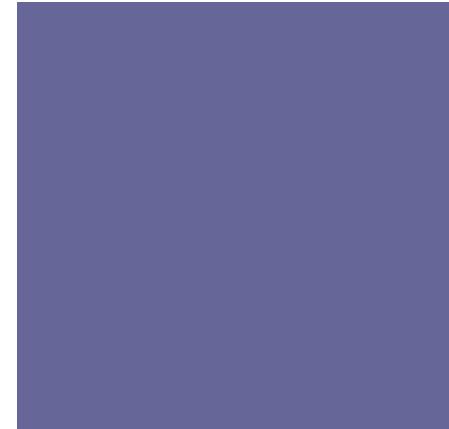




Büyük ölçekli video analizi



Pinar Duygulu



Çok fazla sayıda video



flickr

YAHOO!
SCREEN

- Sadece YouTube'da
 - Her ay 1 milyardan fazla erişim
 - 6 milyar saatten fazla izlenme
 - Her dakika yüklenen yüzlerce saat video

https://youtu.be/IPq_LanED60

<http://www.youtube.com/yt/press/statistics.html>

Uygulamalar

- Analyzing video archives



First appearance of
N. Sarkozy on TV



Sociology research:
Influence of character
smoking in movies



Education: How do I
make a pizza?

- Surveillance



Where is my cat?

Pinar Duygulu, Eylul 2016, Ankara



Predicting crowd behavior
Counting people

- Graphics



Motion capture and animation

Alıntı: I.Laptev



Araştırma için kullanılan veri kümeleri

Veri kümesi
KTH
Weizmann
IXMAS
Hollywood
UCF Sports
Hollywood2
UCF YouTube
MSR
Olympic
UCF50
HMDB51

Sınıf sayısı

6
9
11
8
9
12
11
3
16
50
51





Daha güncel veri kümeleri

Veri Kümesi

UCF-101

ActivityNet

FCVID

Sports-1M

#Sınıf #Video

101 13320

200 20000

239 91223

487 1M



Yahoo! Flickr
100M veri kümesi
418.507 etiketli video

+

Sınırsız ve kısıtlamasız doğal(!) videolar



Harlem Shake : <http://www.youtube.com/watch?v=4hpEnLtqUDg>

+

Hareketleri nasıl tanırız?

Zamansal değişim



Poz



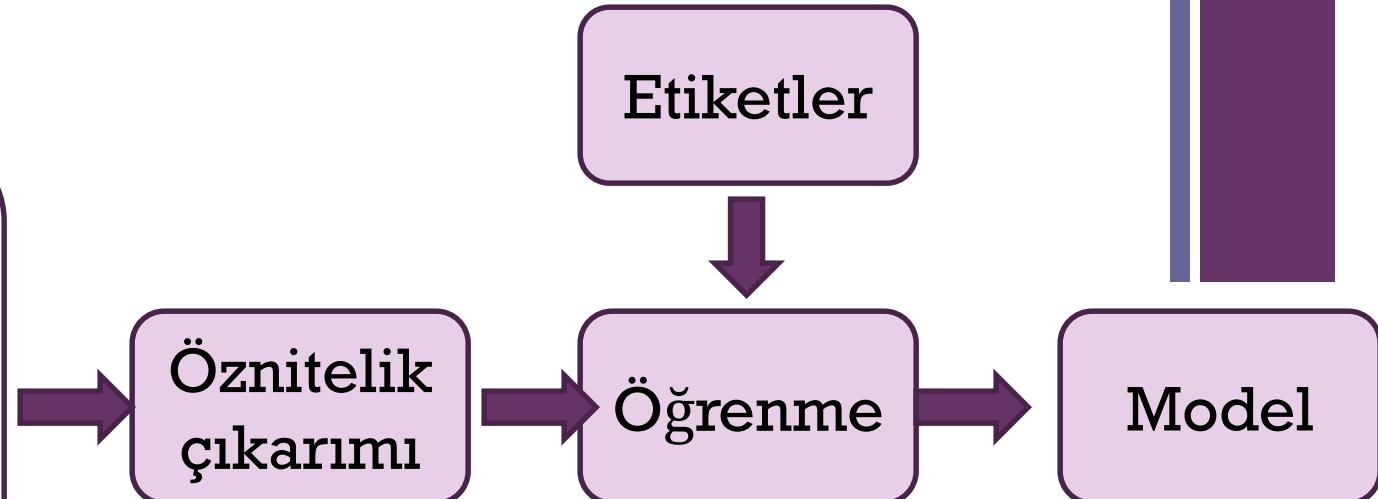
Kullanılan
nesneler



İçinde
bulunulan
ortam

Öğrenme

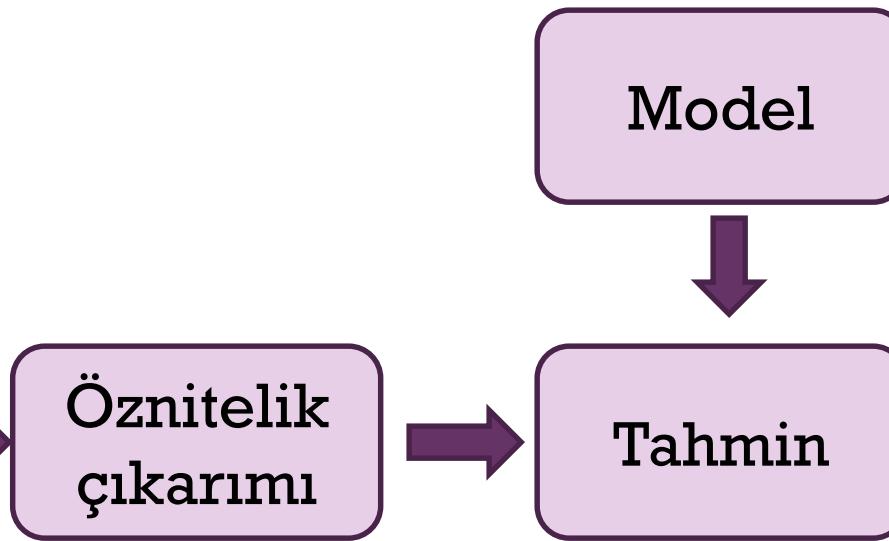
Etiketlenmiş veri



Deneme



Yeni veri





Tek resimden poz bilgisi ile hareket tanıma



Movies



TV



YouTube



Movies

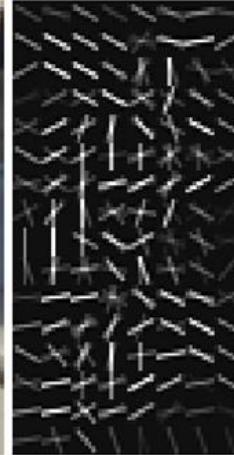


TV



40%

YouTube

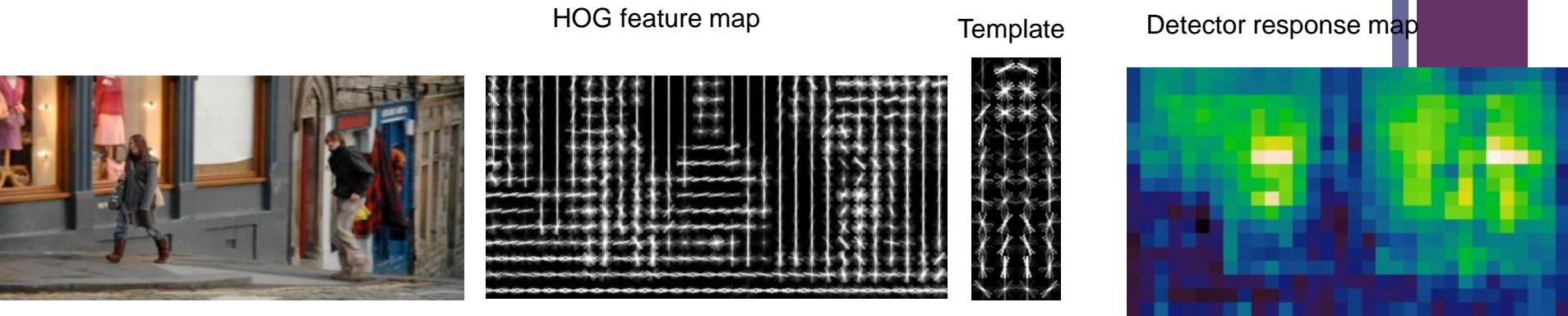


Silhouette description [Sullivan & Carlsson, 2002]

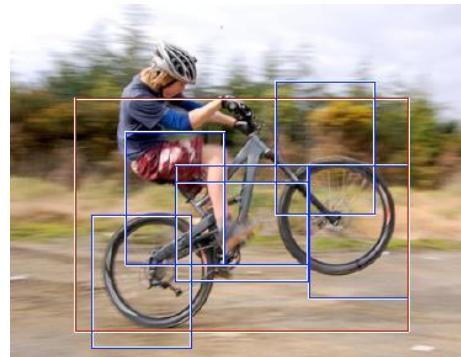
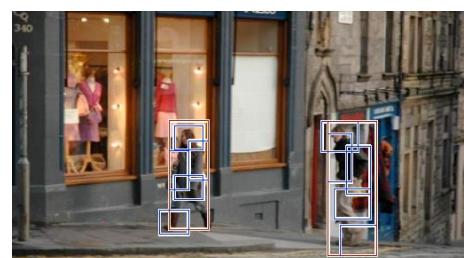
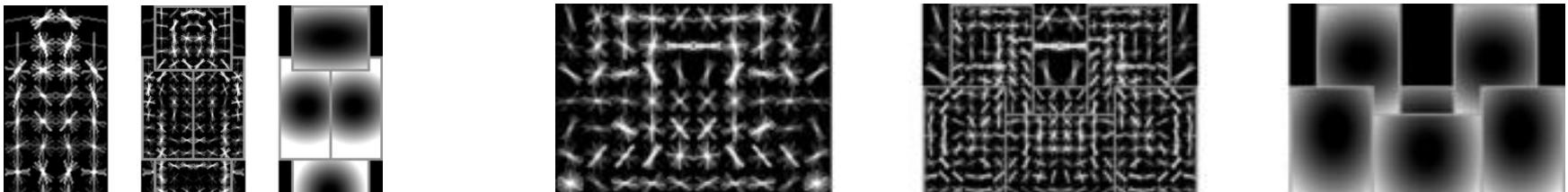
Histogram of gradients (HOG) [Dalal & Triggs 2005]

Human body part layout [Felzenszwalb & Huttenlocher, 2000]

+ Histogram of Oriented Gradients Deformable Part based models



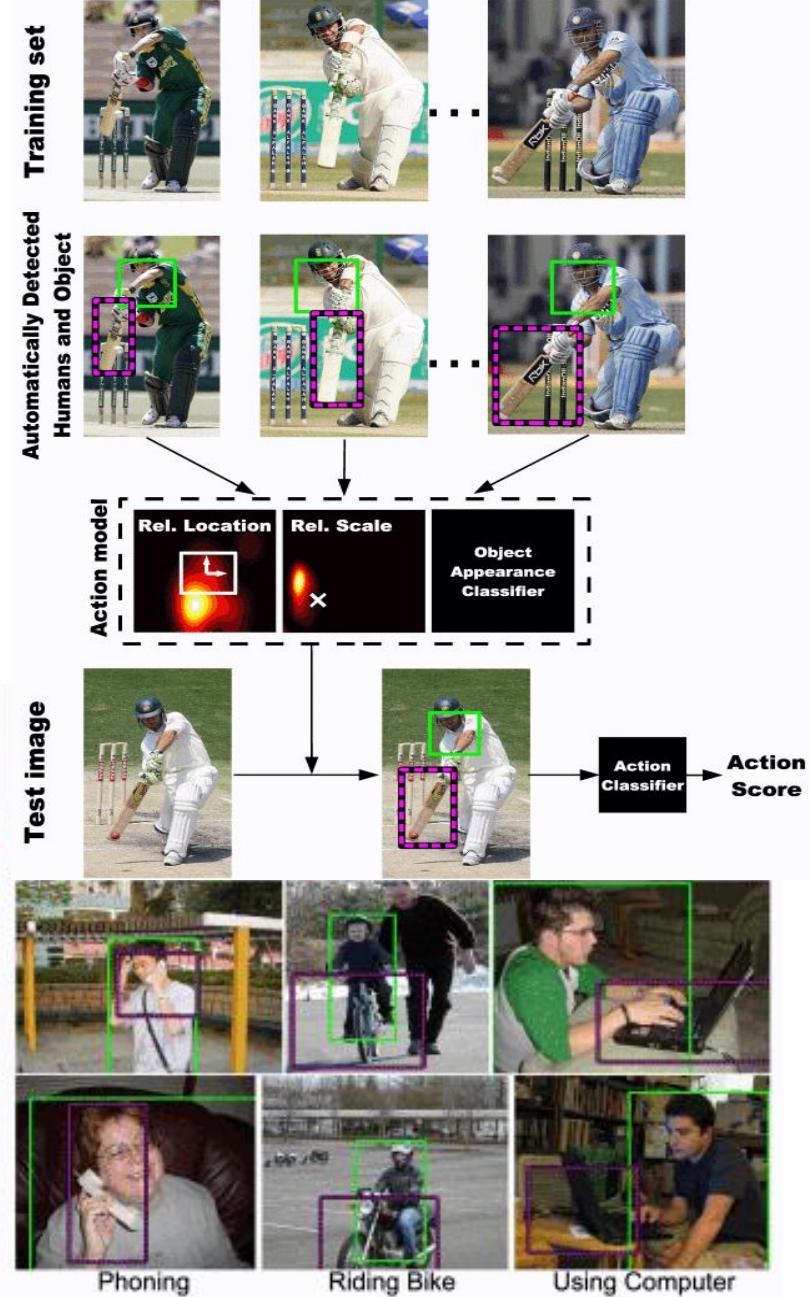
N. Dalal and B. Triggs, [Histograms of Oriented Gradients for Human Detection](#), CVPR 2005



P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, [Object Detection with Discriminatively Trained Part Based Models](#), PAMI 32(9), 2010



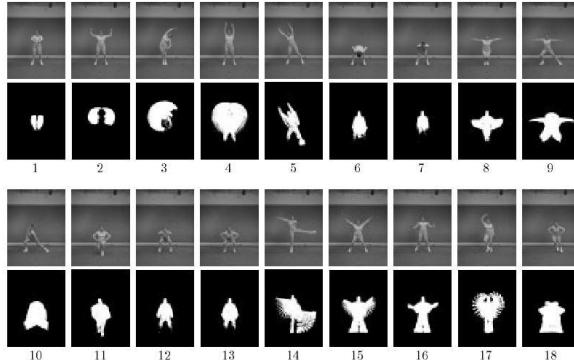
Kullanılan nesne ile hareketi tanıma



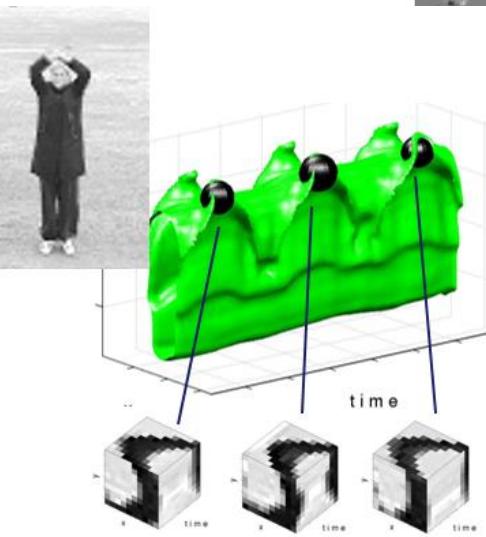
Prest, Schmid & Ferrari , Weakly-supervised learning of interactions between humans and objects, 2011



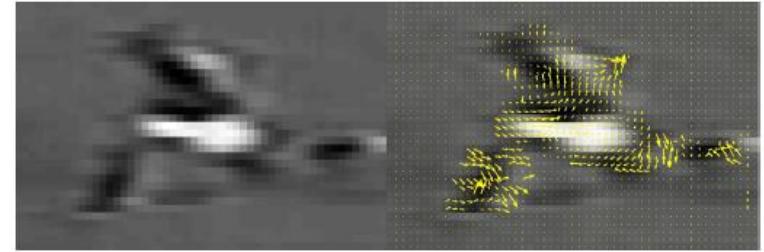
Hareket tanıma için zamansal değişim



Motion Energy and Motion History

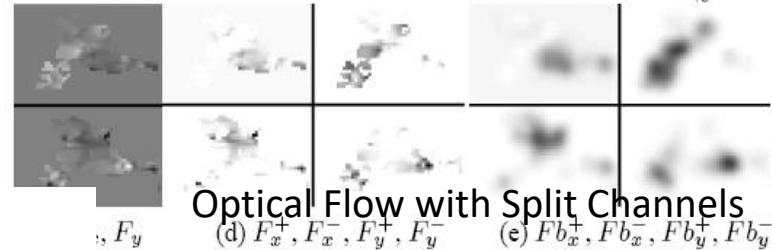


Space Time Interest Points



(a) original image

(b) optical flow $F_{x,y}$

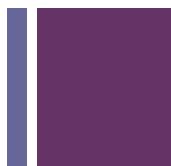


Optical Flow with Split Channels
(c) F_x, F_y (d) $F_x^+, F_x^-, F_y^+, F_y^-$ (e) $Fb_x^+, Fb_x^-, Fb_y^+, Fb_y^-$



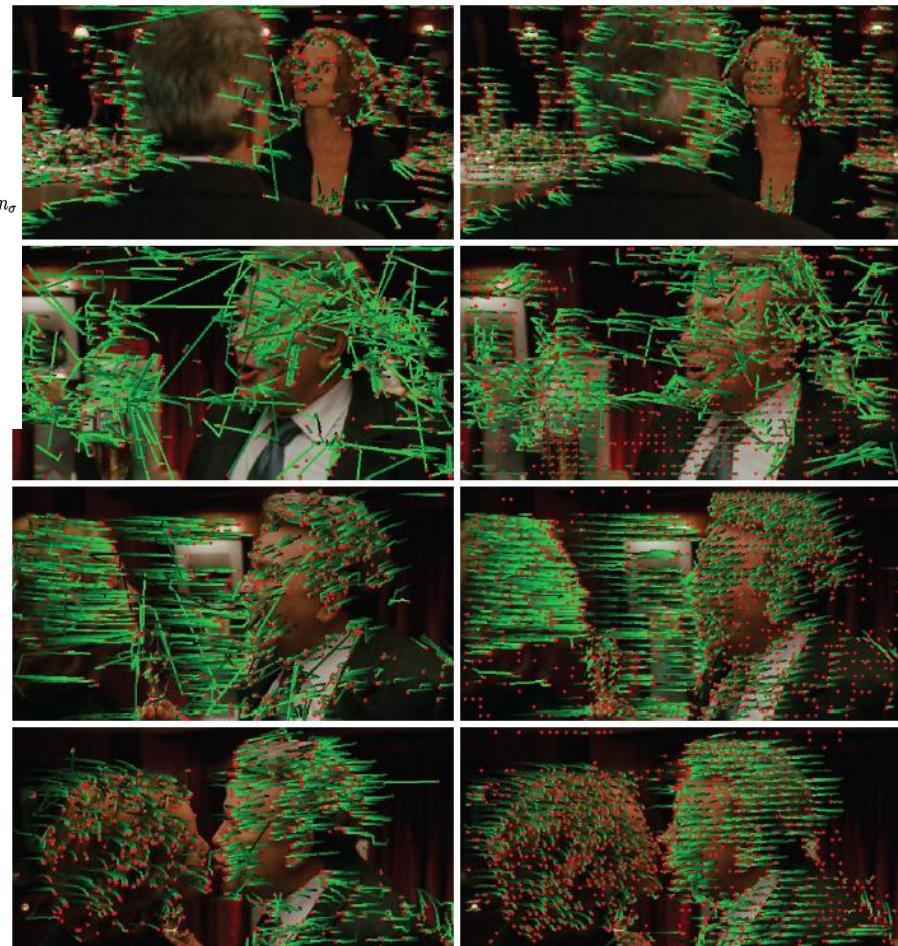
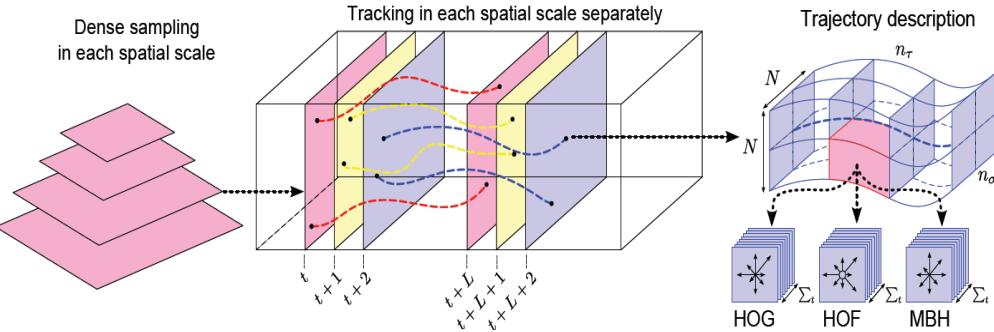
Space-Time Volumes

- A. F. Bobick, J. W. Davis, "The Recognition of Human Movement Using Temporal Templates", PAMI 2001
A. A. Efros, A. Berg, G. Mori, J. Malik, Recognizing Action at a Distance, ICCV 2003
I. Laptev, "On Space-Time Interest Points", IJCV, 2005



+

Yoğun İz takibi



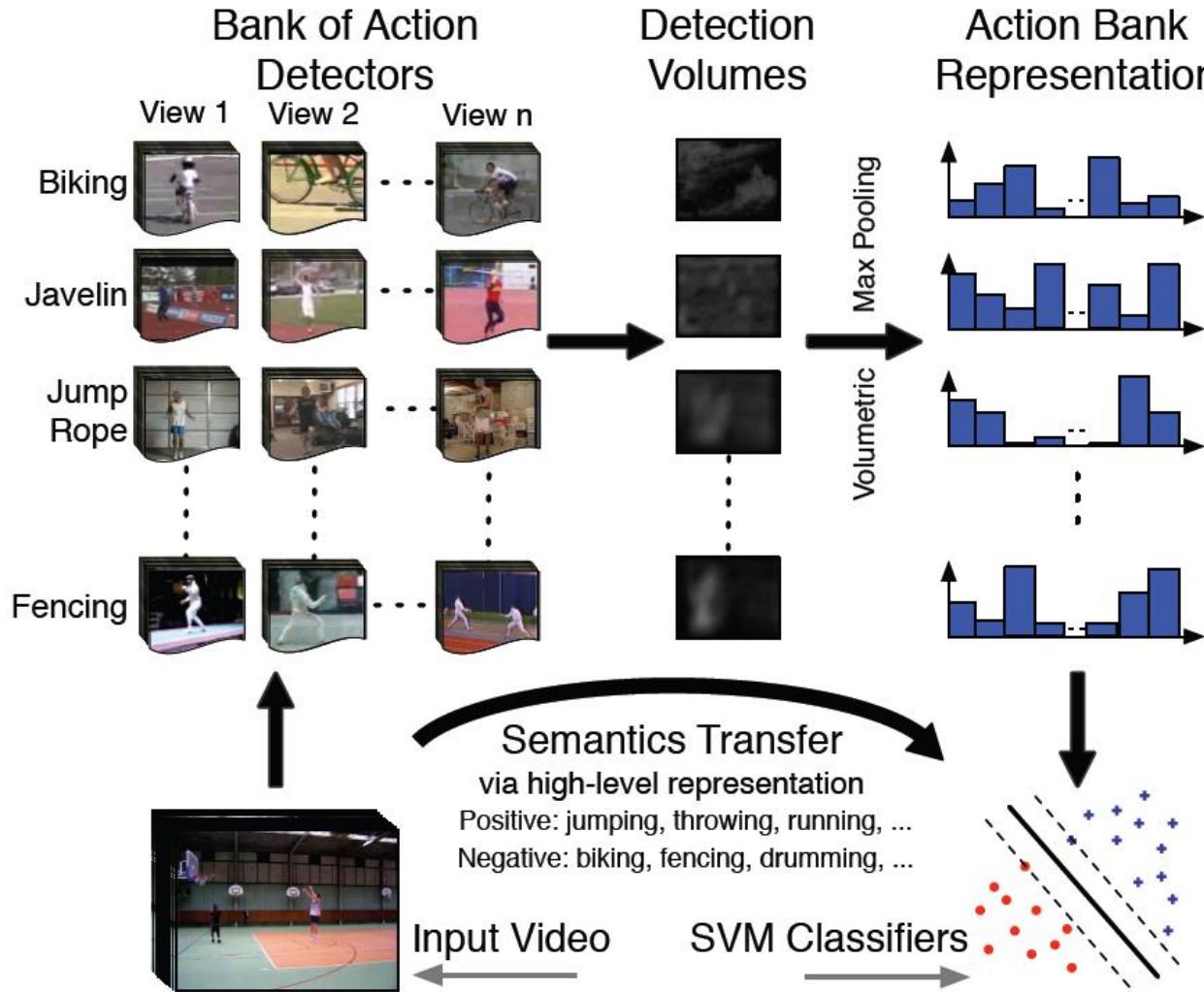
KLT

Dense trajectories

H. Wang, A. Klaser, C. Schmid, C.-L. Liu. Action Recognition by Dense Trajectories. CVPR, 2011.



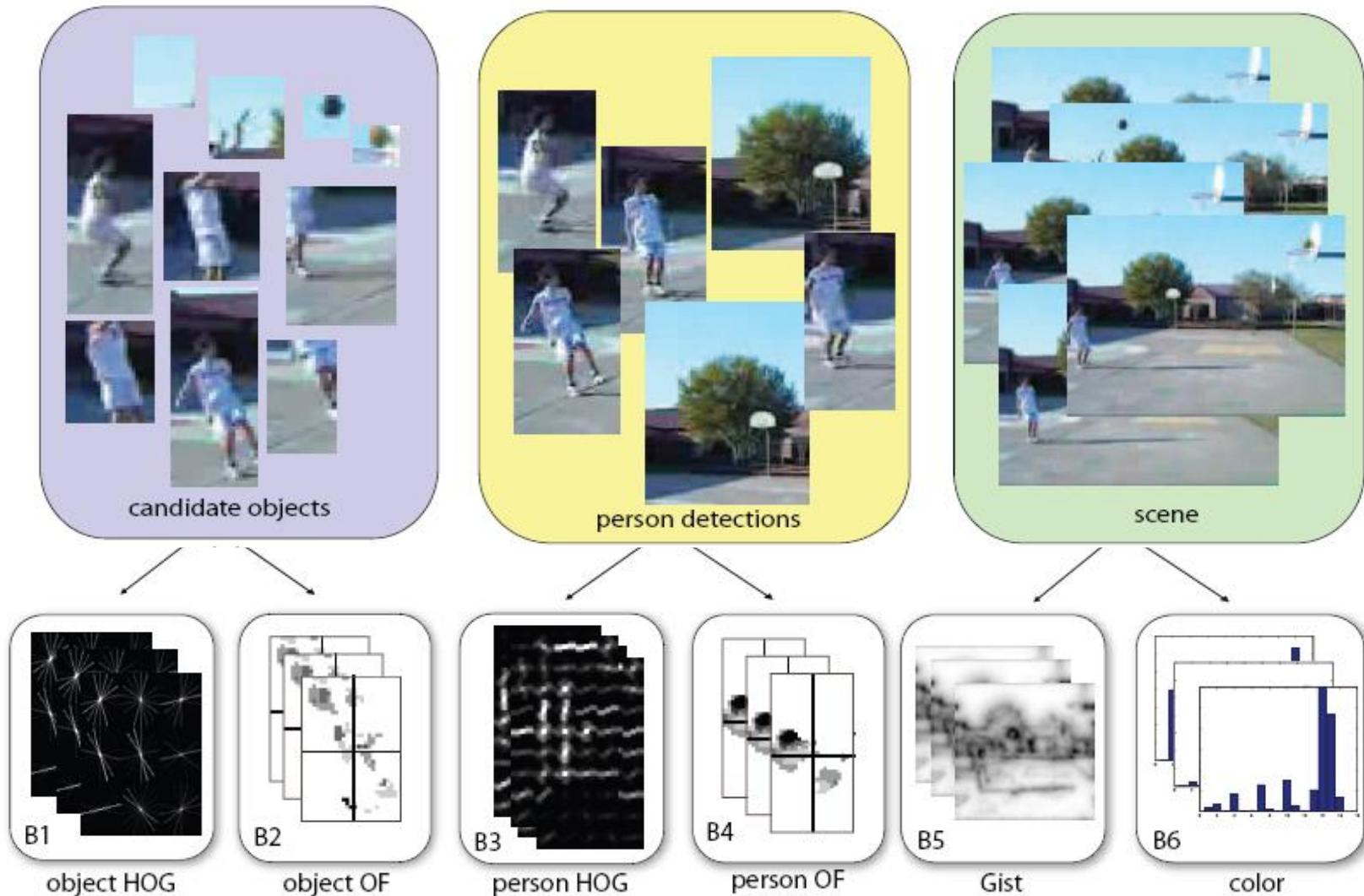
Hareket Bankası (Action bank)



S. Sadanand, J. J. Corso. Action Bank: A High-Level Representation of Activity in Video, CVPR, 2012.



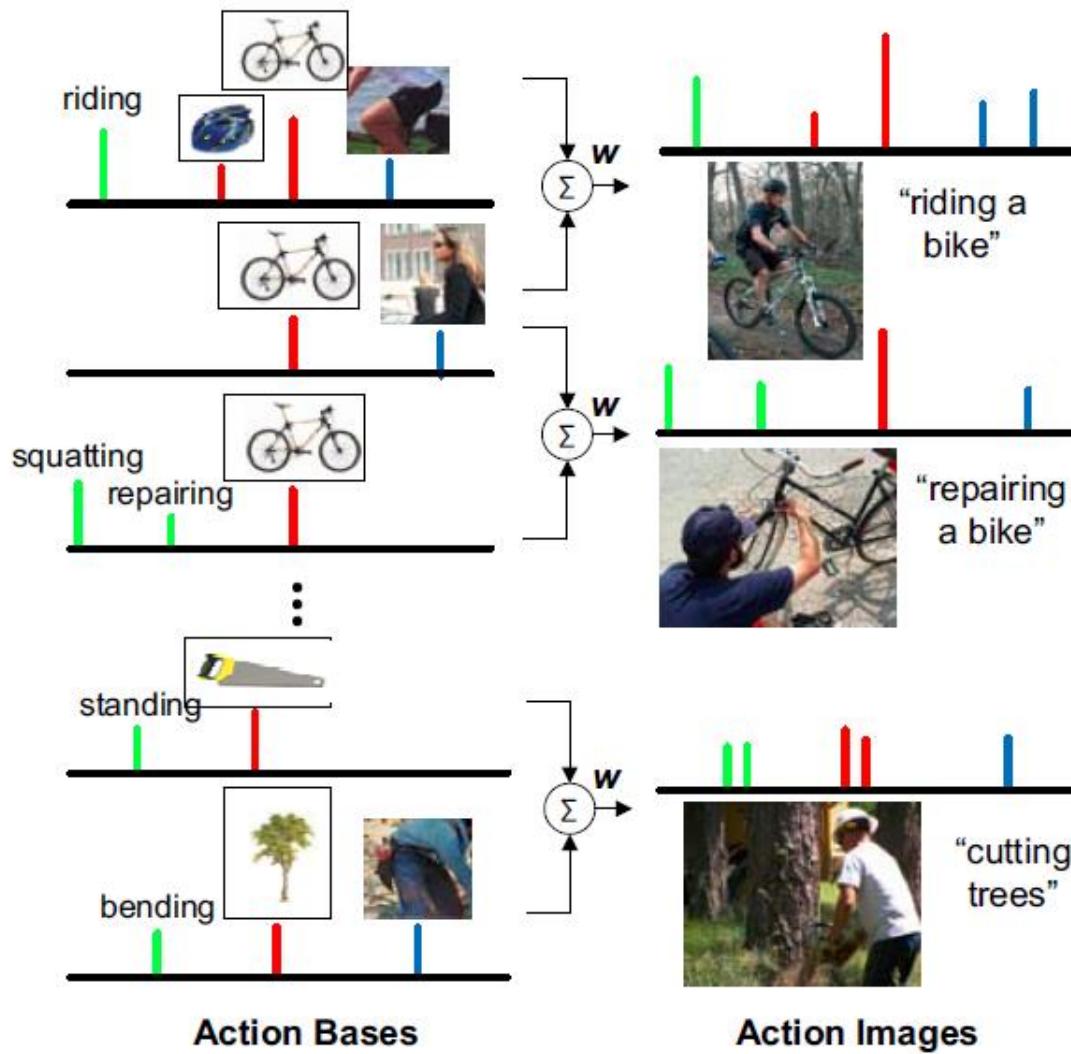
Sahne, nesne ve hareket birlikte



N. Ikizler-Cinbis, S. Sclaroff, Object, Scene and Actions: Combining Multiple Features for Human Action Recognition, ECCV, 2010.



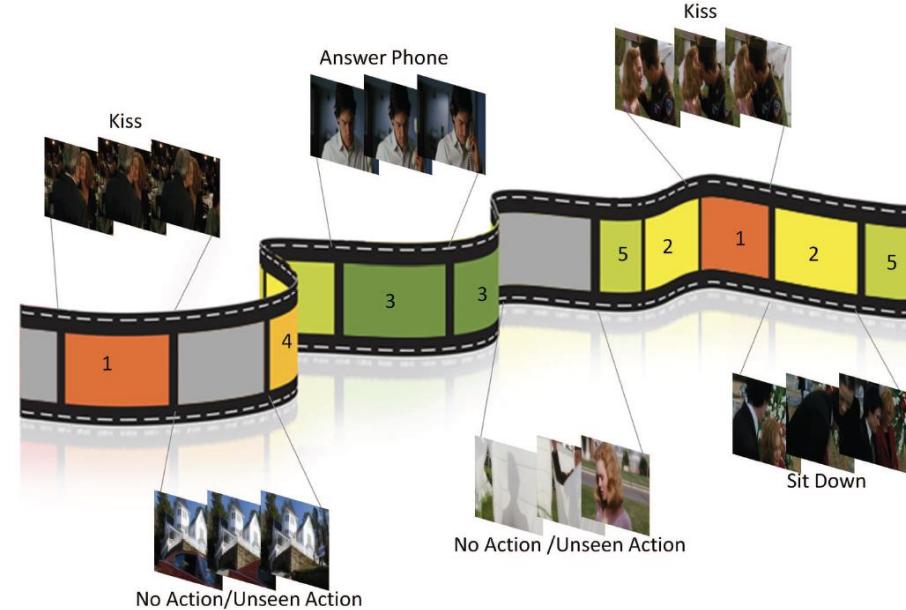
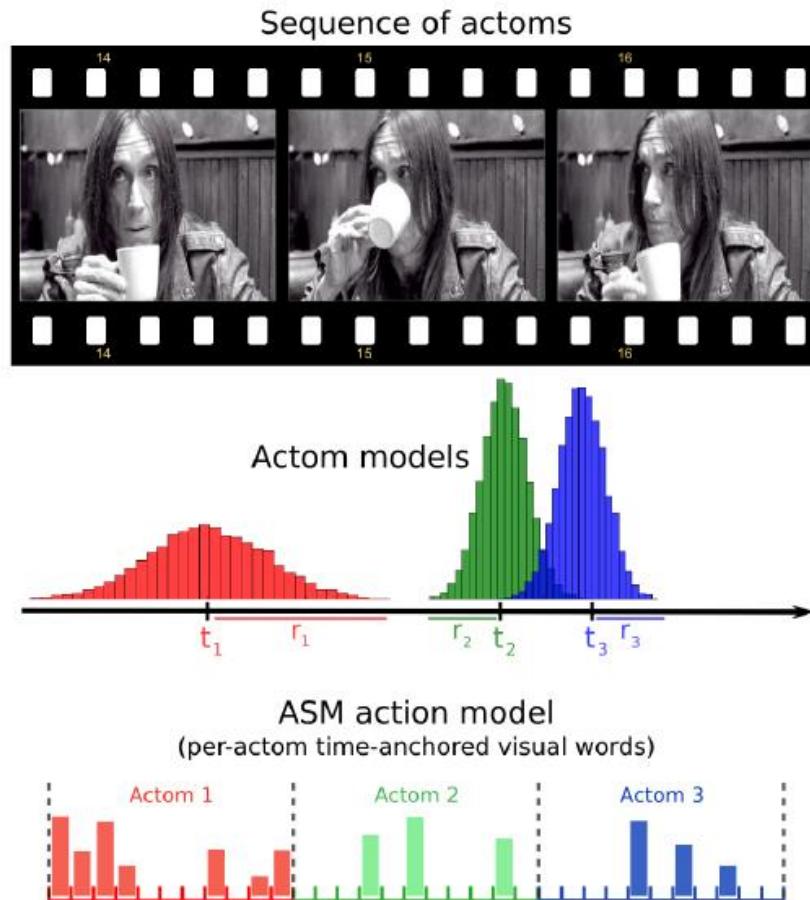
Hareket tanımlayıcı nitelikler



B. Yao, X. Jiang, A. Khosla, A. L. Lin, L. Guibas and L. Fei-Fei, Human action recognition by learning bases of action attributes and parts, ICCV 2011



Bileşik hareketler - Hareket dizileri



ASM action model
(per-actom time-anchored visual words)



A. Gaidon, Z. Harchaoui, and C. Schmid. Actom sequence models for efficient action detection. CVPR, 2011.

M. Hoai, Z. Lan, and F. Torre. Joint segmentation and classification of human actions in video. CVPR, 2011.



Hareket mi Olay mı

- Hareket: içme, oturma



Coffee & Cigarettes dataset

- Olay: sandviç yapma



Hollywood dataset



TRECVID-Multimedia Event Detection



Attempting a board trick



Feeding an animal



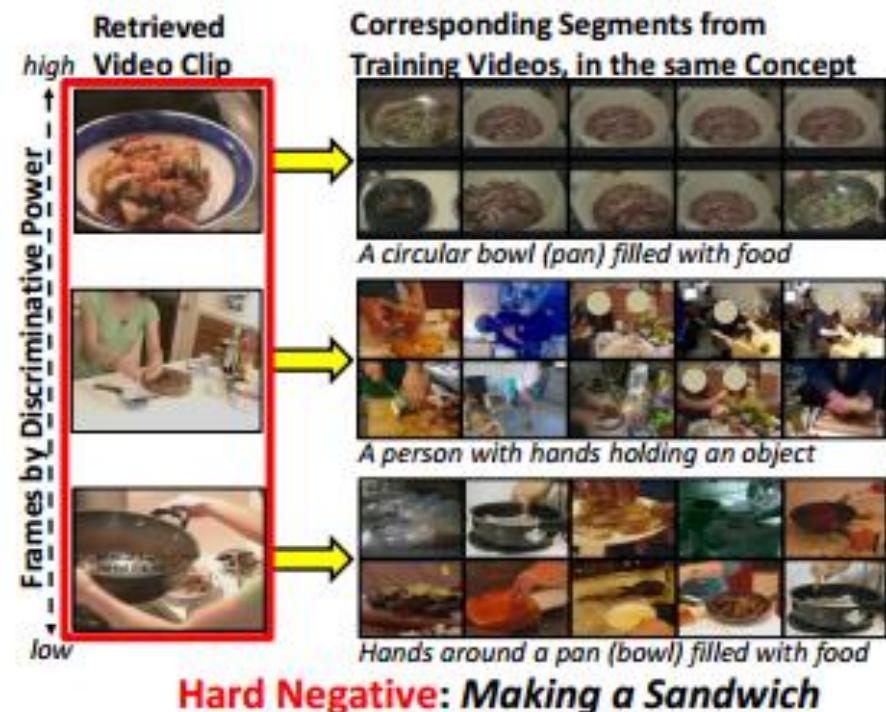
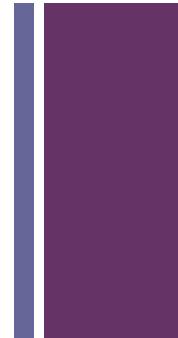
Wedding ceremony

Pinar Duygulu, Eylul 2016, Ankara



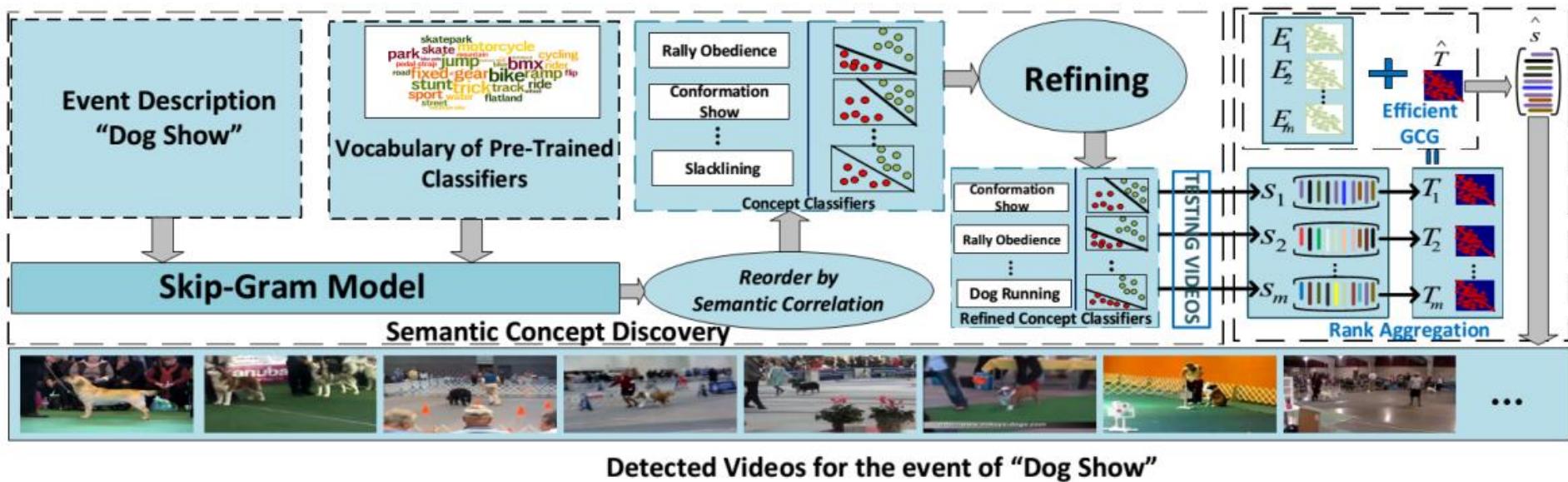
Getting a vehicle unstuck

Alıntı: C. Schmid

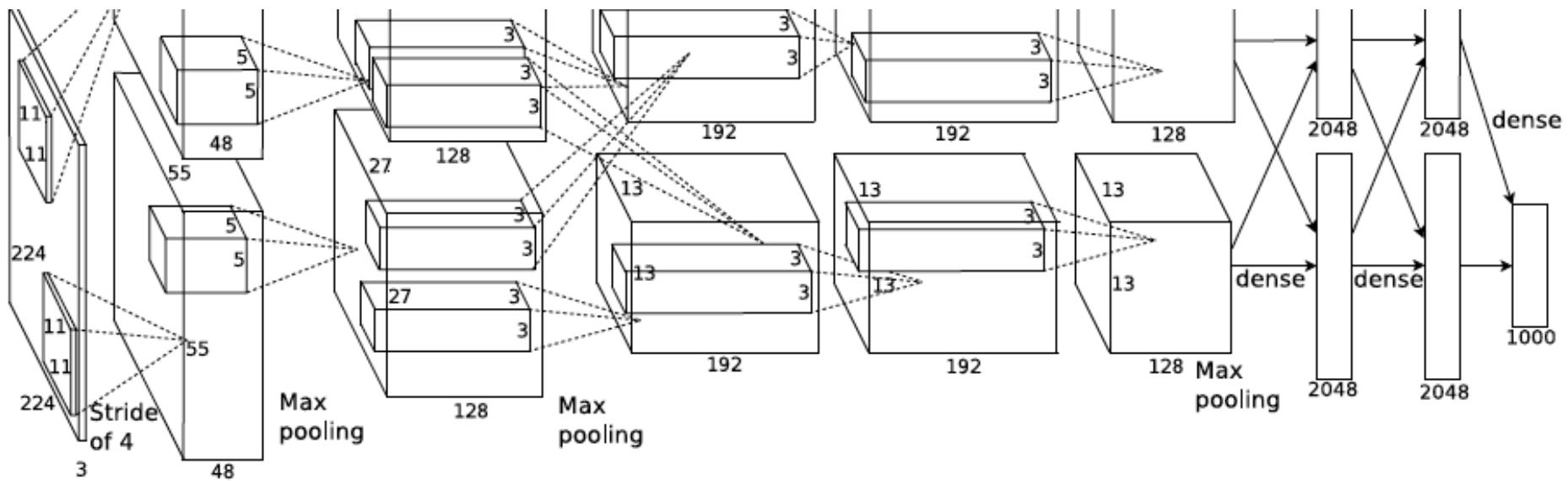
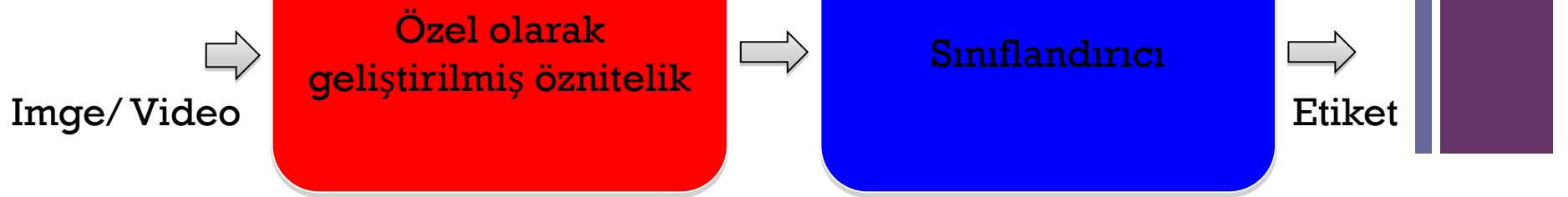


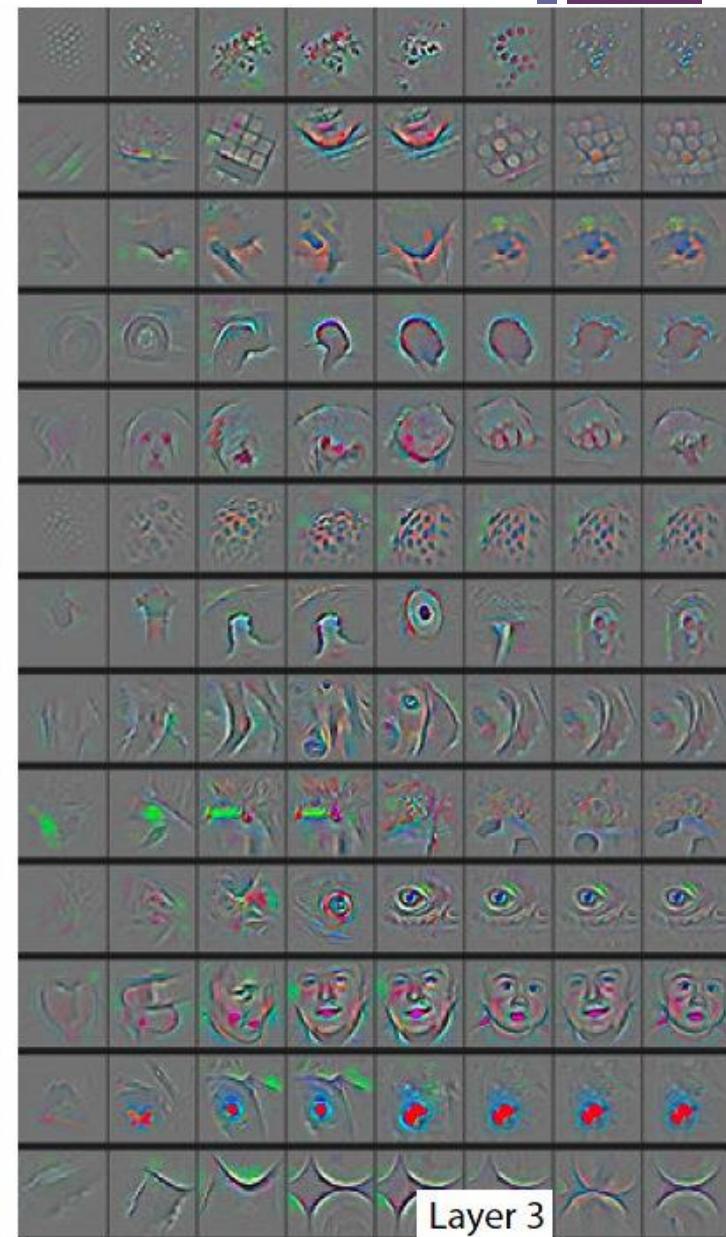
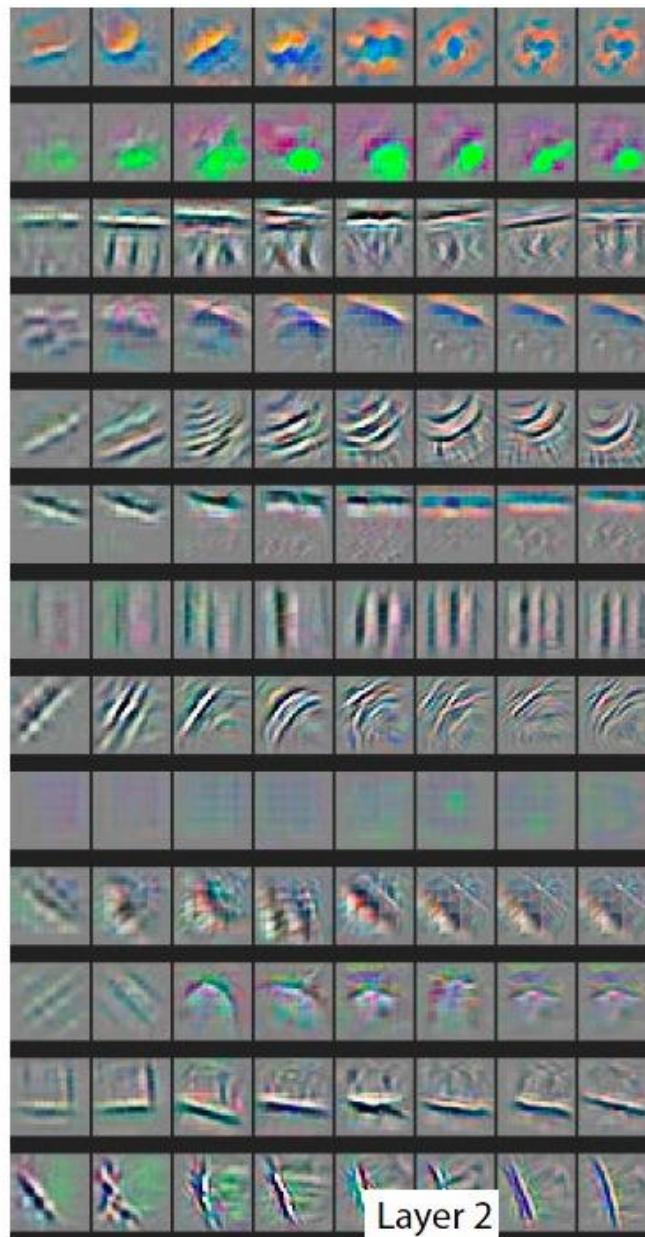
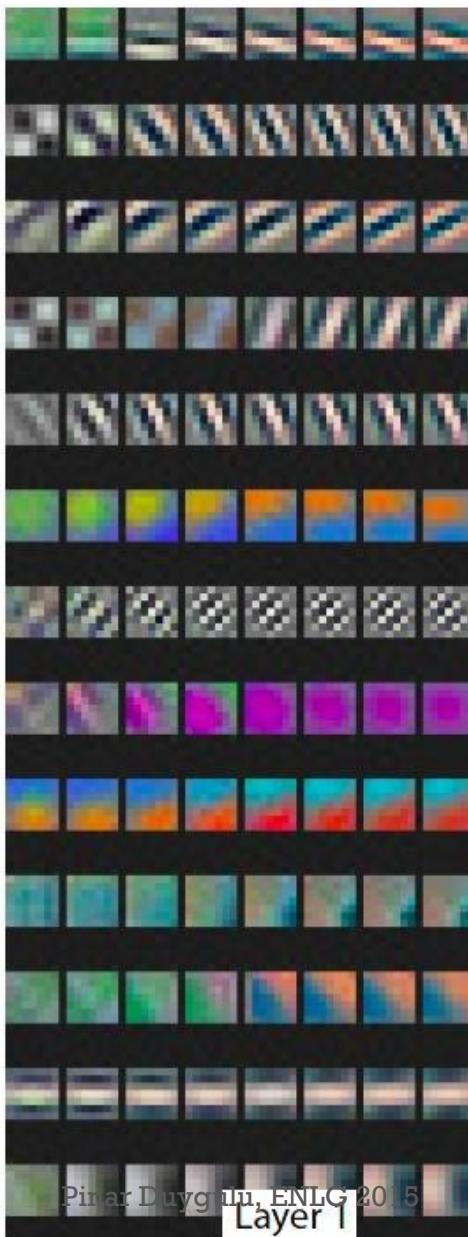


Örnek olmadan olay tanıma (zero-shot event detection)



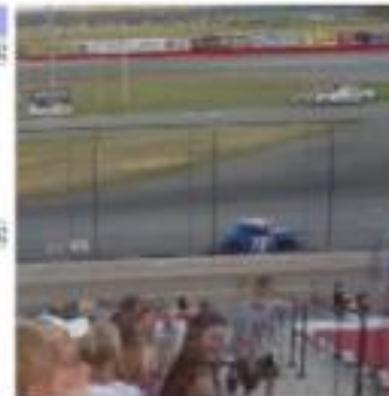
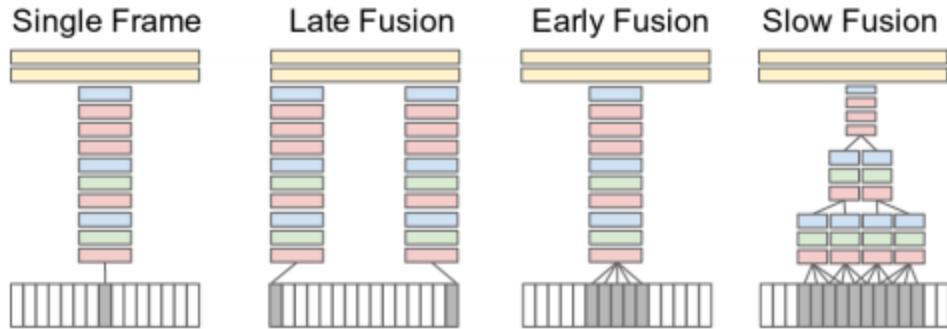
+





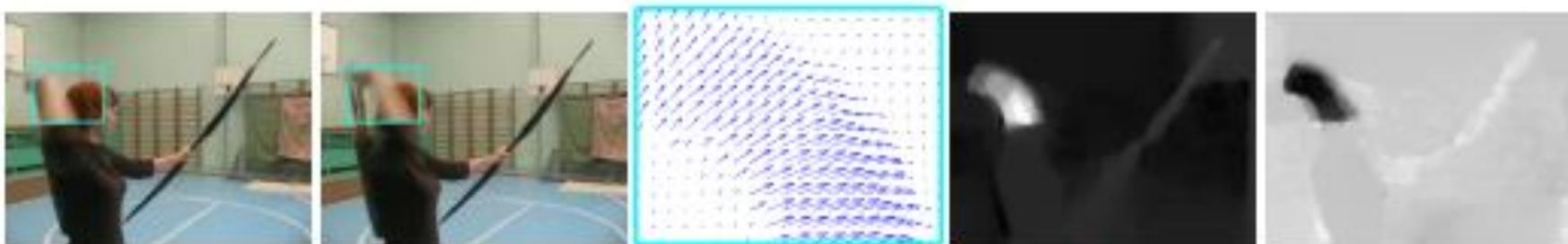
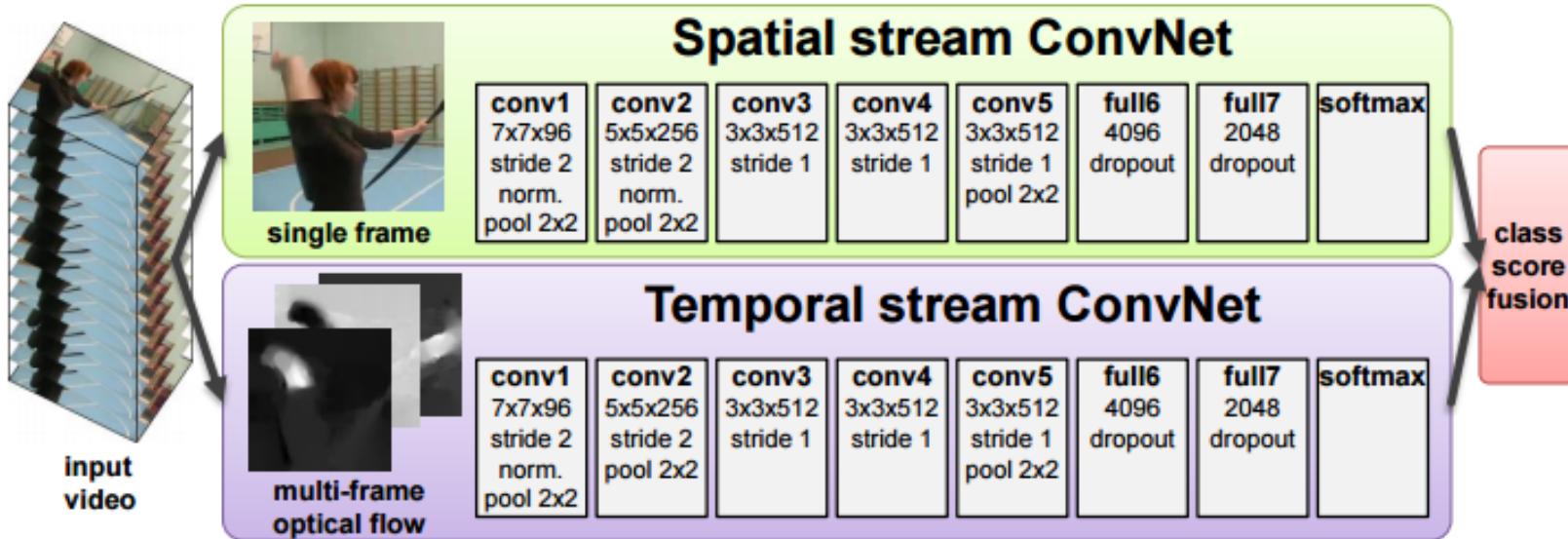


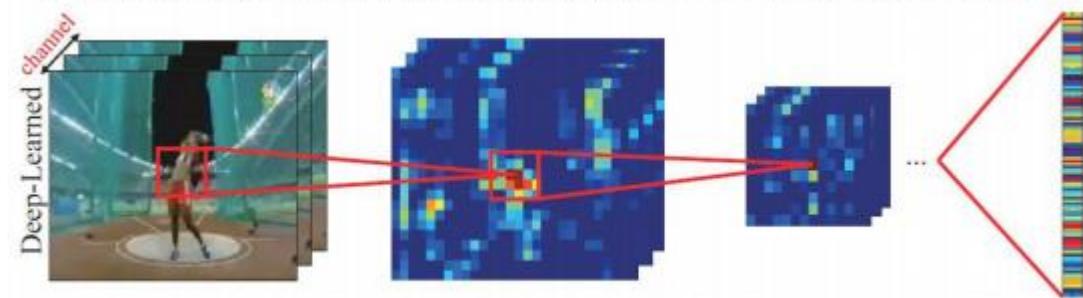
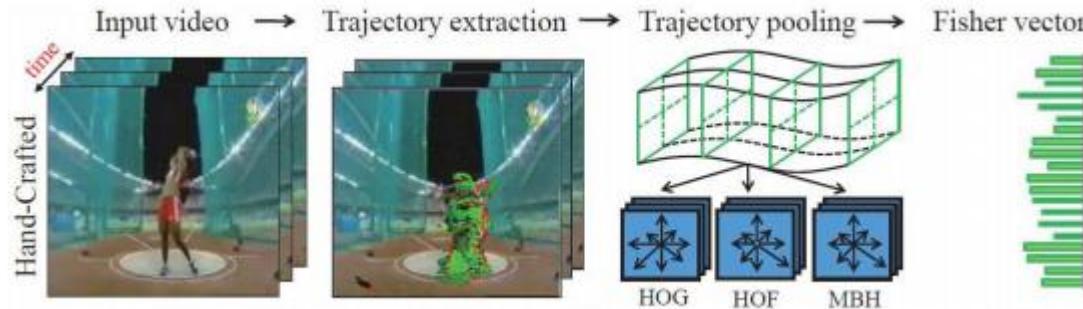
Video için CNN



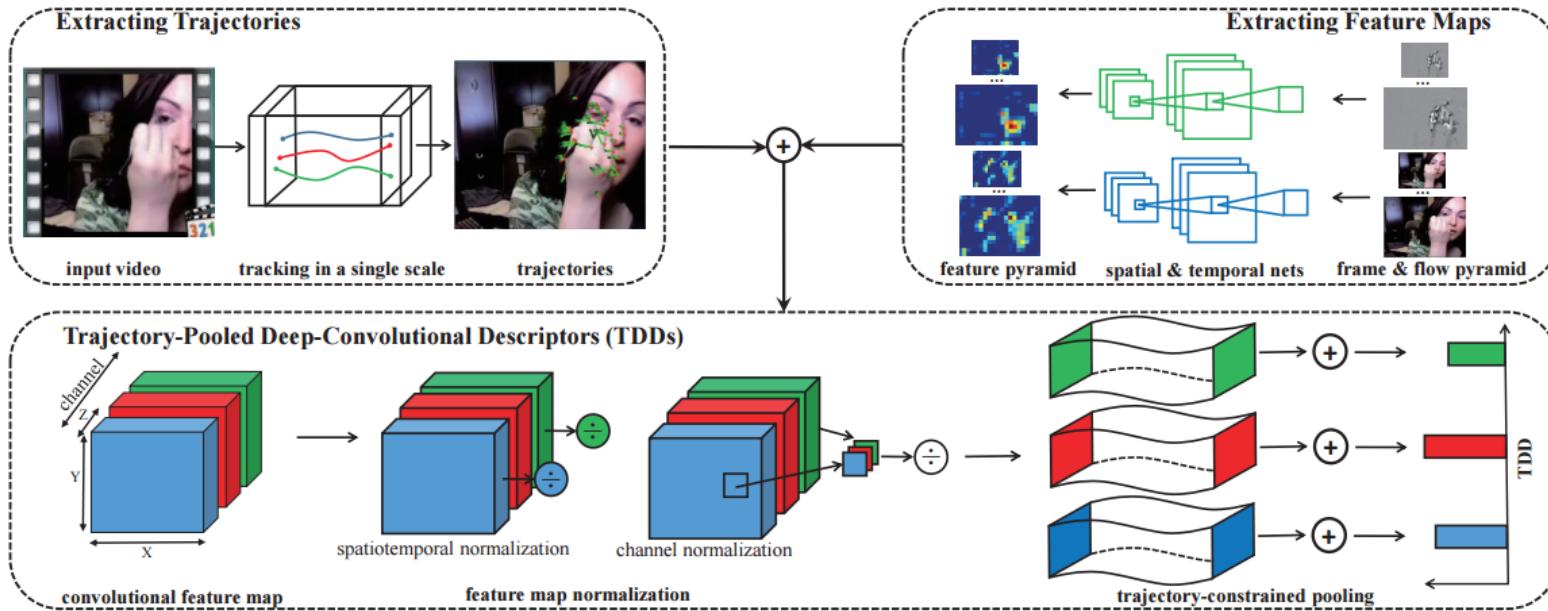


Two stream ConvNets

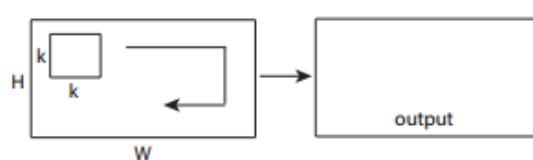




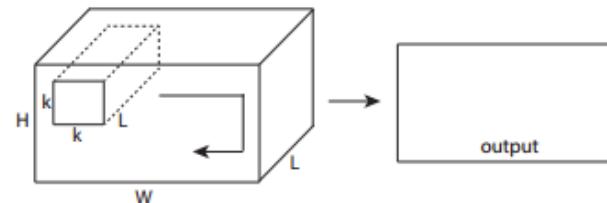
Input video → Convolution Layer → Pooling Layer → ... → Prediction



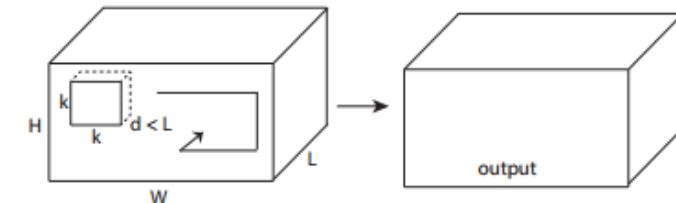
+3D ConvNets



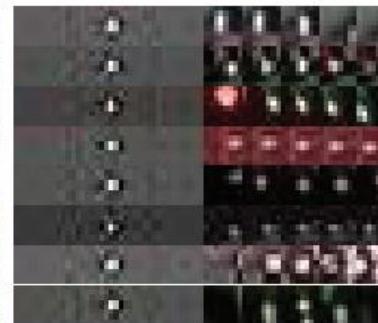
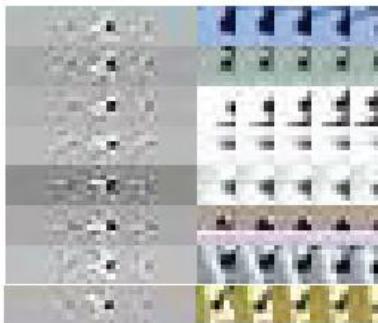
(a) 2D convolution



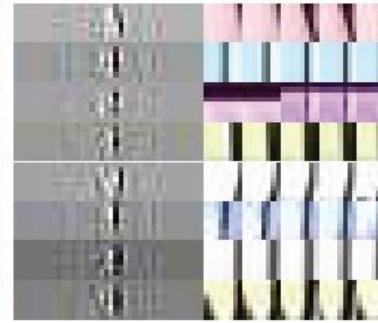
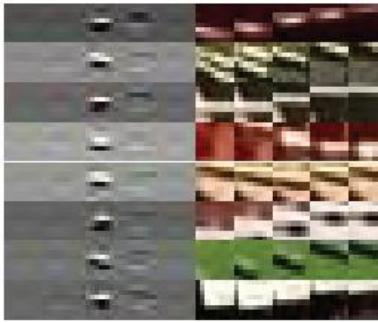
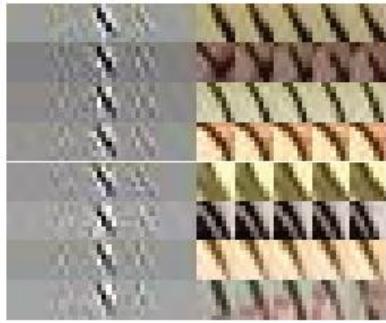
(b) 2D convolution on multiple frames



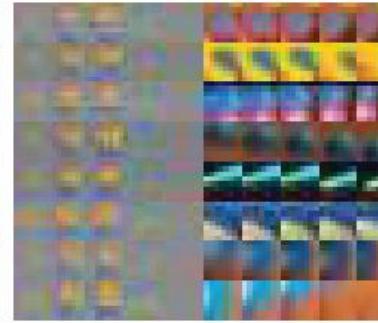
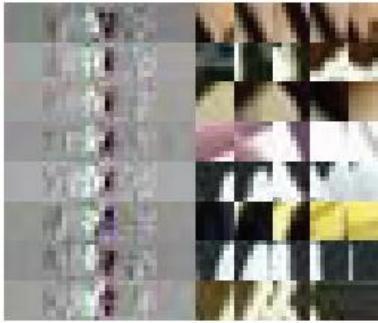
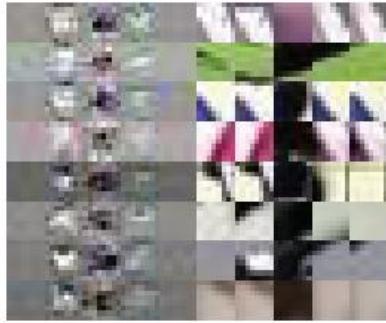
(c) 3D convolution



Moving edges and blobs



Shot changes

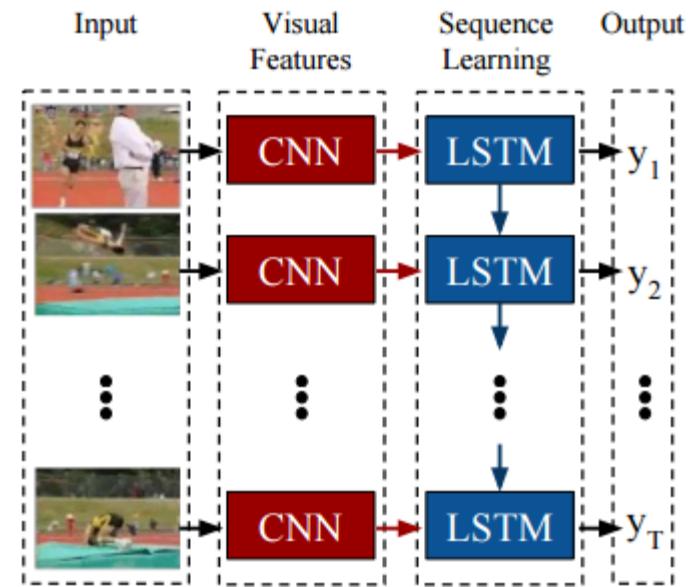
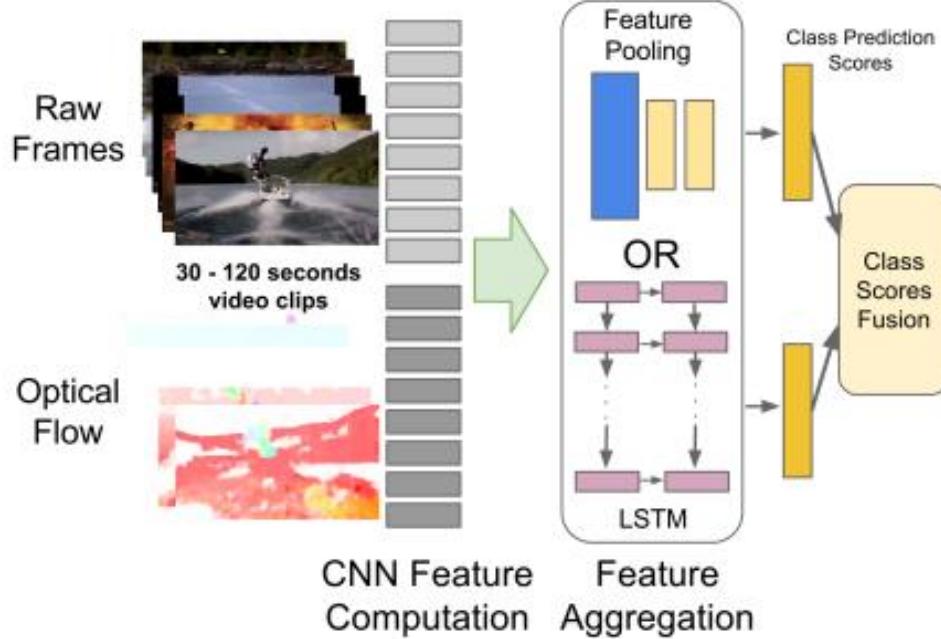


Edge orientation changes

Color changes



CNN ve LSTM





Çözüm Önerileri

- Video karesi temelinde
 - Kavramlarla ifade
 - Arama motorları ile ölçeklendirme
- Videodaki önemli parçaların keşfi
 - Ayırt edici kısımlar
- Hareketleri bulma
 - Daha basit ve hızlı tanımlayıcılar





Internetteki çok sayıda
resim ve videodan
faydalananarak kavram
öğrenme



Google arama sonuçları

Google george colooney



Google angelina jolie



E.Golge, P.Duygulu. FAME: Face Association Through Model Evolution. In CVPR Workshops, 2015

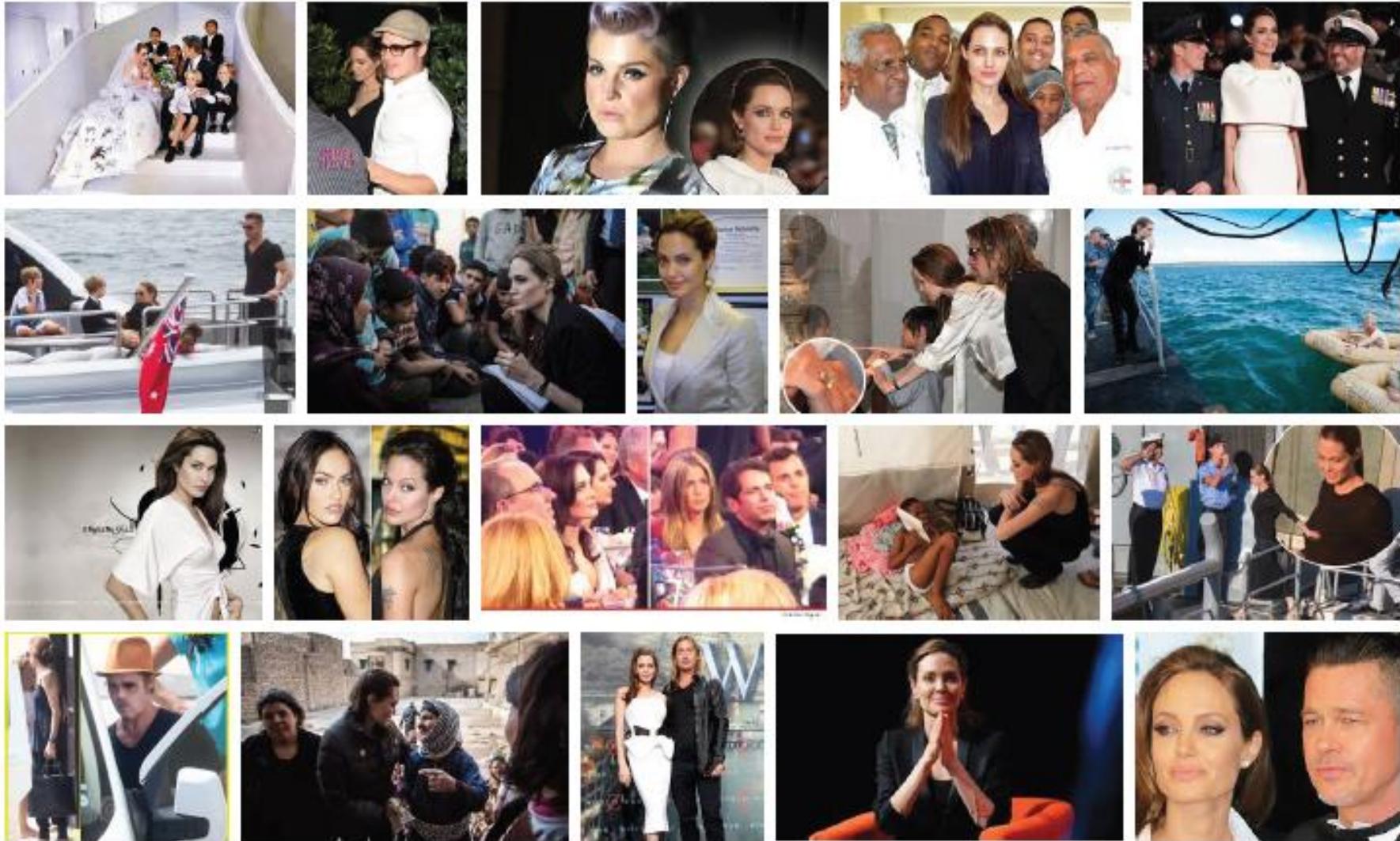
+ Zayıf etiketli öğrenme



+ Sınıf içi farklılıklar



+ Ayrık elemanlar – sınıf dışı örnekler

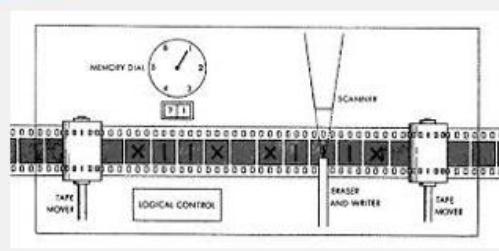
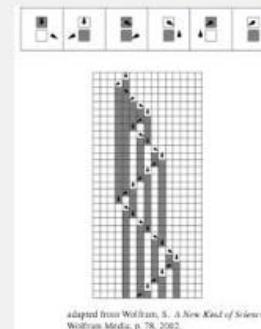
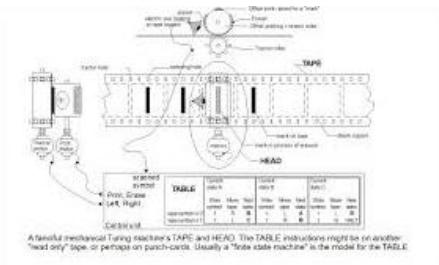
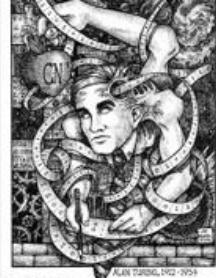


+Daha iyi bir model için veri ayıklama



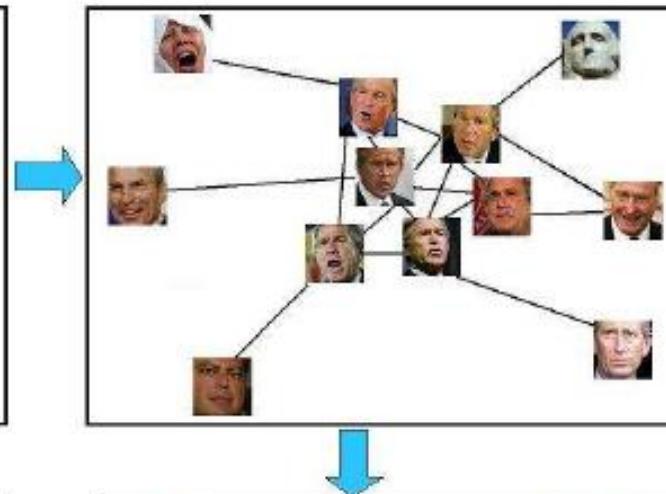
+ Sorgu tek bir sınıfı karşılık geldiğinde

Sorgu : Turing

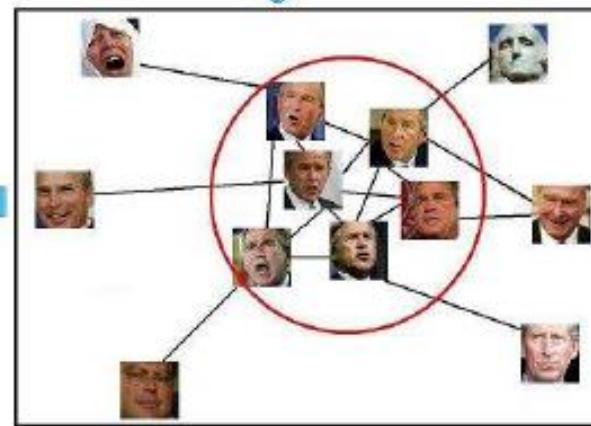
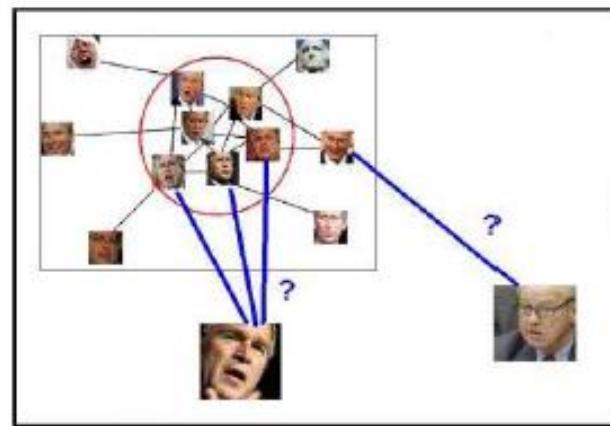




Yüz isimlendirme



Bir isim ile eşleşen yüzler arasında birbirine en yakın alt kümenin bulunması



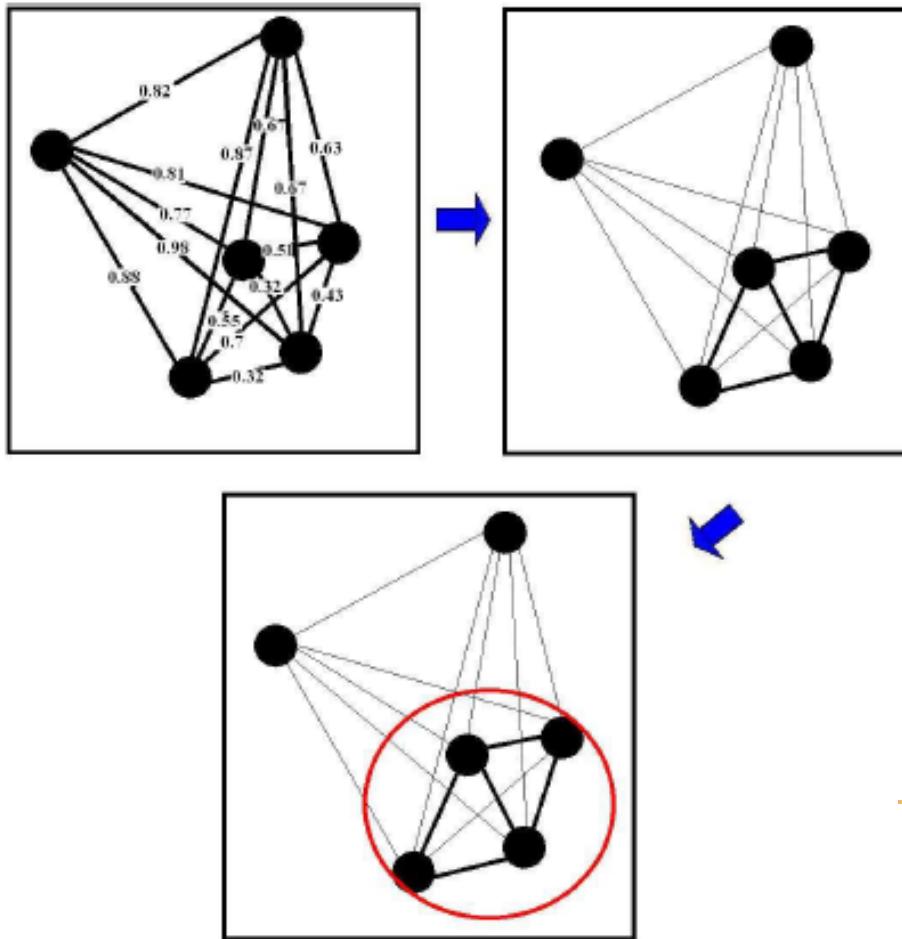
Ozkan, D., Duygulu, P., "Interesting Faces: A Graph Based Approach for Finding People in News", Pattern Recognition, 2010

Ozkan, D., Duygulu, P., "A Graph Based Approach for Naming Faces in News Photos", CVPR, 2006

Ozkan, D., Duygulu, P., "Finding People Frequently Appearing in News", CIVR, 2006



En yoğun alt bileşeni bulma

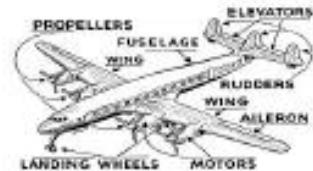
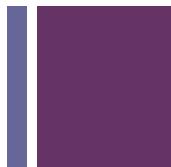


$$f(S) = \frac{|E(S)|}{|S|},$$

Her seferinde minimum dereceye sahip düğüm atılır

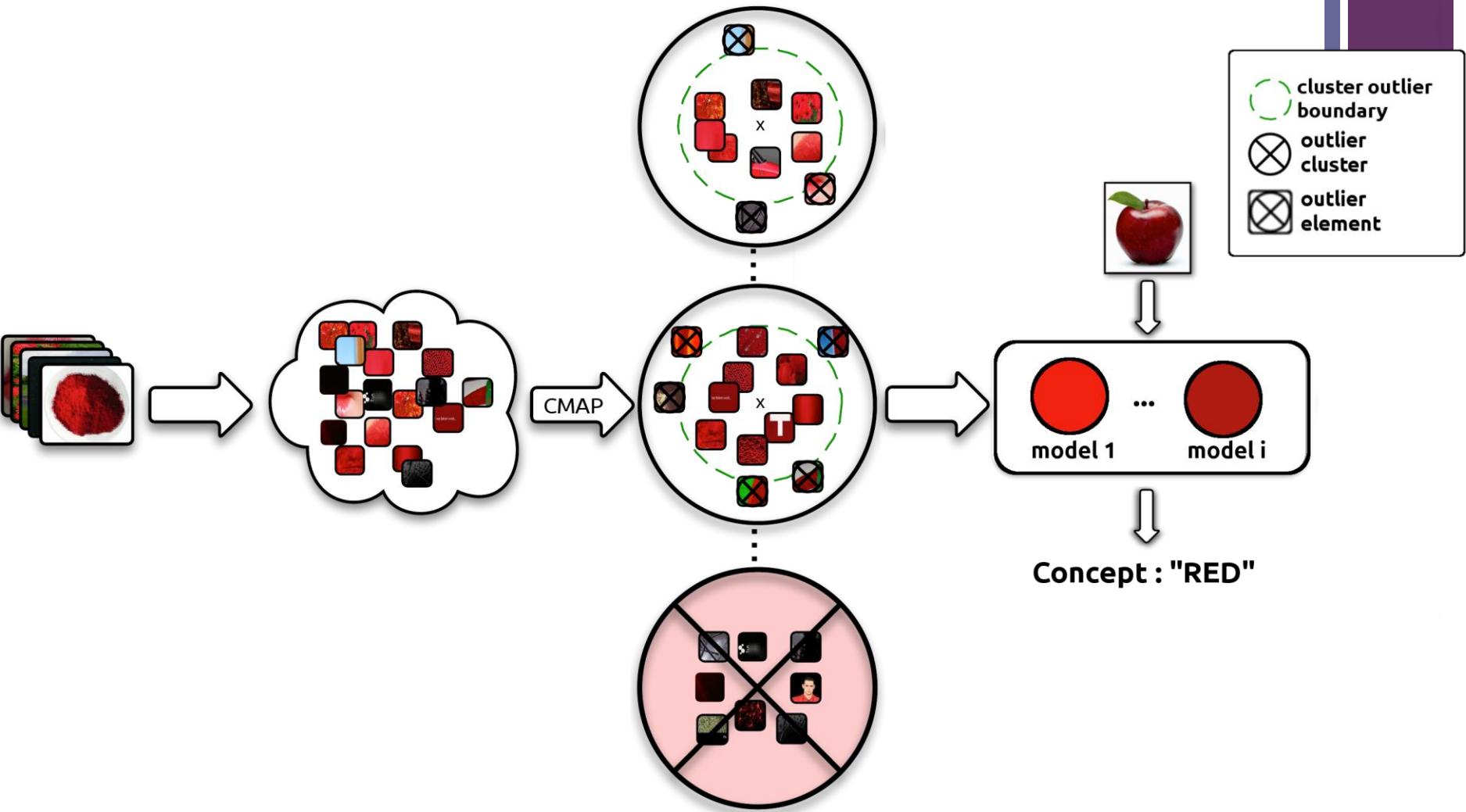
(Charikar, 2000)

+ Birden fazla anlam ya da farklılık olduğunda





Kavram Haritası



Golge, E., Duygulu, P., "Concept Maps: Mining Noisy Web Data for Concept Learning ", accepted to ECCV 2014

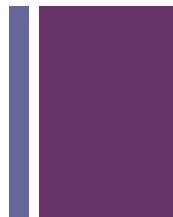


Renk ve doku nitelikleri

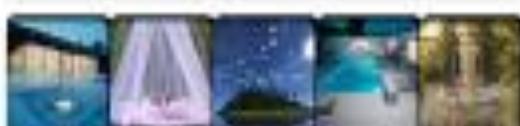
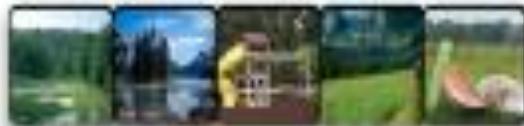




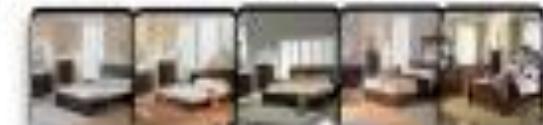
Sahne/Ortam



Outdoor

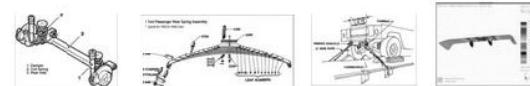
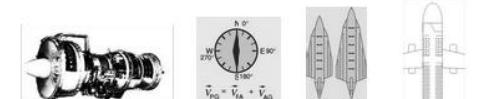
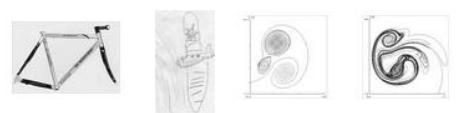


Bedroom



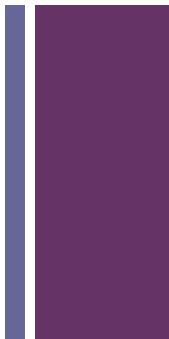


Nesne tanıma

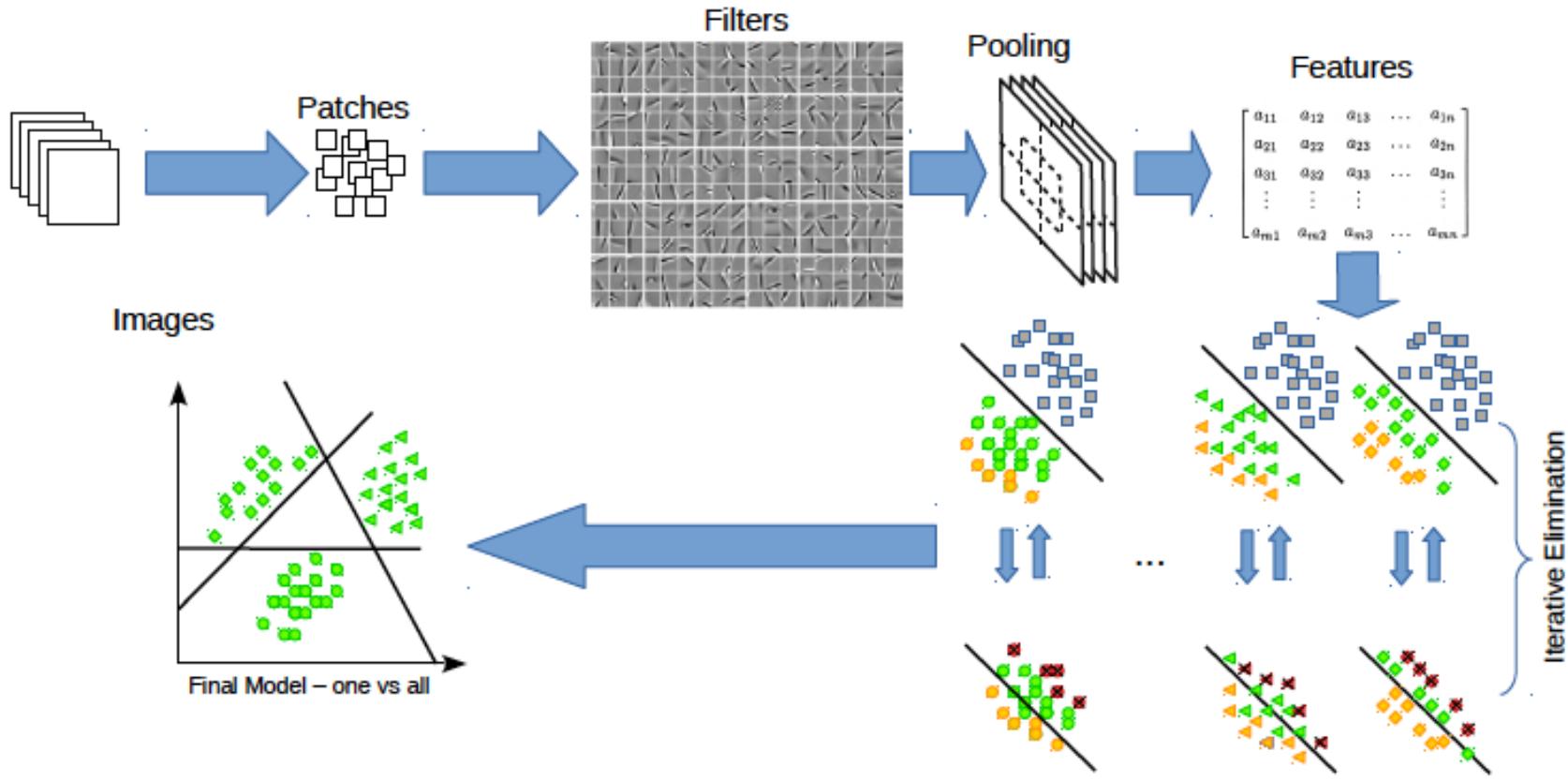
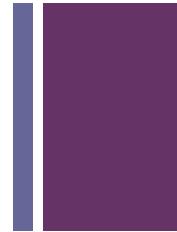




Yüz tanıma



+ Model Evrimi ile Eşleme (AME)



Confident Positives



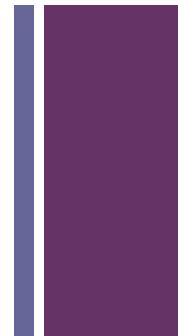
Poor Positives



Final Eliminations



Iter. 1



Iter. 2



Iter. 3



Iter. 4





Zayıf etiketli videolardan öğrenme

basketball

Filters ▾

About 15,800,000 results

Best Crossovers in Basketball History
The Highlight Factory
11 months ago • 4,160,930 views
Song: R.I.C.O. (feat. Drake) - Meek Mill <https://www.youtube.com/watch?v=EgRxFsX538> Copyrights are owned by the NBA, TNT, ...
3:40

NEW BASKETBALL VINES & INSTAGRAM VIDEOS #1 - BEST BASKETBALL MOMENTS
The Best Sports Vines
1 month ago • 2,275,770 views
New Basketball Vines and Instagram Videos with Best Basketball Moments. The Best Sports Vines 2016 #1 ...
12:04

Spiderman Basketball Episode 7 ...Spiderman vs Carnage... SuperHero bball
Professorlive 🎬
8 months ago • 11,548,303 views
Episode Deadpoc
10:02

BEST Basketball Vines of June 2016
Golden State Warriors Basketball
1 month ago • 2,591,091 views
BASKETBALL VINES JUNE 10:13

Mexican Basketball
Barbell Brigade
11 hours ago • 16,200 views
Barbell Brigade: <http://barbellbrigade.com/en/> Barbell Brigade 641 Angeles, CA 90031 Instagram: ...
NEW

INSANE 1 ON 1 WATER BALLOON BASKETBALL!
JesserTheLazer 🎬
17 hours ago • 118,681 views
Me and Mitchell did a one on one with a twist! Mitchell <https://youtu.be/FjblV1A58Qk> Tyler ...
1V1

Giant Basketball Arcade Battle | Dude Perfect
Dude Perfect 🎬
3 months ago • 17,179,995 views
Arcades are more fun when they're GIANT. ~Submit Your Trick Shot Video Here! <http://bit.ly/RufflesDudes> Play our FREE new ...
7:47

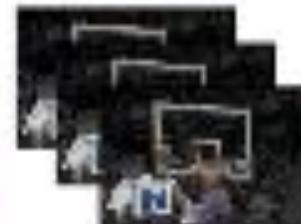
BASKETBALL SHOOTING CHALLENGES!!
TDPresents 🎬
2 weeks ago • 192,190 views
I take on Walker in some IRL basketball challenges. Loser gets spoiled milk (worst forfeit ever) poured on them. NBA 2K16 videos ...
5:58

Epic Basketball Trick Shots & Fails Compilation - Funny Vines 2016
Funny Vines
1 month ago • 1,102,071 views
Basketball Trick Shots and Fails Edition! Check our more Vine Compilations ...
10:00

Watch A Full 5 Minutes of The Basketball Wives LA Season 5 Premiere Episode | VH1
VH1 🎬
23 hours ago • 28,306 views
The ladies of LA pick up right where they left off. See the first 5 minutes of Basketball Wives LA Season 5! Tune in Sunday, July ...
5:11

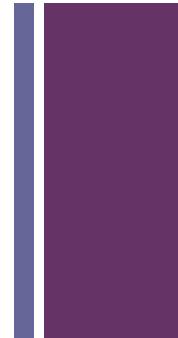
+

Zayıf etiketli videolardan öğrenme



+

Çoğul ortam verisinde olay tanıma Doğumgünü



Mum üfleme



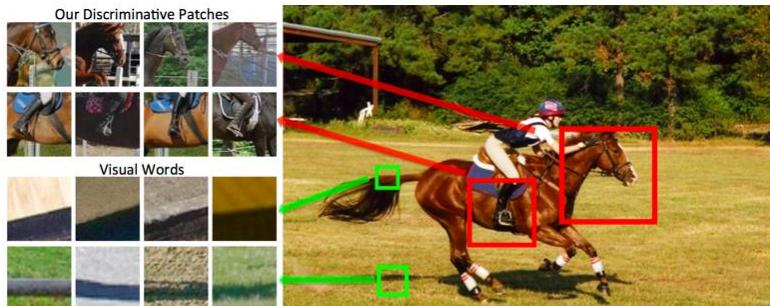
ul 2016, Ankara

+

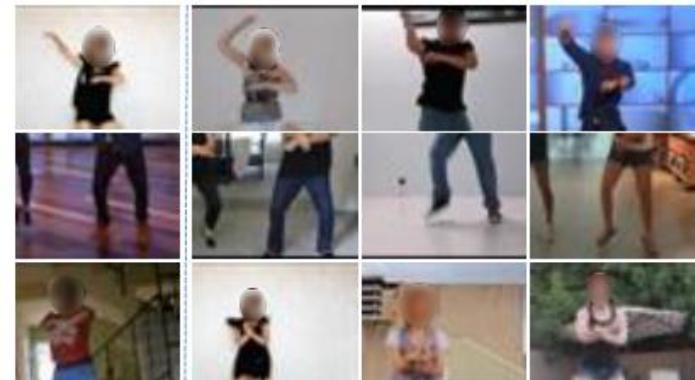
Önemli parçaların bulunması

Singh ECCV 2012

Unsupervised Discovery of Mid-Level Discriminative Patches

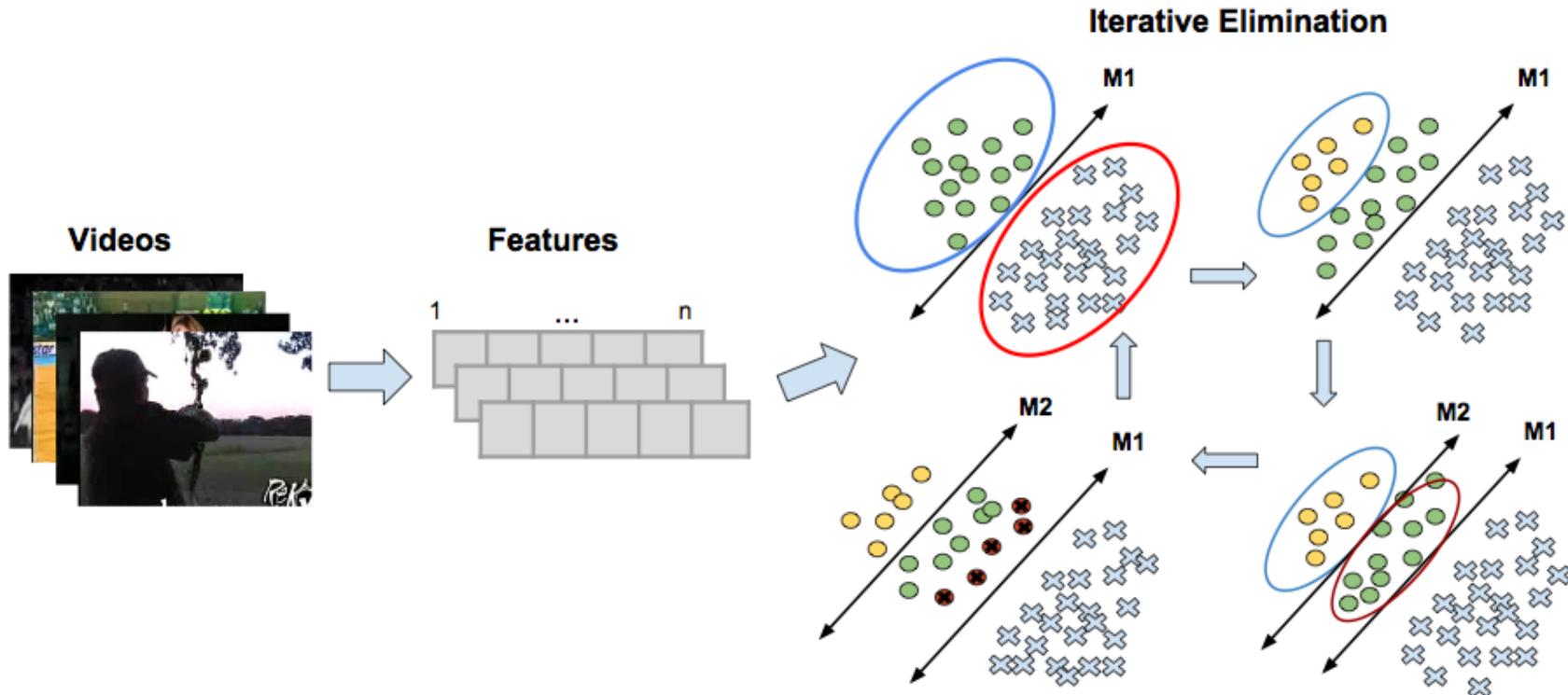
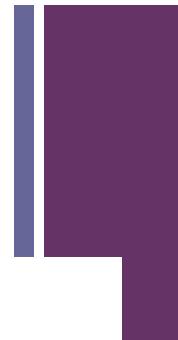


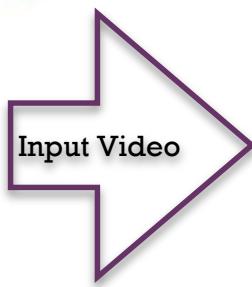
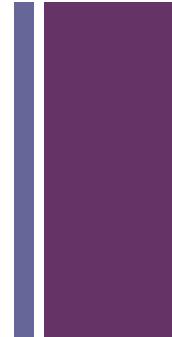
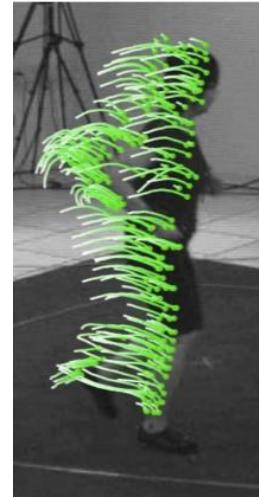
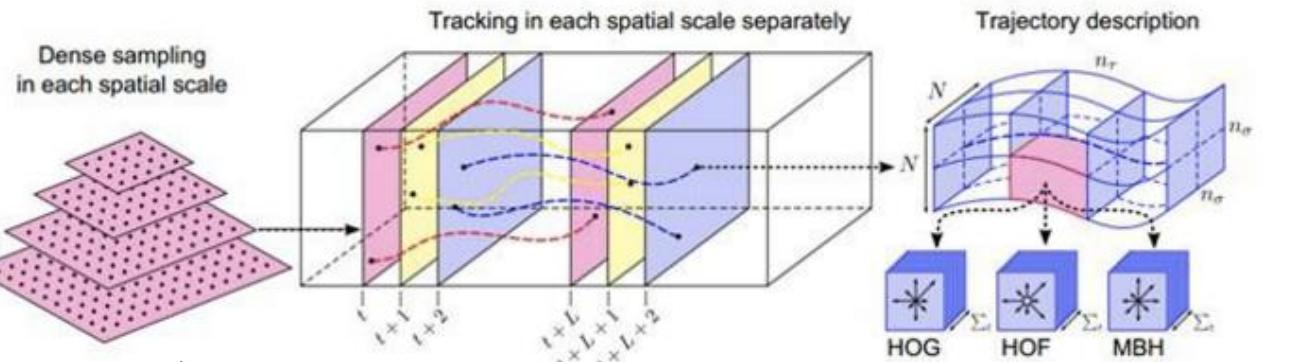
Jain CVPR 2013





AME

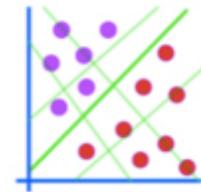
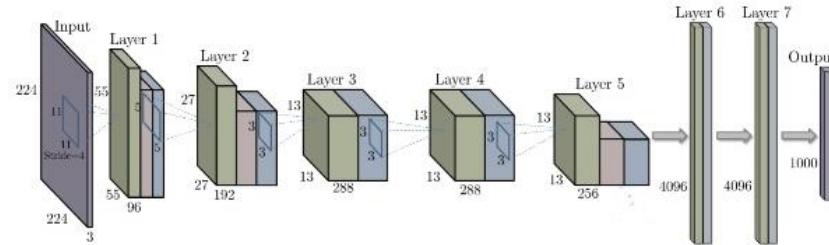




CNN FC-6
Layer
Features

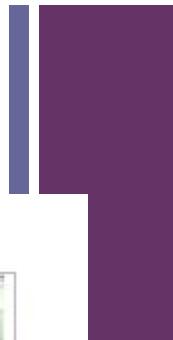


Classification

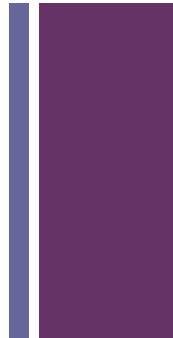


AlexNet [8,9,10]

3D ConvNets [11,12]



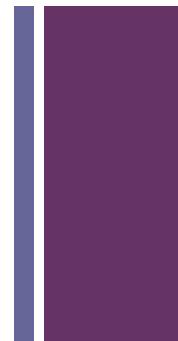
Example of successful eliminated instances by AME^[1] for ActivityNet action classes. (1) "Archery" class. Baseline: 31.57%, AME: 44.73%. (2) "Checking tires" class. Baseline: 26.82%, AME: 41.46%. (3) "Platform diving" class. Baseline: 56.66%, AME: 73.33%

**Remaining****Eliminated**

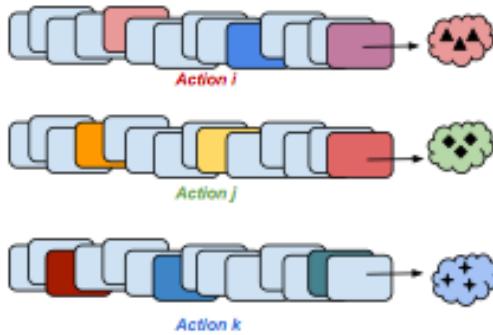
Example of unsuccessful eliminated instances for ActivityNet class "Windsurfing" with AME[1]. Baseline: 74.07%, AME[1]: 66.66%



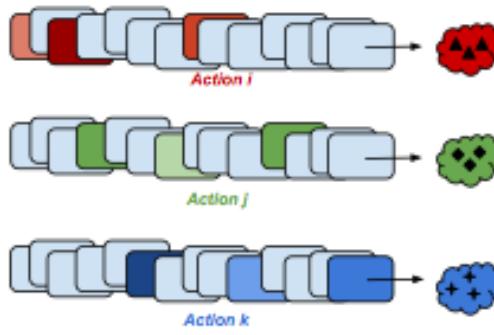
Prototipler ile gösterim



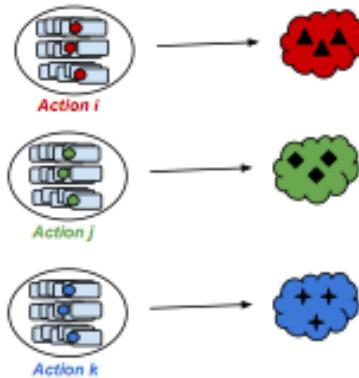
Random Selection



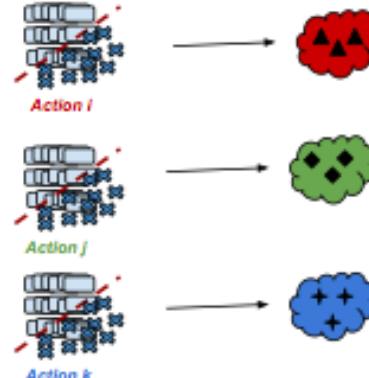
Discriminative Selection



Cluster Center Selection



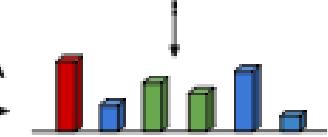
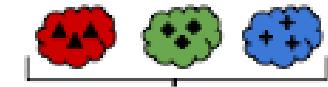
AME Selection



Video clips



Prototypes



Prototype based representation

Euclidean Similarity



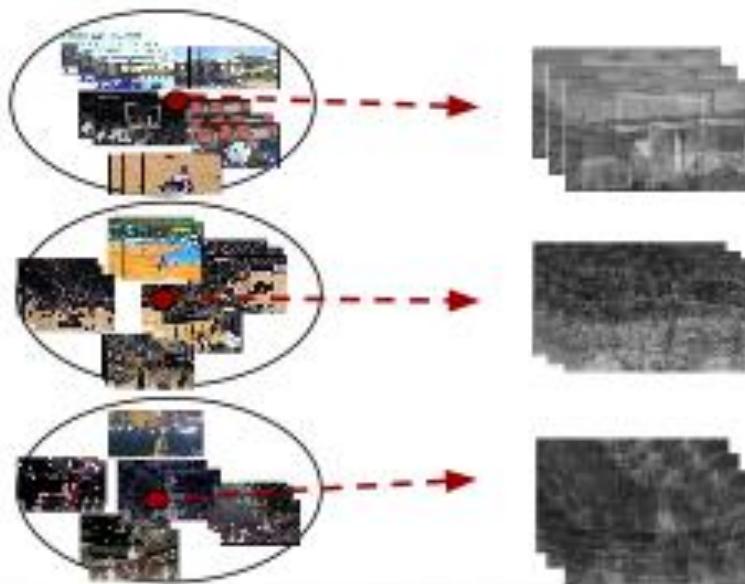
Exemplar SVM Score



Random Selection



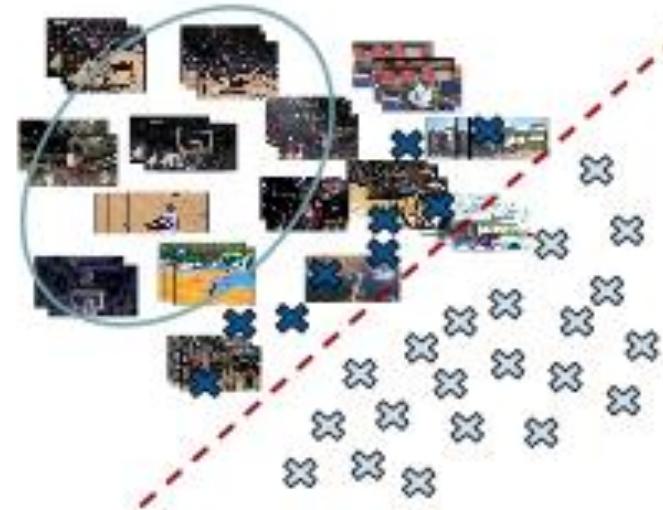
K-means Cluster Center Selection



**Discriminative Selection
(High Scored Cluster Members)**



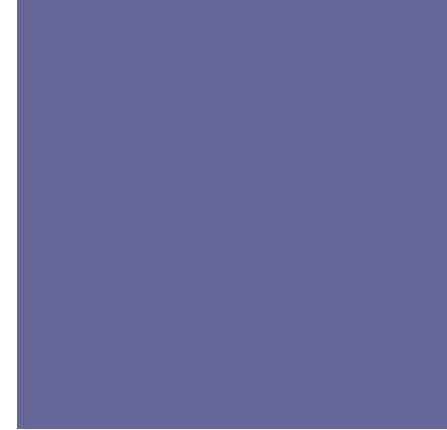
AME Selection



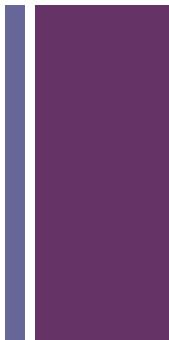


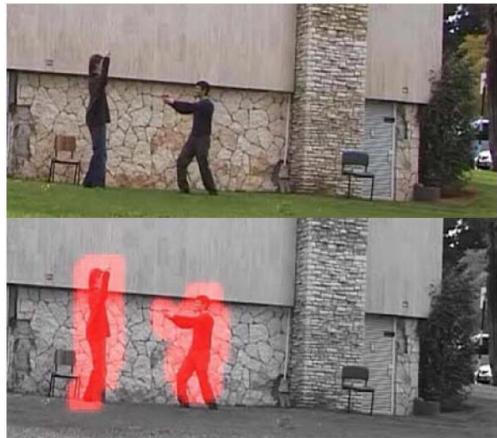
Videolardaki beklenmedik durumların keşfi

Pınar Duygulu, Eylül 2016, Ankara



+

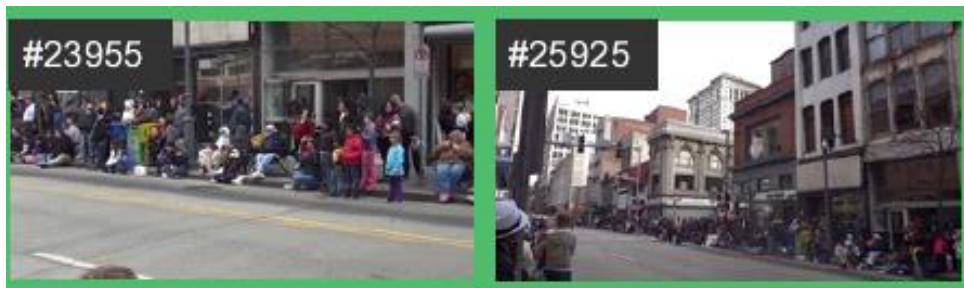




**Boiman and Irani
ICCV 2005**



Roshtkhari and Levine, CVPR 2013



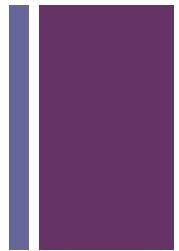
Ito, Kitani, Bagnell, Hebert, 2012



**Zhao, Fei-Fei, Xing,
CVPR 2011**

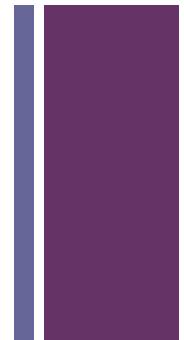
+

Olağan yerine olağandışı





Ani hareketler

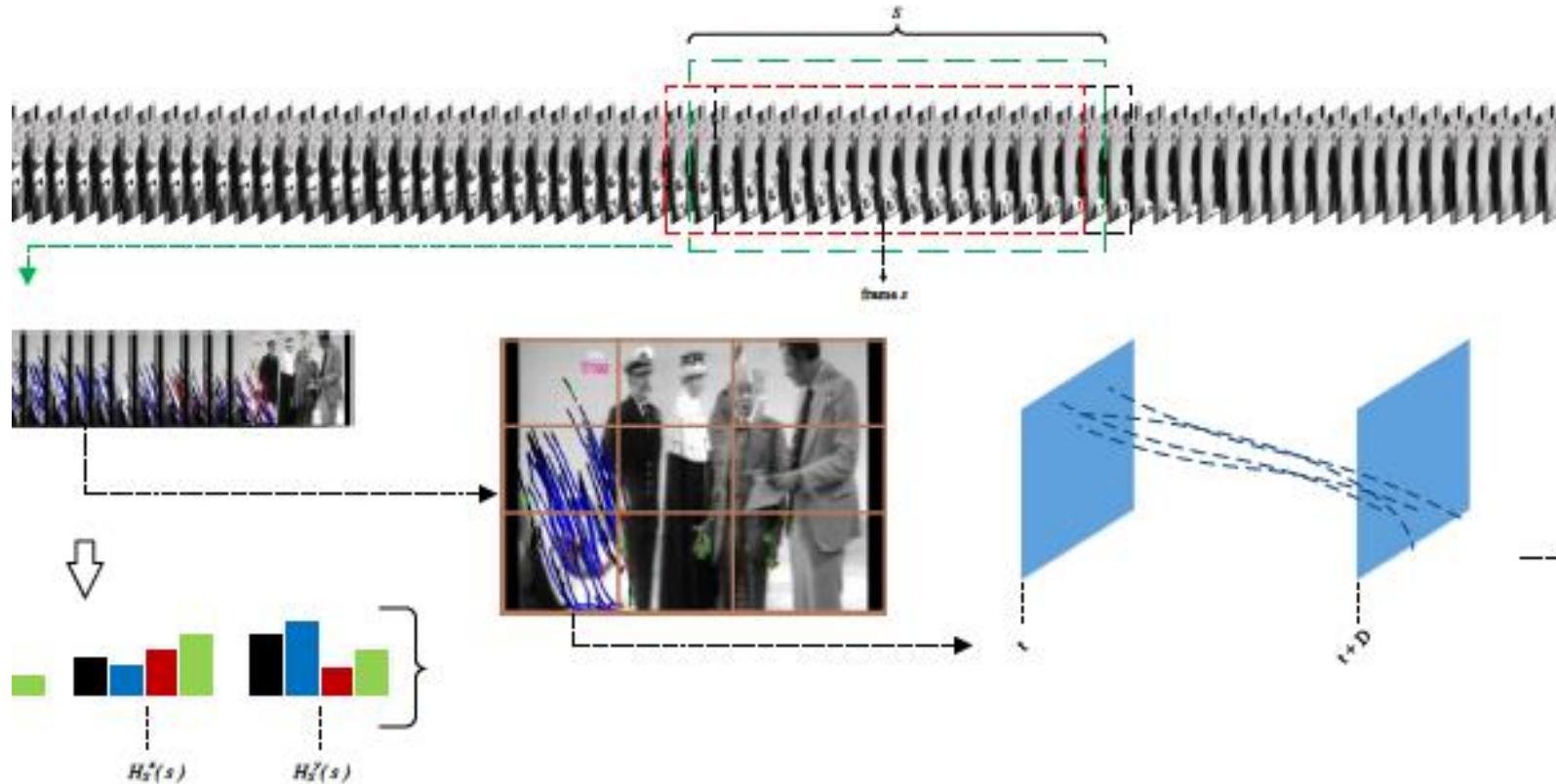


Yüksek hız
Büyük uzamsal kapsam



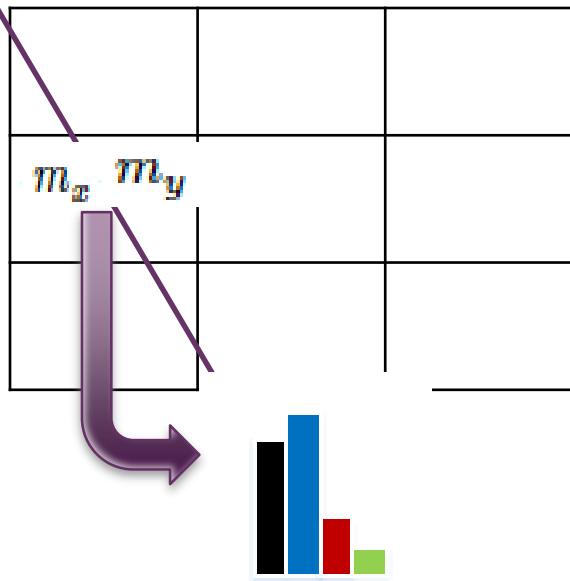
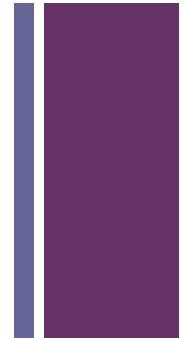


Yoğun İz Dağılımları





Gösterim



$$H_S^l(t) = (H_S^l(t)_{[1,1]}, \dots, H_S^l(t)_{[1,N]}, \dots, H_S^l(t)_{[N,N]})$$



Pınar Duygulu, Eylül 2016, Ankara



$$H_S^l = \sum_{t=s-(\|S\|/2)}^{s+(\|S\|/2)} H_S^l(t) \quad H_S = (H_S^l, H_S^x, H_S^y)$$

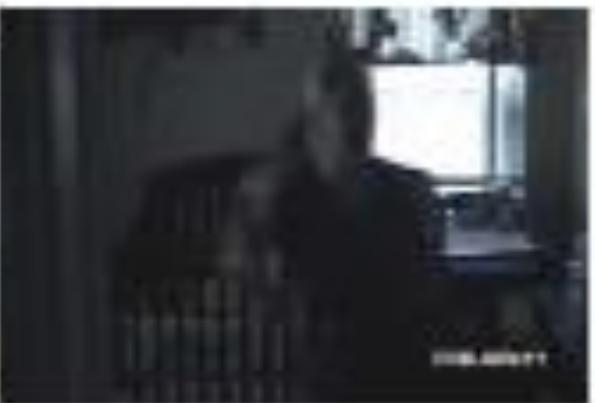
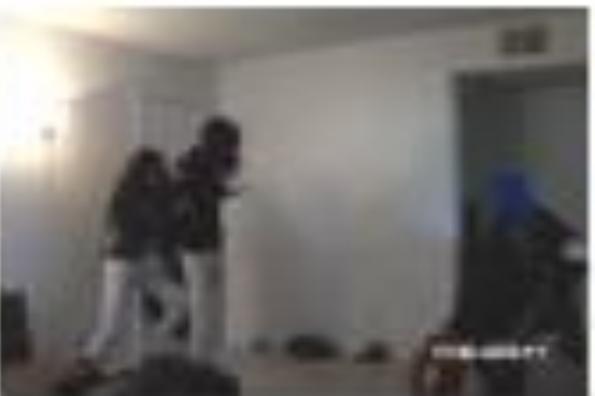
Velocity and spatial extension of the motion

$$\vec{T} = (P_t, \dots, P_{t+D-1}) \quad \vec{P_t} = (x_t, y_t)$$

$$m_x = \frac{1}{D} \sum_t^{t+D-1} x_t, v_x = \frac{1}{D} \sum_t^{t+D-1} (x_t - m_x)^2$$

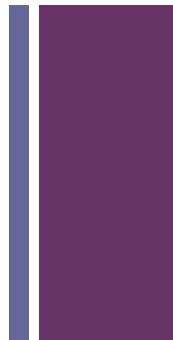
$$m_y = \frac{1}{D} \sum_t^{t+D-1} y_t, v_y = \frac{1}{D} \sum_t^{t+D-1} (y_t - m_y)^2,$$

$$l = \sqrt{\sum_t^{t+D-1} \sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2}}$$



Pınar Duygulu, Eylül 2016, Ankara





Pinar Duygulu, Eylul 2016, Ankara

FAILARMY

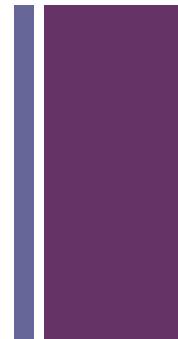
+

Başarısız durumlar



+

Kısıtlar





Hareket analizi

Pinar Duygulu, Eylul 2016, Ankara





Bu insanlar ne yapıyor?



running



walking



throwing



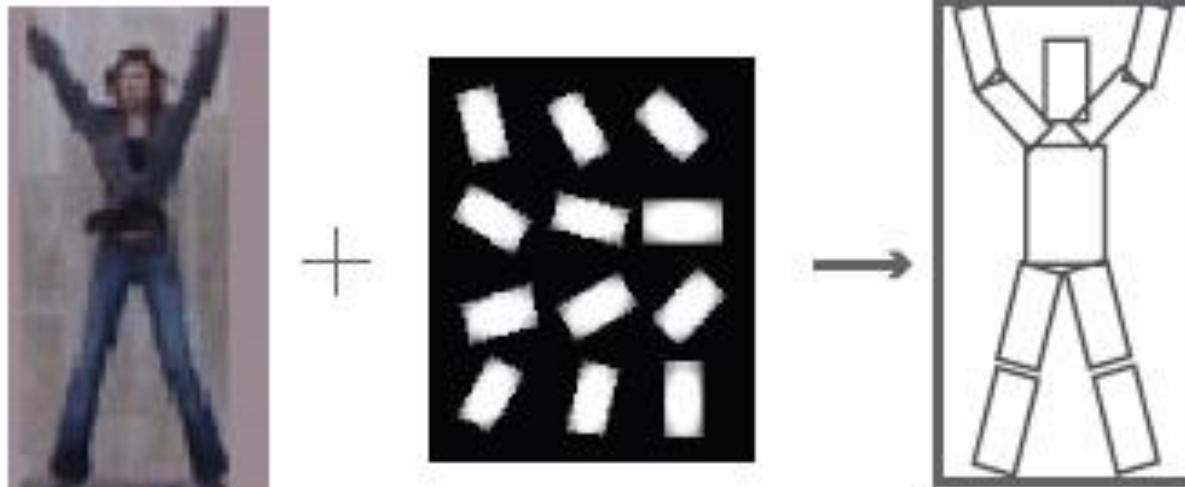
crouching

- Poz çok bilgi içeriyor ama nasıl tanımlamalı?

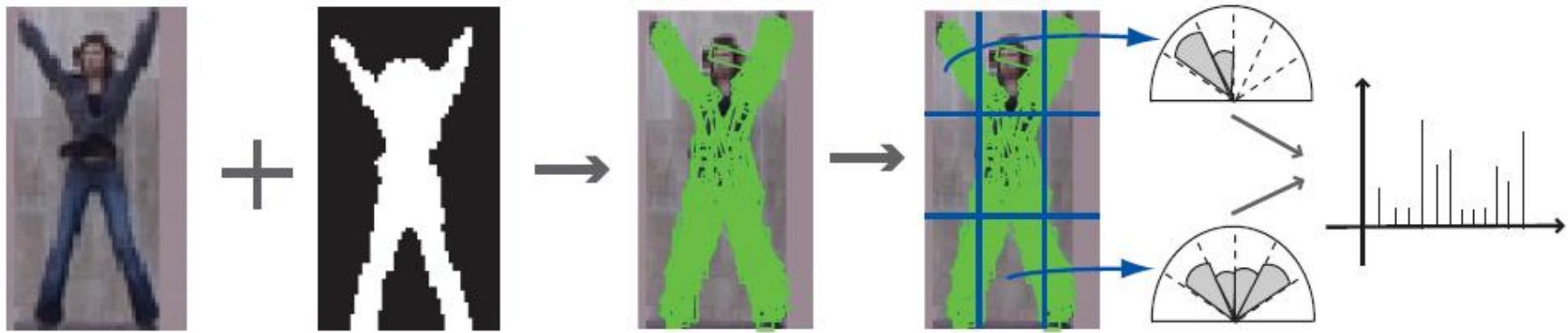
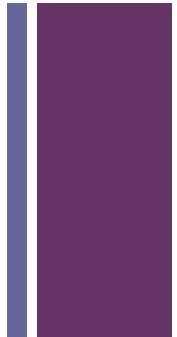


Dikdörtgenlerle pozu tanımlama

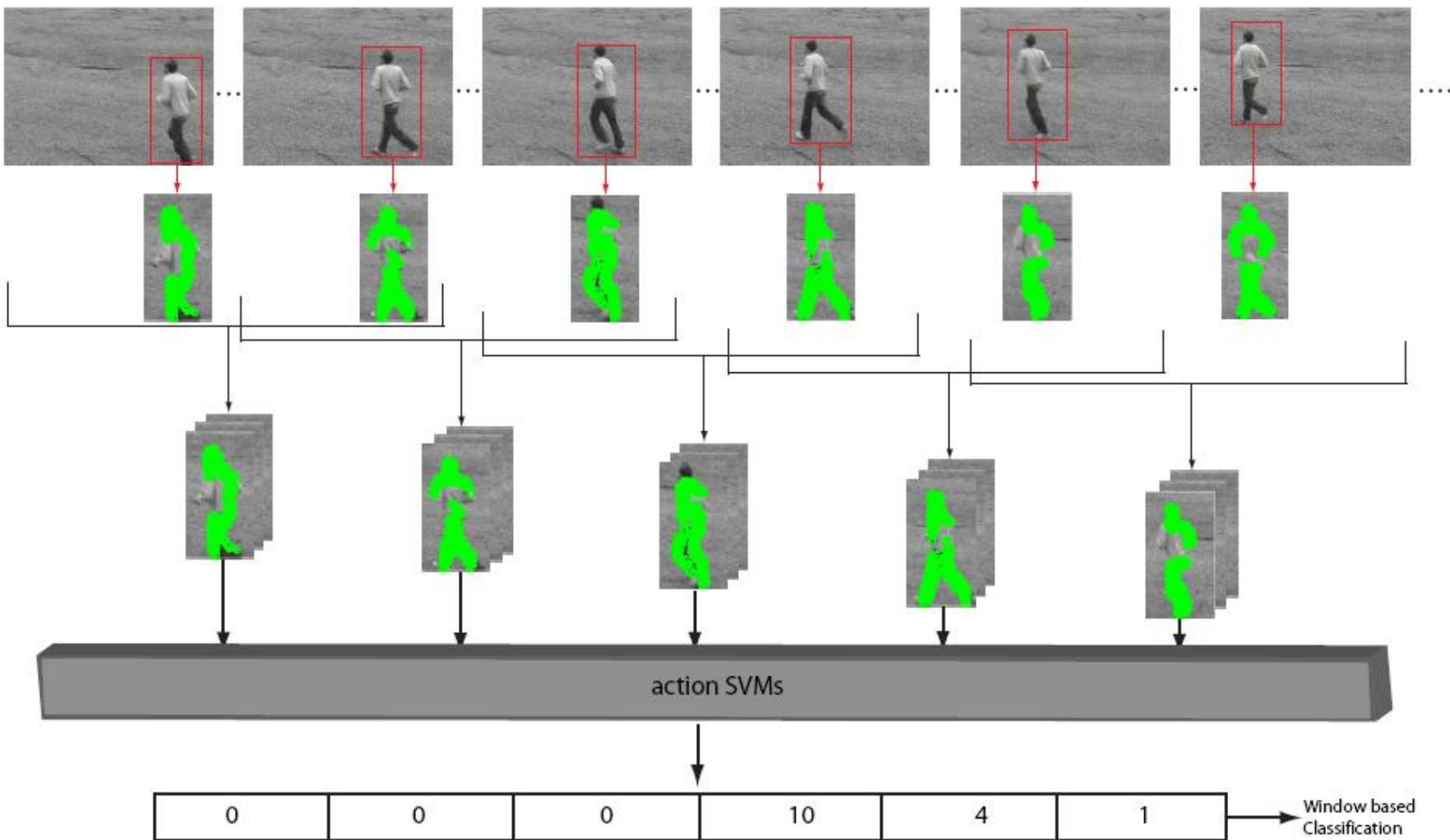
- Silindire benzeyen vücut parçalarının izdüşümü olan dikdörtgenlerin açı ve uzunlukları



+ Histogram of Oriented Rectangles (HOR)



+ Sınıflandırma

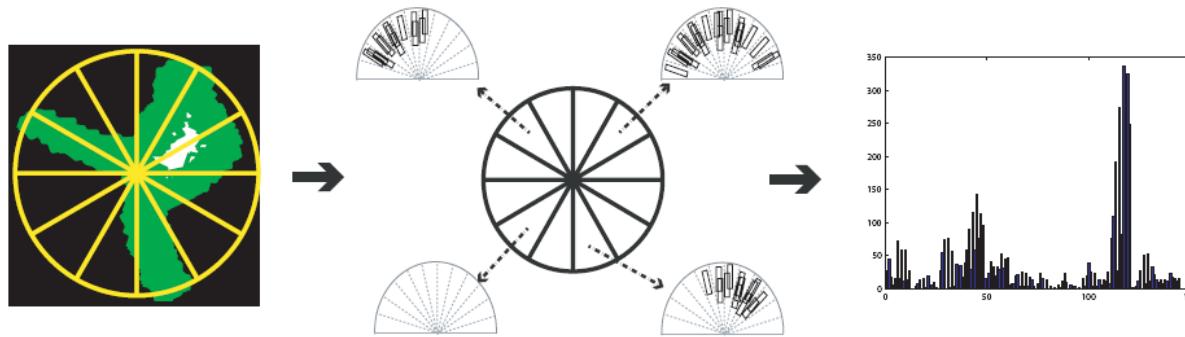




Tek resimde hareket tanıma

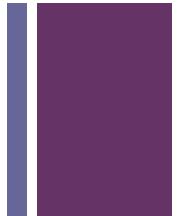


- Ramanan'ın yöntemi ile poz tanıma + HOR



Ikizler, N., Cinbis, R. G., Pehlivan, S., Duygulu, P., "Recognizing actions from still images", Proc. 19th International Conference on Pattern Recognition (ICPR 2008)

+Still Image Results



running



walking



throwing



catching



crouching



kicking

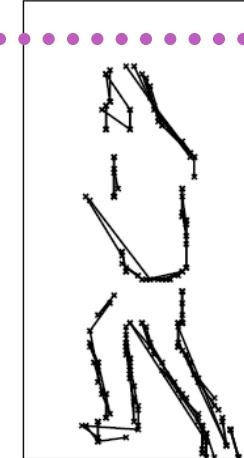
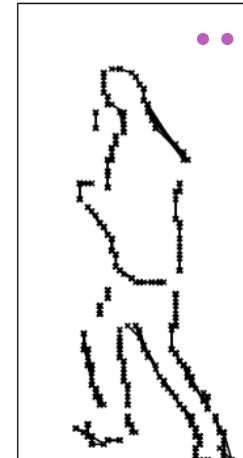
ActionWeb dataset -
467 images collected
from the web



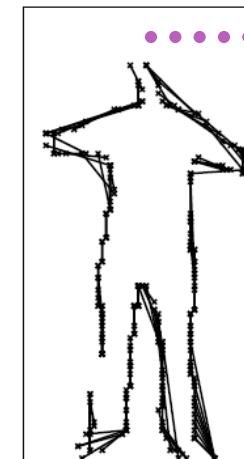
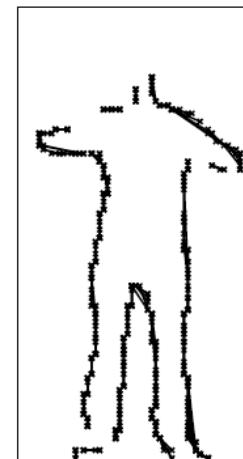
Çevreye oturtulan çizgiler

- Siluet olmadığında çevreye oturtulan çizgiler (Pb)

(Martin PAMI2004) yoluyla tanımlama



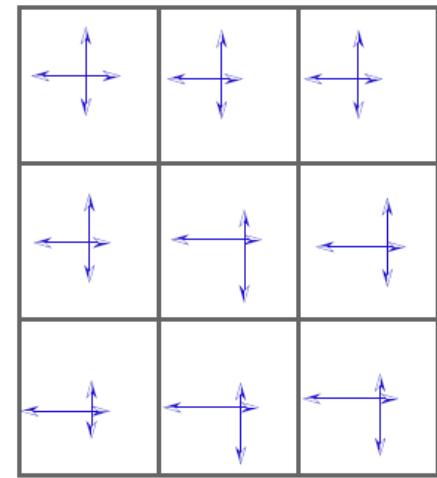
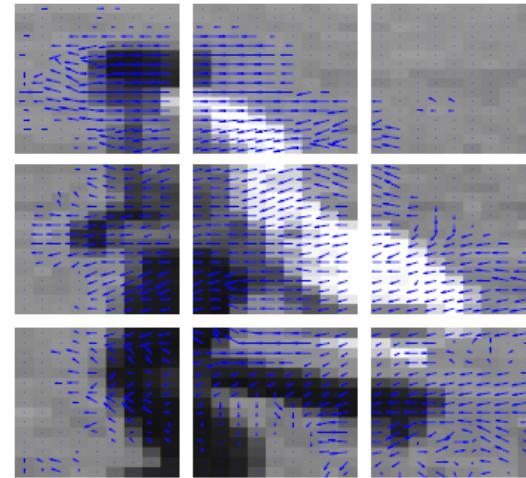
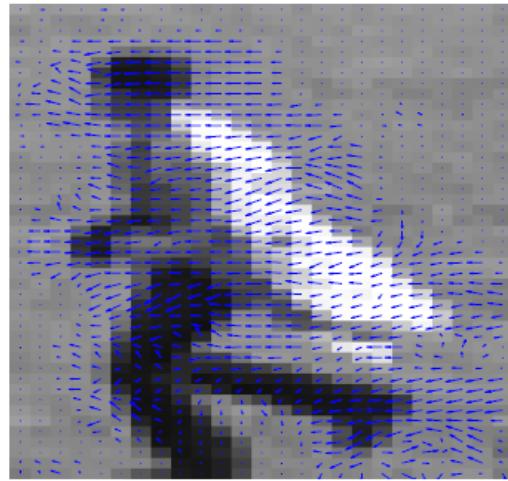
► fine detail



► coarse shape

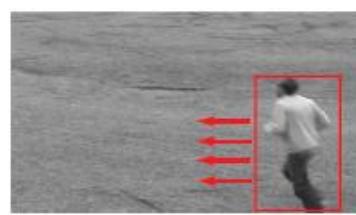


..ve Optik Akış

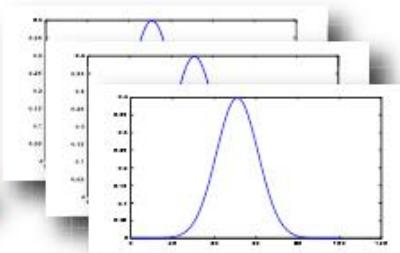




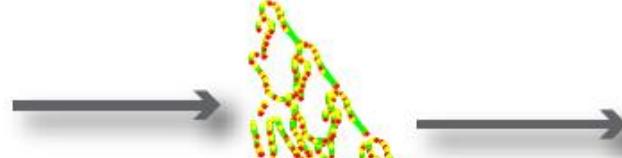
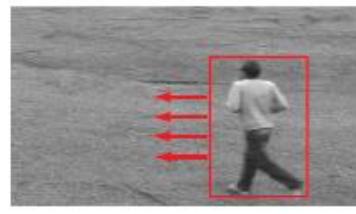
LHist ve OFHist



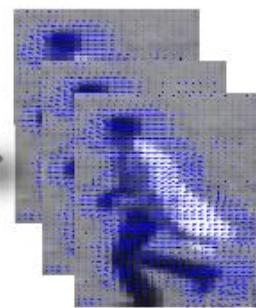
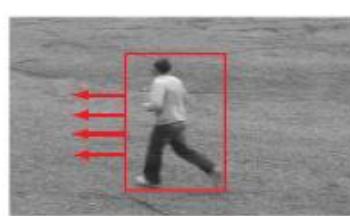
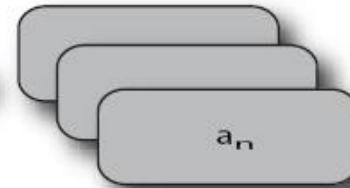
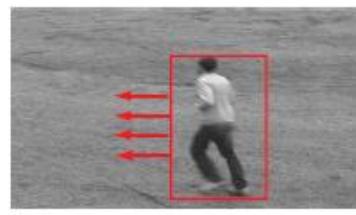
$$\rightarrow dx/dt \rightarrow$$



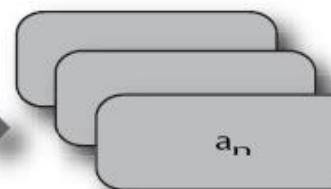
velocity GMs



LHist



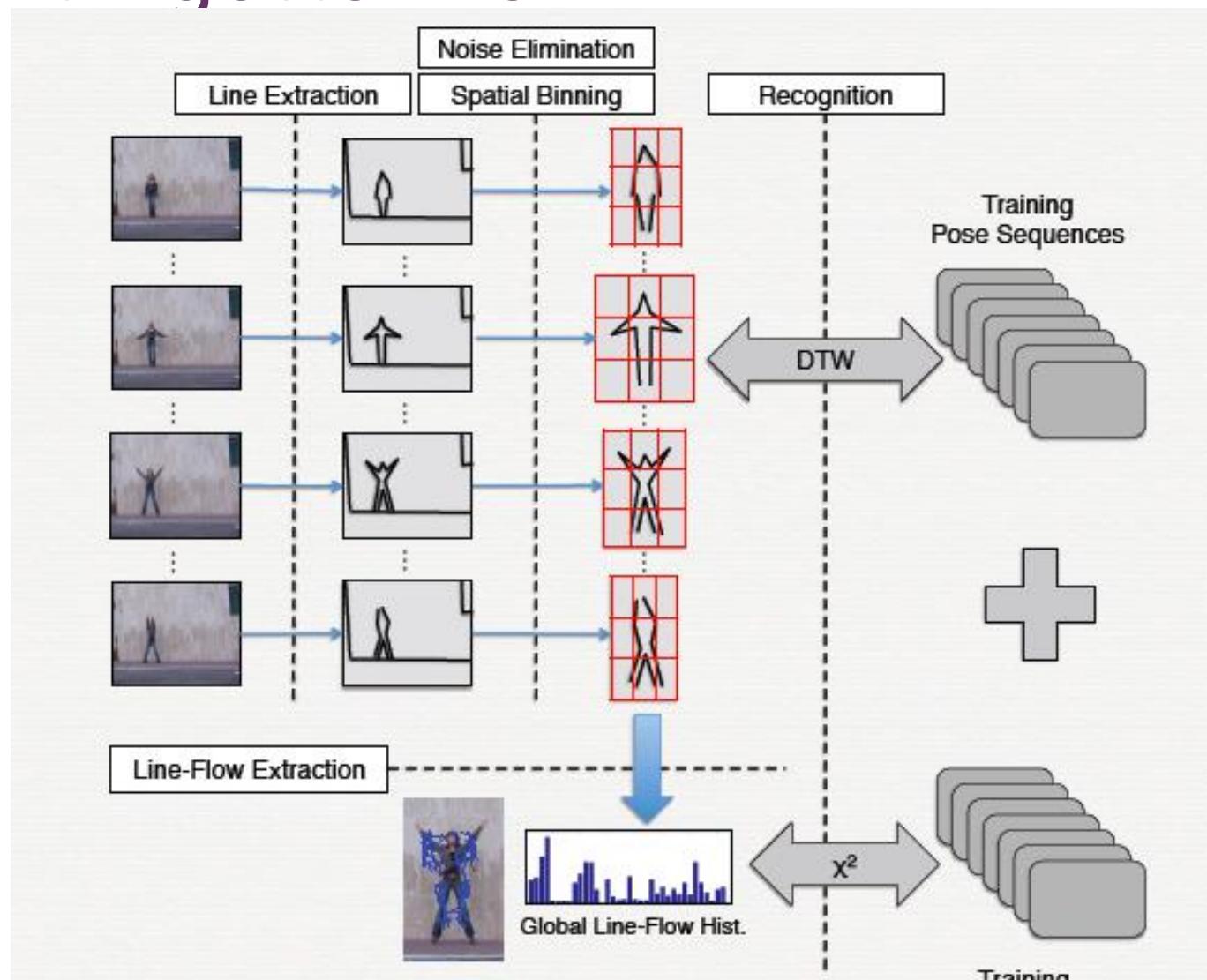
OFHist



SVM models

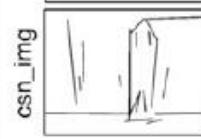
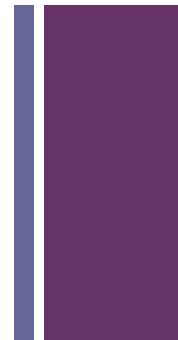


+ Pozu çizgi parçalarının dağılımı olarak göstermek

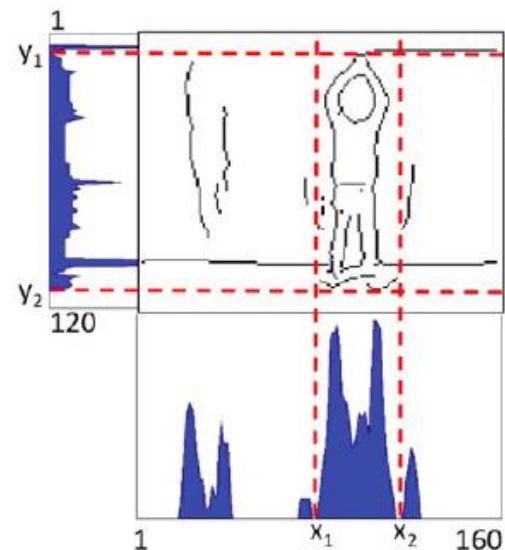




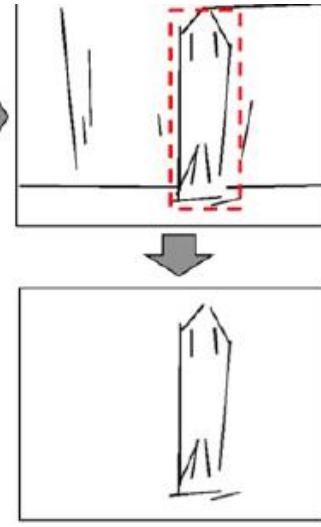
Çizgi eşleri



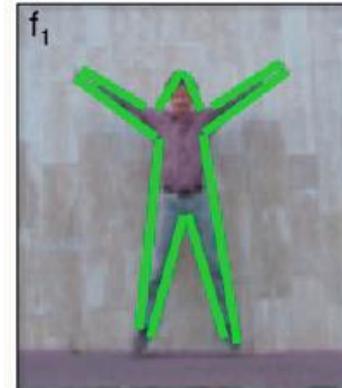
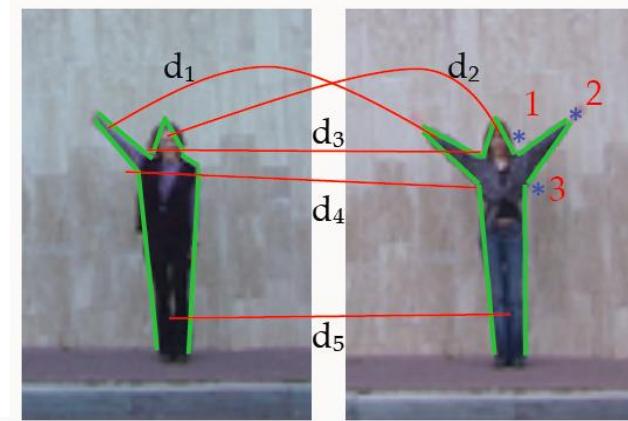
(a)
pose extraction



(b)
projection of edge_img
to extract bounding box



(c)
noise elimination
on esn_img



$$\begin{array}{l} \checkmark = \checkmark \\ \not\equiv = \not\equiv \\ \rangle = \rangle \\ \not\mid = \not\mid \end{array}$$

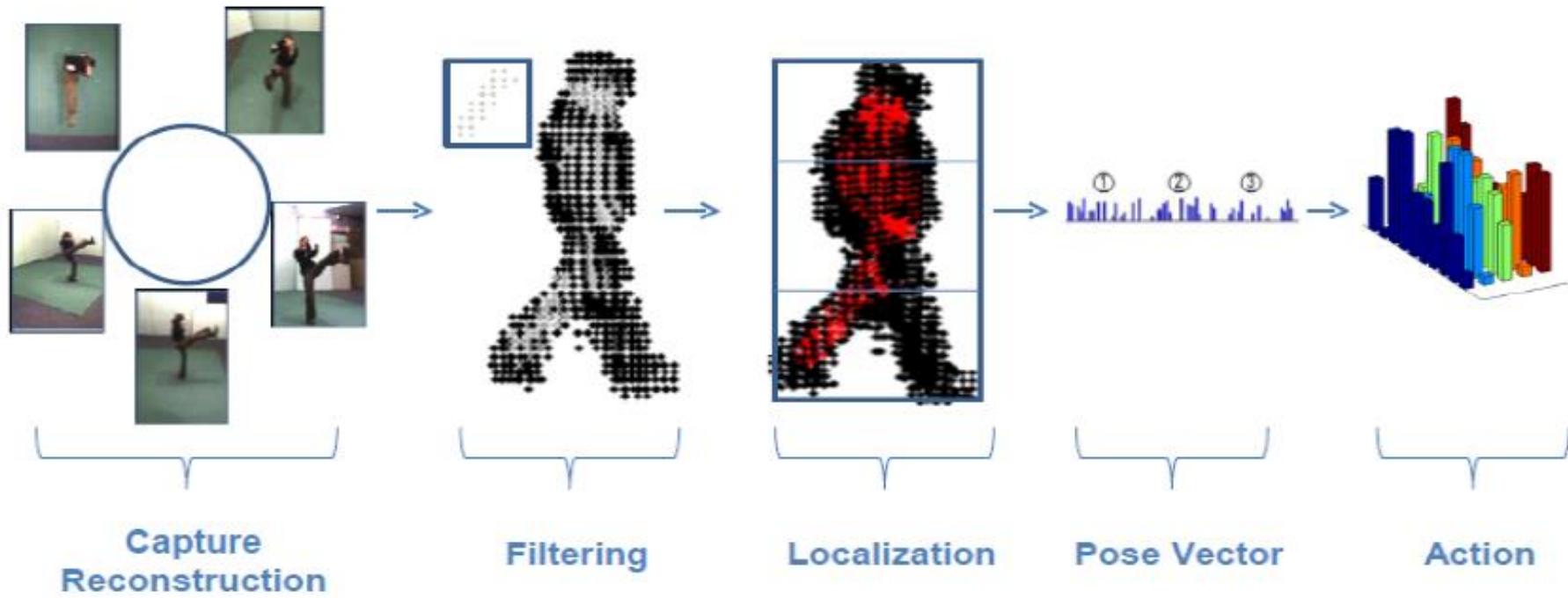
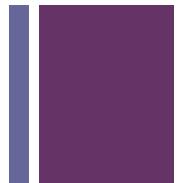
$$\begin{array}{l} \wedge = \wedge \\ \not\equiv = \not\equiv \\ ? = \sqcup \\ \wedge = \wedge \end{array}$$

$$\begin{array}{l} \checkmark = \checkmark \\ \not\equiv = \not\equiv \\ \langle = \langle \\ \not\mid = \not\mid \end{array}$$



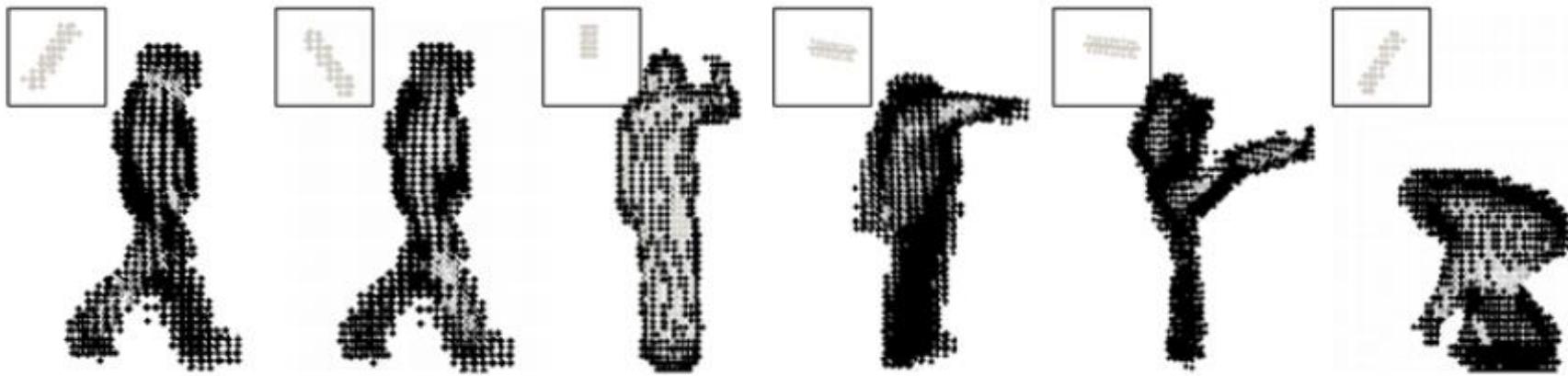
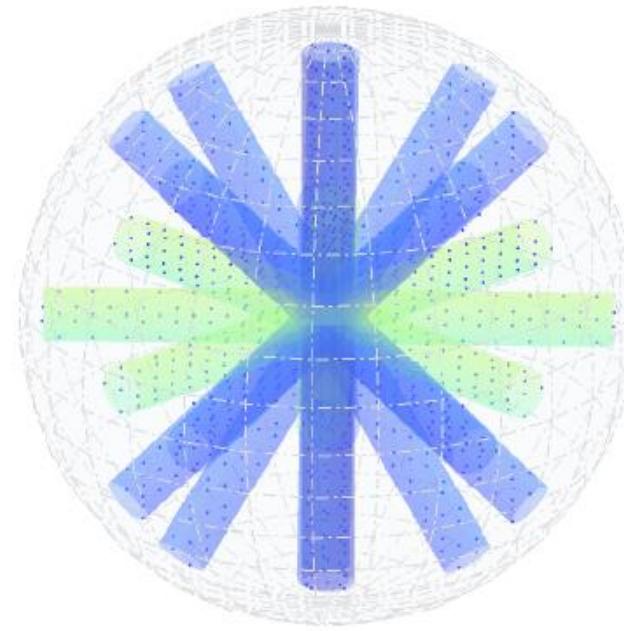
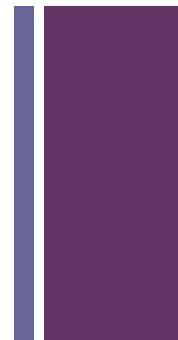


Birden fazla kamera görüntüsü



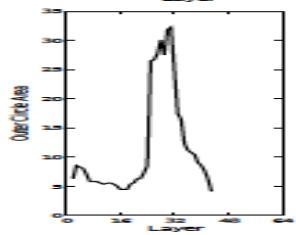
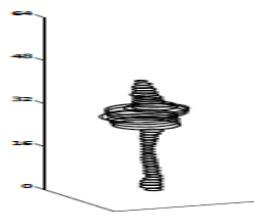
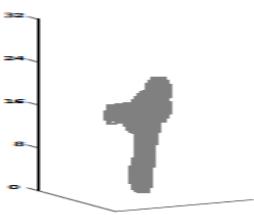
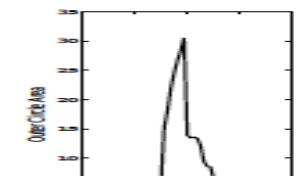
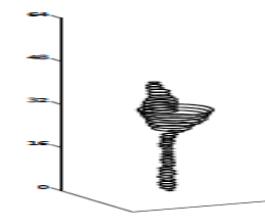
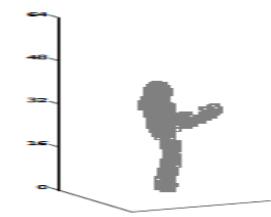
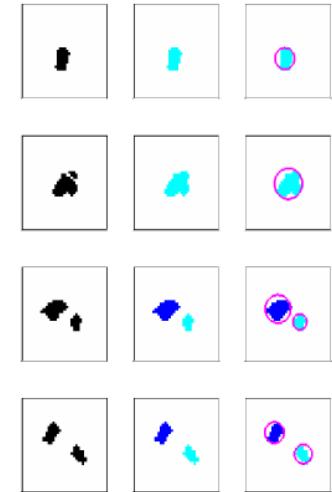
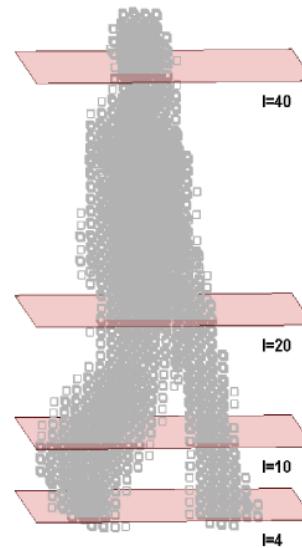
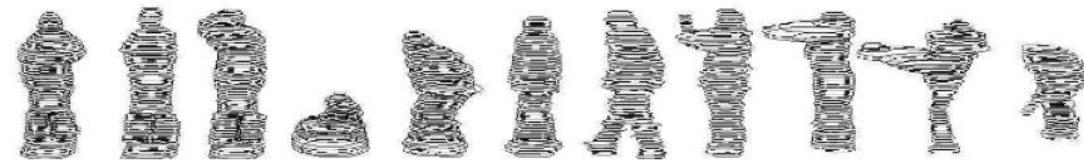
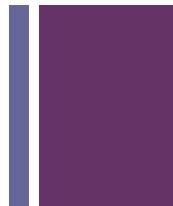


Silindir dağılımı





Daire olarak izdüşüm





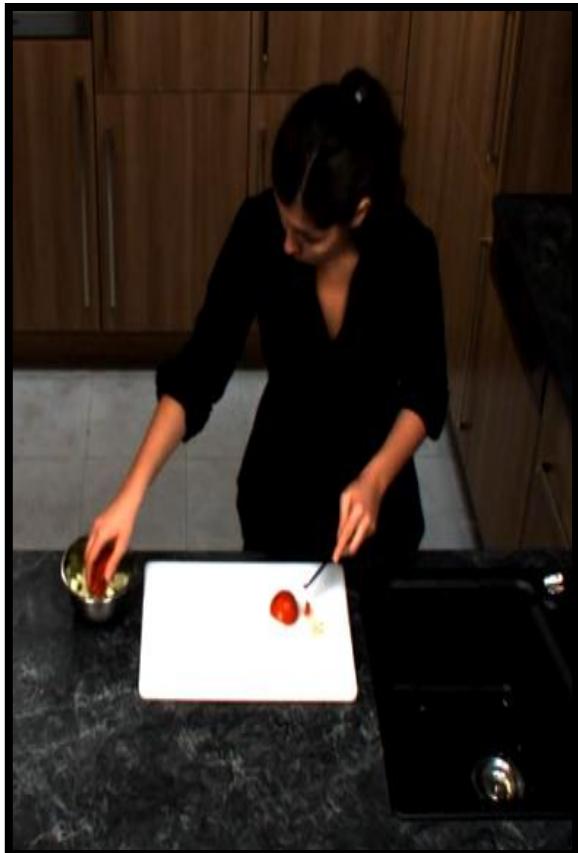
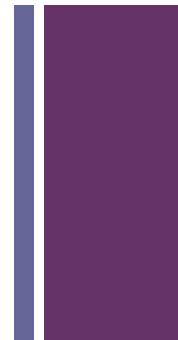
Sağlık uygulamaları

Pinar Duygulu, Eylul 2016, Ankara



Yemek yapma

Yüksek sınıf içi farklılık





Düşük sınıflarası farklılık

Cut apart, cut ends. cut slices, cut stripes, cut dice

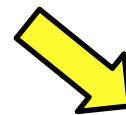
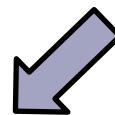


+

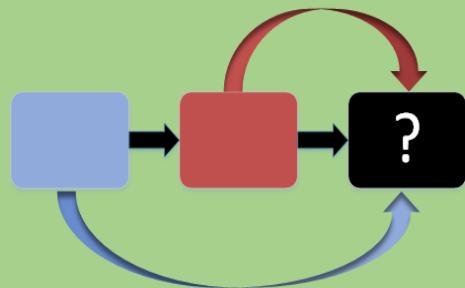
Çözüm

$$y = \operatorname{argmax}_i P(c_i | x)$$

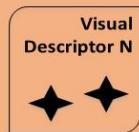
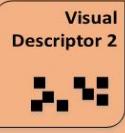
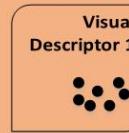
$$P(c_i | x) = \mathbf{T}(c_i) \cdot \mathbf{A}(c_i, x)$$



Temporal Context



Visual Information



+

Tavaya koymak mı kaseye koymak mı?

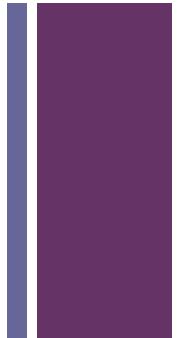


$P(\text{"put in bowl"} \mid \text{"cut dice"}) >$
 $P(\text{"put in pan"} \mid \text{"cut dice"})$



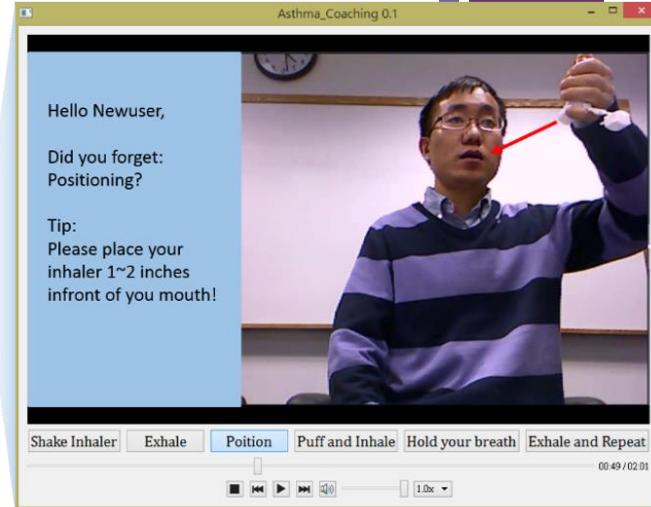
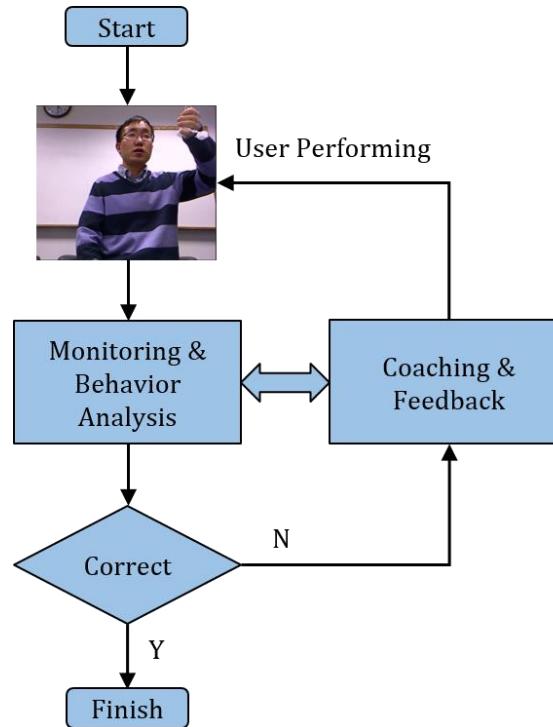
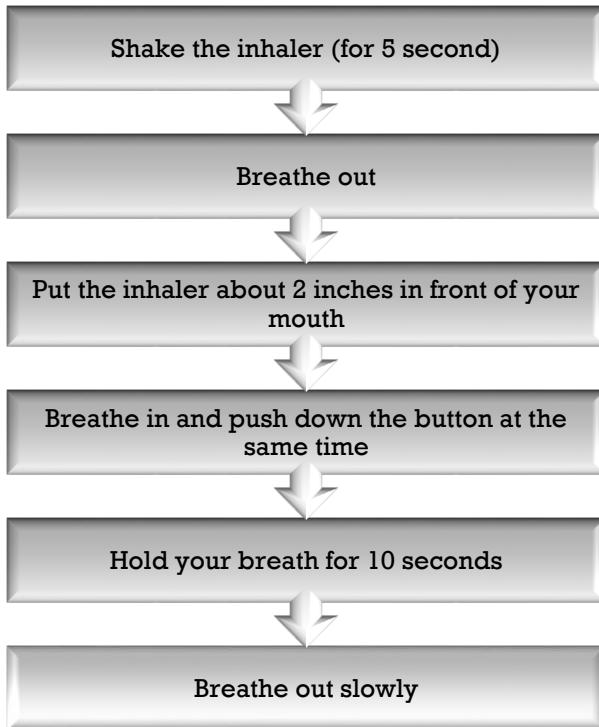
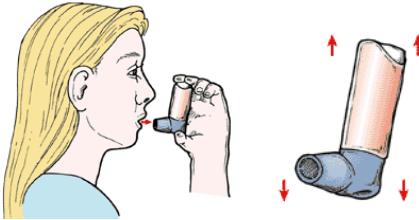
$P(\text{"put in pan"} \mid \text{"spread"}) >$
 $P(\text{"put in bowl"} \mid \text{"spread"})$

+ İlaç ve alet kullanımı



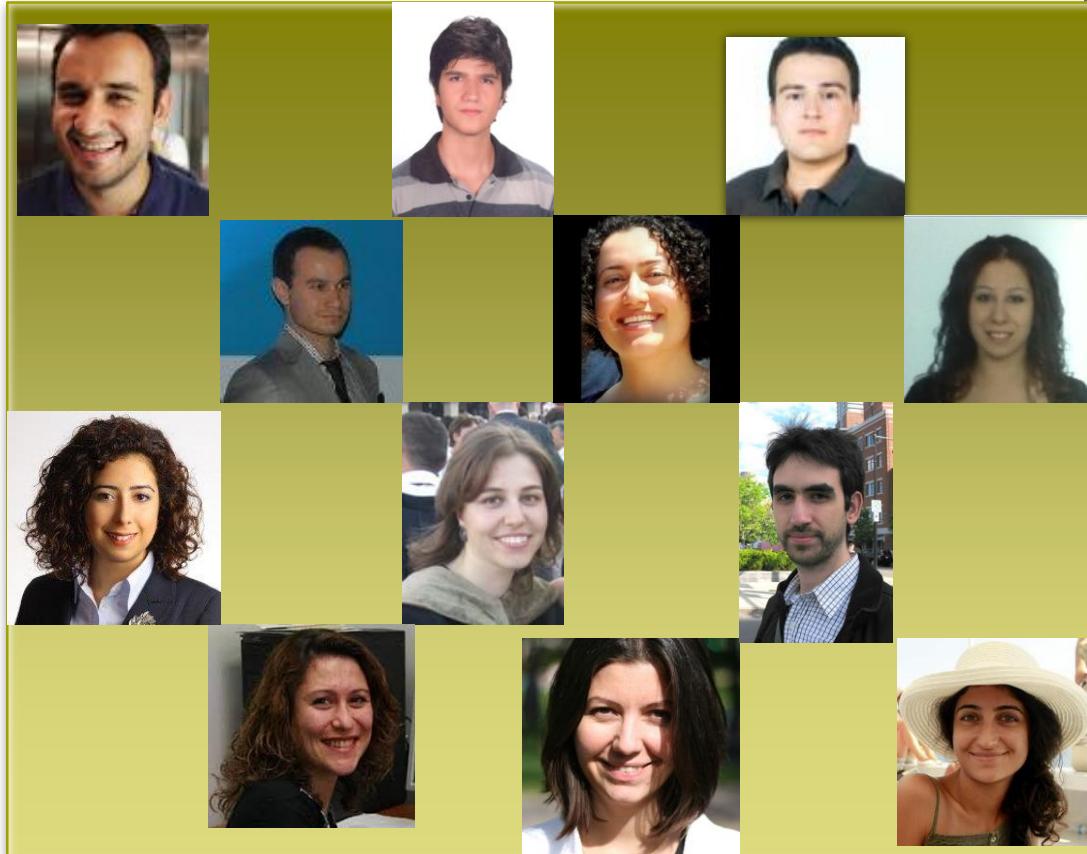


Astım için inhaler kullanımı





- Ahmet Iscen
- Eren Golge
- Anil Armagan
- Sermetcan Baysal
- Fadime Sener
- Hilal Zitouni
- Sare Gul Sevil
- Selen Pehlivan
- Gokberk Cinbis
- Derya Ozkan
- Nazli Ikizler
- Özge Yalçinkaya



+

