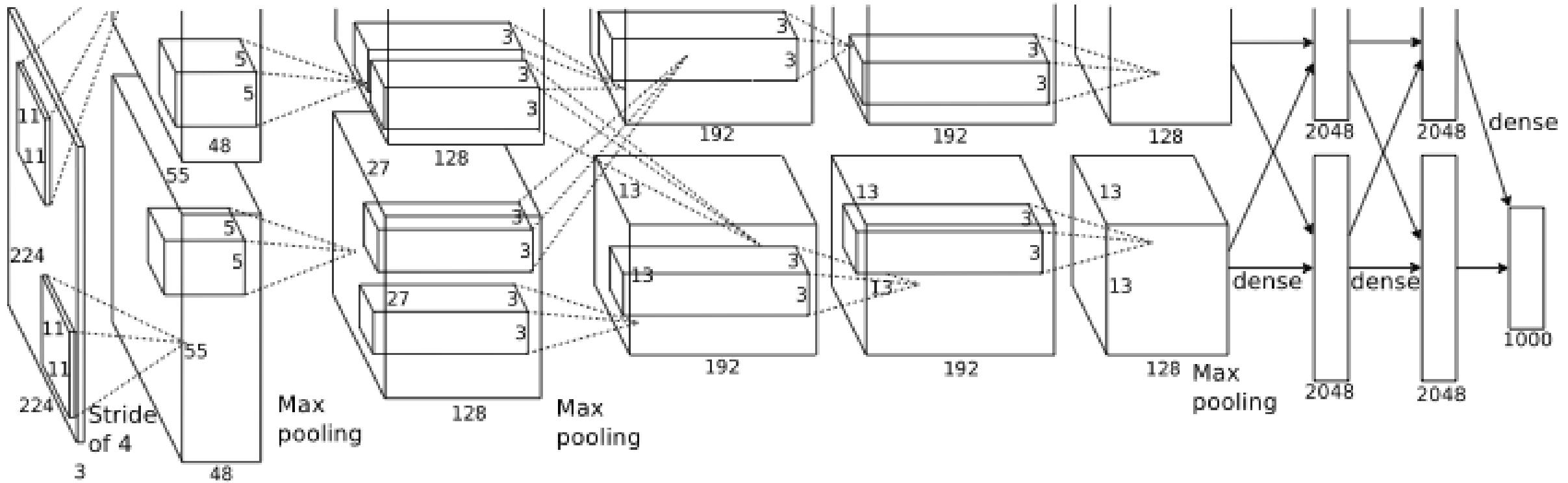


YILLARA GÖRE ILSVRC'YE KATILIM DURUMU

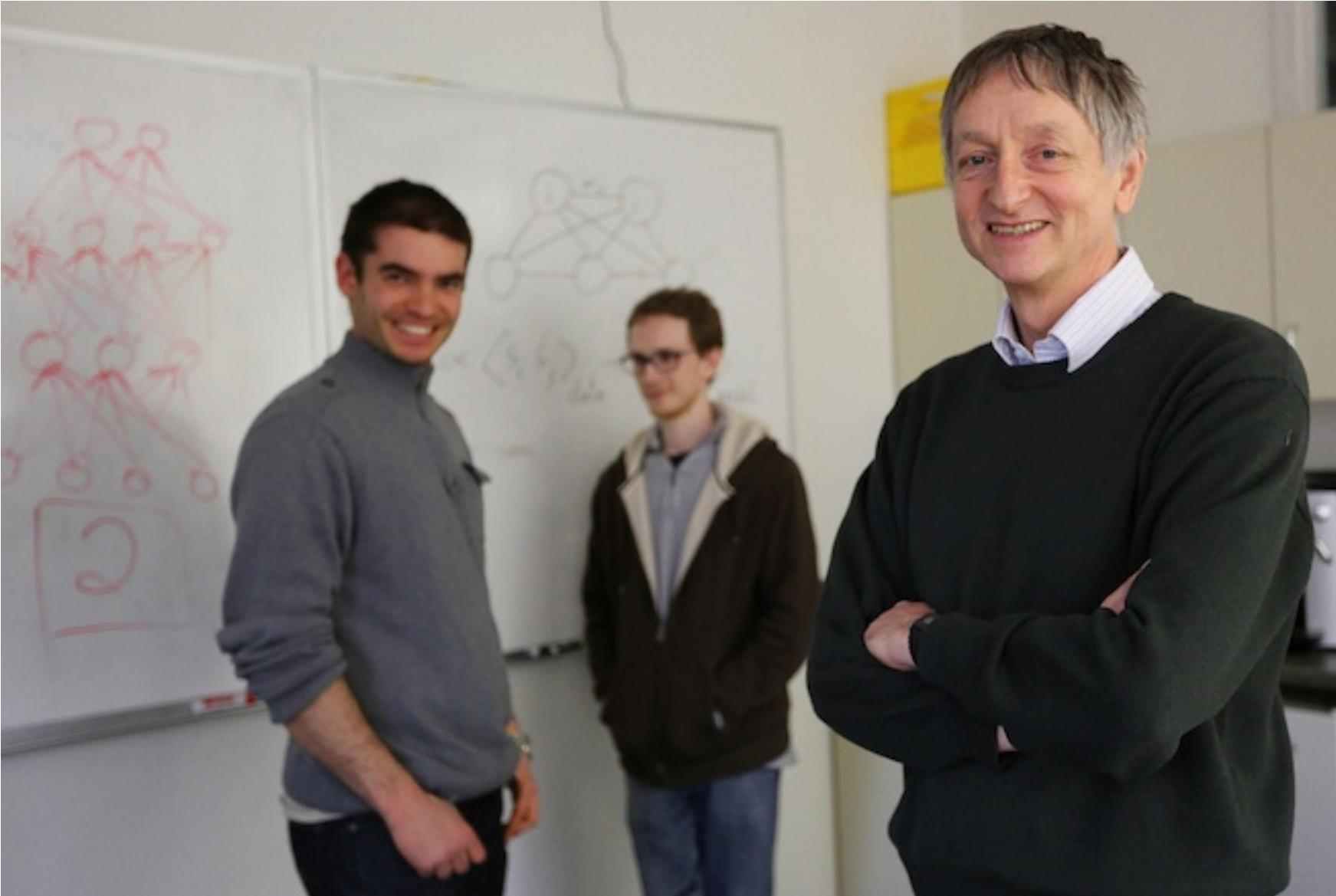


YARIŞMA YILI	Başarı
Imagenet 2011 kazananı	%74,3
Imagenet 2012 kazananı (Krizhevsky)	%83,6
Imagenet 2013 kazananı (Zeiler/Clarifai)	%88,3
Imagenet 2014 kazananı (GoogLeNet)	%93,3
Andrej Karpathy	%94,9
Baidu ArXiv Yayıni 3 Oca 15	%94
MS Research ArXiv yayını 6 Sub 15	%95,1
Google ArXiv yayını 2 Mar 15	%95,2

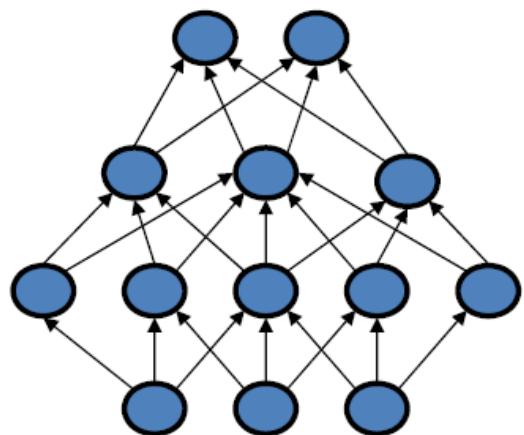
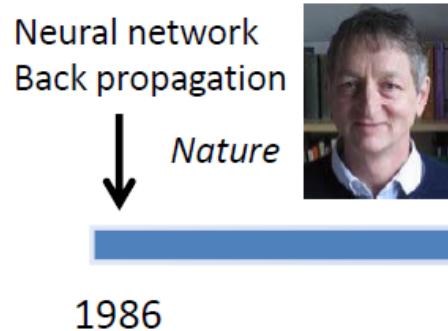




A. Krizhevsky, I. Sutskever and G.E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". NIPS (2012)



Geoffrey Hinton (sağda) Alex Krizhevsky ve Ilya Sutskever (solda). *Fotoğraf: U of T*



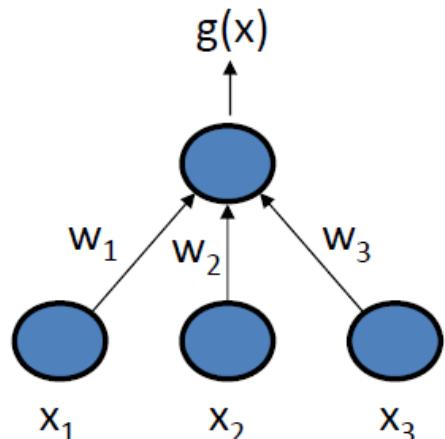
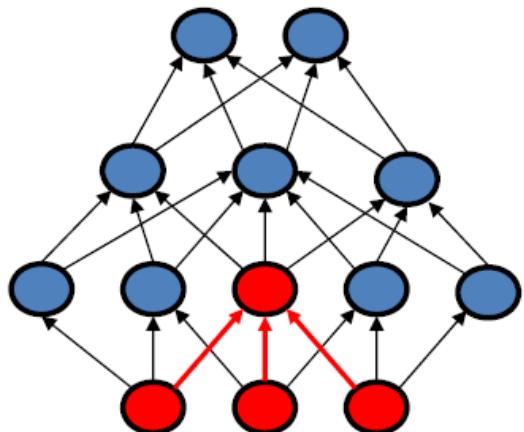
- Solve general learning problems
- Tied with biological system

Neural network
Back propagation

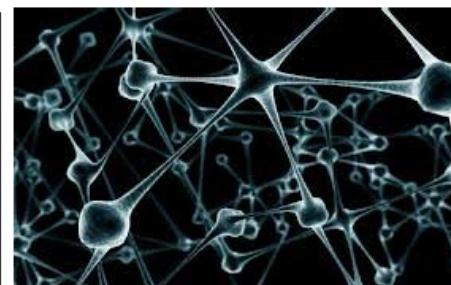
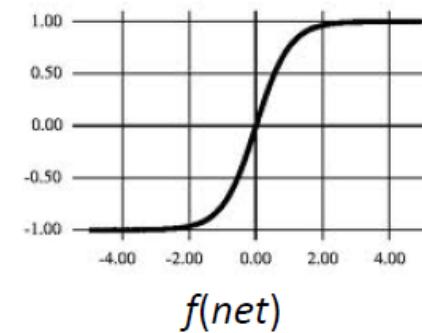


↓ *Nature*

1986

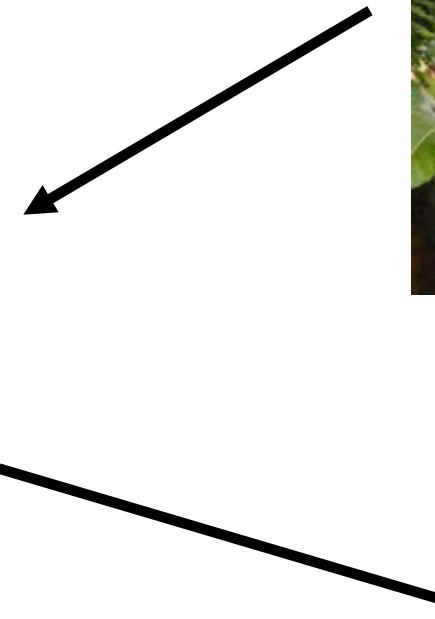
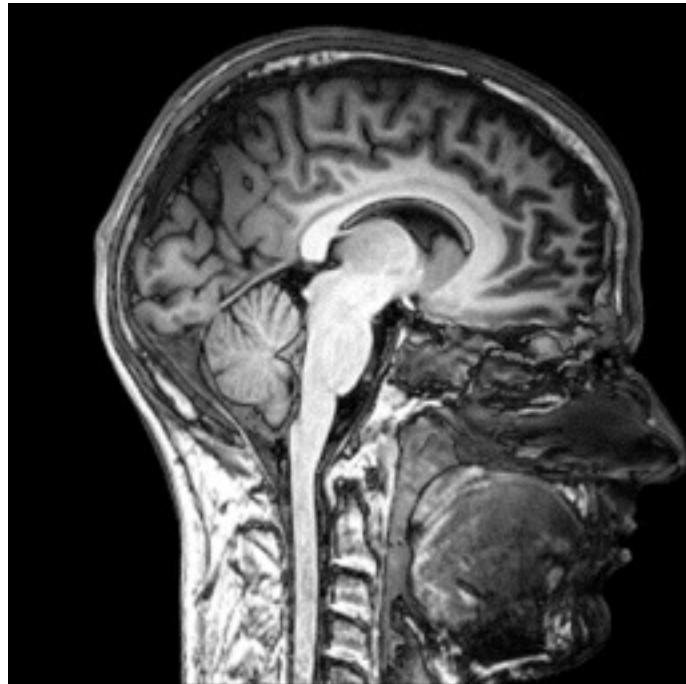


$$g(\mathbf{x}) = f\left(\sum_{i=1}^d x_i w_i + w_0\right) = f(\mathbf{w}^t \mathbf{x})$$



Yapay Sinir Ağı Nasıl Çalışıyor?

0.1 sec:
neurons
can fire
only 10
times!



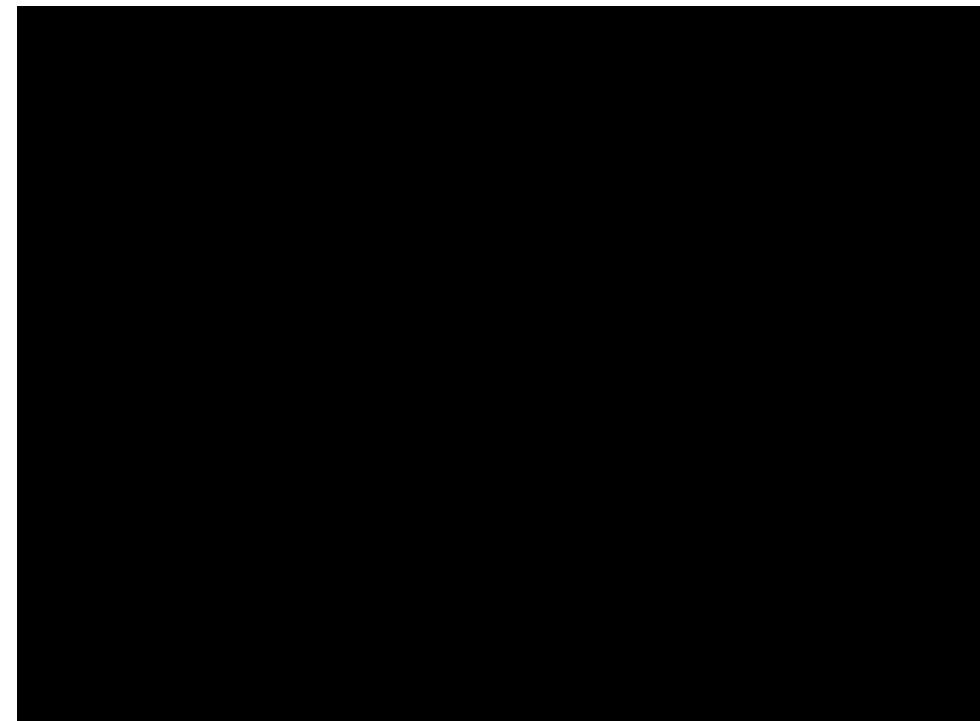
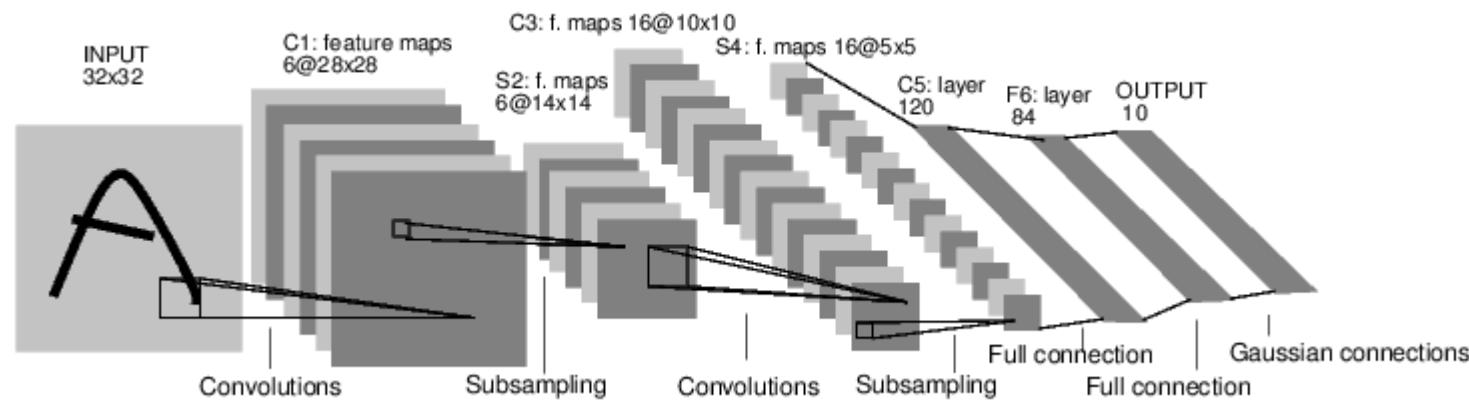
Kedi

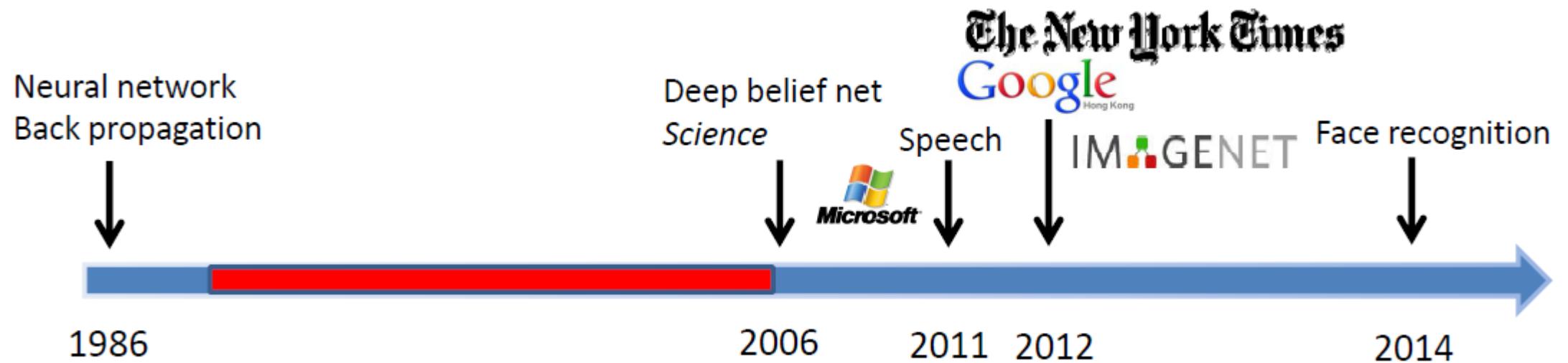


İnsanın 0,1 saniyede yapabildiği herşeyi,
10 katmanlı büyük bir ağa yapabilir.

Eğer kedi ise tıkla

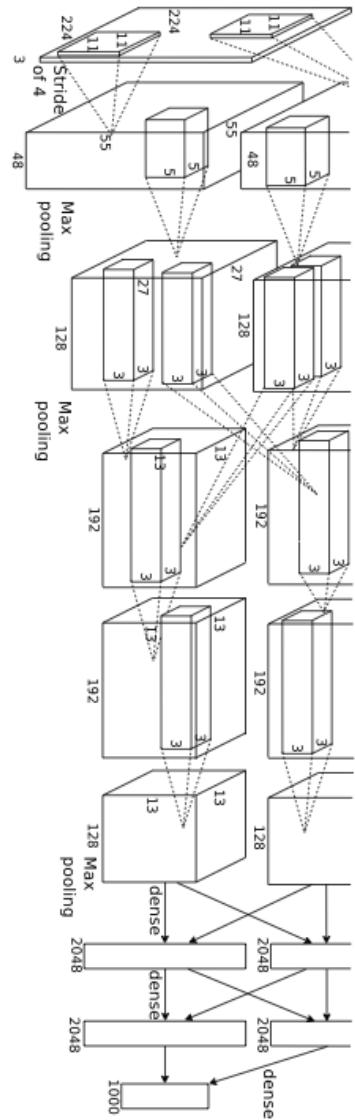
LeNet (LeCun et al. 1998)





2012

SuperVision

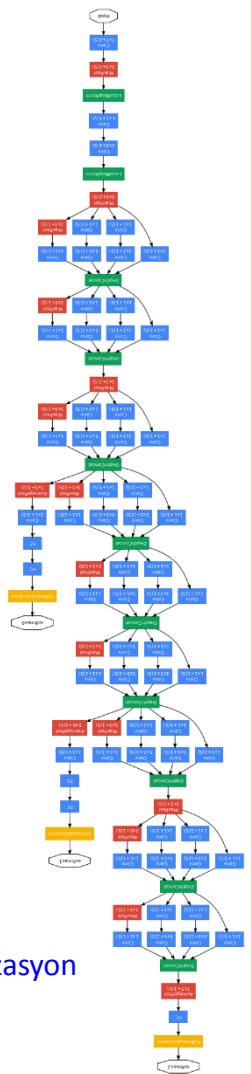


[Krizhevsky NIPS 2012]

2014

GoogLeNet

Konvolüzasyon
Pooling
Softmax
Diğer



[Szegedy arxiv 2014]

VGG

image

conv-64

conv-64

MSRA



[Simonyan arxiv 2014]

[He arxiv 2014]

35/36 takım
derin öğrenme
kullanmıştır.

20/36 takım
açık-kaynaklı
Caffe
uygulamasını
kullanmıştır.

GELENEKSEL MAKİNE ALGISI – Hand Crafted Özellikler

Ham veri



Özellik çıkarıcı

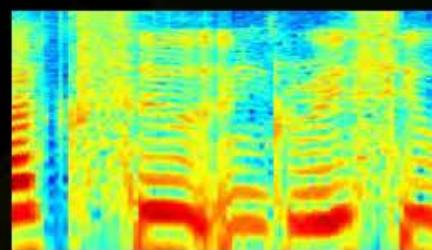


Lineer
Sınıflandırıcı

e.g. SVM



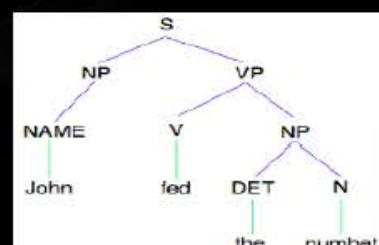
Sonuç



e.g. HMM



Speaker ID,
speech transcription, ...



e.g LSA

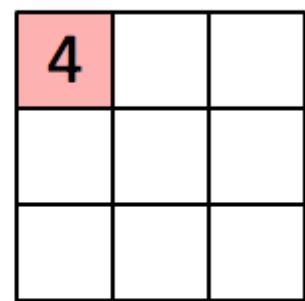


Topic classification,
machine translation,
sentiment analysis...

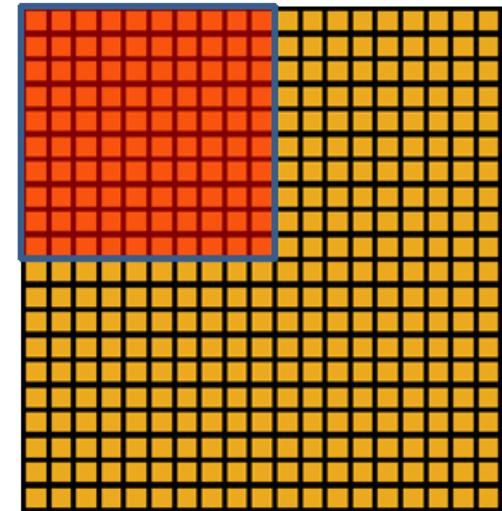
Konvolüzasyonel Sinir Ağları

1 x1	1 x0	1 x1	0	0
0 x0	1 x1	1 x0	1	0
0 x1	0 x0	1 x1	1	1
0	0	1	1	0
0	1	1	0	0

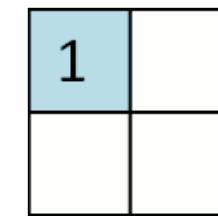
Image



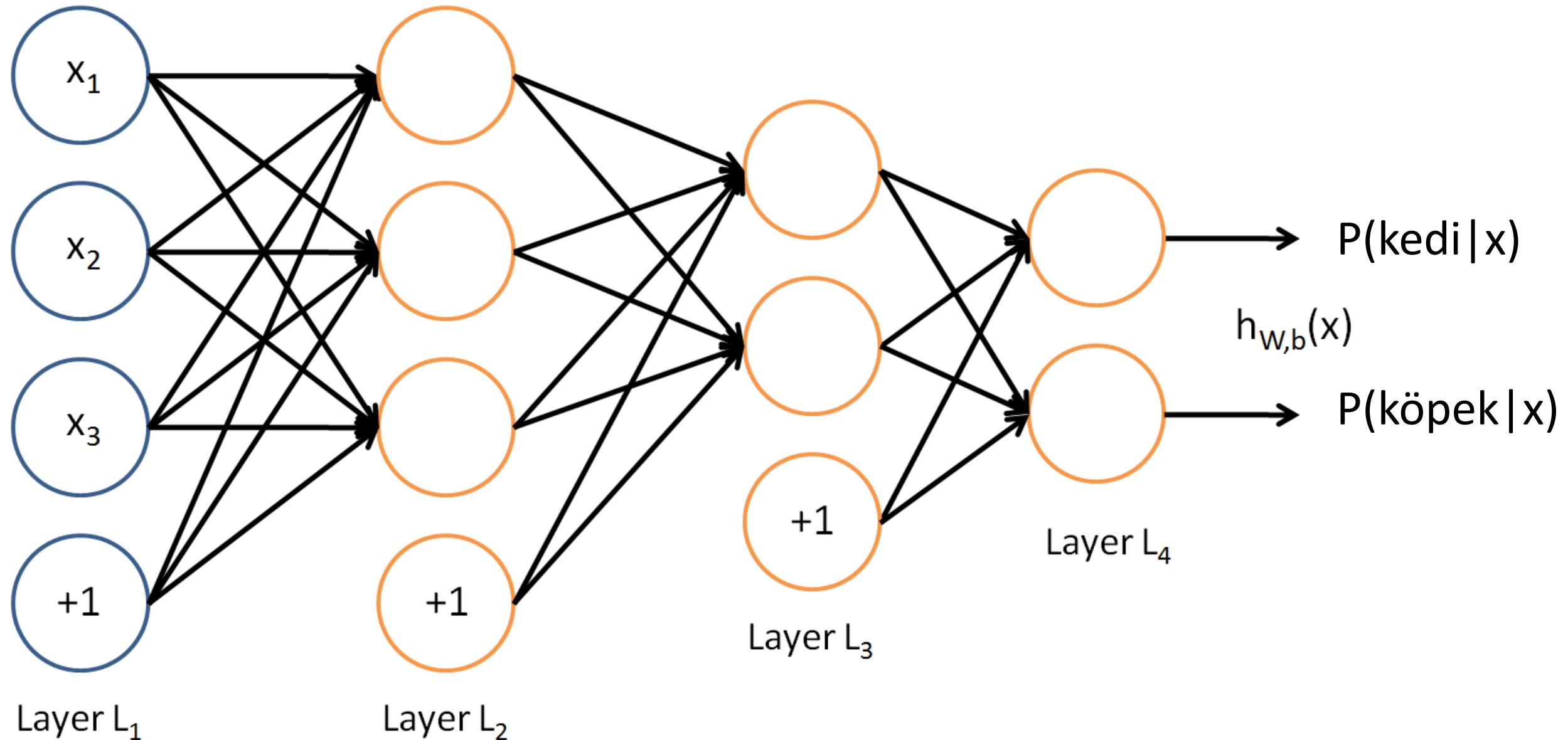
Convolved
Feature

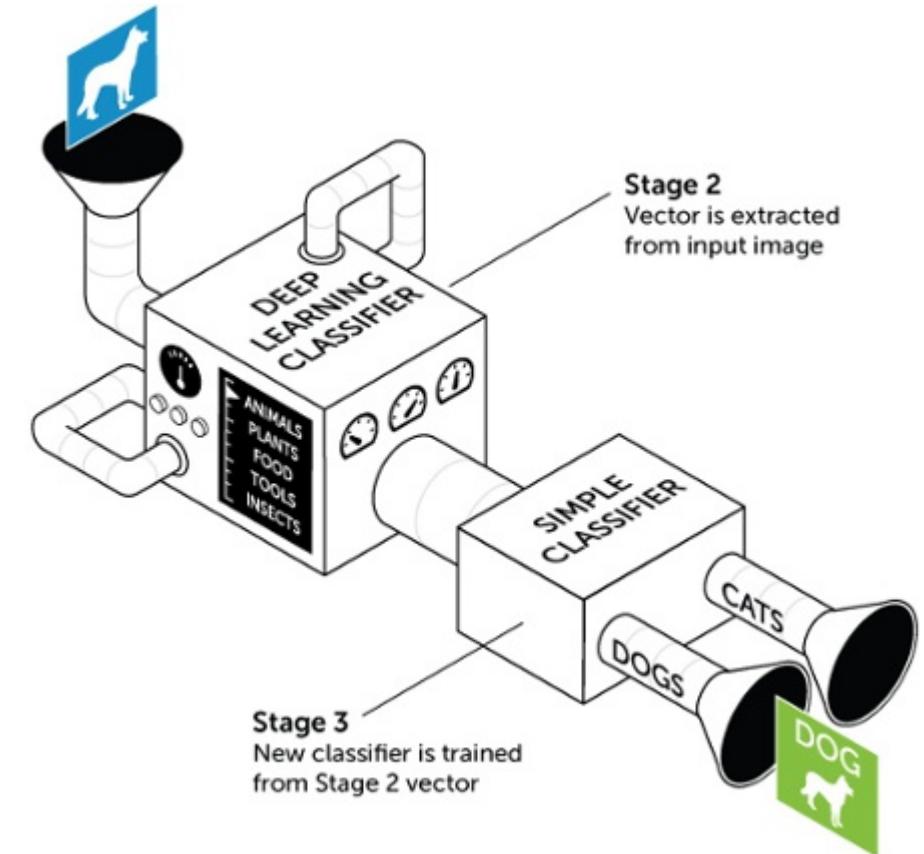
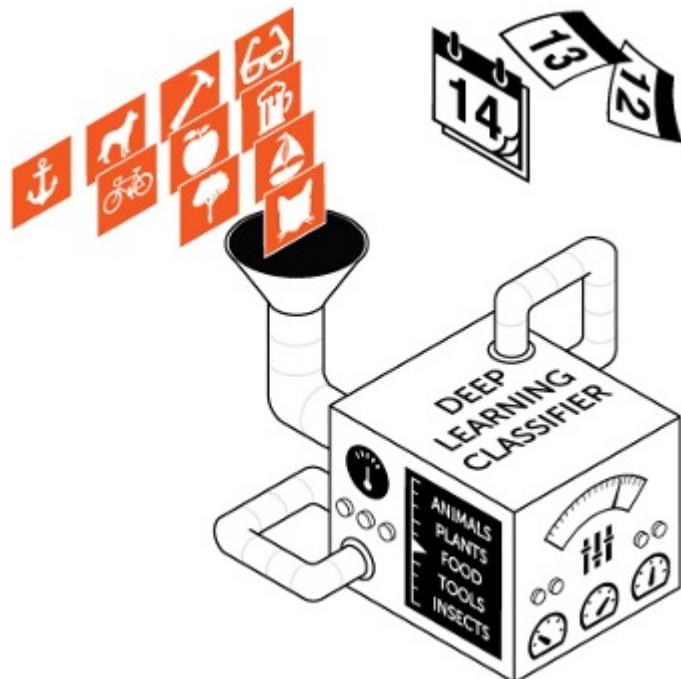
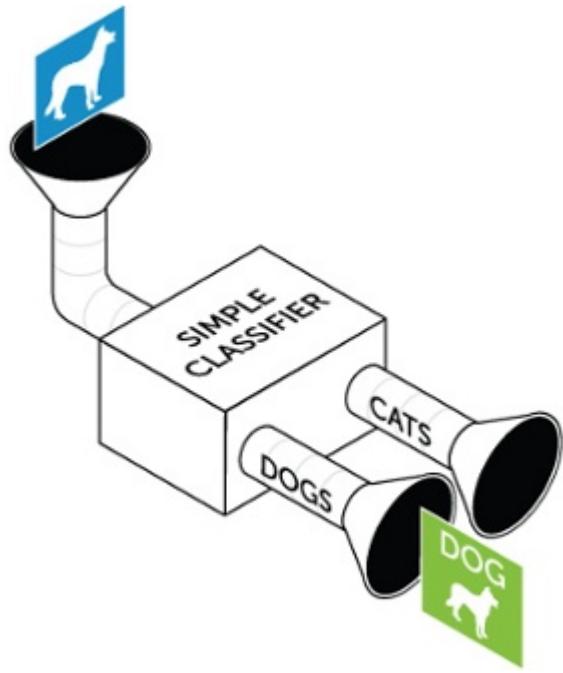


Convolved
feature



Pooled
feature





<https://dato.com>

MULTIMODAL RECURRENT NEURAL NETWORK



«siyah beyaz köpek bar
üzerinden atlıyor.»



«iki genç kız lego
oyuncağı ile oynuyor.»



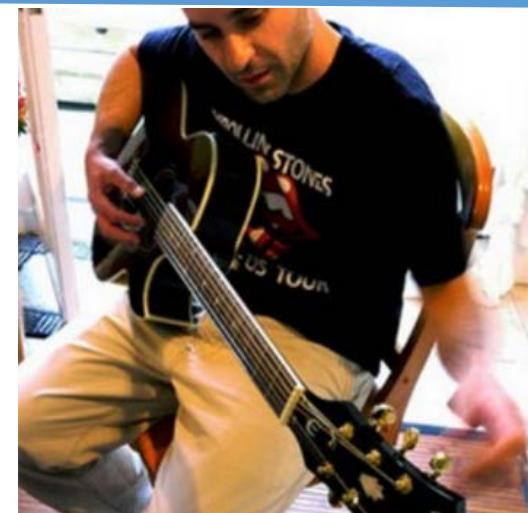
«pembe kıyafetli kız
havada zıplıyor.»



«turuncu güvenlik yeleği giyen
yapı işçisi yolda çalışıyor»



«mavi dalış kıyafetli adam dalga
üstünde sörf yapıyor.»



«siyah tişört giyen
adam gitar çalışıyor.»



© secureborderintel.org

Sırtında sigara kutusu taşıyan 5 kişi kuru otların arasından sınıra doğru yaklaşıyor.

Büyük Veriye Erişim

facebook

Günlük 350 milyon
resim yükleme

YouTube

Her bir dakikada
100 saatlik video

Walmart

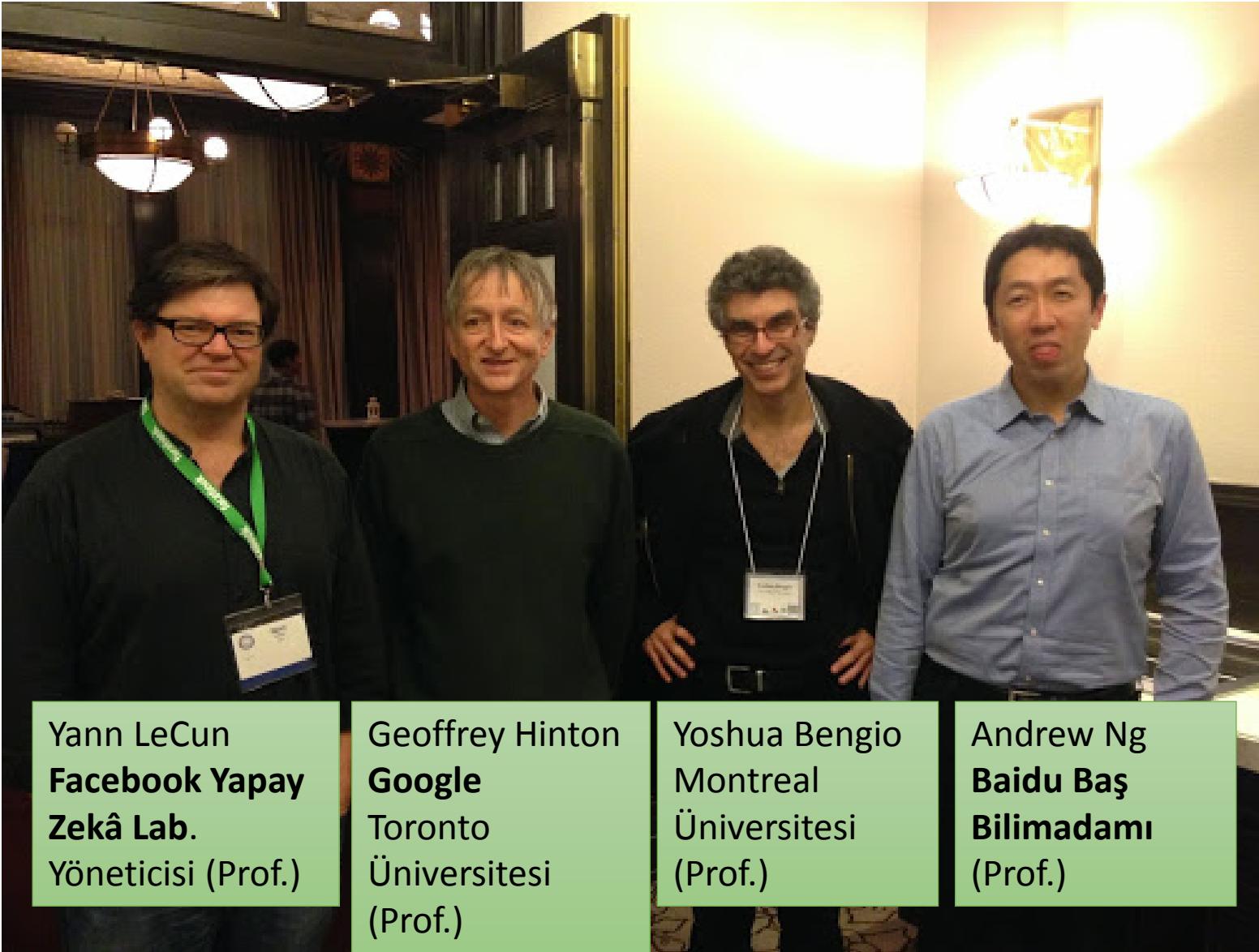
Saatte 2,5 Petabyte
Müşteri verisi

Yeni Derin Öğrenme Teknikleri



NVIDIA

GPU Hızlandırıcıları



Yann LeCun
**Facebook Yapay
Zekâ Lab.
Yöneticisi (Prof.)**

Geoffrey Hinton
**Google
Toronto
Üniversitesi
(Prof.)**

Yoshua Bengio
Montreal
Üniversitesi
(Prof.)

Andrew Ng
**Baidu Baş
Bilimadamlı
(Prof.)**

KARŞILAŞTIRMA

GOOGLE DATACENTER



1,000 CPU Servers
2,000 CPUs • 16,000 cores

600 kWatts
\$5,000,000

WIRE

STANFORD AI LAB



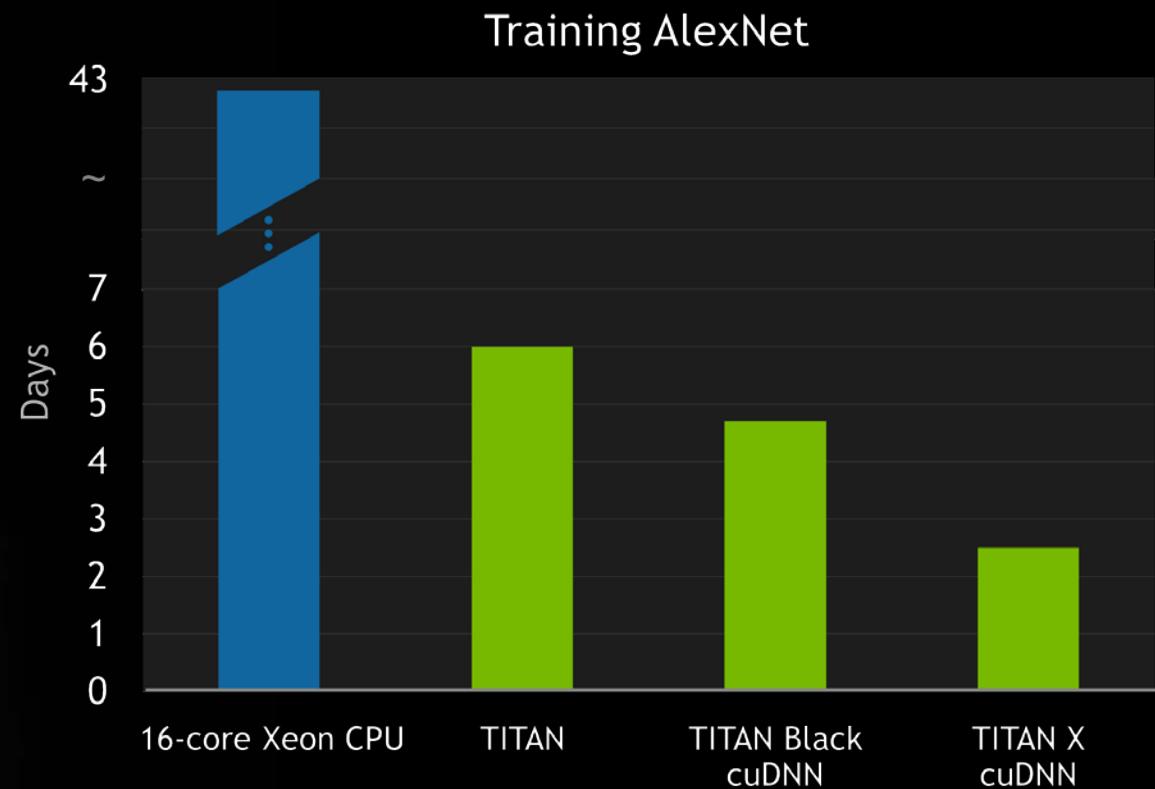
3 GPU-Accelerated Servers
12 GPUs • 18,432 cores

4 kWatts
\$33,000

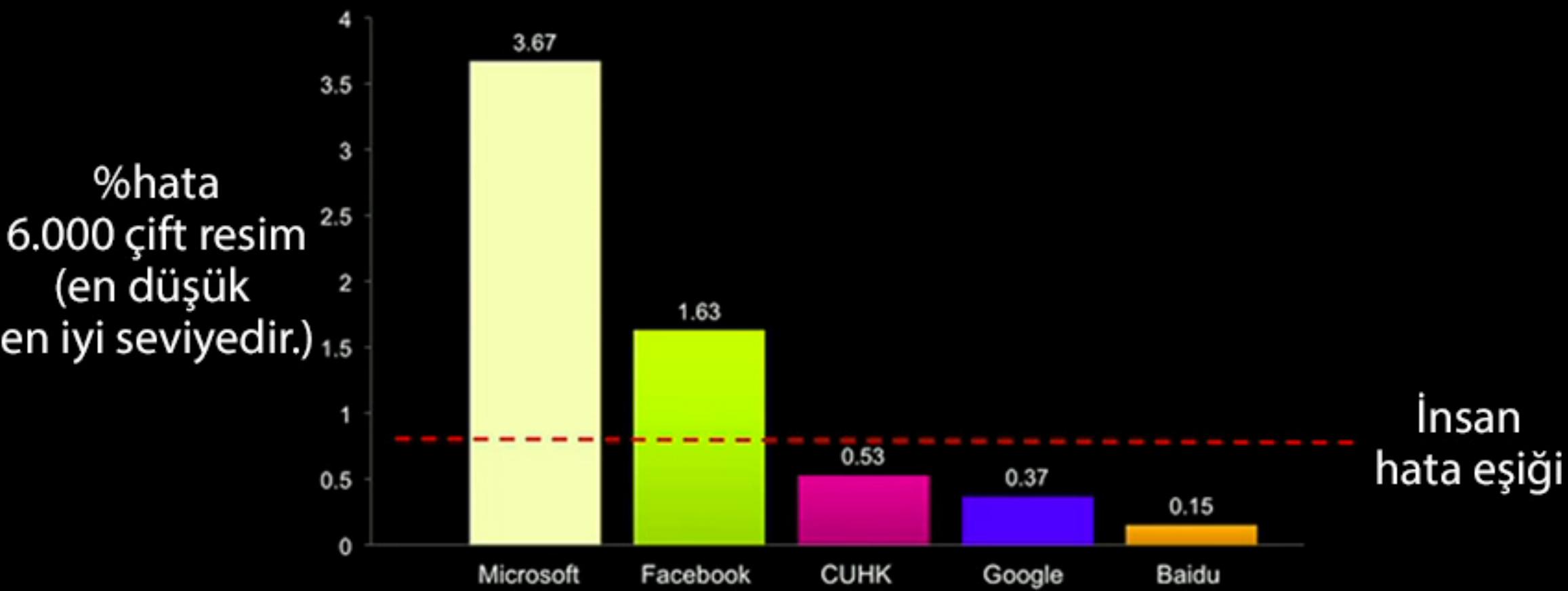
TITAN



8 Billion Transistors
3,072 CUDA Cores
7 TFLOPS SP / 0.2 TFLOPS DP
12GB Memory



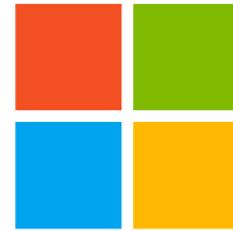
Yüz tanıma hata oranları (en düşük seviye en iyidir)





Speech Milestones





Microsoft

skype™

translator



Microsoft
HoloLens

30

Windows 10



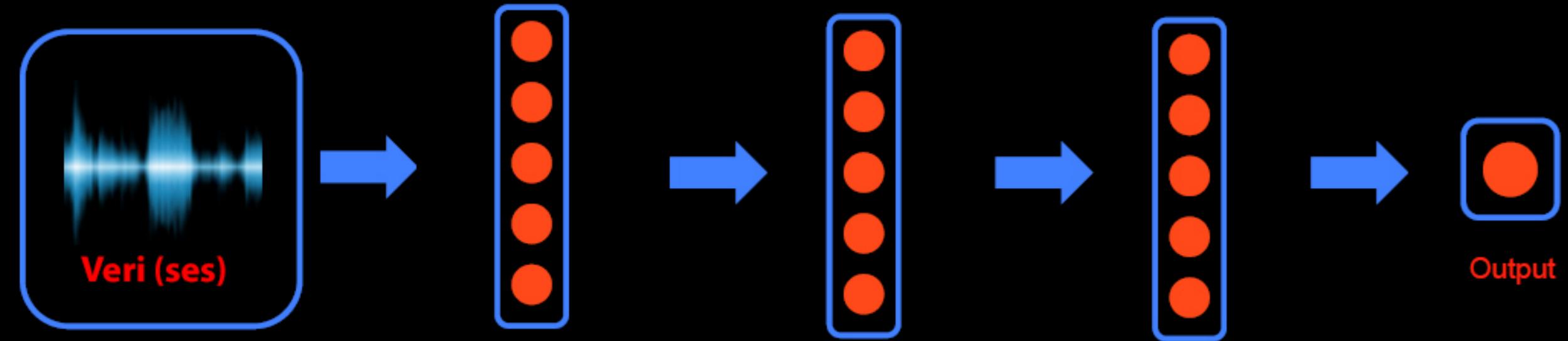
Konuşma Tanıma



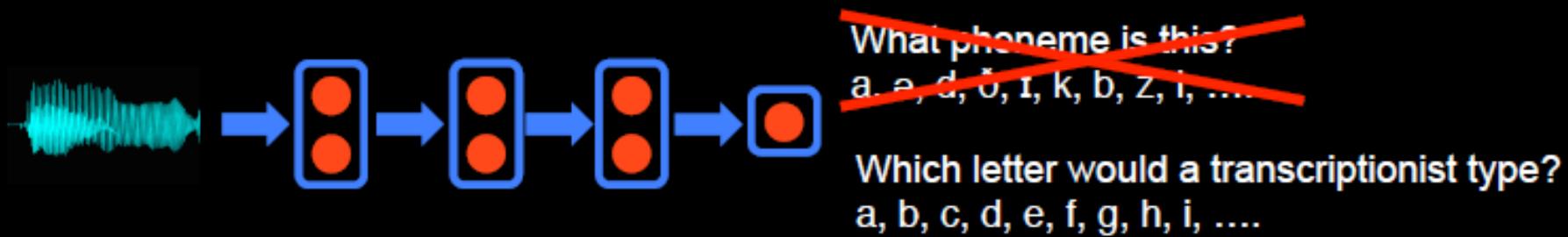
Konuşma Tanıma



Konuşma Tanıma

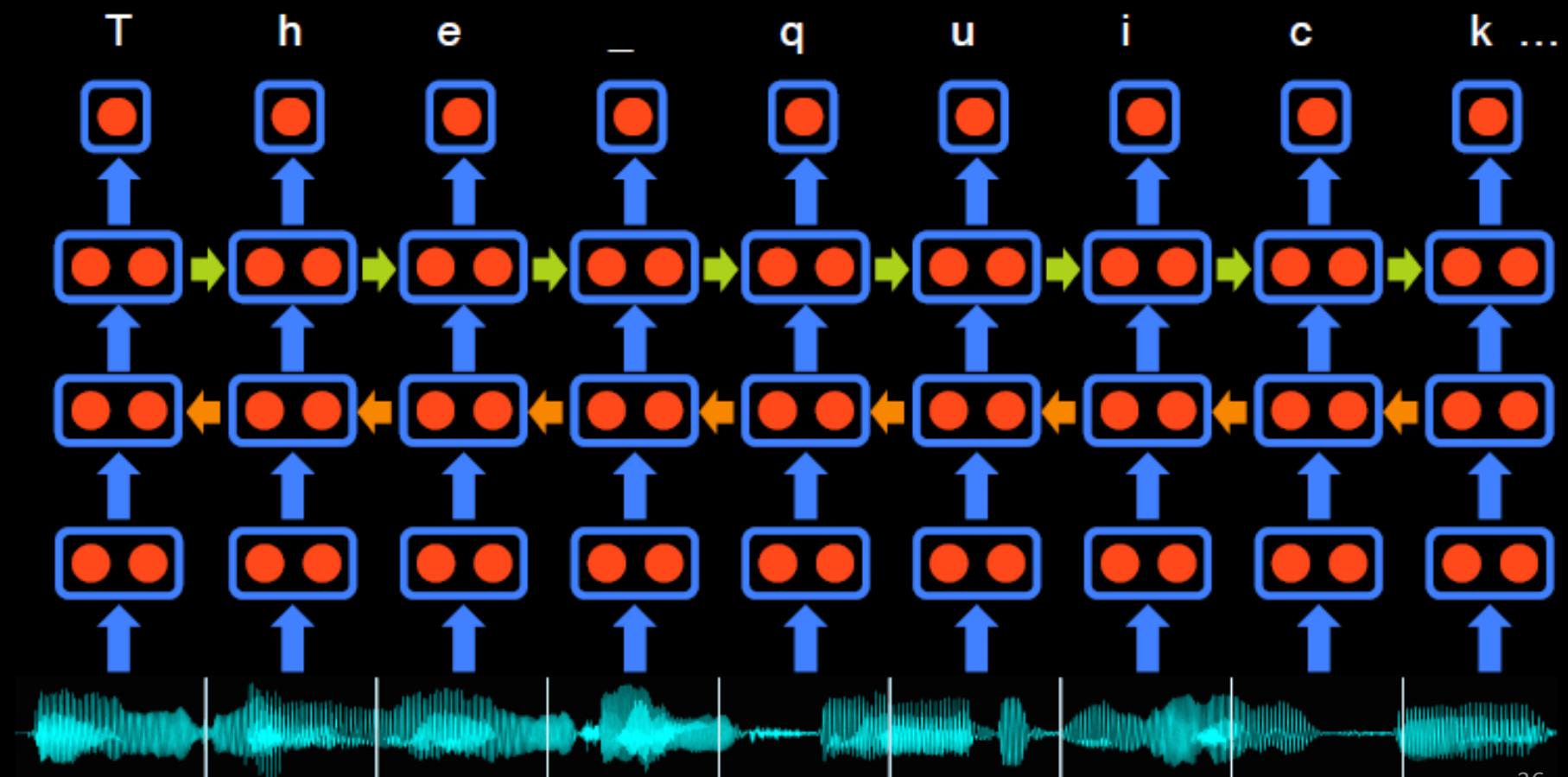


Günümüzde Kullanılan Ses Tanıma Sistemleri

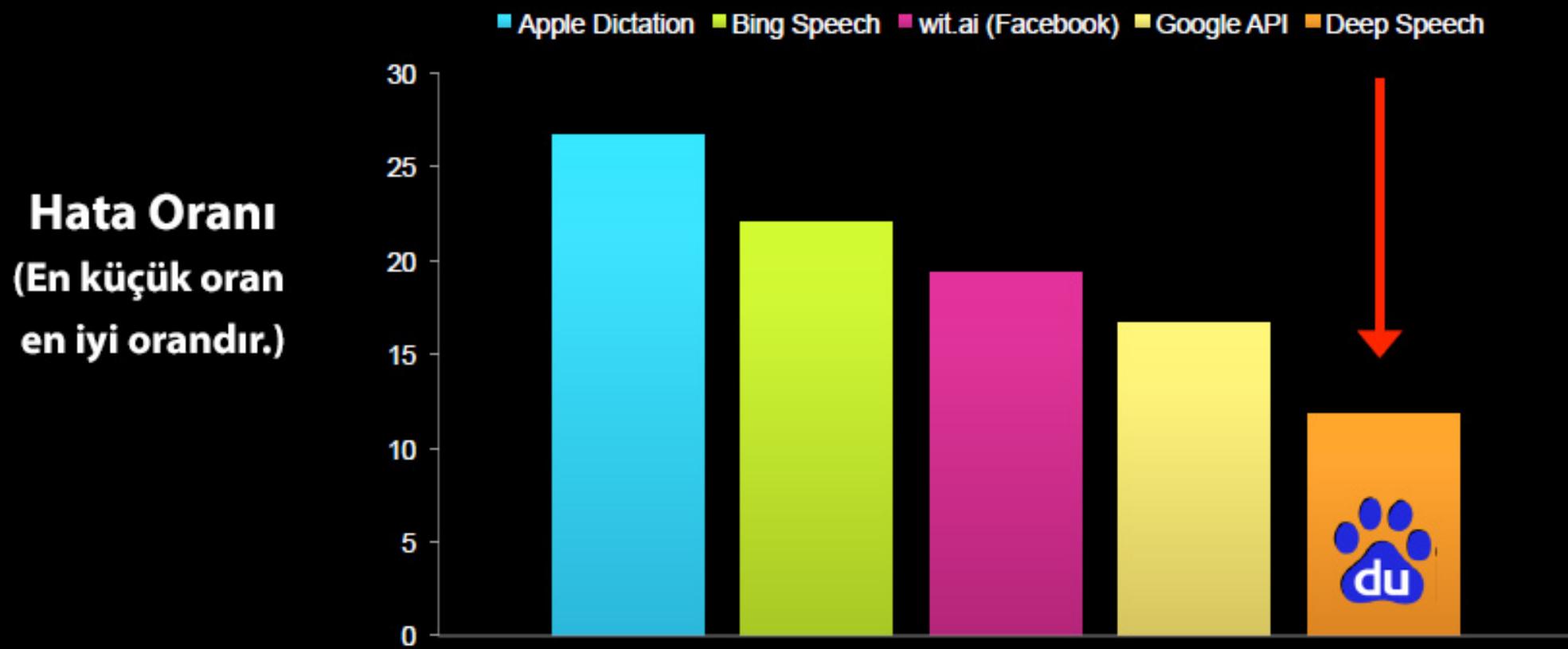


Baidu Deep Speech

Bi-directional Recurrent
Neural Network (BDRNN)



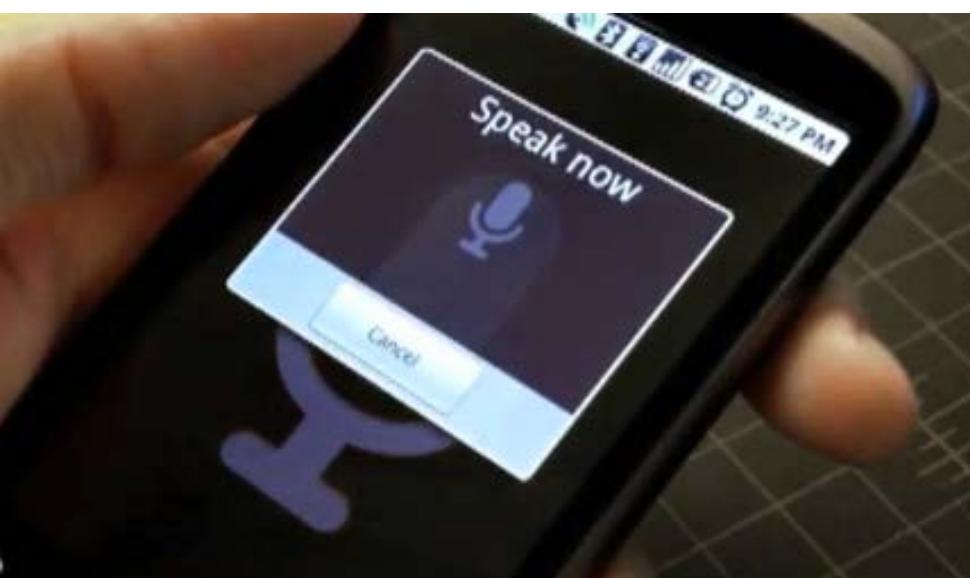
Konuşma Tanıma Hata Oranları



Çoğu insan %95 doğruluk ve
%99 doğruluk arasındaki farkı
anlamamaktadır.

%99 oyun değiştiren bir orandır.

%99 doğruluk ile akıllı
telefonlar sesle kullanılacak
hale getirilebilir.



Ses Nesnelerin İnternetinin Dönüşümünü Sağlayacaktır.



Araç Arayüzleri

Ev Gereçleri

Giyilebilir Araçlar



25 Eylül 2008 - DARPA Deep Learning Workshop

2009 - DARPA insansız hava araçlarının düşman toprakları üzerinde elde ettiği görüntü ve videoların karargâha aktarımıyla oluşturulan büyük veri (BigData) yiğinıyla baş edebilmek maksadıyla daha iyi bir istihbarat katmanı geliştirilmesi kapsamında 2009 yılında derin öğrenme çalışmalarına destek vermeye başlamıştır.

Hava Kuvvetleri ile Leland Stanford Junior Üniversitesi (Stanford University)arasında **15 Mart 2010 - 15 Haziran 2014** sonlanmıştır. Bu anlaşma kapsamında **\$2,693,119** ödenmiştir.

2015 - Amerikan Hava Kuvvetleri Araştırma Laboratuarı [Deep Learning Analytics](#) ile Target Recognition and Adaption in Contested Environments (TRACE) programı kapsamında 6 milyon dolarlık bir sözleşme imzaladı.

2015 - Deep Learning for Actionable Intelligence Discovery and Exploitation ABD Hava Kuvvetleri, derin öğrenme alanındaki devrimsel nitelikteki başarıları gördükten sonra görüntü analizinde derin öğrenme ve doğal dil işleme kullanımına yönelik ihale duyurusunu yaptı.



Uzaktan algılama firmalarının sayısı ve kapasitelerinin **hızlı artışı**

Derin Öğrenmenin Coğrafi Analizde Kullanılması

Başarılması gereken her özel bilgi ayıklama görevi için algoritma yazmaya ne yeteri kadar zaman ne de uzmanlık bulunmaktadır.



Günlük 350 milyon resim yükleme



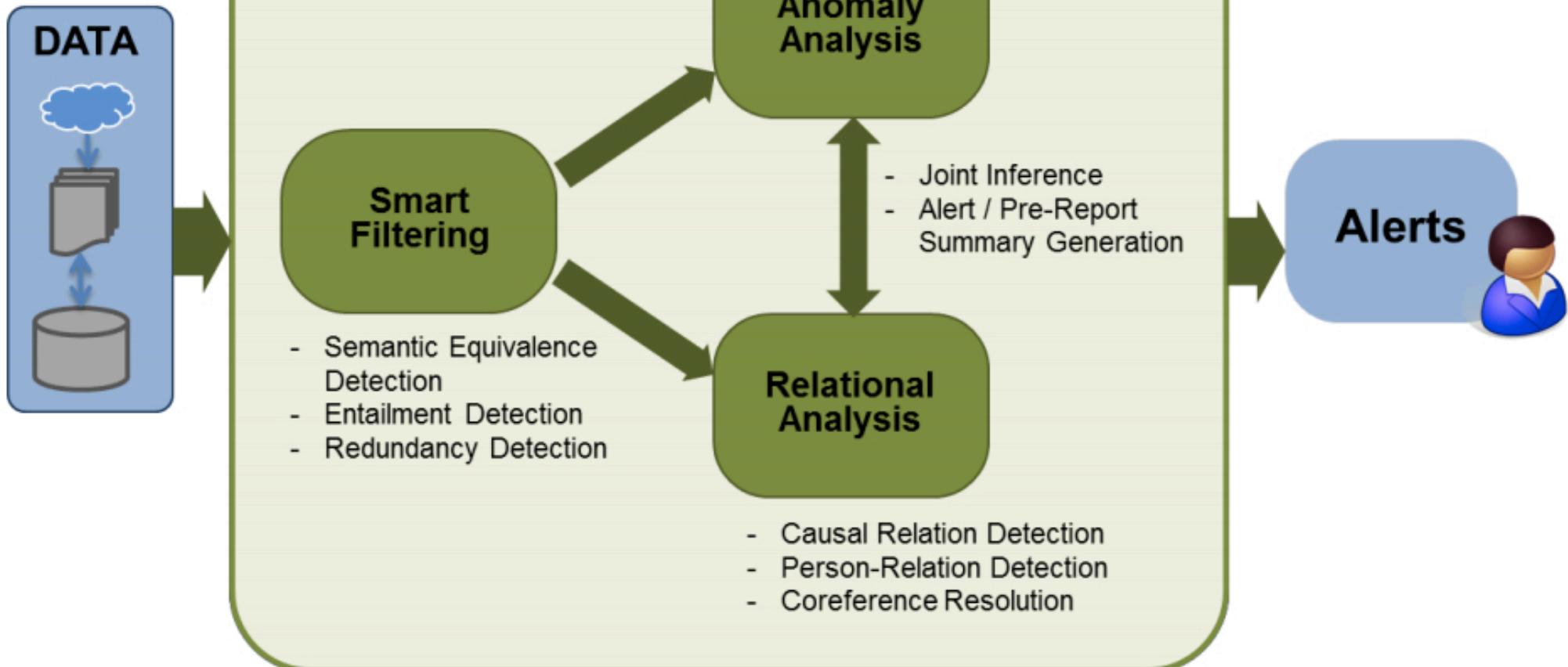
Her dakika 100 saatlik video yüklemesi

The GDELT Project

On binlere sosyal ve politik olay günlük olarak indekslenmektedir.

Derin Öğrenme, görevle ilgili içerik ve örüntüleri ham veri olarak makine hızında tanımlayan genel algoritmalar sunmaktadır.

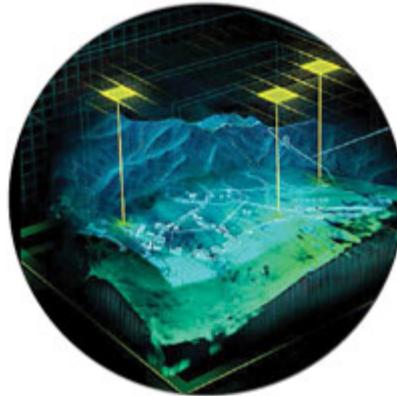
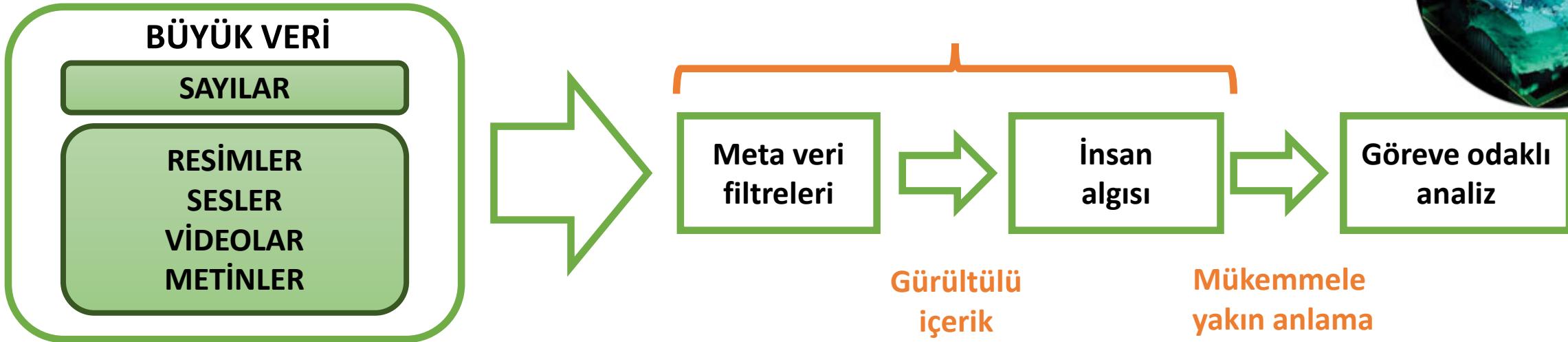
DEFT System



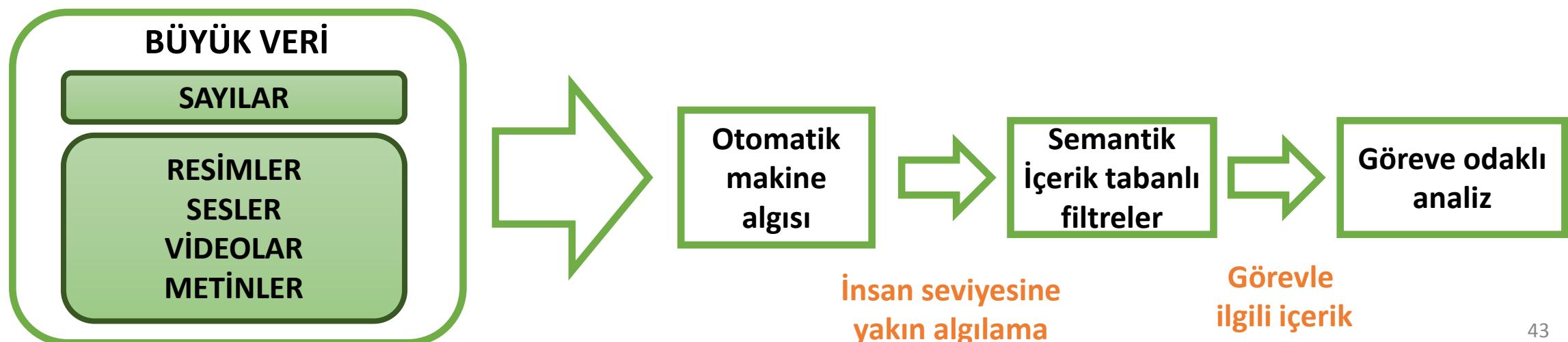
Kaynak: DARPA

Coğrafi İstihbarat Analizi İş Akışı

BUGÜN

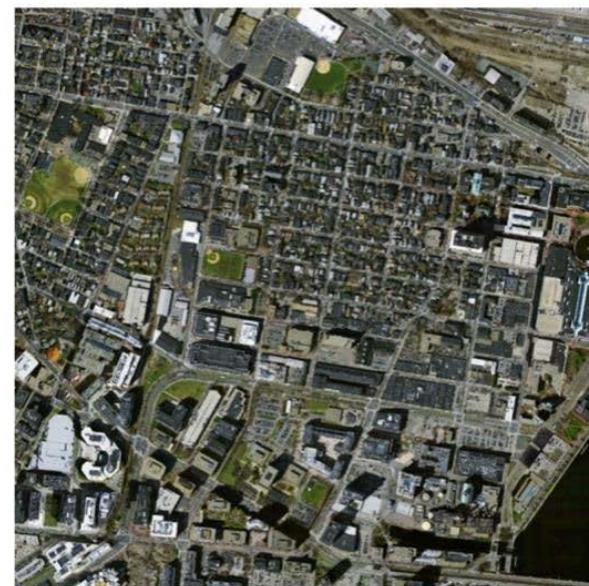


VİZYON

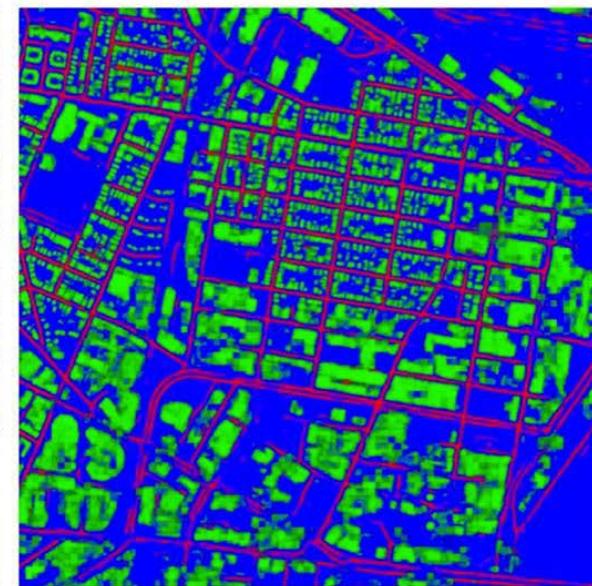


Uzaktan Görüntü Algılamanın Kullanımı

- Obje tespiti ve sınıflandırma
- Sahne bölütleme (segmentasyon)
- Arazi kullanımının sınıflandırılması
- Jeolojik özellik sınıflandırma
- Değişiklik Tespiti (Change detection)
- Ürün verimi tahmini
- Yüzey su tahmini
- Nüfus yoğunluk tahmini
- Süper çözünürlük
- Fotogrametri

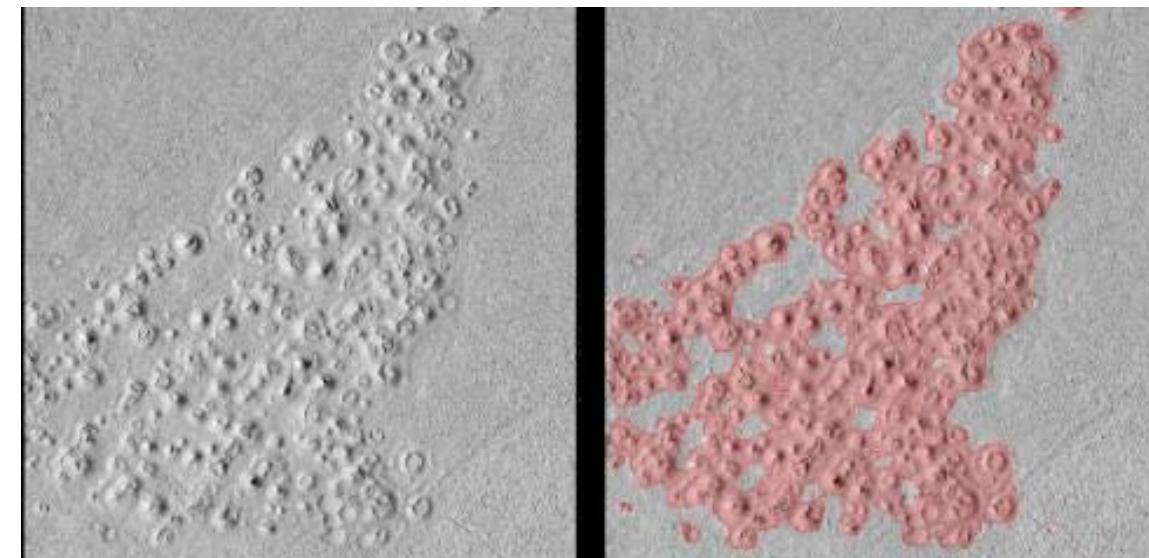


Input aerial image

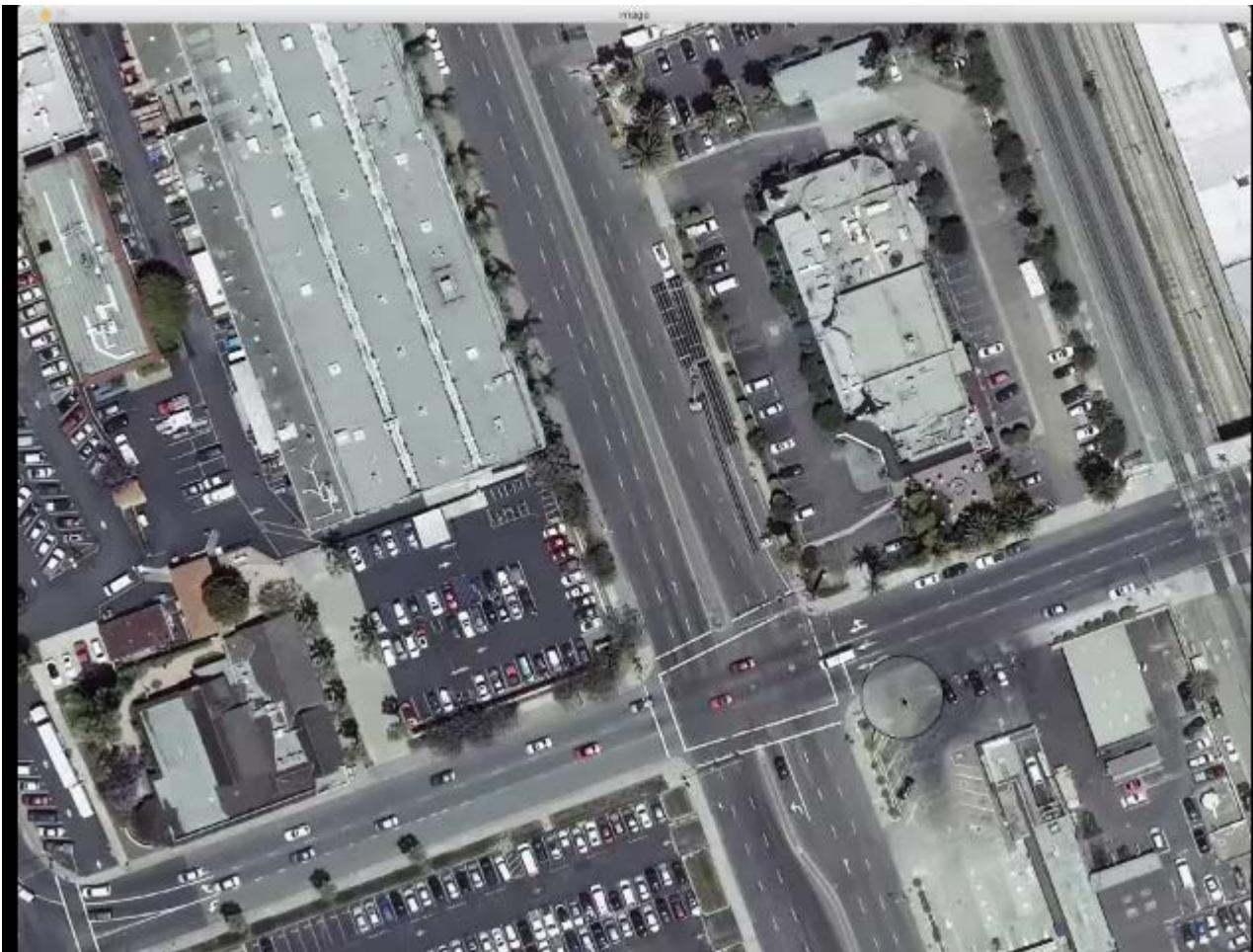


Predicted 3-channel label image

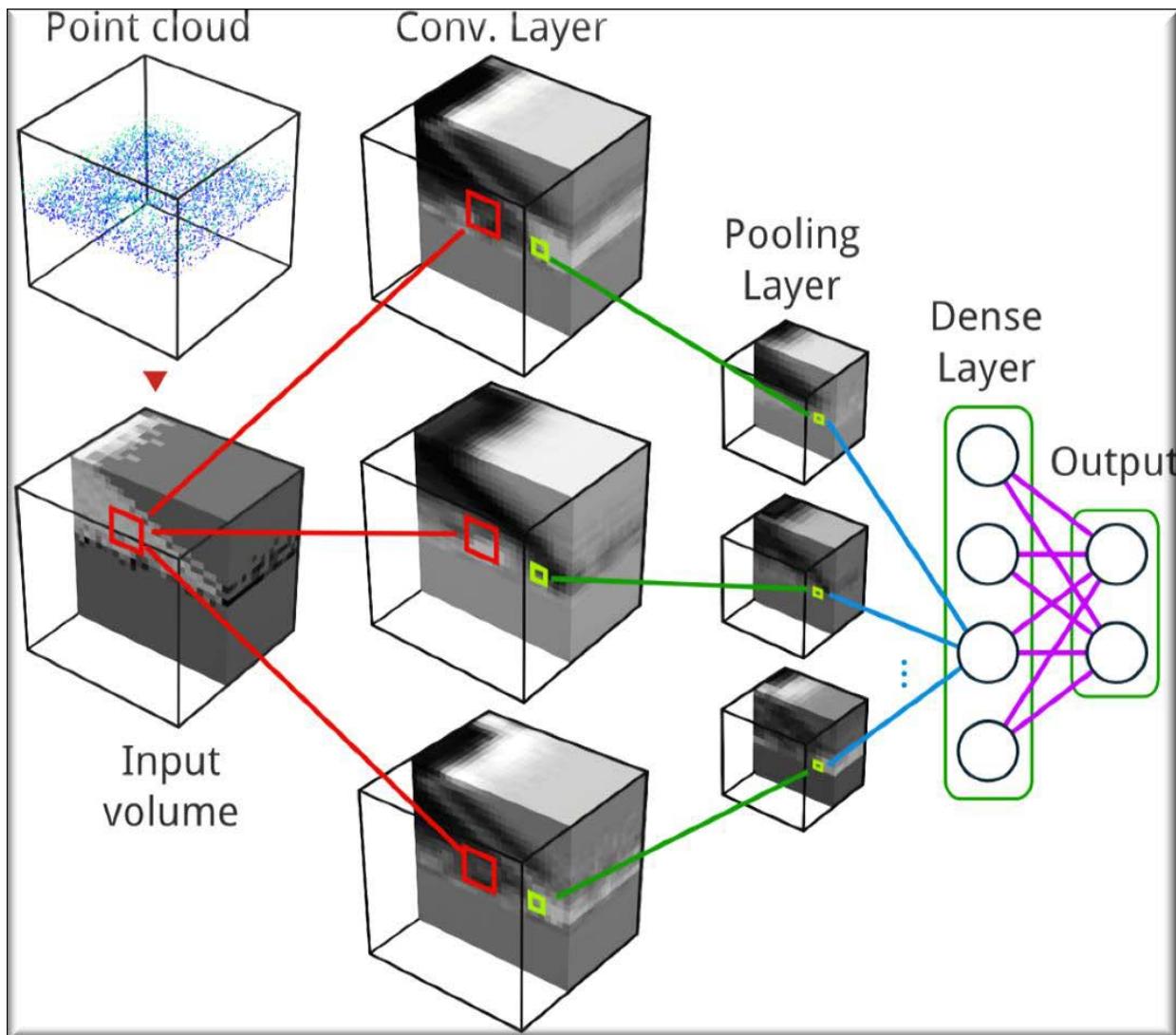
Keio Üniversitesi, Japonya – SPIE EI 2015



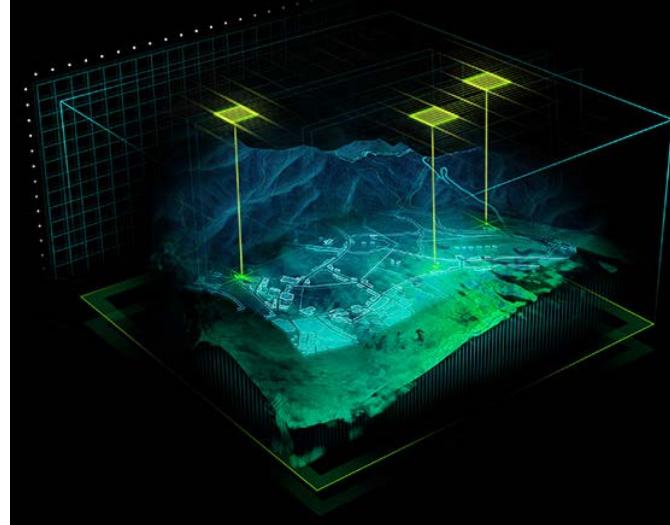
Derin Öğrenme Yapısıyla Analiz İşlemi

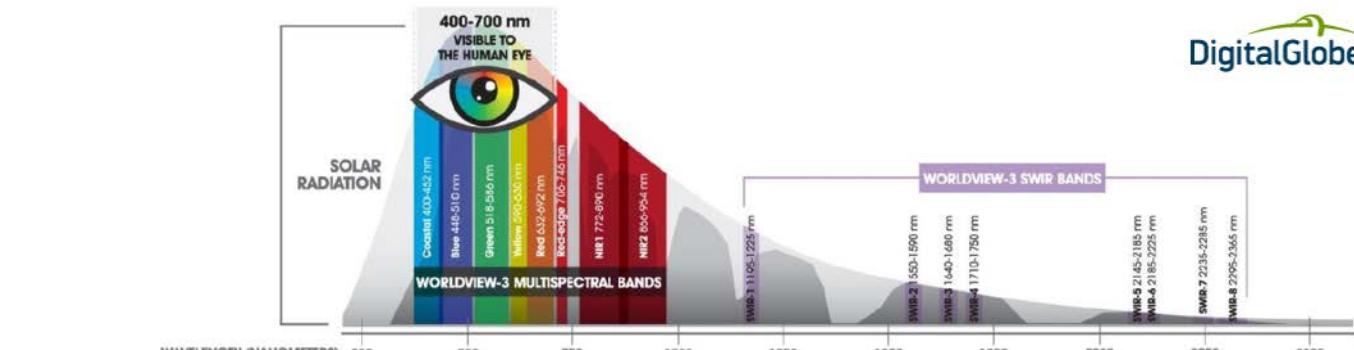


GELİŞMİŞ GÖRÜNTÜLEME YÖNTEMİ



- Konvolüzasyonel Sinir Ağları yapısı aşağıdaki veri türlerini desteklemektedir:
 - MSI/HSI data cubes
 - SAR imagery
 - Volumetric data, e.g.
- LIDAR Low-TRL research topics







Kazalarda
Her yıl 1,3 milyon,
günde 3.287 kişi ölüyor.

Ayrıca 20-50 milyon
insan yaralı veya engelli
oluyor.

Trafik kazaları genel
ölüm oranına (%2,2)
göre 9'uncu sırada yer
alıyor.

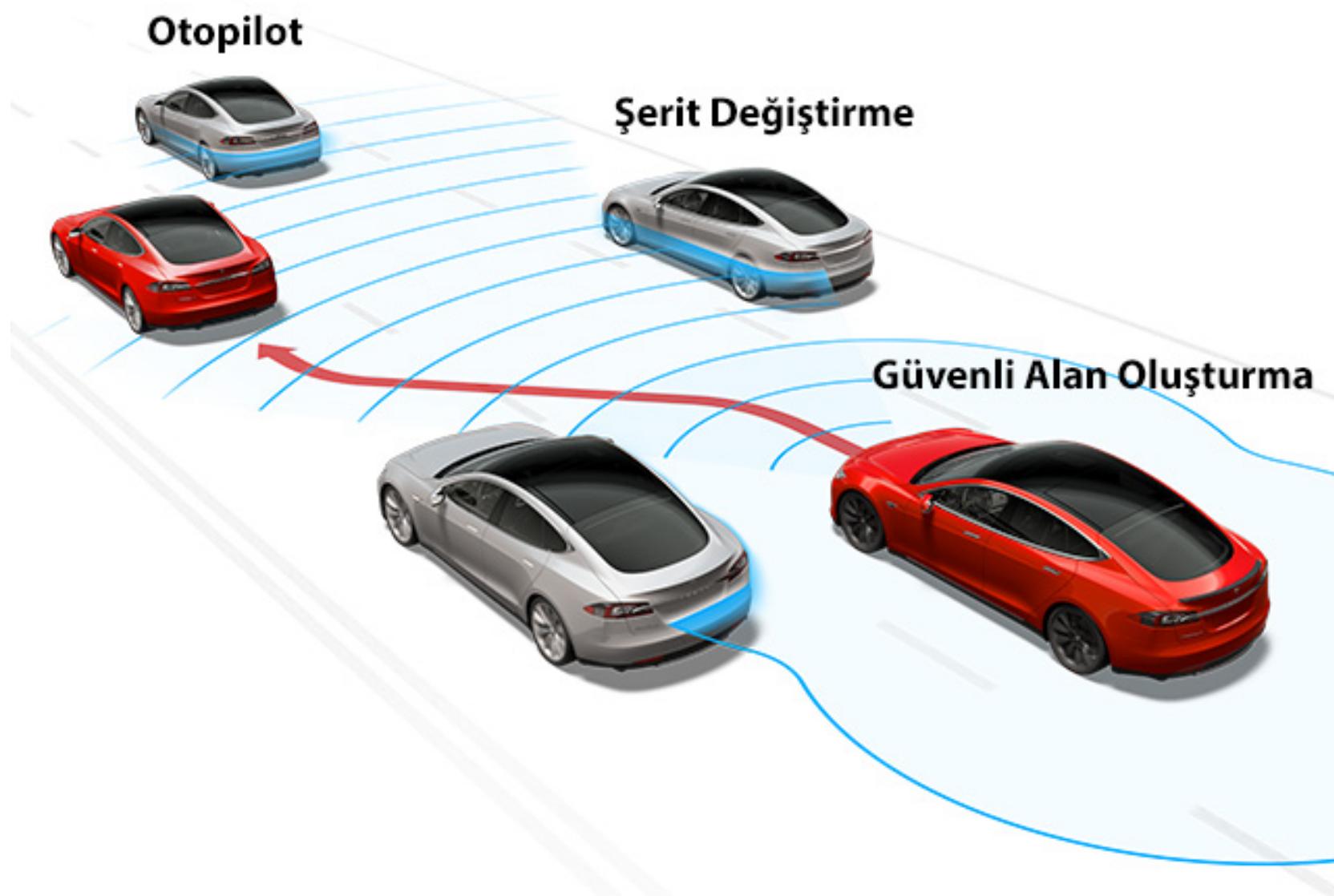
Her yıl 25 yaşın altında
400.000 kişi yaşamını
yitiriyor (Günde 1.000
kişi).

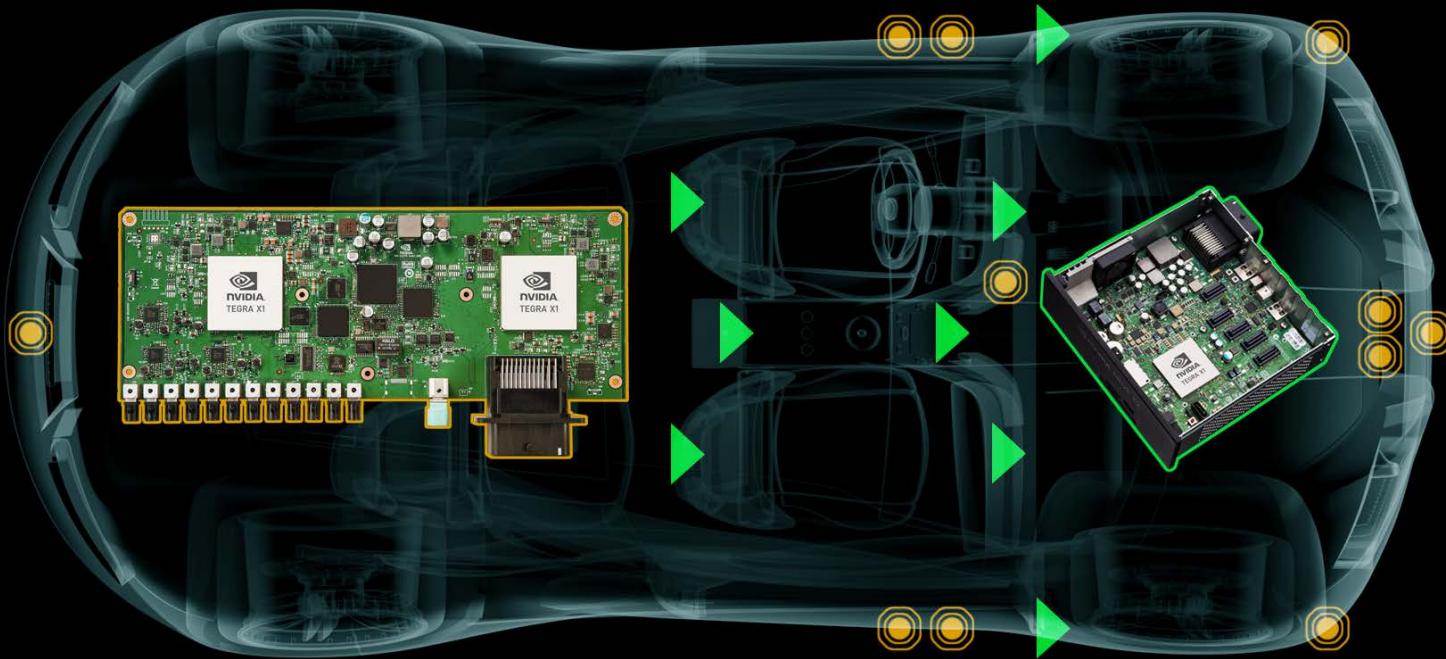
Yol kazalarının genel
maaliyeti 518 milyar
dolar (GSMH'nin %1-2
sine karşılık)

Tedbir alınmazsa 2030
yılında insanların ölüm
nedeni sıralamasında 5.
sıraya yükselecek.

Otonom







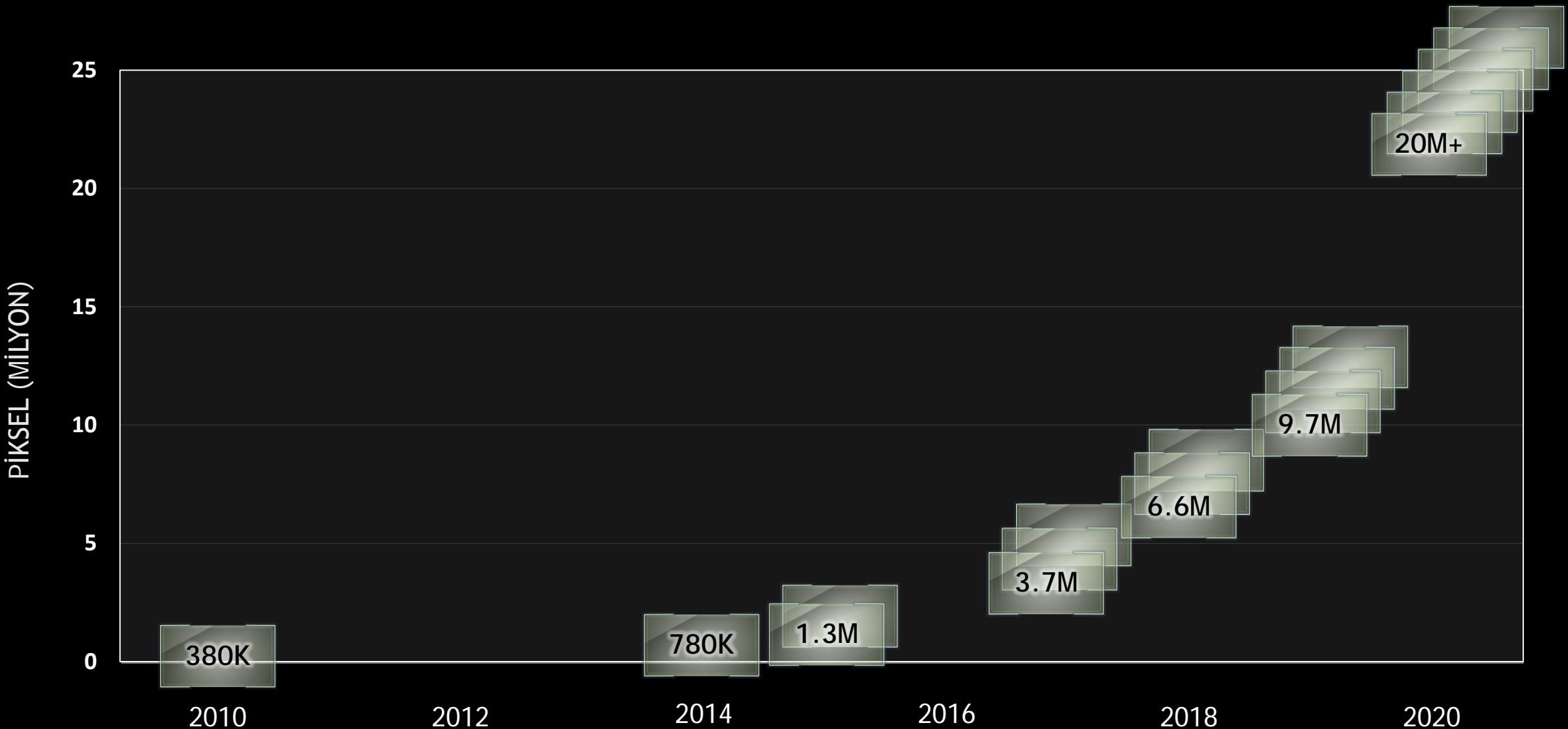
NVIDIA TEGRA X1
DSTEKLİ

DRIVE CX
KOKPİT GÖRÜNTÜLEME BİLGİSAYARI

DRIVE PX
OTO-PİLOT BİLGİSAYARI



ARAÇ GÖRÜNTÜLEME PİKSEL ARTIŞI



NVIDIA DRIVE™ CX

GELİŞMİŞ KOKPİT GÖRÜNTÜLEME BİLGİSAYARI

16.6M piksel maksimum çözünürlük

- Maxwell — NVIDIA'nın en yeni GPU mimarisi
- NVIDIA DRIVE Studio

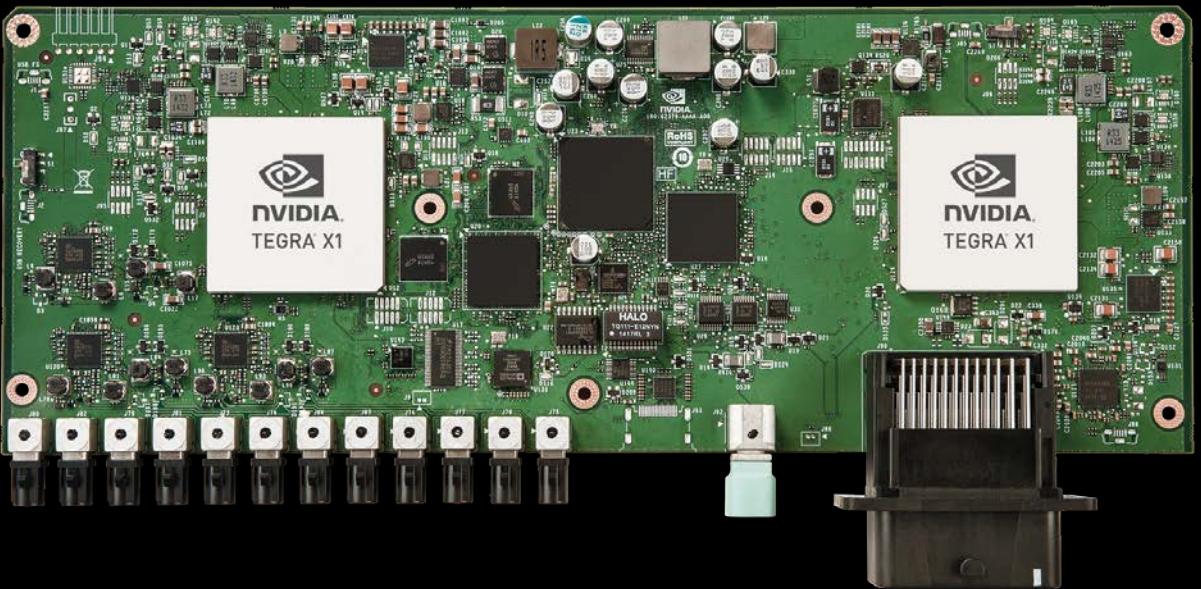


NVIDIA DRIVE™ PX

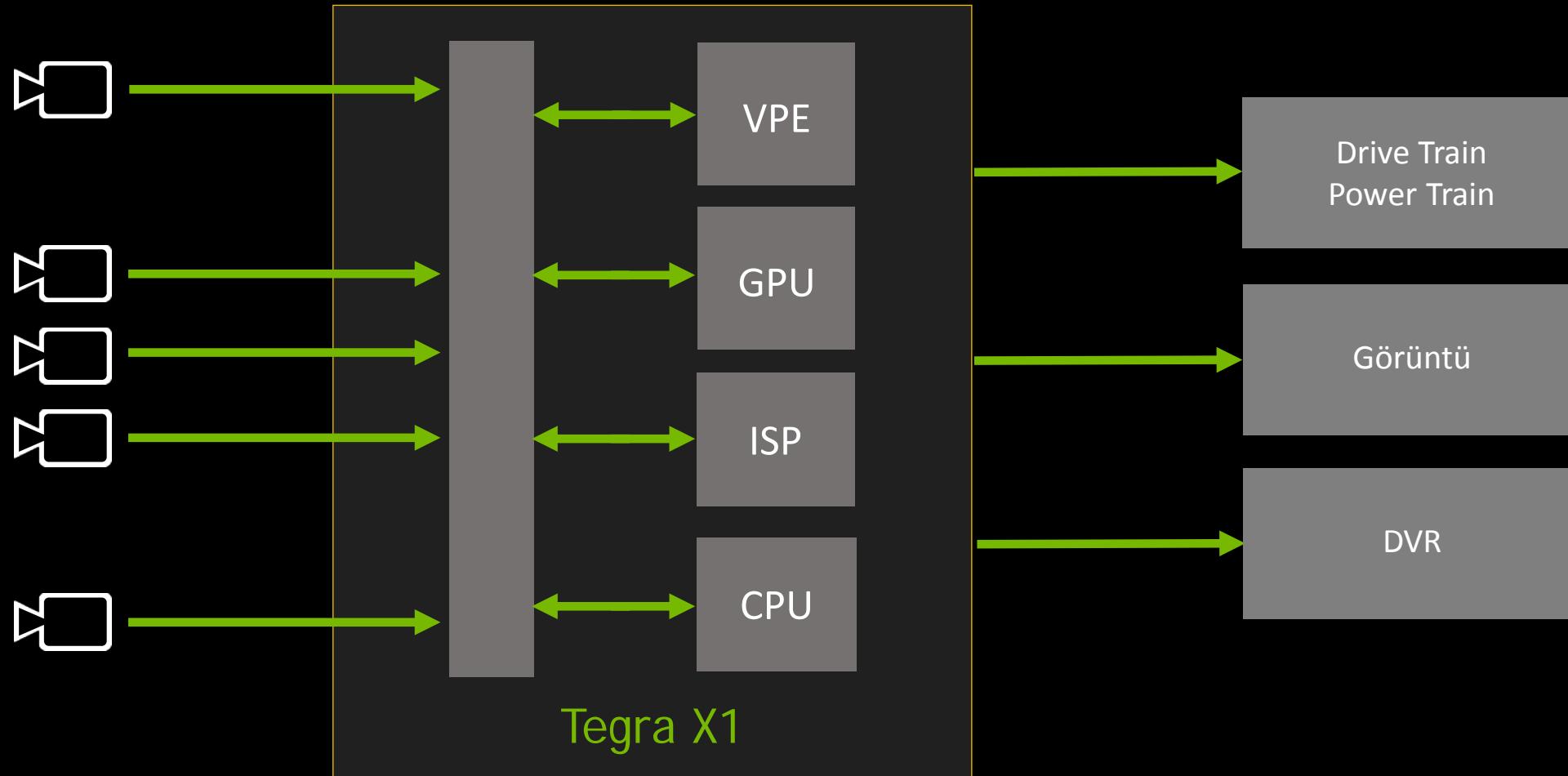
OTO-PİLOT BİLGİSAYAR

Dual Tegra X1 • 12 kamera girişi • 1.3 GPix/sec

- ▶ 2.3 Teraflops mobile supercomputer
- ▶ CUDA programmability
- ▶ Derin Sinir Ağları Yapılı Bilgisayarlı Görü
- ▶ Çepeçevre Görünüm

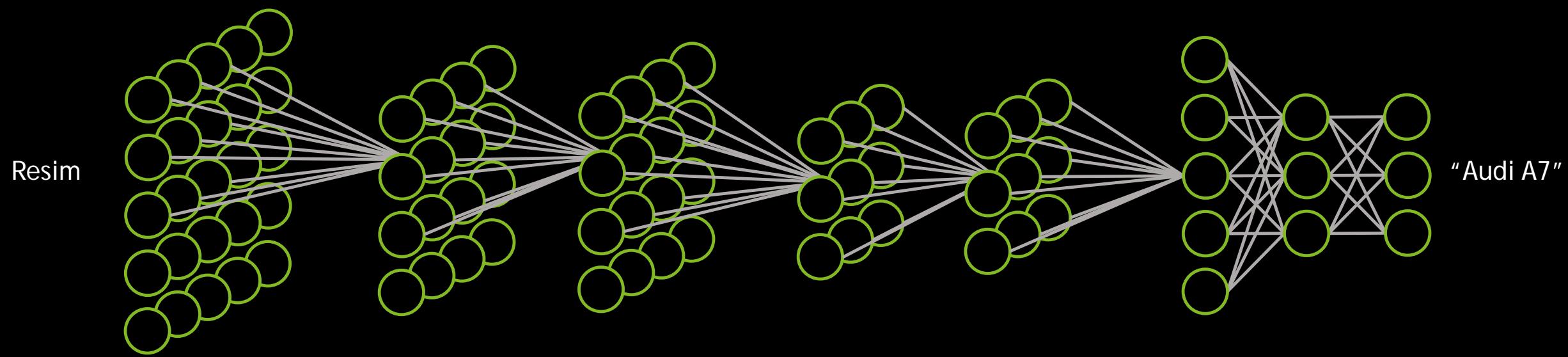
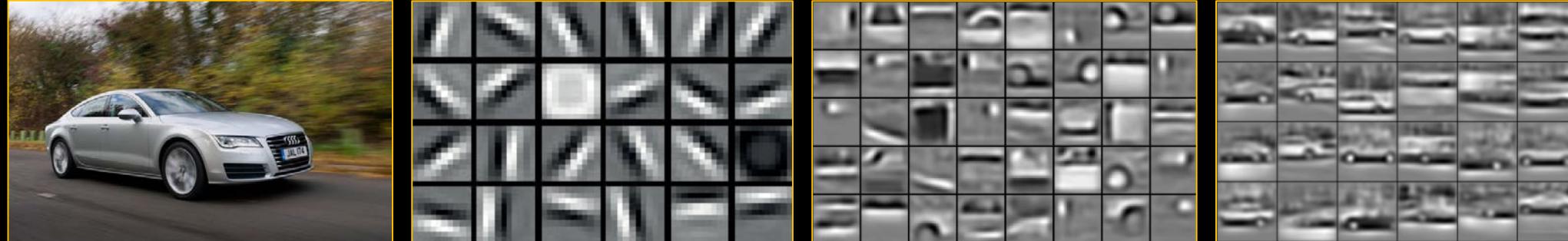


NVIDIA DRIVE PX PIPELINE

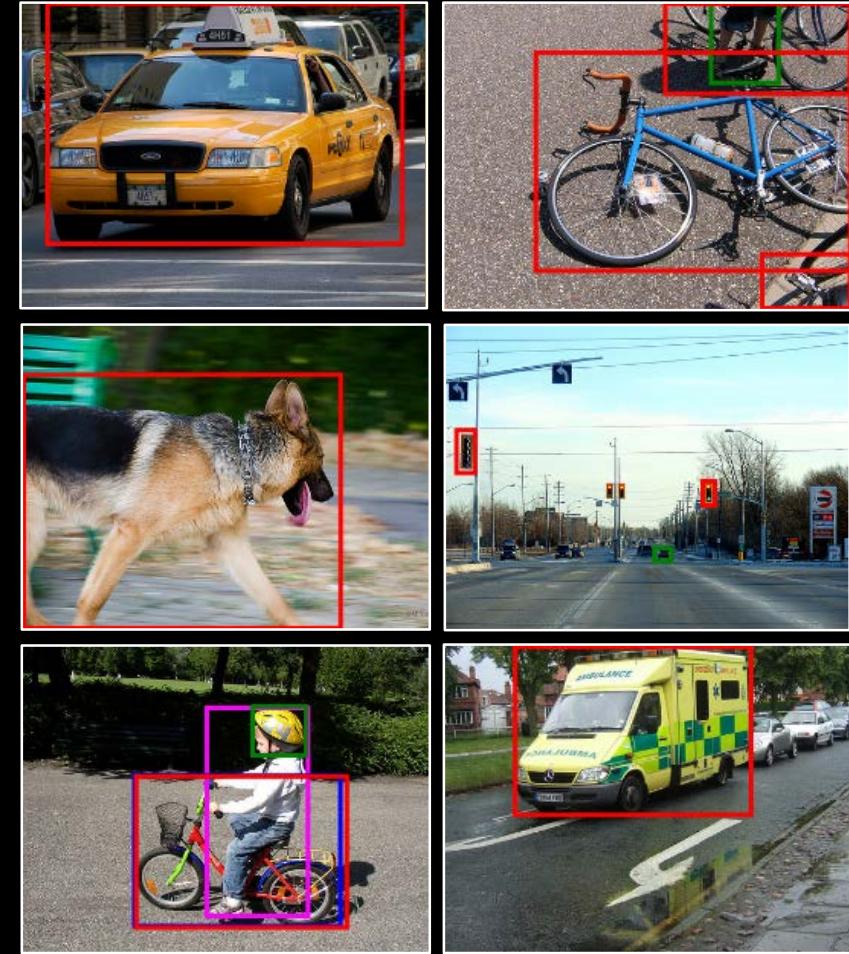
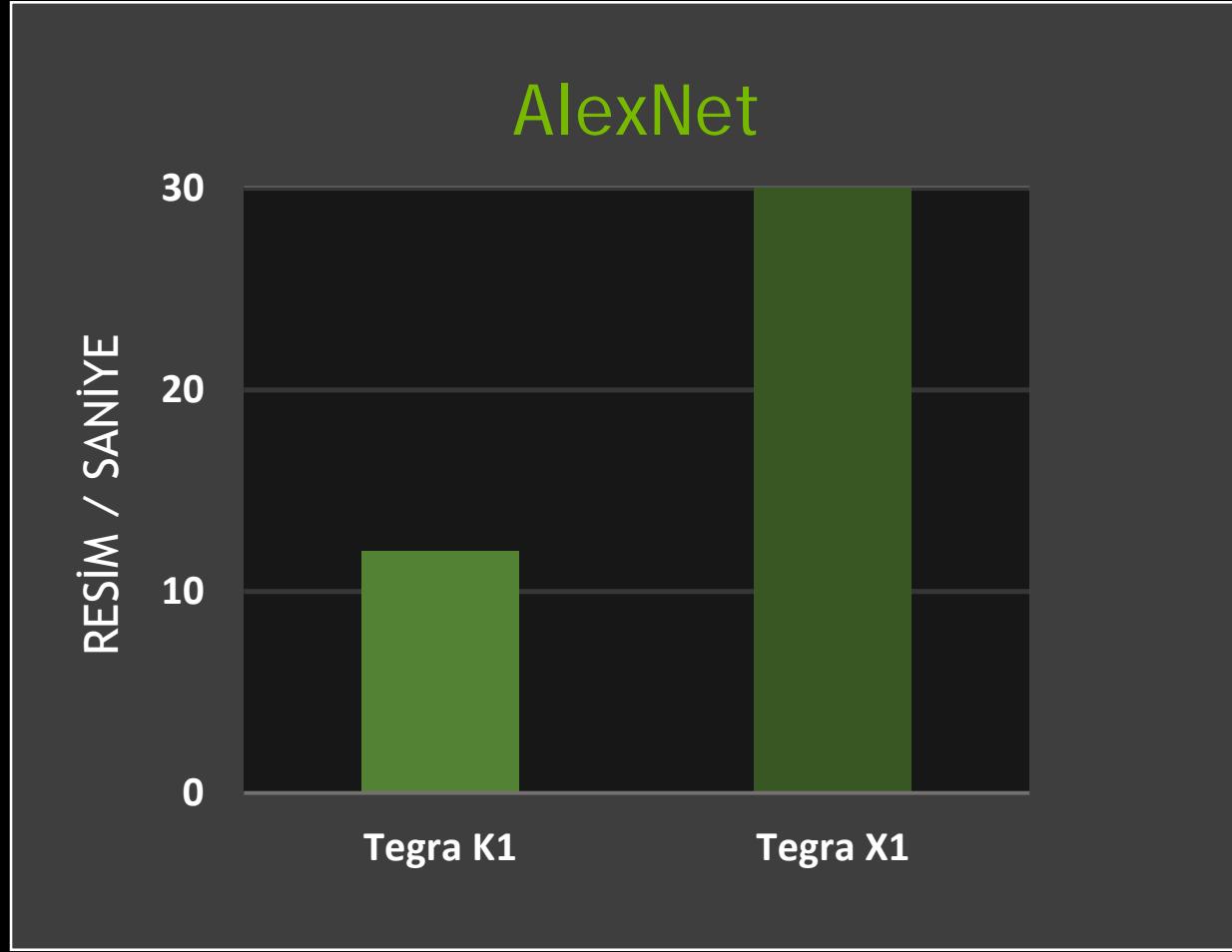


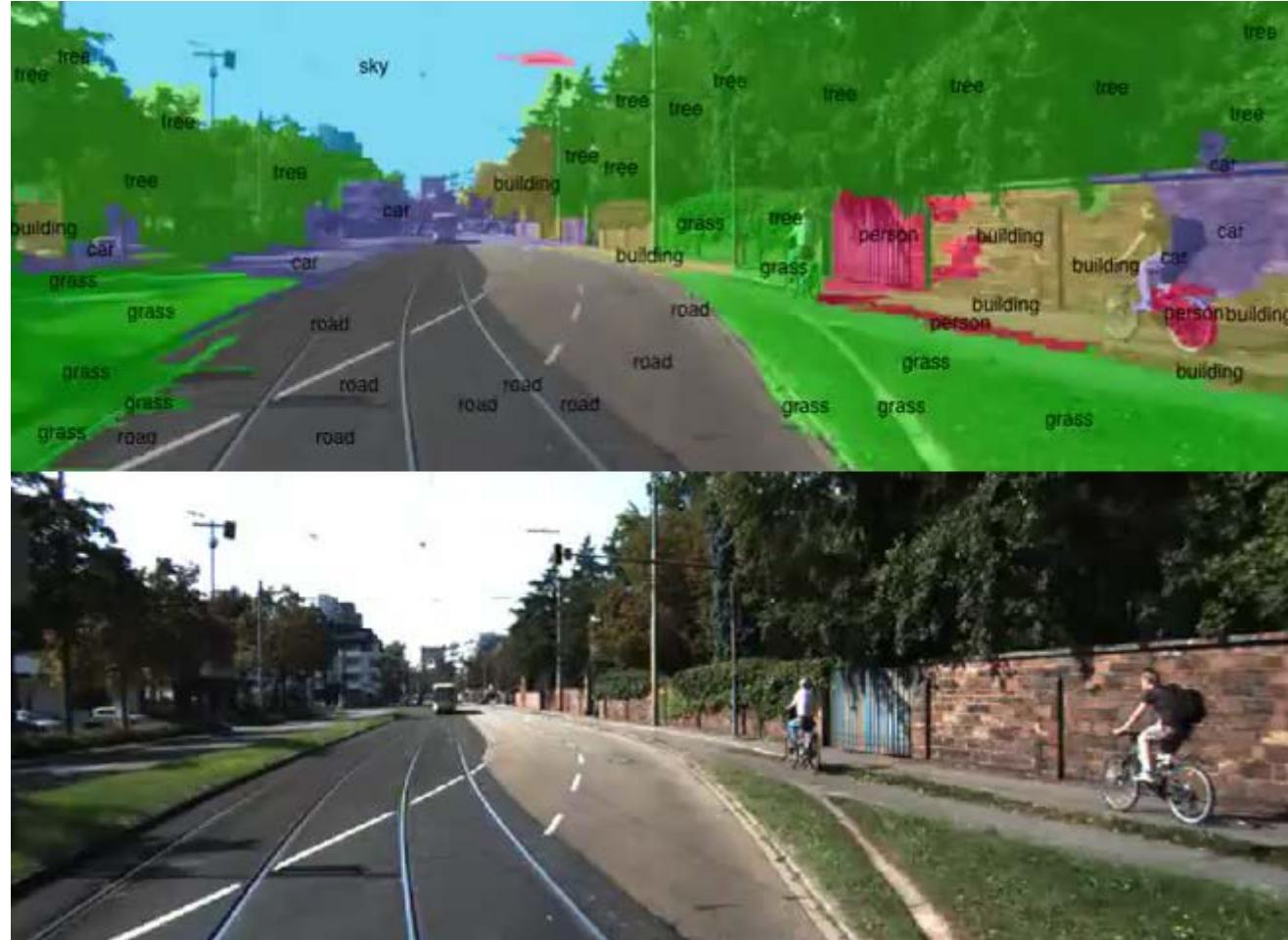
Çevre Modeli • Durumsal Farkındalık • Yol Gösterimi • Öğrenme

DERİN SİNİR AĞI NASIL GÖRÜR?



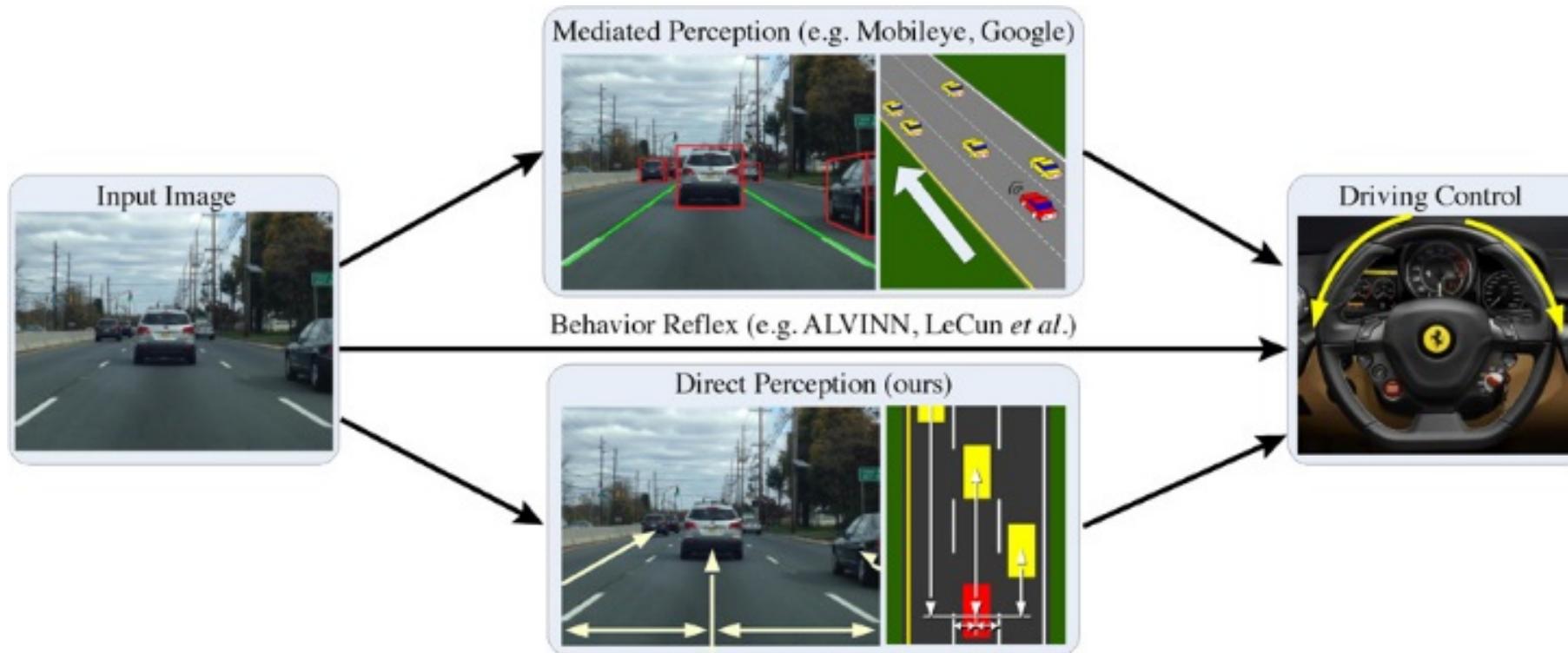
TEGRA X1 SINIFLANDIRMA PERFORMANSI

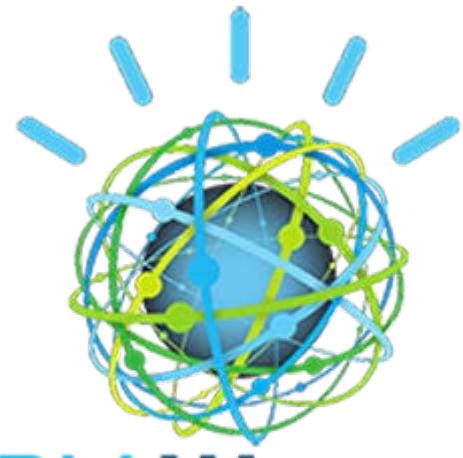




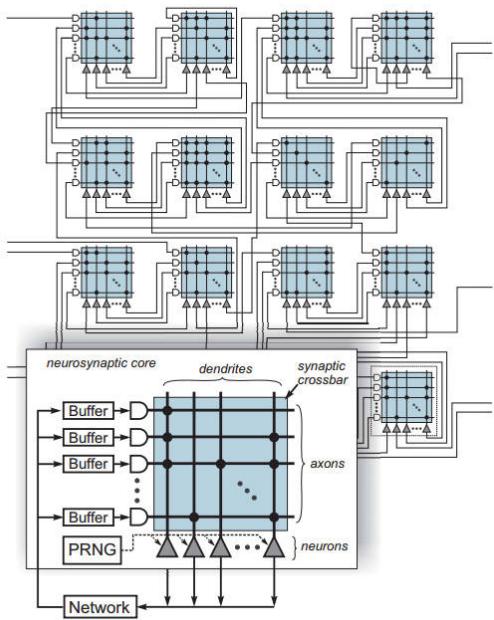


PRINCETON VISION GROUP





IBM **Watson**



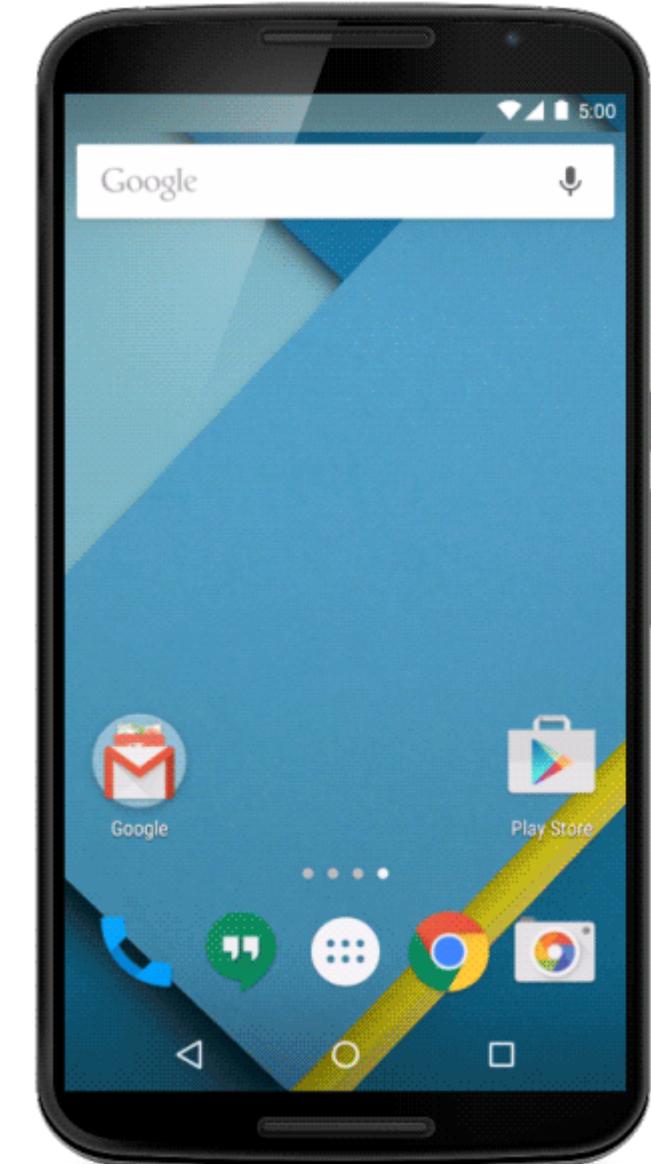
Google

Stanford

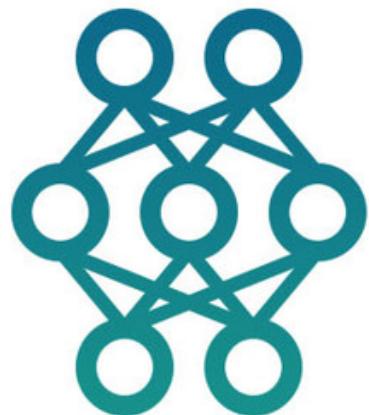
University



Google arama
motorunda **her 20**
aramadan birisi
sağlık ile ilgili bilgi
aramaya yönelik
yapılıyor.



DEEP LEARNING

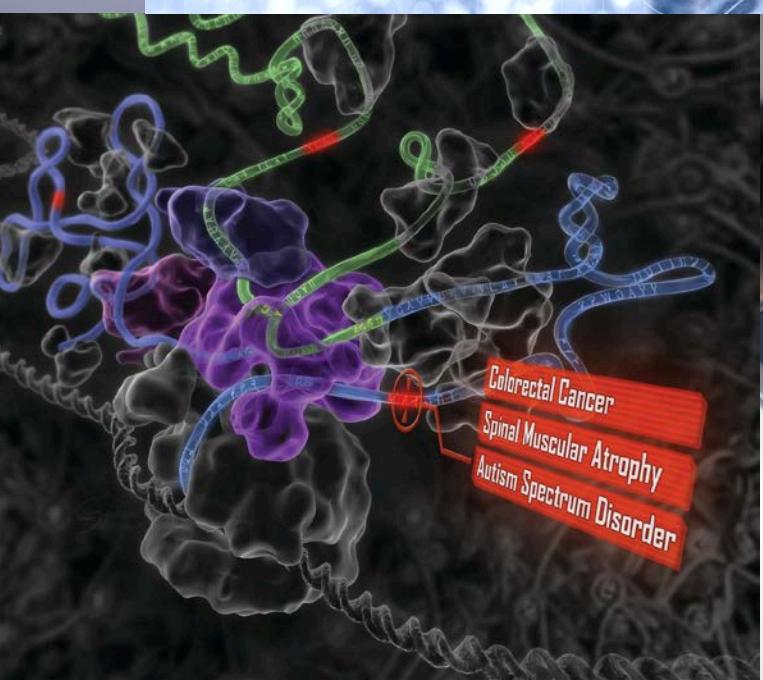


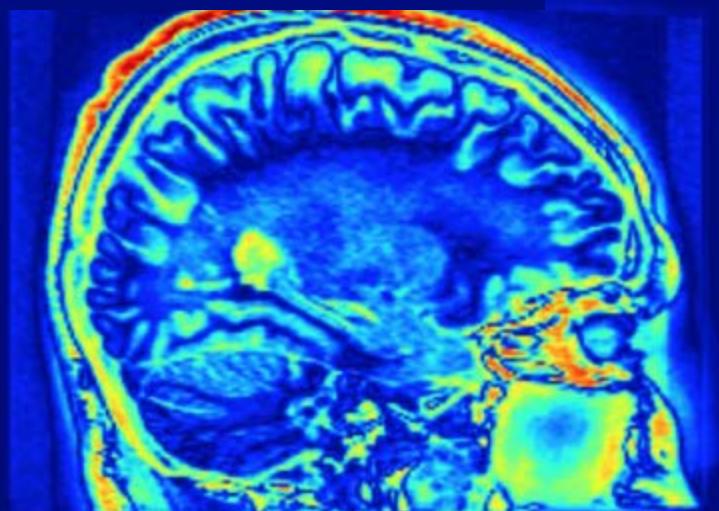
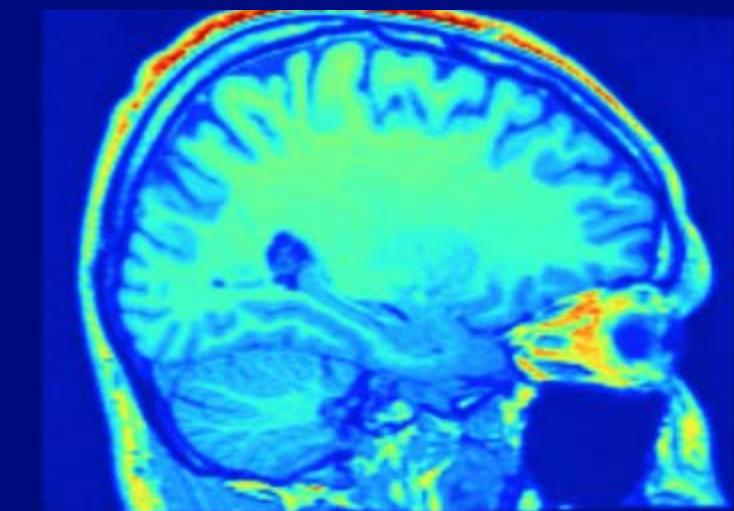
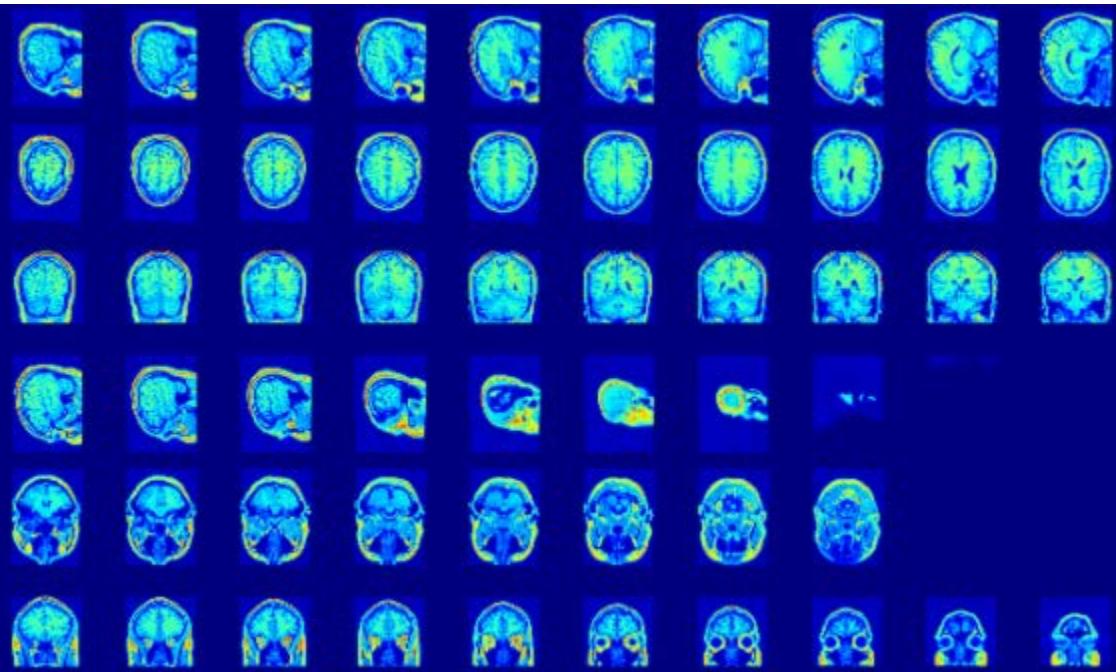
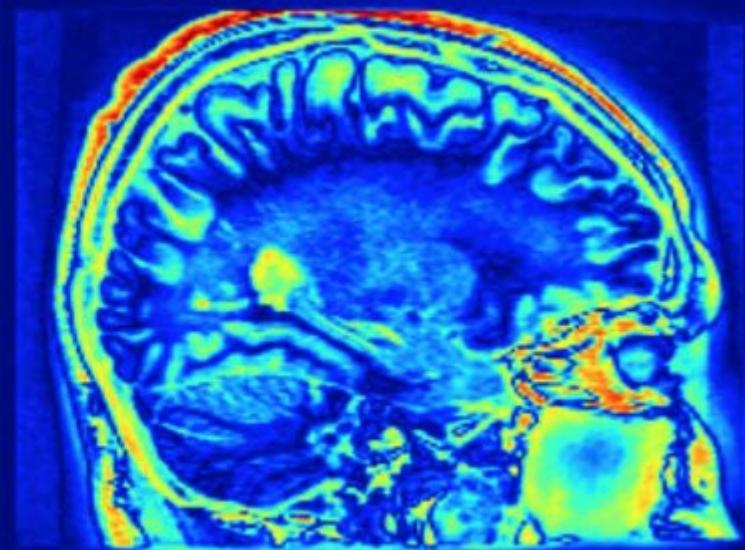
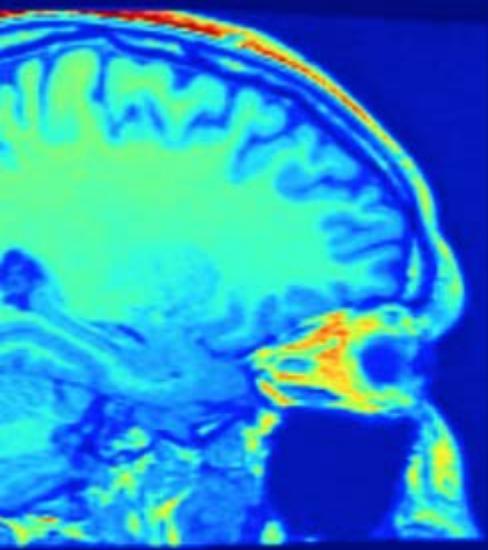
GENOMICS



PRECISION MEDICINE







Google Ne Yapıyor?

- Deepmind Firmasını 400 Milyon dolar karşılığında bünyesine katmıştır.
- Google derin öğrenme uzmanlarını işe alma konusunda açık bir kısıtlarının olmadığını, bu alanda 10'larda yıllık bilimsel çalışmaya eş değer bir çalışmanın çok kısa sürede gerçekleştiğini ifade etmiştir.
- Milyarlarca veri örneğinden öğrenmek için paralel olarak binlerce CPU ve GPU'lara yatırım yapmıştır.
- Resim tanıma, konuşma tanıma gibi akademik yarışmalarda sıkılıkla birinci olmuş ve makaleler yayımlamıştır.
- Yoğun ve artan bir şekilde gerçek uygulamalar üzerinde derin öğrenme tecrübesini kullanmıştır. Son iki yıl içinde 50 ürününde derin öğrenme kullanmıştır. (Resim arama, Android konuşma tanıma, StreetView, Ads yerleştirme...)



2014 yılında işe almıştır.

- 2013** – Görsel arama motoru
- 2013** – Silikon Vadisinde derin öğrenme lab. Kurulumu
- 2014** – Andrew Ng'nin Google'dan transfer edilmesi
- 2015** – Ses-metin dönüşümü
- 2015** – AskADoctor, Bir Doktara Sor Uygulaması

.....



2013 - New York Üniversitesi’nde görev yapan Yann LeCun'u işe almıştır. Şu anda Facebook Yapay Zekâ Laboratuarı'nın yöneticiliğini yapmaktadır.

Siri ses tanıma yazılımında derin öğrenme kullanıyor.



Ağustos 2015 - Apple, Ne Düşündüğünüzü Tahmin Etmek için derin öğrenme uzmanı işe alım duyurusu yaptı.

Ekim 2015 - Apple, Siri'yi daga akıllı hale getirmek için **VocalIQ girişim firmasını satın aldı.**

Ekim 2015 – Apple, telefonlarda kullanıcı verisine çok fazla ihtiyaç duymayan ileri yapay zeka sistemi geliştiren **Perceptio'yu satın aldı.**

DERİN ÖĞRENME TAVSİYE MOTORU

TECRÜBE



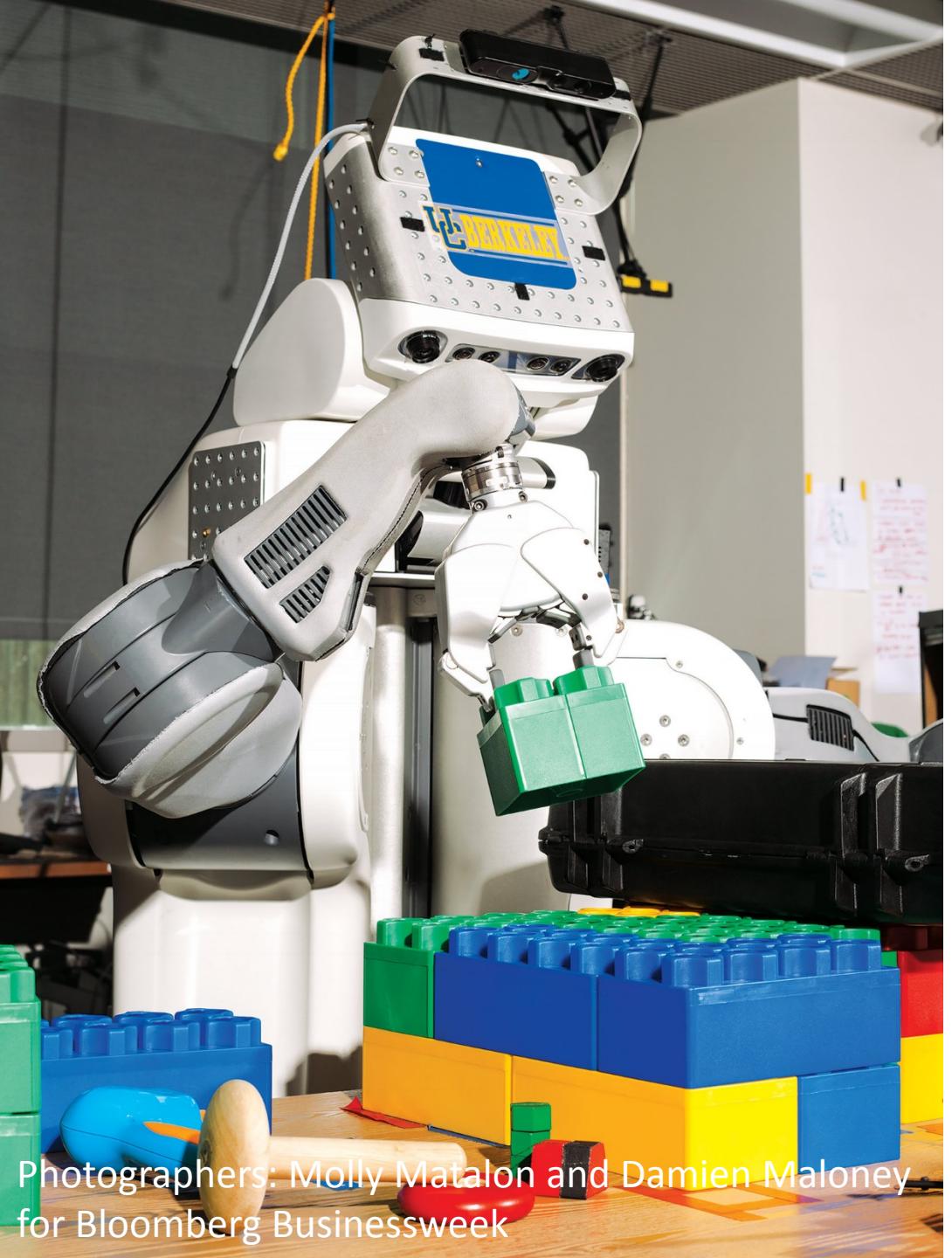
TECRÜBE +



TECRÜBE ++







Photographers: Molly Matalon and Damien Maloney
for Bloomberg Businessweek



[Chelsea Finn](#), Xin Yu Tan, [Yan Duan](#), [Trevor Darrell](#), [Sergey Levine](#), [Pieter Abbeel](#). Learning Visual Feature Spaces for Robotic Manipulation with Deep Spatial Autoencoders.

Preprint: [\[PDF\]](#) [\[arXiv\]](#)

[Sergey Levine*](#), [Chelsea Finn*](#), [Trevor Darrell](#), [Pieter Abbeel](#). End-to-End Training of Deep Visuomotor Policies. Presented at the IEEE International Conference on Robotics and Automation (ICRA) 2015 Late Breaking Results Session, Thursday May 28, 2015, Seattle WA.

Preprint: [\[PDF\]](#) [\[arXiv\]](#)

[Sergey Levine](#), Nolan Wagener, [Pieter Abbeel](#). Learning Contact-Rich Manipulation Skills with Guided Policy Search. Presented at the IEEE International Conference on Robotics and Automation (ICRA) 2015, Wednesday May 27, 2015, Seattle WA.

Best Robotic Manipulation Paper Award.

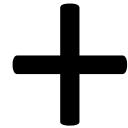
Preprint: [\[PDF\]](#) [\[website\]](#)



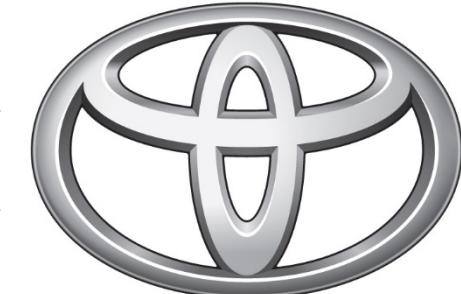
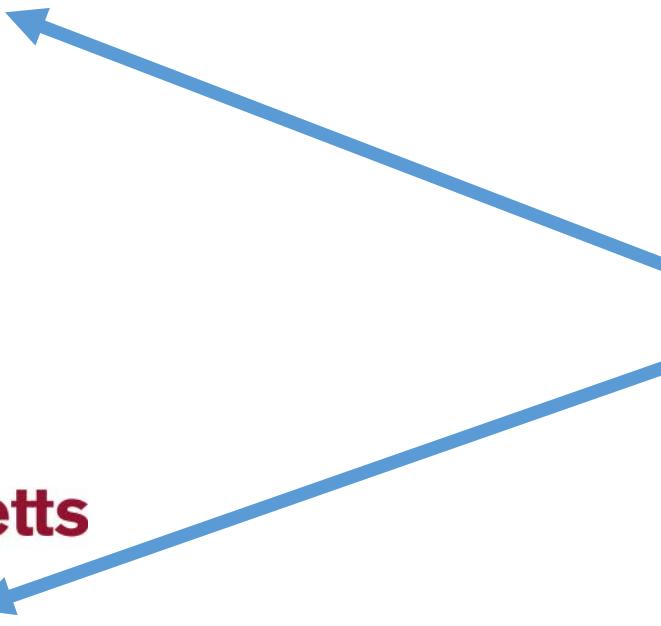
2014 yılında derin öğrenme uygulamaları geliştirmek üzere kurulmuştur.

Robot kollarda derin öğrenme kullanmak maksadıyla 7,3 milyon dolarlık bir anlaşma imzalanmıştır.

Stanford University



Massachusetts
Institute of
Technology



TOYOTA

50 Milyon Dollar

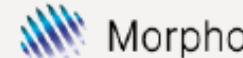
1 Milyar Dollar

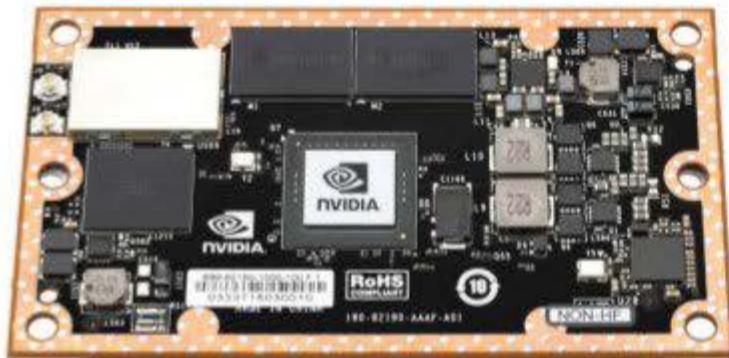


Dr. Gill Pratt (eski DARPA Program Yöneticisi)

embedded VISION SUMMIT







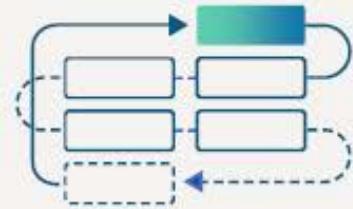
JETSON TX1	
GPU	1 TFLOP/s 256-core Maxwell
CPU	64-bit ARM A57 CPUs
Memory	4 GB LPDDR4 25.6 GB/s
Storage	16 GB eMMC
Wifi/BT	802.11 2x2 ac/BT Ready
Networking	1 Gigabit Ethernet
Size	50mm x 87mm
Interface	400 pin board-to-board connector



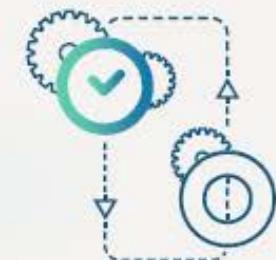
Security and user privacy



Low Latency



Efficient use of network bandwidth



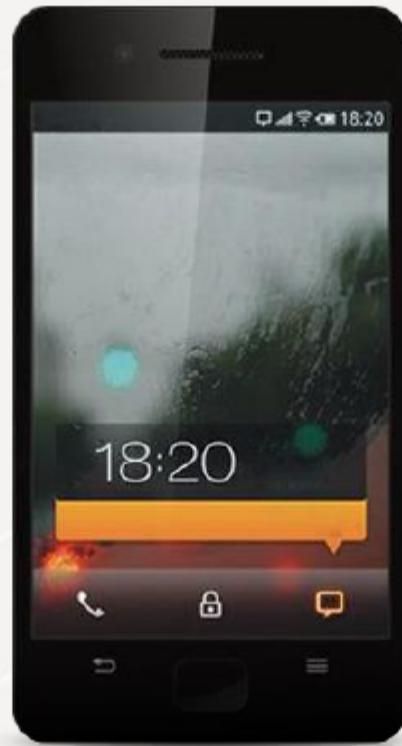
Reliability

Process data closest to the source, complement cloud





QUALCOMM®



Görsel algı iş yükü

Yoğun hesaplama
Büyük ve karmaşık
sinir ağ modelleri

Qualcomm Technologies, Inc. yaklaşımı

- Başlangıç SVD yaklaşımında NYU'dan Denton, et. al. makalesi temel alınmış,
- Tek katmanları çoklu katmanlarla değiştirme yaklaşımı
- Yaklaşım sadece sıkıştırılmış katmanların üzerindekileri değil tüm katmanlarda iyileştirmeye olanak sağlamıştır.

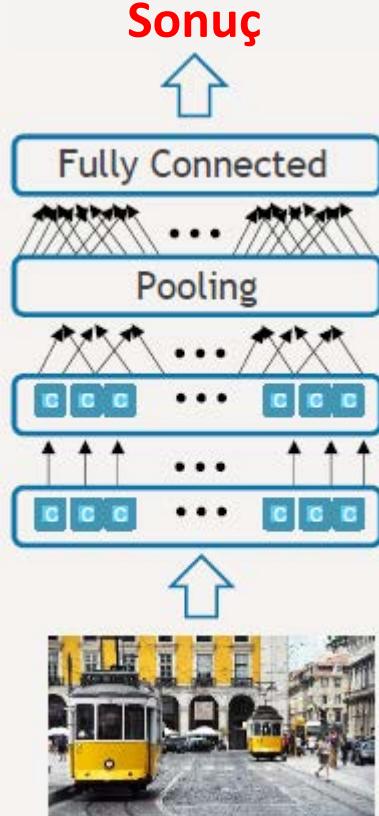


Mobil çevre kısıtları

Güç ve termal verim
Depolama ve hafıza bant genişliği kısıtları
Batarya

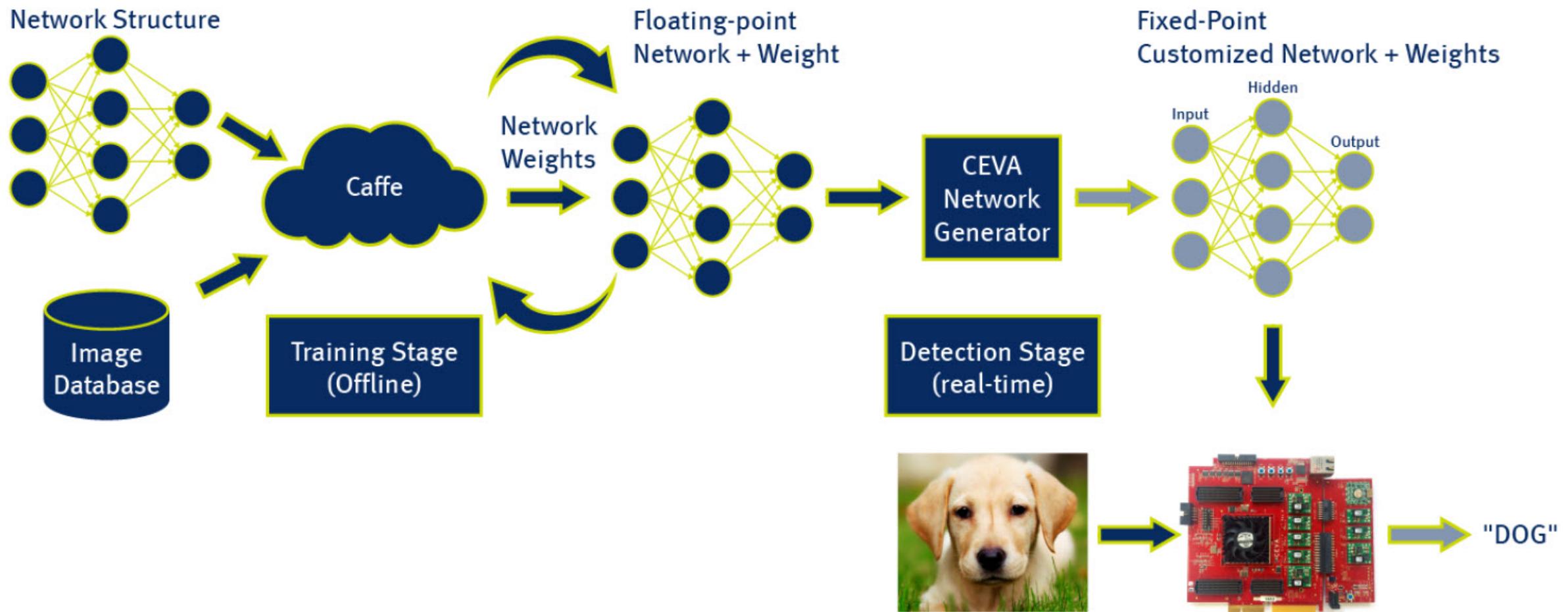
Sonuç

- Fiziksel model boyutunda 10X azalma
- Minimum hassasiyet kaybıyla 35% oranında MAC işlemlerinde azalma.



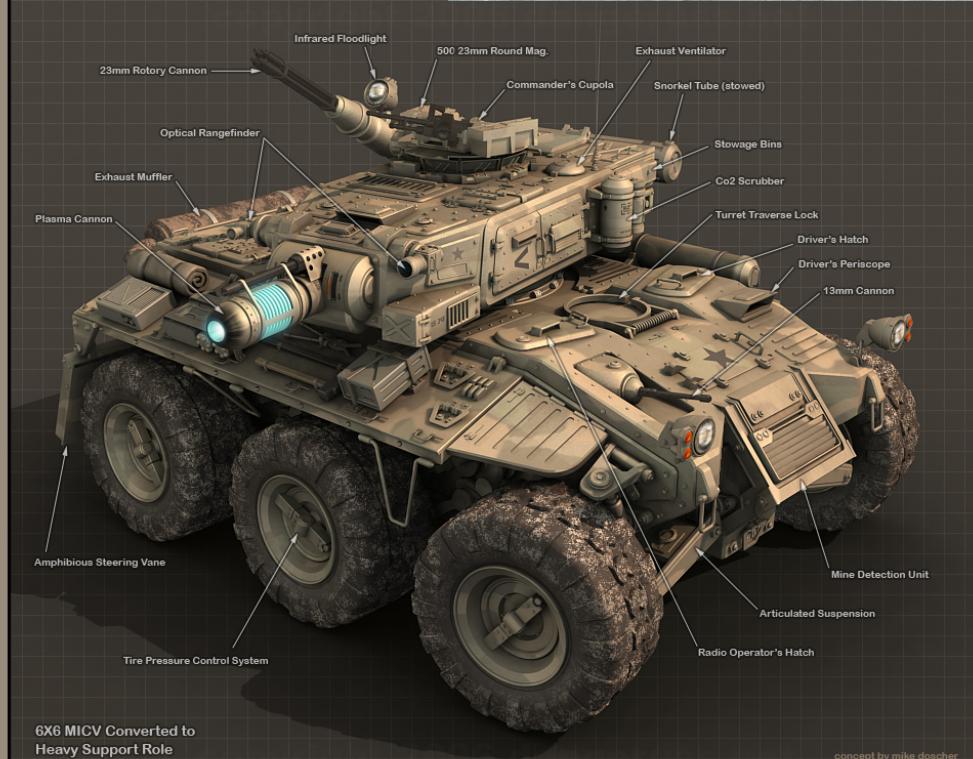
Derin Ağ

CNN Usage Flow with Caffe & CDNN



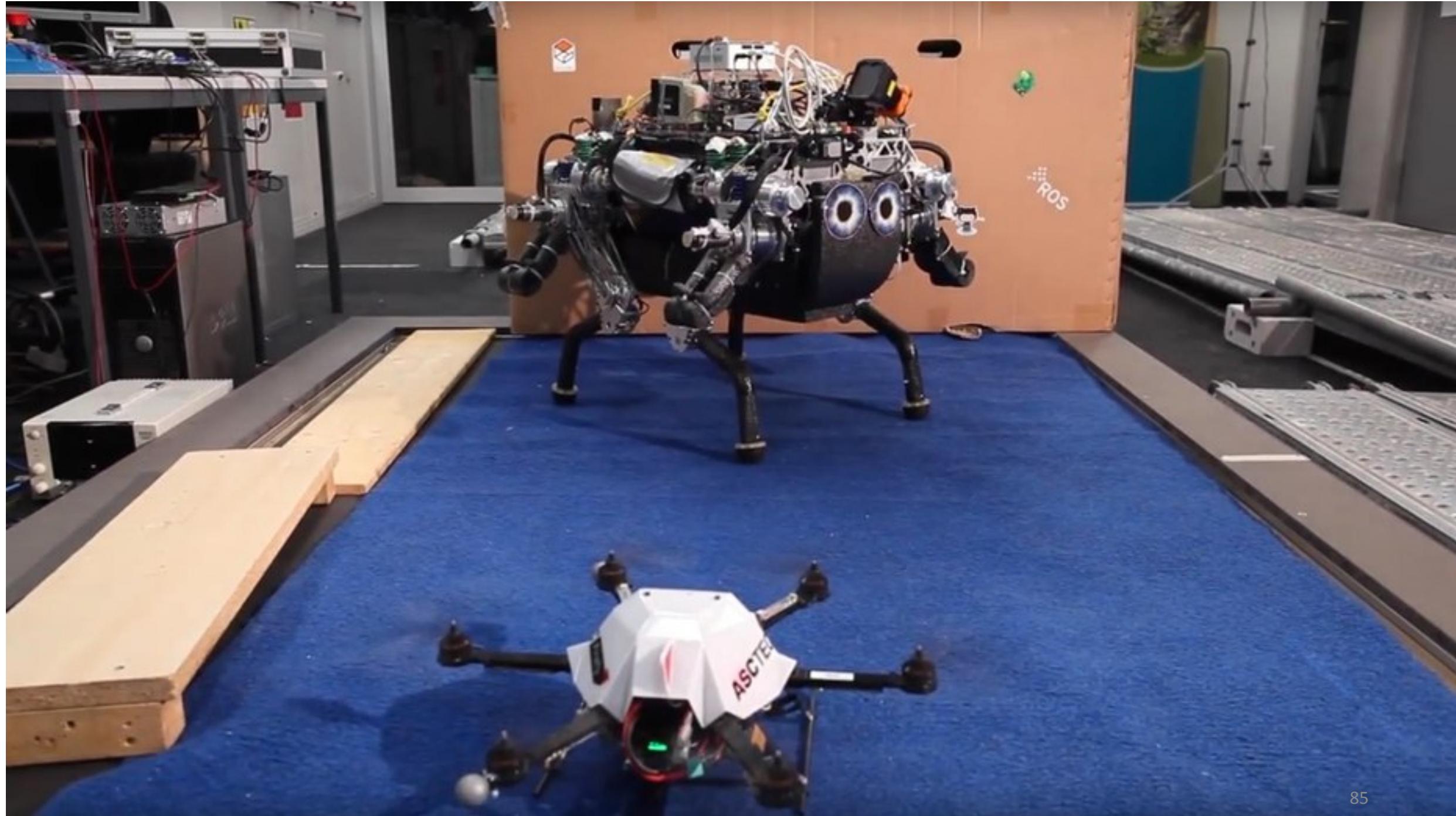


6X6 ARMORED CAR CONCEPT



concept by mike dascher
3D model by Dangeruss.net





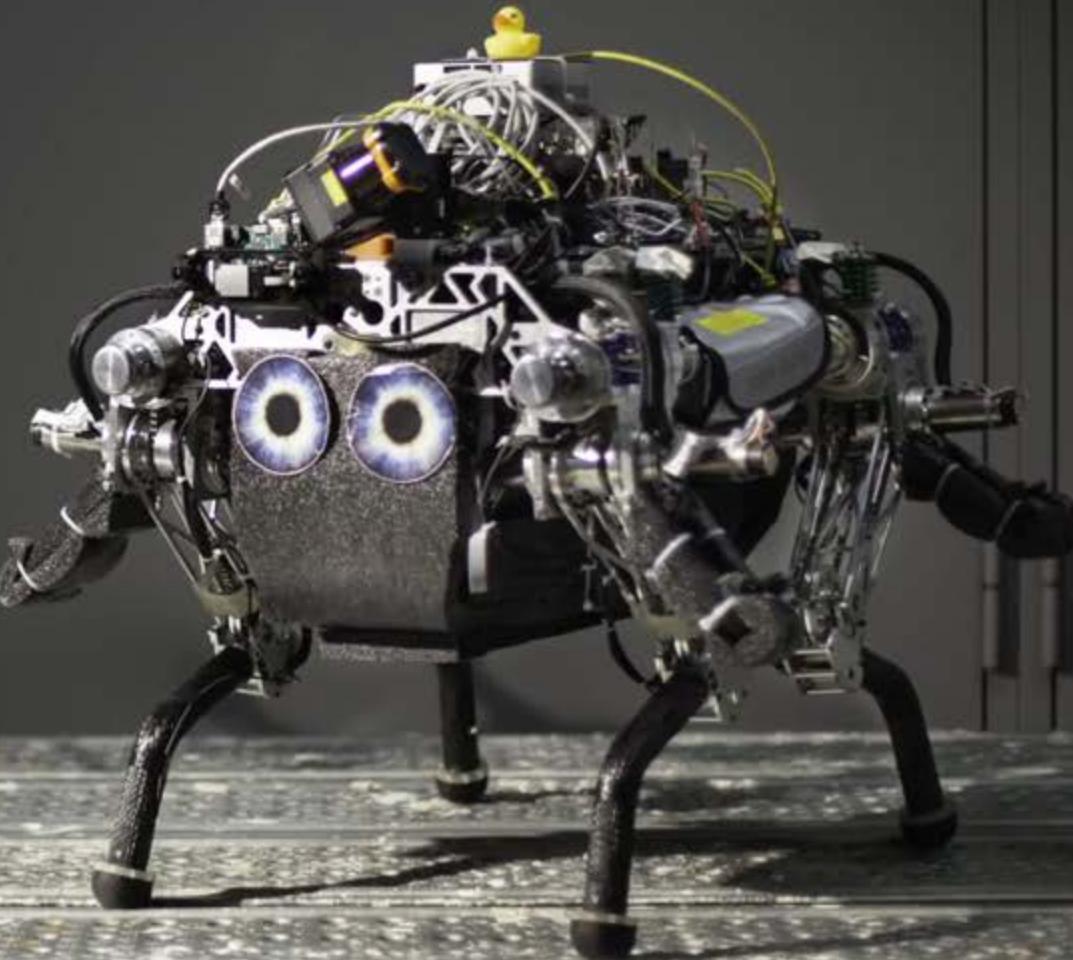
September 2015

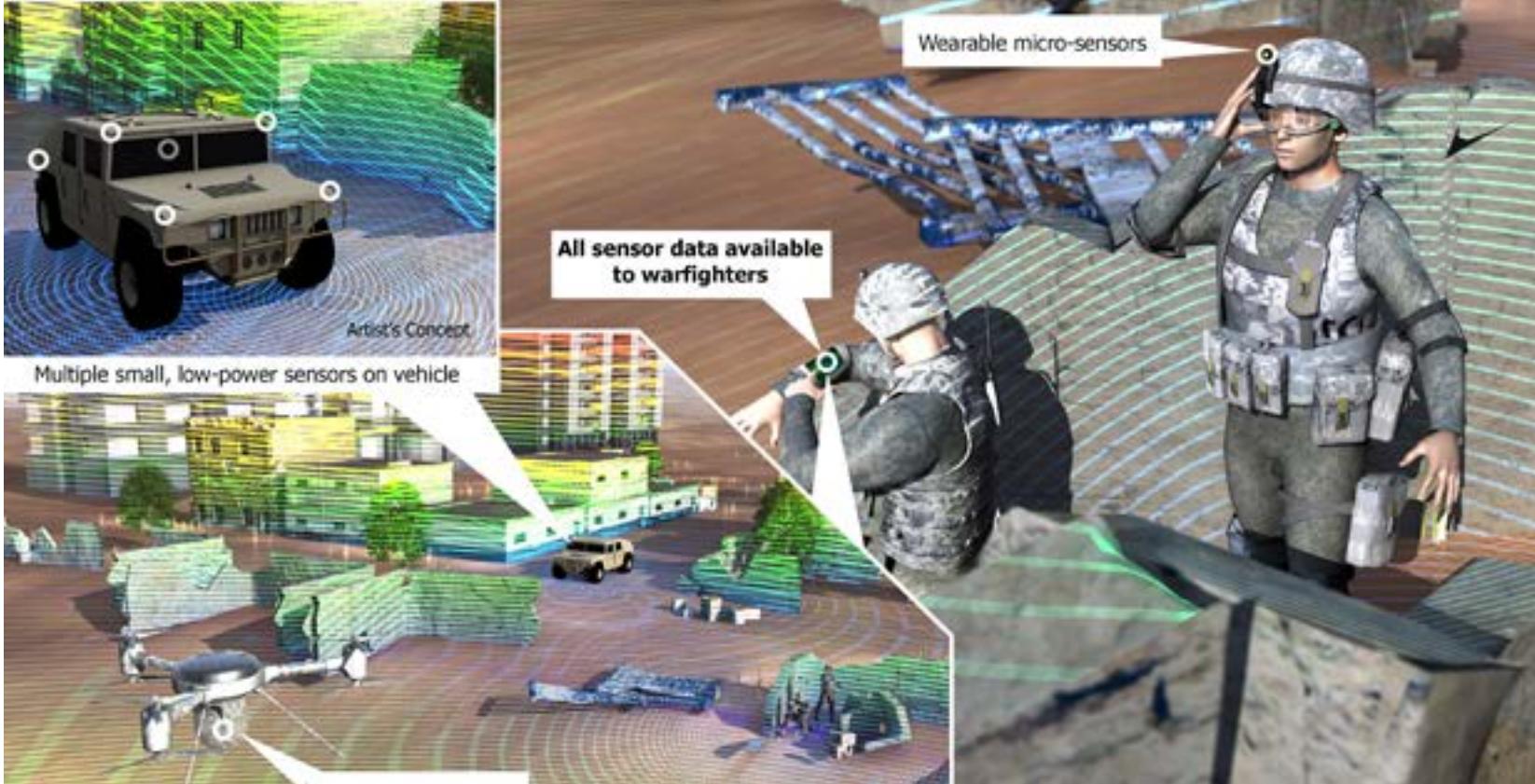
Collaborative Navigation for Flying and Walking Robots

Péter Fankhauser, Remo Diethelm, Martin Wermelinger,
Thomas Schneider, Marcin Dymczyk, Michael Burri,
Michael Bösch, Dario Bellicoso, Christian Gehring,
Simon Lynen, Philipp Krüsi, Marco Hutter, Roland Siegwart

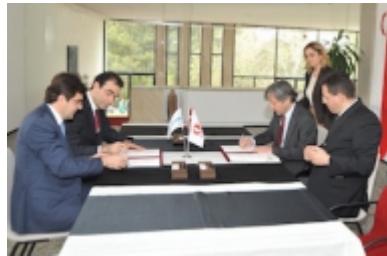


Autonomous Systems Lab





Derin Öğrenme Farkındalığı



ODTÜ Görüntü Analizi Uygulama ve Araştırma Merkezi (OGAM) ile ASELSAN Araştırma Merkezi arasında **“Derin Öğrenme Tabanlı Görüntü Analizi Algoritmaları Geliştirme Projesi”** sözleşmesi imzalandı.

Derin Öğrenme, insan beyninin algusal problemlerin çözümündeki üstünlüğünü sağlayan katmanlı biyolojik yapısını taklit eden bir makine öğrenmesi yaklaşımıdır. Derin Öğrenme teknolojisi ile otonom sistemler görsel analiz açısından insanın algılama yeteneğine yaklaşan bir kabiliyet kazanabilmektedir. Projede geliştirilecek Derin Öğrenme Teknolojisi, ASELSAN ürün ailesinde bulunan başta kara, hava ve deniz insansız araçları olmak üzere, görüntü sensörü içeren ve otomatik sınıflandırmaya ihtiyaç duyan askeri ve sivil sistemlerde kullanılacaktır.

<http://www.derinogrenme.com>



Türkiye Derin Öğrenme Grubu Sayfası: <https://www.linkedin.com/grp/home?gid=8334641>



Ankara Derin Öğrenme Meetup Sayfası: <http://www.meetup.com/Ankara-Deep-Learning>



Derin Öğrenme Grup Sayfası: <https://www.facebook.com/groups/derin.ogrenme>

TEŞEKKÜRLER.

“If we knew what it was we were doing, it would not be called research, would it?” Einstein