컴퓨터지능 및 패턴인식 연구실

(Machine Intelligence & Pattern Recognition Lab.)

곽노준

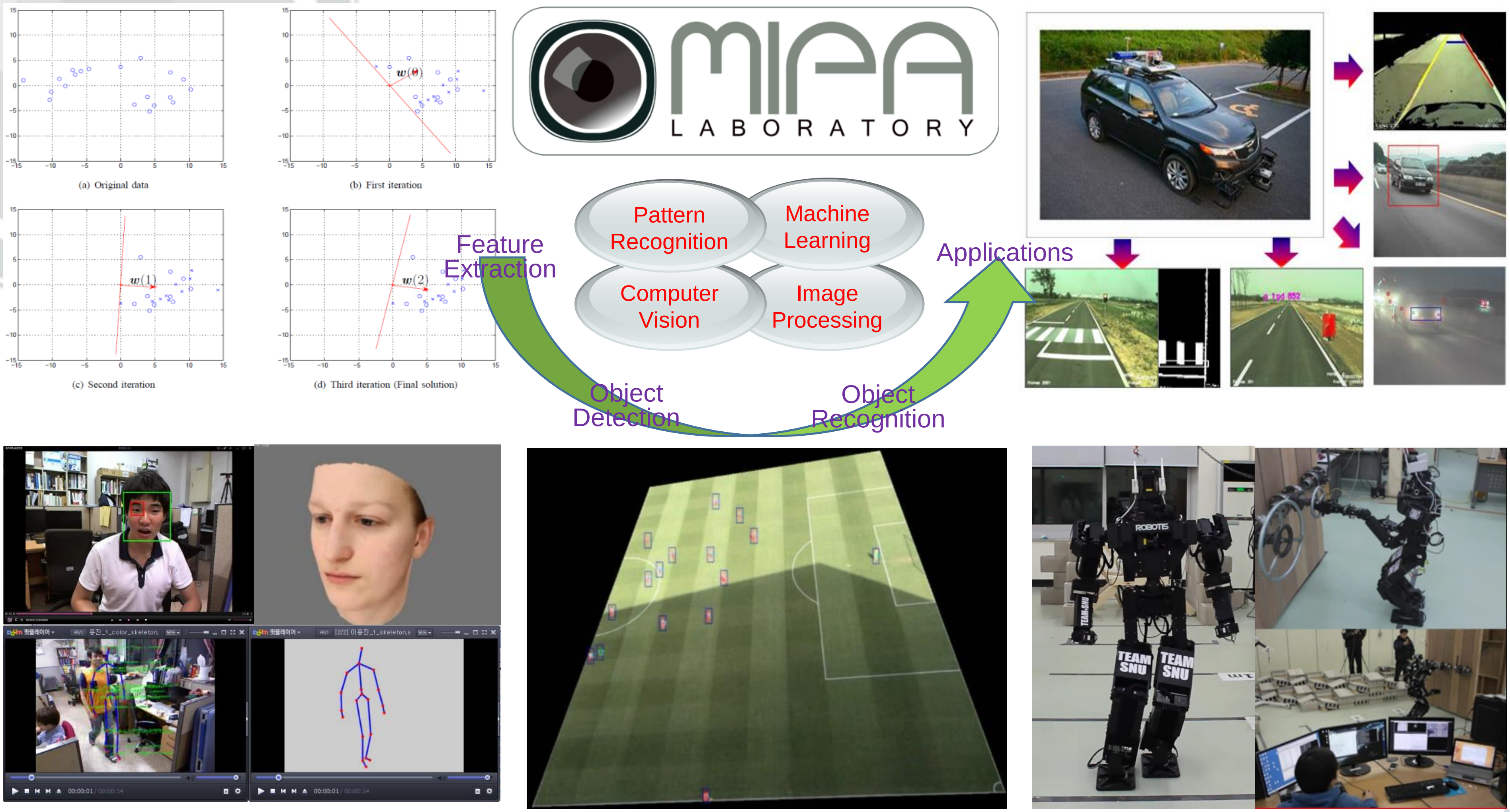
서울대학교 융합과학부

nojunk@snu.ac.kr

http://mipal.snu.ac.kr

MIPA LABORATORY

서울대학교 SEOUL NATIONAL UNIVERSITY



MIPA
LABORATORY

서 울 대 학 교
SEOUL NATIONAL UNIVERSITY

Short CV (곽노준)

- 소속: 서울대학교 융합과학기술대학원
- 연락처
  - 연구실: 031-888-9166
  - 휴대전화: 010-9530-5809
  - e-mail: nojunk@snu.ac.kr
  - Homepage: http://mipal.snu.ac.kr
- 학력
  - 서울대 전기공학부 93-97 (학사)
  - 서울대 전기공학부 97-99 (석사, 최종호 교수)
  - 서울대 전기공학부 99-03 (박사, 최종호 교수)
- 경력
  - 01~02: 방문대학원생 (UIUC, Narendra Ahuja)
  - 03~06: 삼성전자 통신연구소
  - 06~07: 서울대학교 정보기술사업단 BK조교수
  - 07~11: 아주대학교 전자공학부 조교수
  - 11~13: 아주대학교 전자공학과 부교수
  - 13~현재: 서울대학교 융합과학기술대학원 부교수
- 학협회활동
  - IEEE Senior member
  - 정보과학회 인공지능소사이어티 이사, 자동차공학회 전기전자ITS 부문 위원, 제어로봇시스템학회 편집위원, 대한전자공학회 회원

MIPA LABORATORY

서울대학교 SEOUL NATIONAL UNIVERSITY

Short CV (곽노준)

- 학위논문
  - Nojun Kwak, "*Feature selection and extraction based on mutual information for classification*," Ph.D Thesis, Seoul National Univ., Seoul, Korea, Feb. 2003.
  - Nojun Kwak, "*Input feature selection for neural networks using mutual information and Taguchi method*," Mater Thesis, Seoul National Univ., Seoul, Korea, Feb. 1999
- 연구분야
  - **패턴인식, 기계학습, 컴퓨터비전, 영상처리**
    - **특징추출 방법 (PCA/ICA/LDA 등 개선)**
    - **Kernel methods (NPT – nonlinear projection trick)**
  - **딥러닝 알고리즘 및 응용**
    - **CNN 기반 물체 검출 알고리즘**
    - **CNN 기반 Human pose estimation**
    - **CNN 기반 segmentation**
    - **CNN + autoencoder 기반 image enhancement**
    - **CNN + RNN 기반 VQA (Visual Q&A) / TQA (textbook Q&A)**
    - **RNN (LSTM/GRU) 기반 video 분석**
    - **CNN 성능향상 알고리즘 (BCN – broadcasting convolutional network)**
    - **CNN 경량화**
    - **Generative model (GAN – generative adversarial network) 연구**
    - **RL (Reinforcement Learning) 알고리즘**

MIPA LABORATORY

서울대학교 SEOUL NATIONAL UNIVERSITY

Short CV (곽노준)

- 수행 연구과제
  - 얼굴인식, 시선검출 (삼성전자 08, 수자원공사 10~11, 현대차 10~11, SKT 16)
  - **물체 검출** (만도 11, 와이즈오토모티브 09, 산업부 16~18, 과기정통부 17~21, 삼성전자 16~17, SL 16, SKT 14, SKT 15, ETRI 15)
  - **패턴인식/기계학습 이론** (학진 07~08, 08~09, 연구재단 10~13, 14~17)
  - 차선검출 (현대차 16, 항우연/파인디지털 13)
  - 문자인식 (듀얼아이 11, 중기청 12~13)
  - 악보인식 (삼성전자 13)
  - 운동선수 인식/트래킹 (티브이로직/문체부 12~13)
  - 공연용 3차원 영상 복원 (이지위드 13~14)
  - 무인자동차 경진대회 (현대차 08~10, 11~12, 지경부 12~13)
  - DARPA Robotics Challenge (산업부 14~15)
  - 반도체 패키지 납땜볼의 3D depth 측정 (삼성전자 14~15)
  - Target tracking (삼성탈레스 15)

Short CV (곽노준)

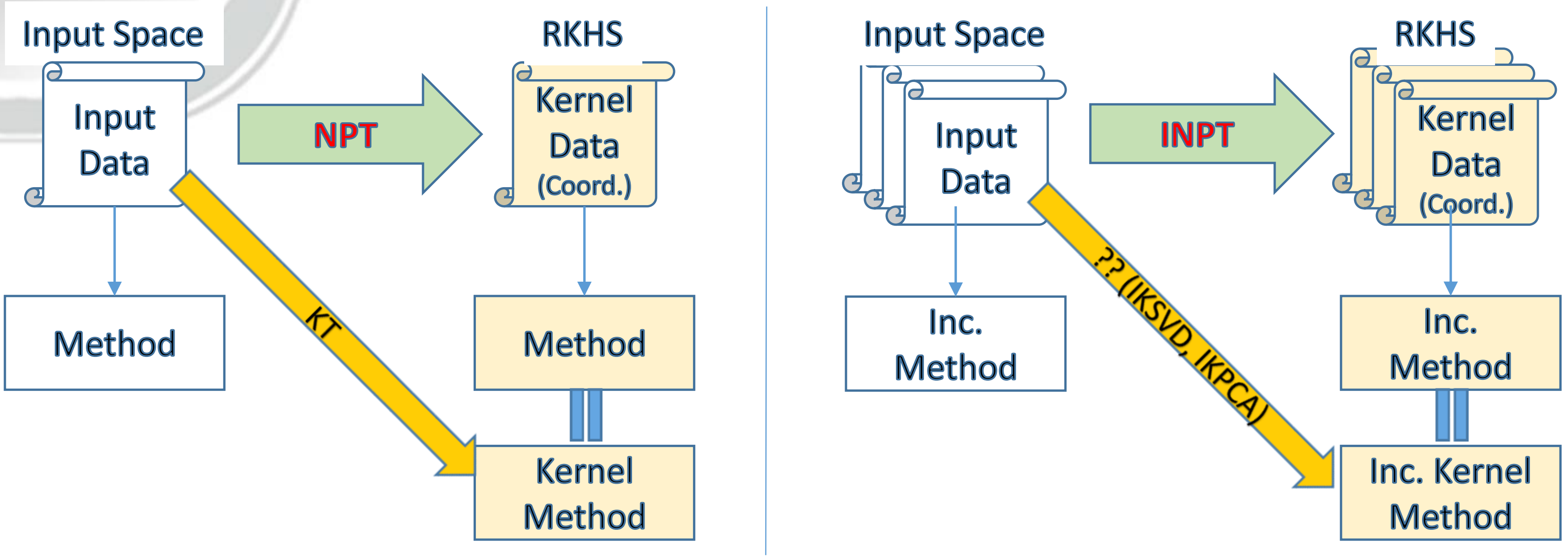
- 논문/특허
  - 대표 논문 (http://scholar.google.com 참조)
    - (SCI) Nojun Kwak, "Principal component analysis based on L1 norm maximization", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 9, pp. 1672 – 1680, Sep. 2008. (**320회 인용**)
    - (SCI) Nojun Kwak and Chong-Ho Choi, "Input feature selection for classification problems," IEEE Transactions on Neural Networks, vol. 13, no. 1, pp. 143 – 159, Jan. 2002. (**700회 인용**)
    - (SCI) Nojun Kwak and Chong-Ho Choi, "Input feature selection by mutual information based on Parzen window," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 12, pp. 1667-1671, Dec. 2002. (**490회 인용**)
  - 특허 현황 (http://www.kipris.or.kr 참조)
    - 국내외 특허 **30여건**

Short CV (곽노준)

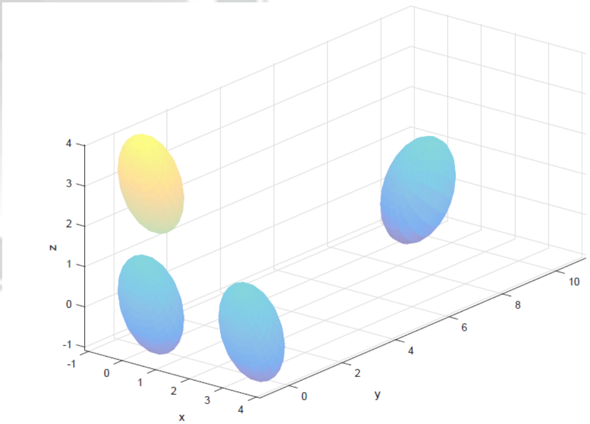
- 수상/기타
  - 2009 서봉 연구장학재단 우수연구자 선정 (SL 그룹) – 3000만원 연구기금
  - Marquis Who's Who in the World, Cambridge Blue Book of Foremost International Scientists 등 등재
  - Textbook Question Answering Challenge (CVPR 2017 Workshop, 4th both in Text Question and Diagram Question)
  - Charades Action Recognition Challenge (CVPR 2017 Workshop, 5th place)
  - Kaggle Data Science Bowl 2017 Challenge (35th place among 1,972 teams)
  - Imagenet Challenge (localization) (ICCV 2015 Workshop, Poster Presentation)
  - ISBI 2015 Challenge on Automatic Polyp Detection in Colonoscopy Videos (3rd place)
  - ChaLearn Looking at People - Cultural Event Classification Challenge (CVPR 2015 Workshop, 3rd place)
  - DARPA Robotics Challenge Finals 2015 (12th place among 25 teams, co-leaded Team-SNU)
  - Korean Autonomous Vehicle Contest 2013 (4th place among 10 final teams, co-leaded Team-ADAM)

기계학습 이론

Nonlinear Projection Trick



Feature extraction with Generalized Mean

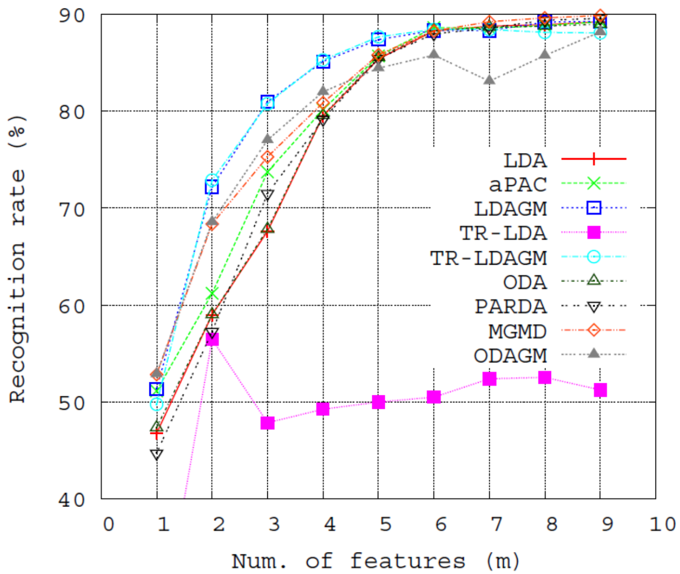


$$\mathbf{W}_{LG} = \underset{\mathbf{W}^T \mathbf{S}_W \mathbf{W} = \mathbf{I}}{\arg\max} \; J_{LG}(\mathbf{W})$$

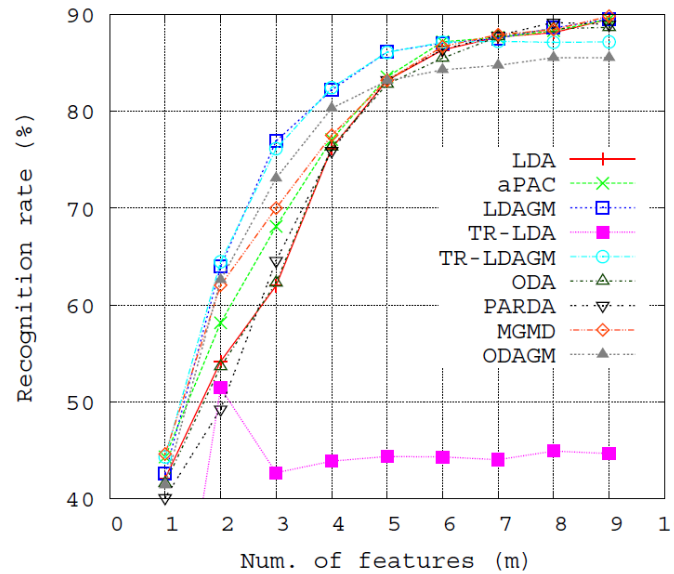
$$= \underset{\mathbf{W}^T \mathbf{S}_W \mathbf{W} = \mathbf{I}}{\arg\max} \left(\frac{1}{q} \sum_{i=1}^{C-1} \sum_{j=i+1}^{C} q_i q_j \left[E_{ij}(\mathbf{W}) \right]^p \right)^{1/p}$$

$$E_{ij}(\mathbf{W}) = tr \left[\mathbf{W}^T \mathbf{M}_{ij} \mathbf{W} \right],$$

$$= (\mathbf{m}_i - \mathbf{m}_j)^T \mathbf{W}\mathbf{W}^T (\mathbf{m}_i - \mathbf{m}_j)$$



(a) Results of ML classifier on USPS data set

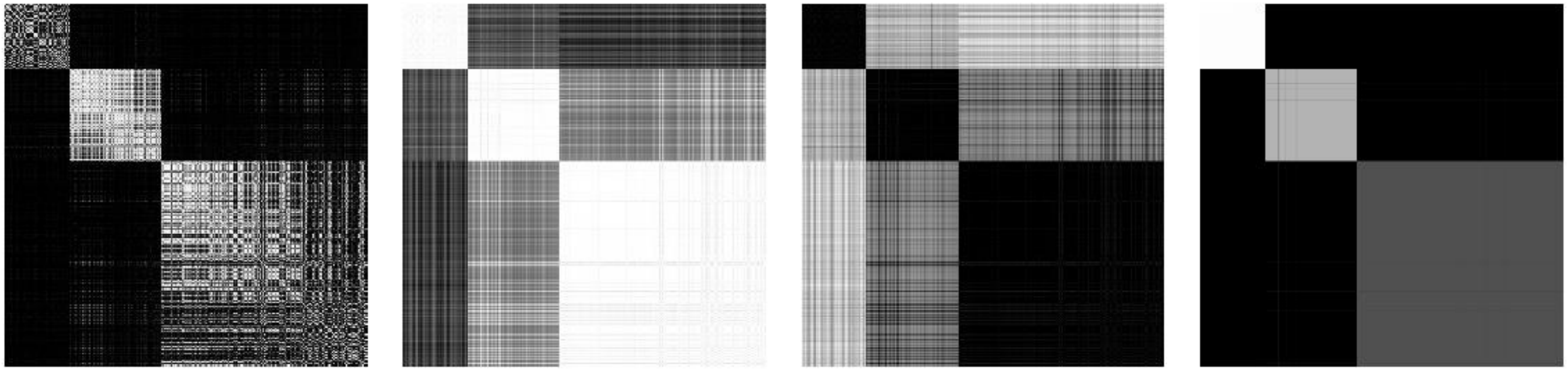


(b) Results of NN classifier on USPS data set

MIPA LABORATORY

서울대학교 SEOUL NATIONAL UNIVERSITY

Membership Representation for LRR & SSC



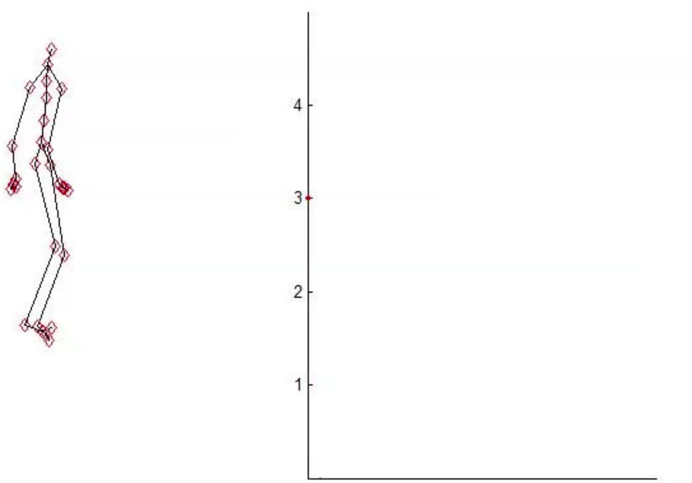
(a) Input (latent matrix $\mathbf{W}$ = output of SSC/LRR)

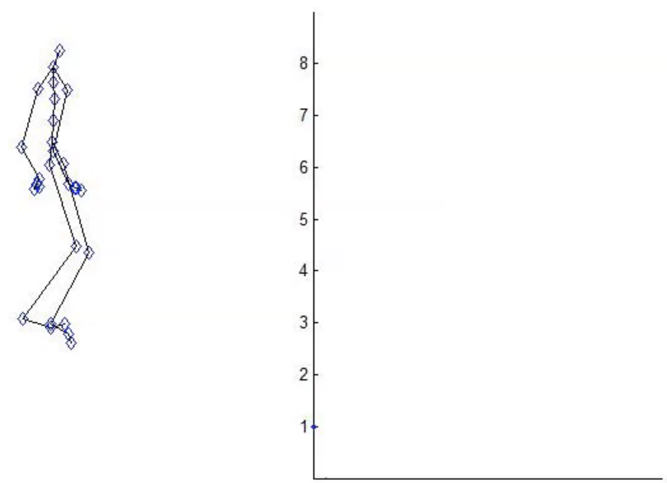
(b) Output of the 1st stage: Approximation of membership matrix ($\check{\mathbf{M}}$)

(c) Dissimilarity matrix ($\mathbf{H} = \mathbf{1}\mathbf{1}^T - \check{\mathbf{M}}$)

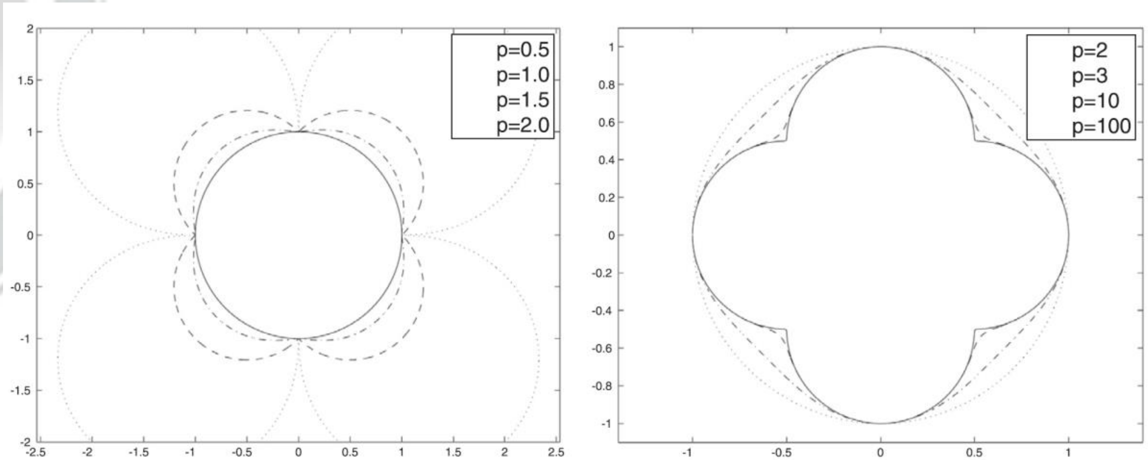
(d) Output of the 2nd stage: Approximation of normalized membership matrix ($\check{\mathbf{F}}$)

Figure 1. An example of the intermediate results of the proposed algorithm (Hopkins155, $K = 3$, $n = 336$).





ICA by Lp-norm optimization



(a) Super-Gaussian sources ($p < 2$) (b) Sub-Gaussian sources ($p > 2$)

ICA by l_p-norm optimization.

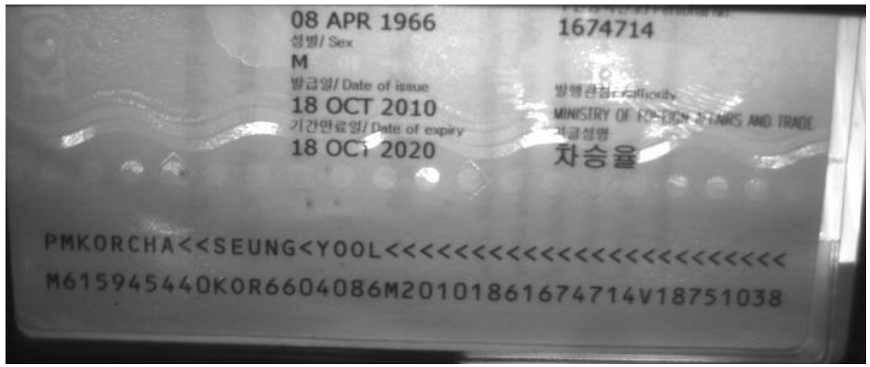
	minimization	maximization
$p < 2$	super-G. (see Section 4 and 5)	sub-G. (PCA-L1 [23] or Lp-PCA[24])
	e.g. $\mathbf{w}^* = \arg\min_{\mathbf{w}:\mathbf{w}^T\mathbf{w}=1} \sum_{i=1}^{n} \lvert\mathbf{w}^T\mathbf{x}_i\rvert^1$	e.g. $\mathbf{w}^* = \arg\max_{\mathbf{w}:\mathbf{w}^T\mathbf{w}=1} \sum_{i=1}^{n} \lvert\mathbf{w}^T\mathbf{x}_i\rvert^1$
$p > 2$	sub-G. (see Section 4 and 5)	super-G. (Lp-PCA [24])
	e.g. $\mathbf{w}^* = \arg\min_{\mathbf{w}:\mathbf{w}^T\mathbf{w}=1} \sum_{i=1}^{n} \lvert\mathbf{w}^T\mathbf{x}_i\rvert^3$	e.g. $\mathbf{w}^* = \arg\max_{\mathbf{w}:\mathbf{w}^T\mathbf{w}=1} \sum_{i=1}^{n} \lvert\mathbf{w}^T\mathbf{x}_i\rvert^3$

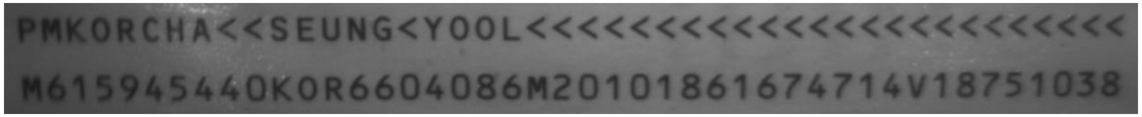
MIPA LABORATORY

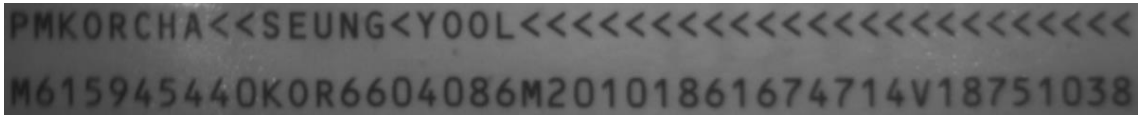
서울대학교 SEOUL NATIONAL UNIVERSITY

영상처리/컴퓨터비전

OCR


08 APR 1966
성별/ Sex
M
발급일/ Date of issue
18 OCT 2010
기간만료일/ Date of expiry
18 OCT 2020
1674714
차승율
PMKORCHA<<SEUNG<YOOL<<<<<<<<<<<<<<<<<<<<<<<<
M615945440KOR6604086M2010186167471 4V18751038


PMKORCHA<<SEUNG<YOOL<<<<<<<<<<<<<<<<<<<<<<<<
M615945440KOR6604086M20101861674714V18751038


PMKORCHA<<SEUNG<YOOL<<<<<<<<<<<<<<<<<<<<<<<<
M615945440KOR6604086M20101861674714V18751038


PMKORCHA<<SEUNG<YOOL<<<<<<<<<<<<<<<<<<<<<<<<
M615945440KOR6604086M20101861674714V18751038

Illumination robust optical flow

(c) TV-L1 gray (d) Spherical RGB (e) MDP-flow (f) Our method

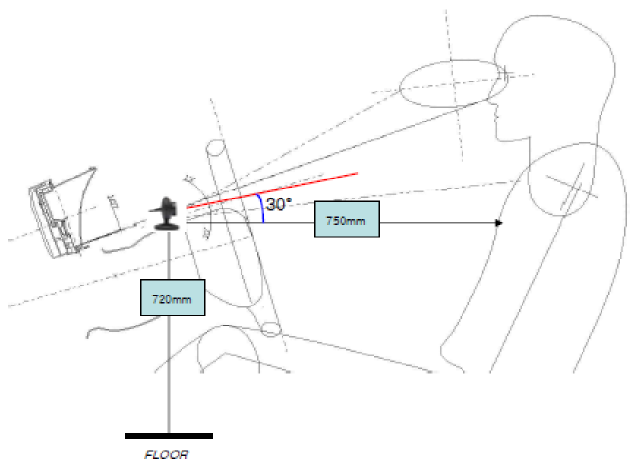
Lane detection

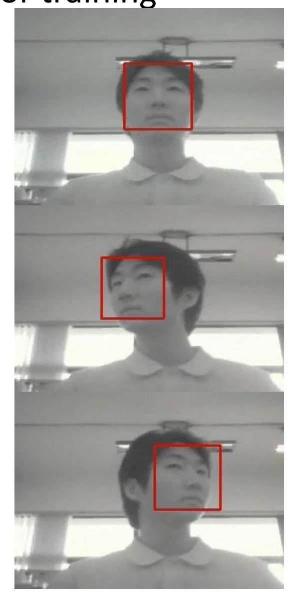
MIPA LABORATORY

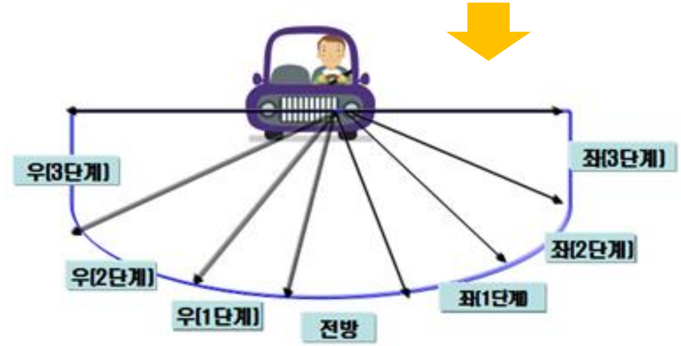
서울대학교
SEOUL NATIONAL UNIVERSITY

Driver's gaze recognition

- Jae Hyun Oh and Nojun Kwak, "Recognition of a driver's gaze for vehicle headlamp control", IEEE Trans. on Vehicular Technology, vol. 61, no. 5, pp. 2008–2017, June 2012.
  - **2DLDAr** (2D Linear Discriminant Analysis for regression) was used for training

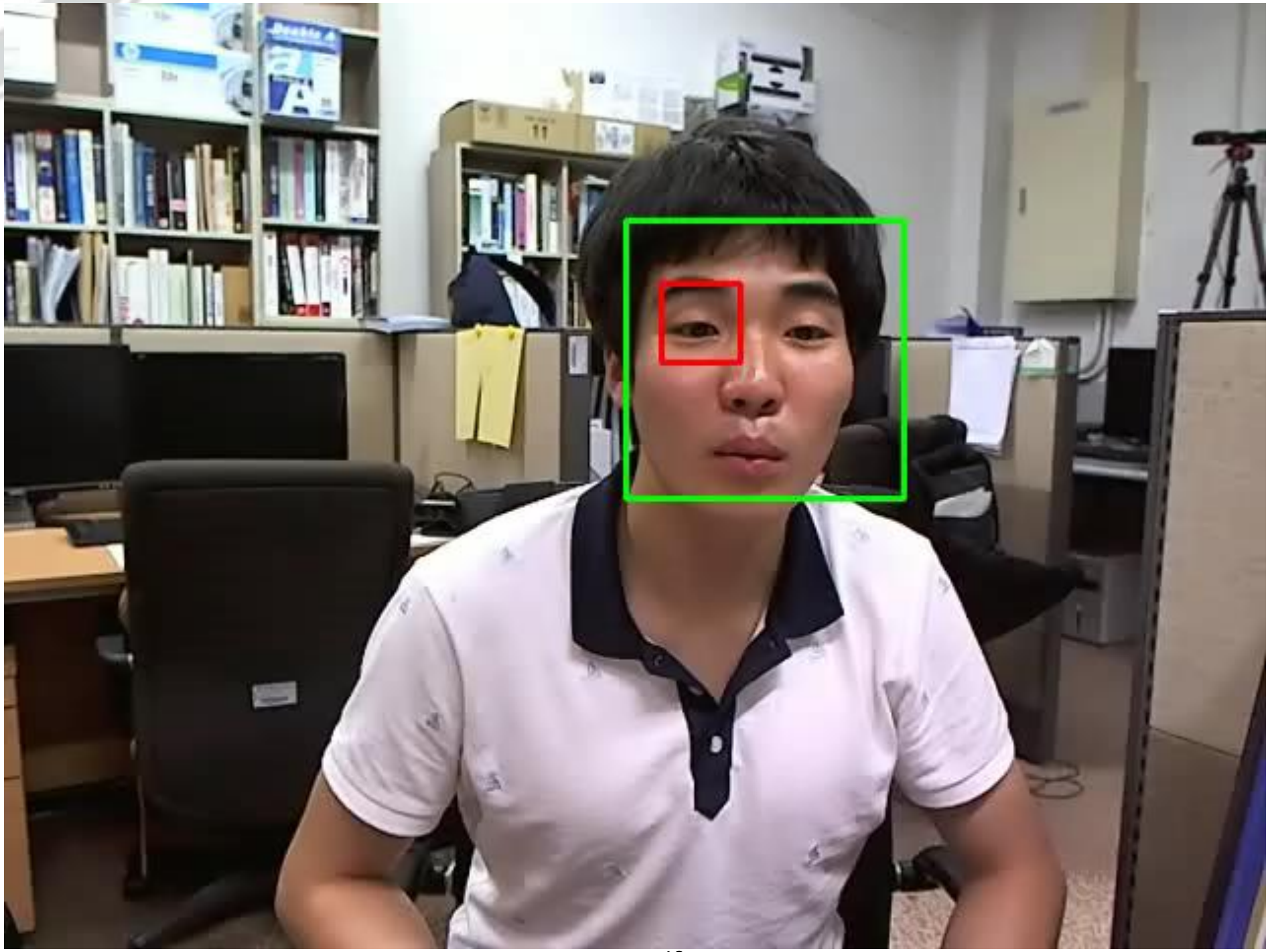






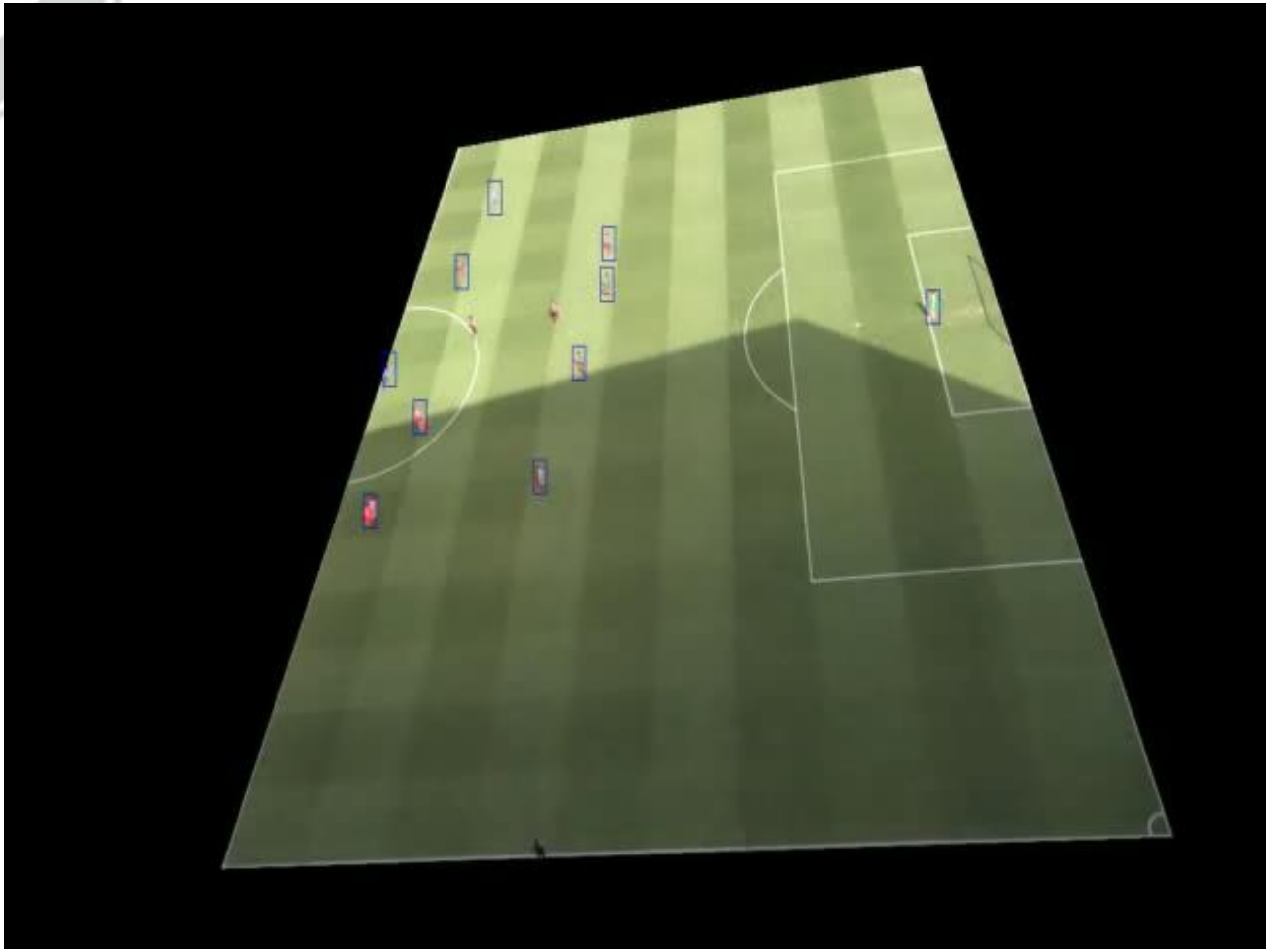
Drowsy driver monitoring

- Speeded up performance by incorporating
  V-J detector + Template Matching



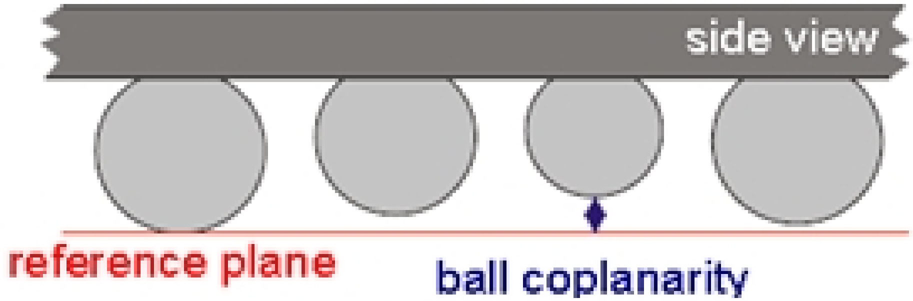
Athlete detection & tracking

- Because of changing environment (illumination), detection is challenging.

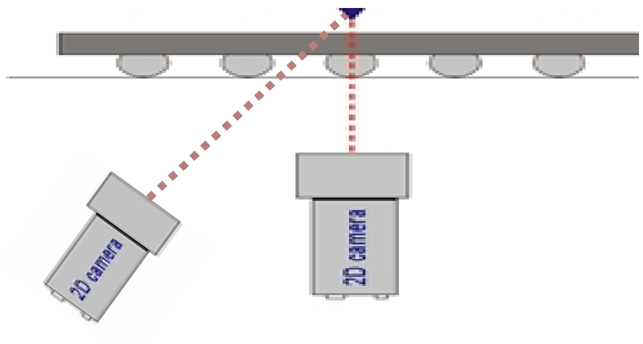


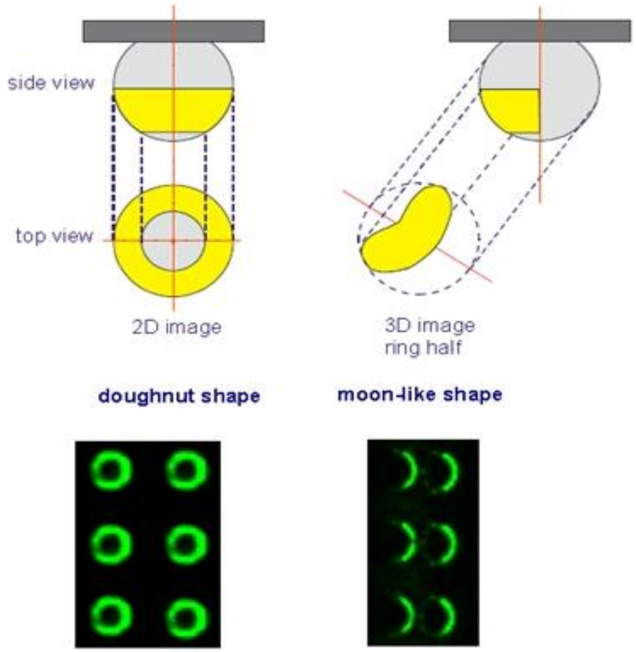
스테레오 비전을 이용한 납볼의 3D depth측정

- Coplanarity



- 카메라 Setting





Non-Rigid Structure from Motion (2D → 3D)

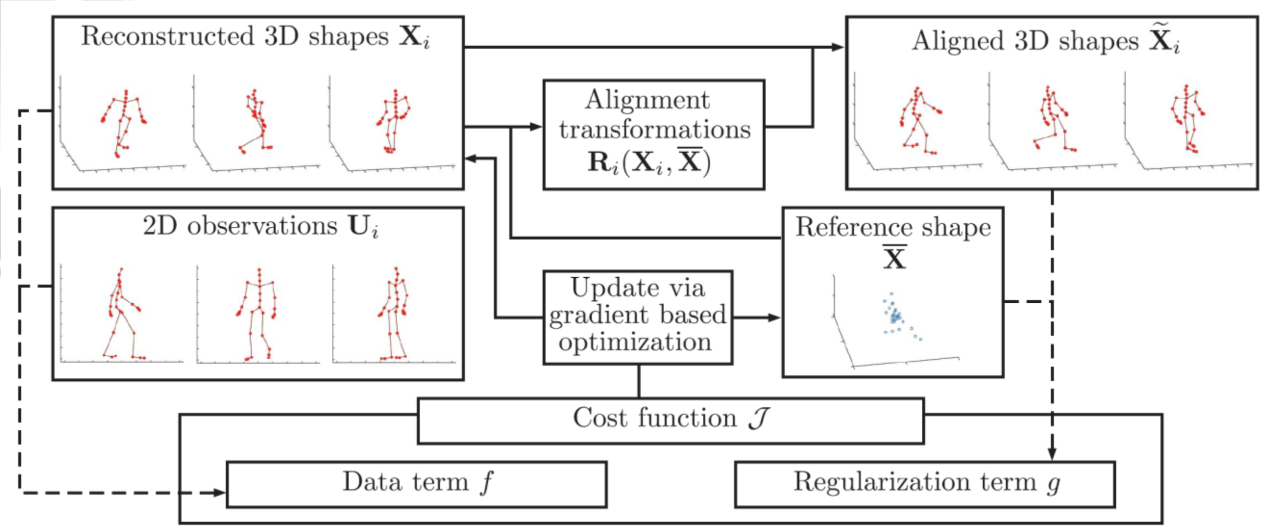
Supplementary to Scalable Non-Rigid Structure from Motion with Procrustean Constraint by a Proximal Gradient Method

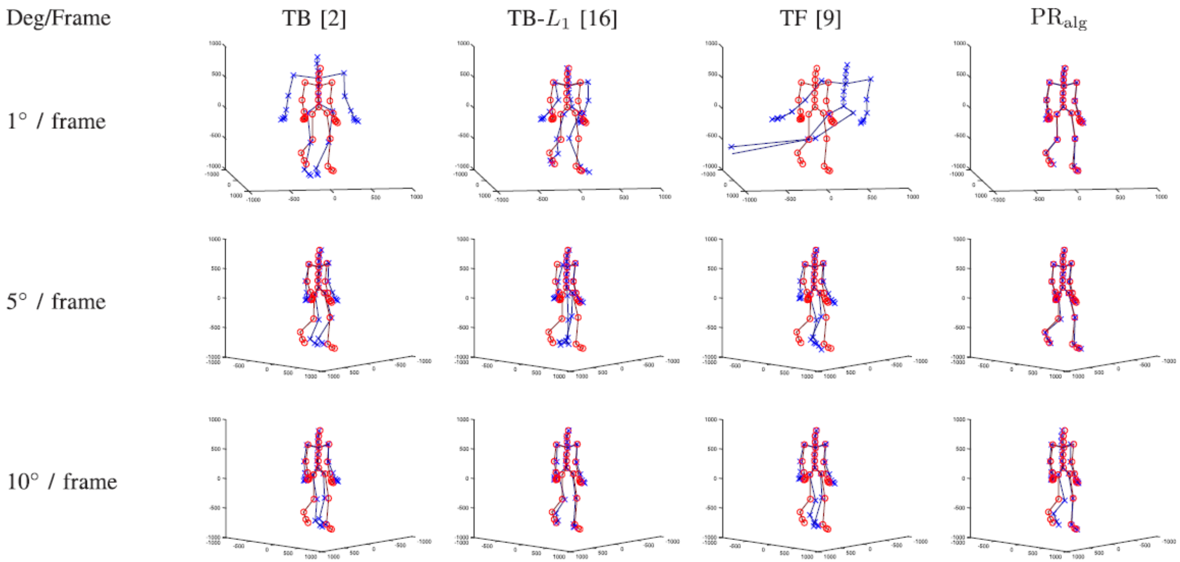
ACCV Submission ID #867

MIPA LABORATORY

서울대학교
SEOUL NATIONAL UNIVERSITY

Procrustean Regression: NRSfM



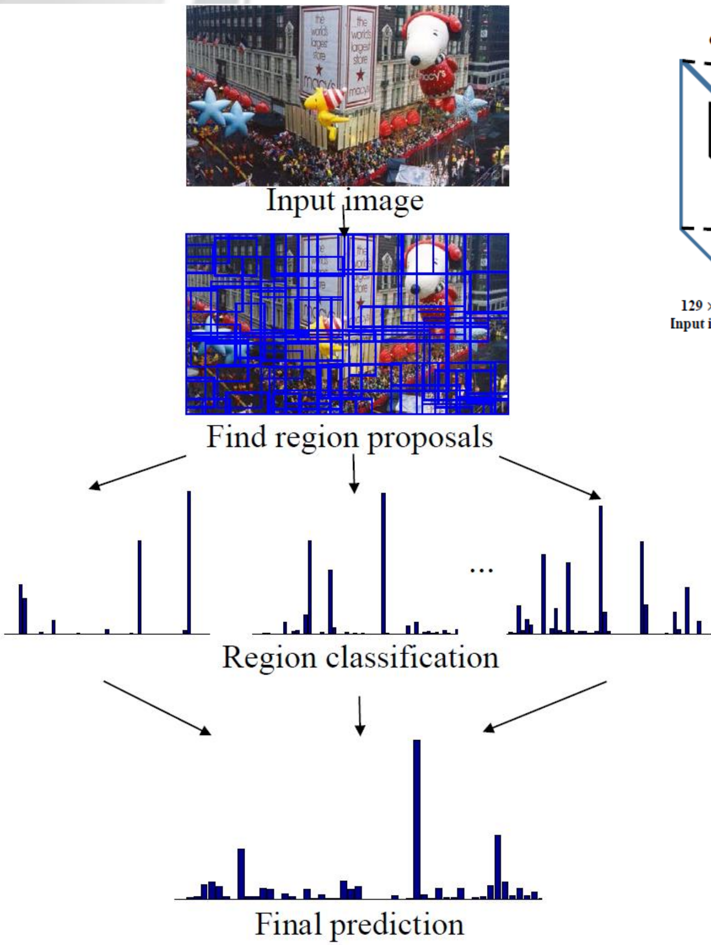


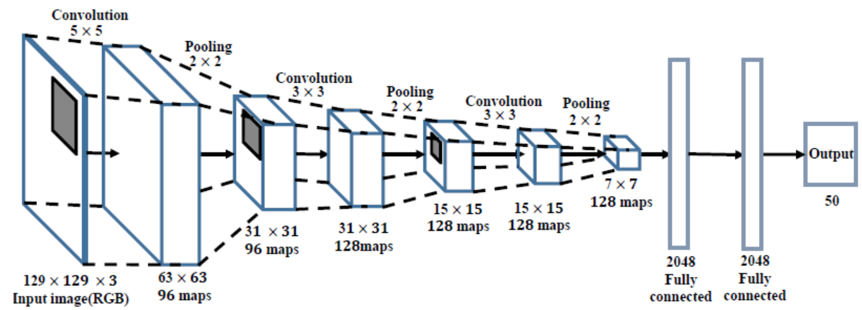
MIPA LABORATORY

서 울 대 학 교
SEOUL NATIONAL UNIVERSITY

Deep Learning

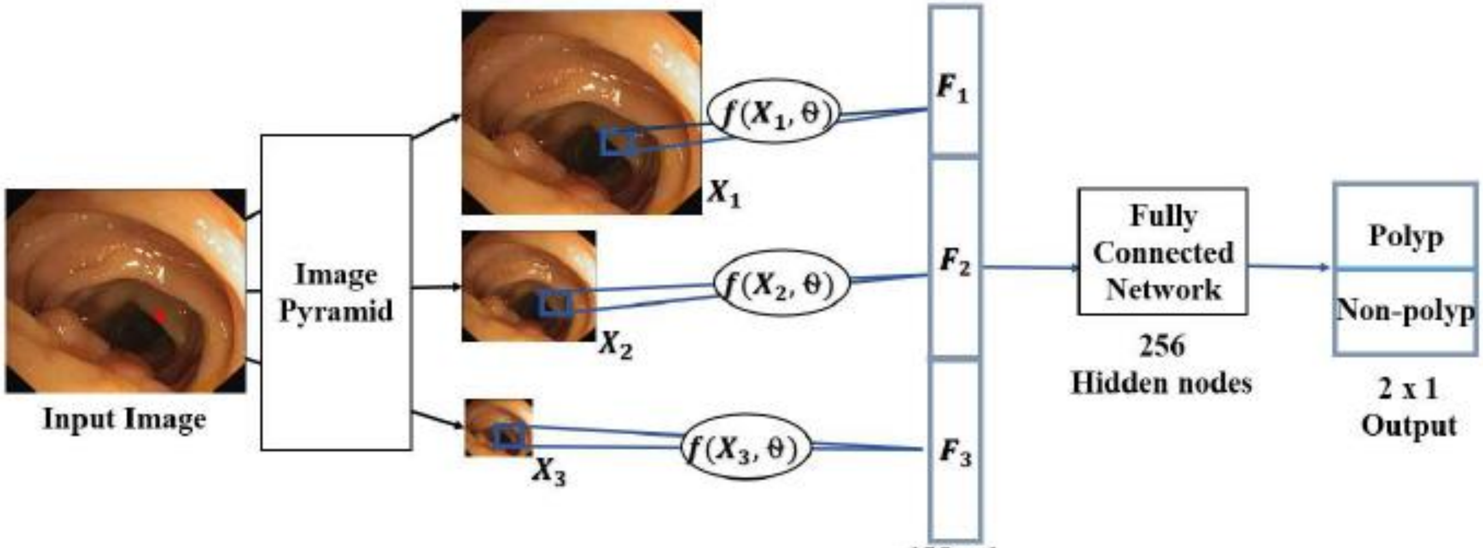
Cultural event recognition

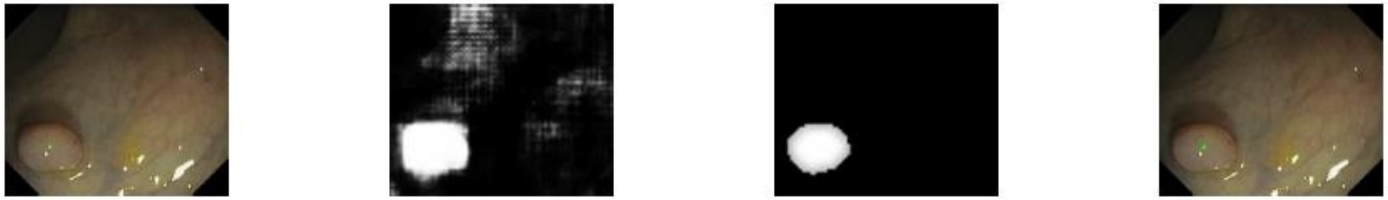




Team name	mAP
MMLAB	0.855
UPC-STP	0.767
MIPAL_SNU	**0.735**
SBU_CS	0.610
MasterBlaster	0.582
Nyx	0.319

Polyp detection in Colonoscopy Videos

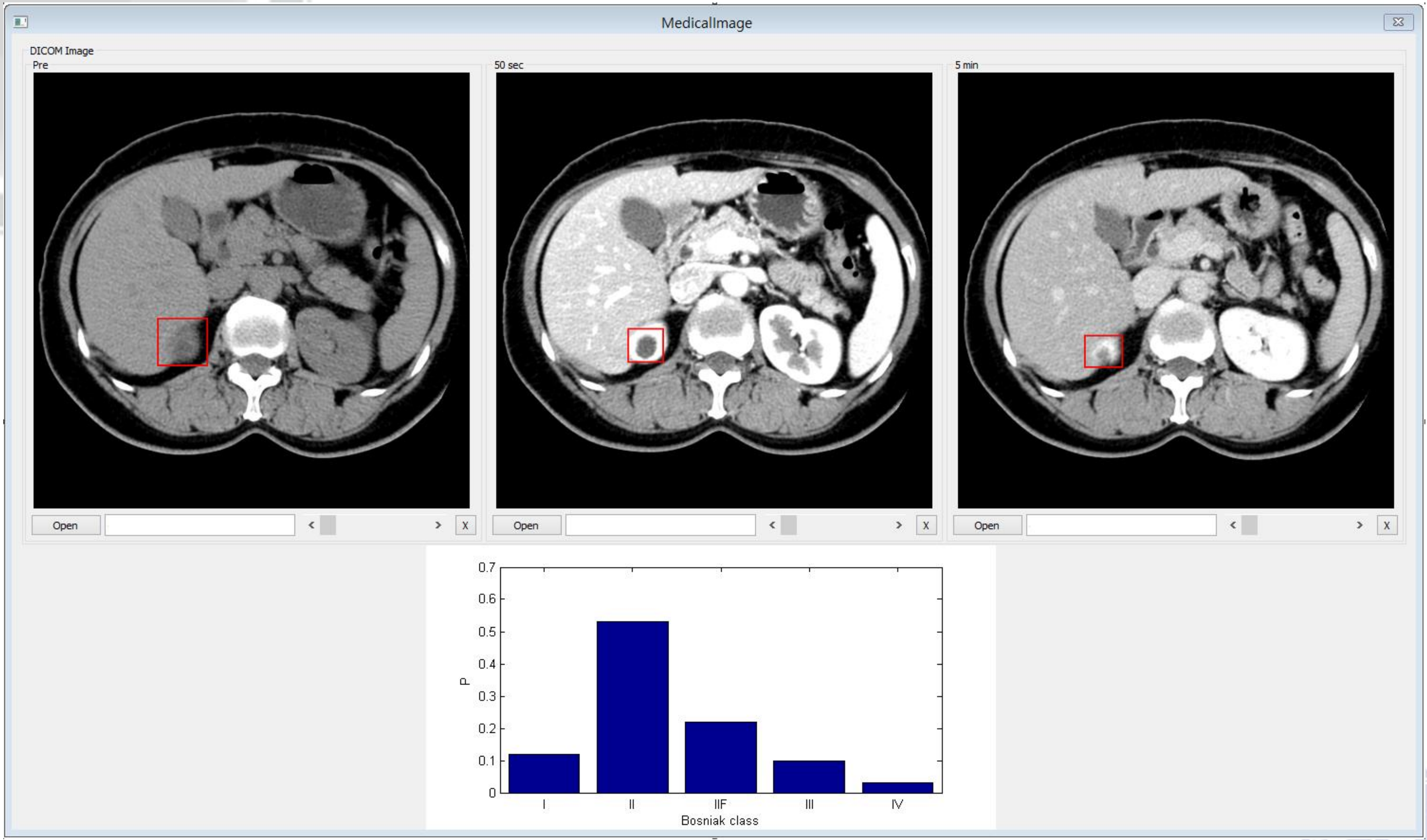




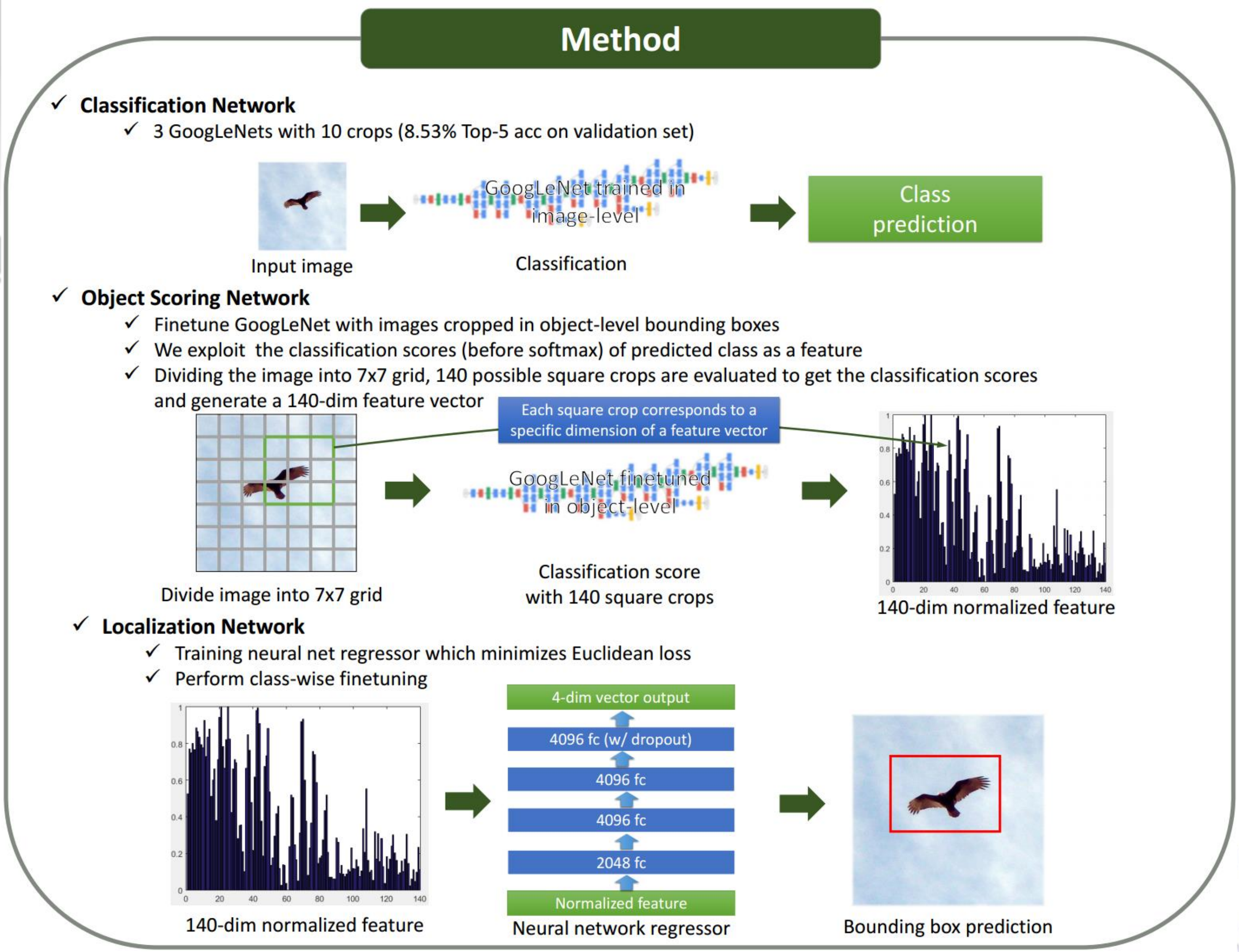
(a) (b) (c) (d)

Figure 3: An example of polyp detection result. (a) Input image. (b) Grayscale probability map. Black is mapped to probability of 0 and white is mapped to probability of 1. (c) Smoothed probability map after post-processing (d) Final result. The location of the polyp is marked with a green dot.

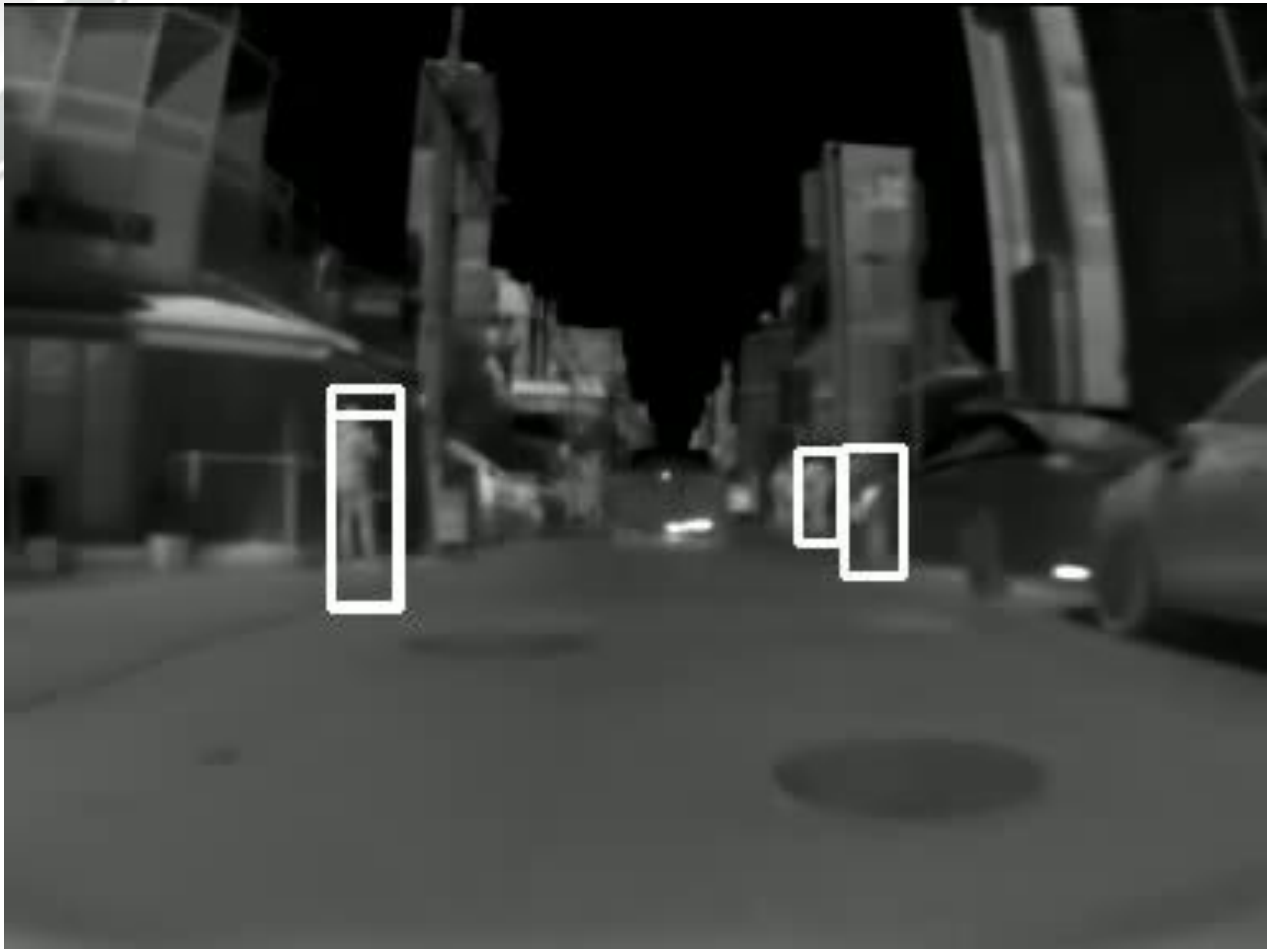
Bosniak Classification in renal cystic lesion



Imagenet localization



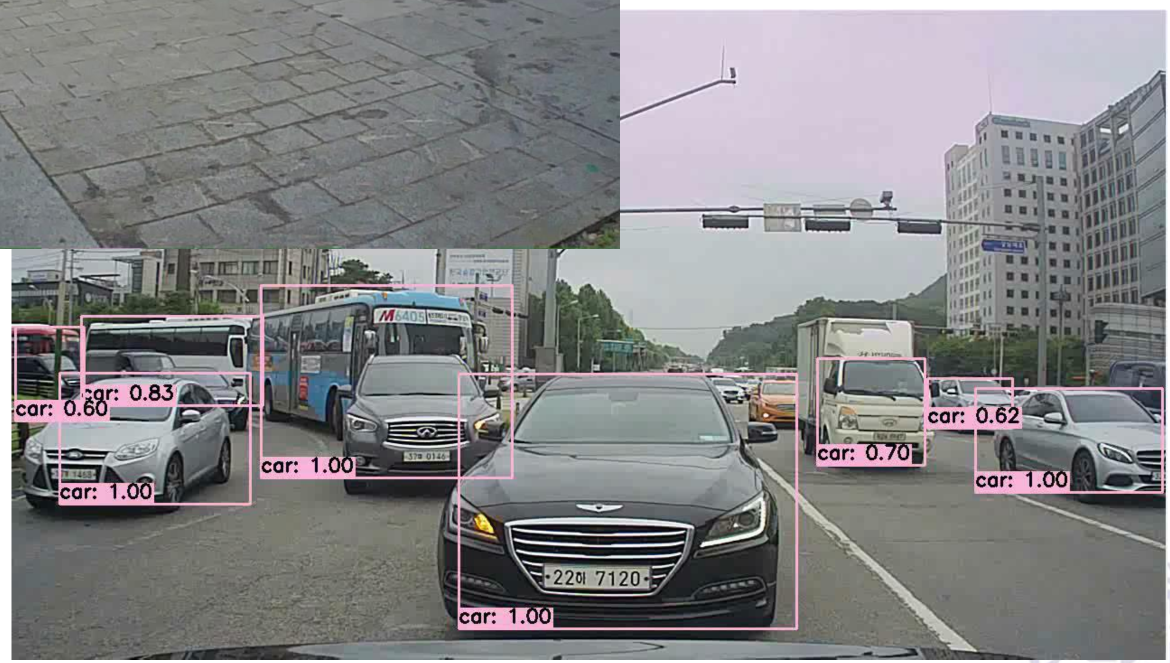
Pedestrian Detection



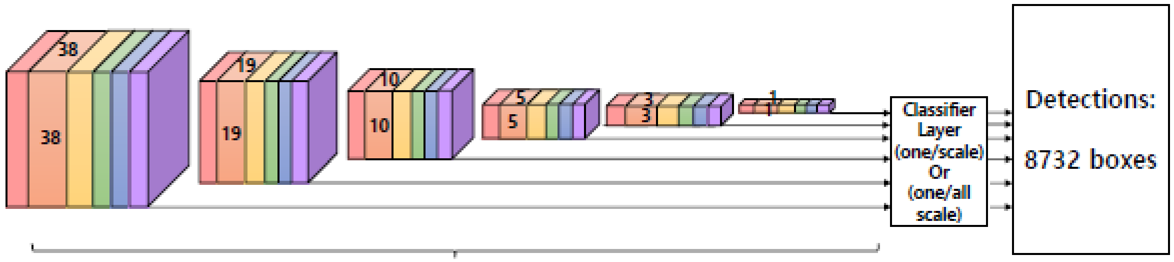
MIPA LABORATORY

서울대학교 SEOUL NATIONAL UNIVERSITY

Object Detection

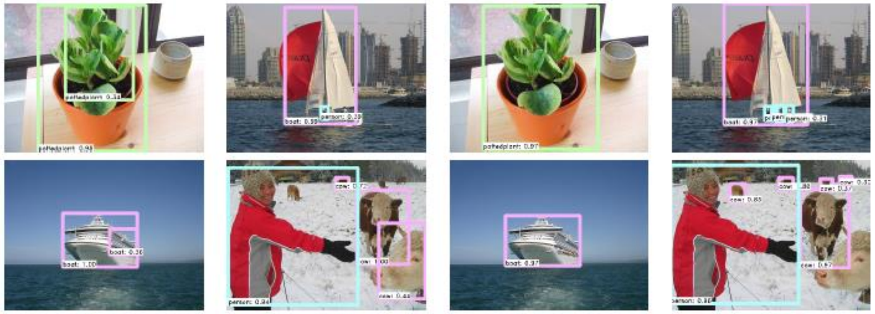


Object Detection (Rainbow SSD)



	Input	Train	Test	mAP	FPS
YOLO[14]	448	VOC2007+2012	2007	63.4	45
YOLOv2[13]	416	VOC2007+2012	2007	76.8	67
YOLOv2 544x544[13]	544	VOC2007+2012	2007	78.6	40
Faster R-CNN[15]		VOC2007+2012	2007	73.2	5
R-FCN (ResNet-101)[8]		VOC2007+2012	2007	**80.5**	5.9
SSD*[11]	300	VOC2007+2012	2007	**77.7**	61.1
DSSD (ResNet-101)[3]	321	VOC2007+2012	2007	78.6	9.5
ISSD*	300	VOC2007+2012	2007	78.1	26.9
ours (SSD pooling)*	300	VOC2007+2012	2007	77.1	48.3
ours (SSD deconvolution)*	300	VOC2007+2012	2007	77.3	39.9
ours (R-SSD)*	300	VOC2007+2012	2007	**78.5**	35.0
ours (R-SSD one classifier (4 boxes))*	300	VOC2007+2012	2007	**76.2**	34.8
ours (R-SSD one classifier (6 boxes))*	300	VOC2007+2012	2007	**77.0**	35.4
SSD*[11]	512	VOC2007+2012	2007	79.8	25.2
DSSD (ResNet-101)[3]	513	VOC2007+2012	2007	81.5	5.5
ours (R-SSD)*	512	VOC2007+2012	2007	**80.8**	16.6

Table 3: VOC2007+2012 training and VOC 2007 test result (* is tested by ourselves)



(a) (b) (c) (d)

Figure 2: Conventional SSD vs. the proposed Rainbow SSD (R-SSD). Boxes with objectness score of 0.3 or higher is drawn: (a) SSD with two boxes for one object; (b) SSD for small objects; (c) R-SSD with one box for one object; (d) R-SSD for small objects

MIPA LABORATORY

서울대학교 SEOUL NATIONAL UNIVERSITY

CNN – Dropout

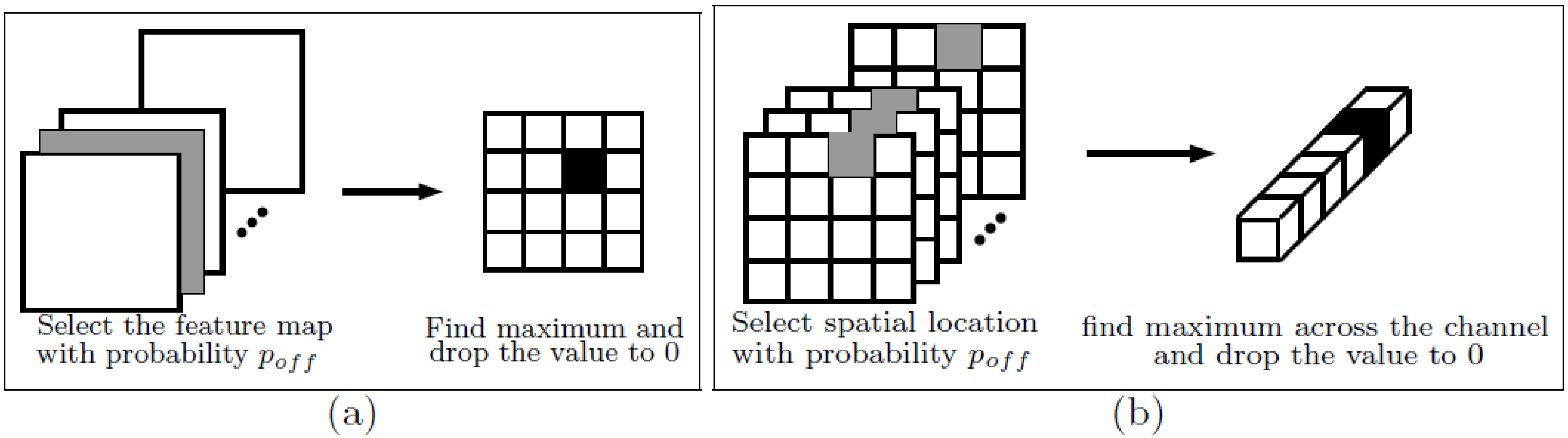
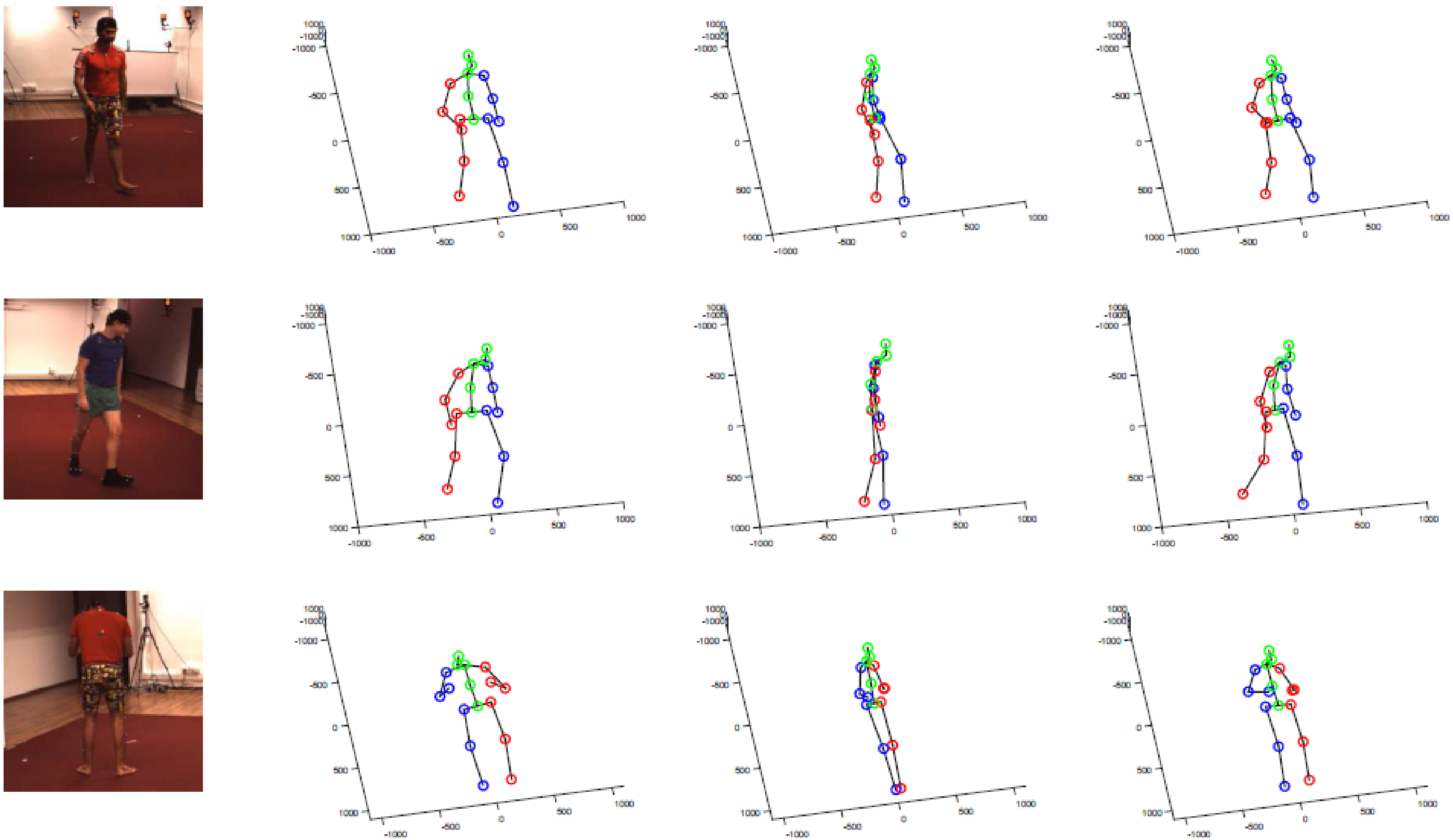


Fig. 3. Illustration of max-drop layer. Two different ways to find maximum value is proposed in this paper: (a) Feature-wise max-drop finds maximum value within each feature map and drops the maximum values with probability p_{off}. (b) Channel-wise max-drop finds maximum value across each channel in the same spatial position and drops the maximum values with probability p_{off}.

MIPA LABORATORY

서울대학교 SEOUL NATIONAL UNIVERSITY

Pose estimation (2D+3D CNNs)



Pose Estimation

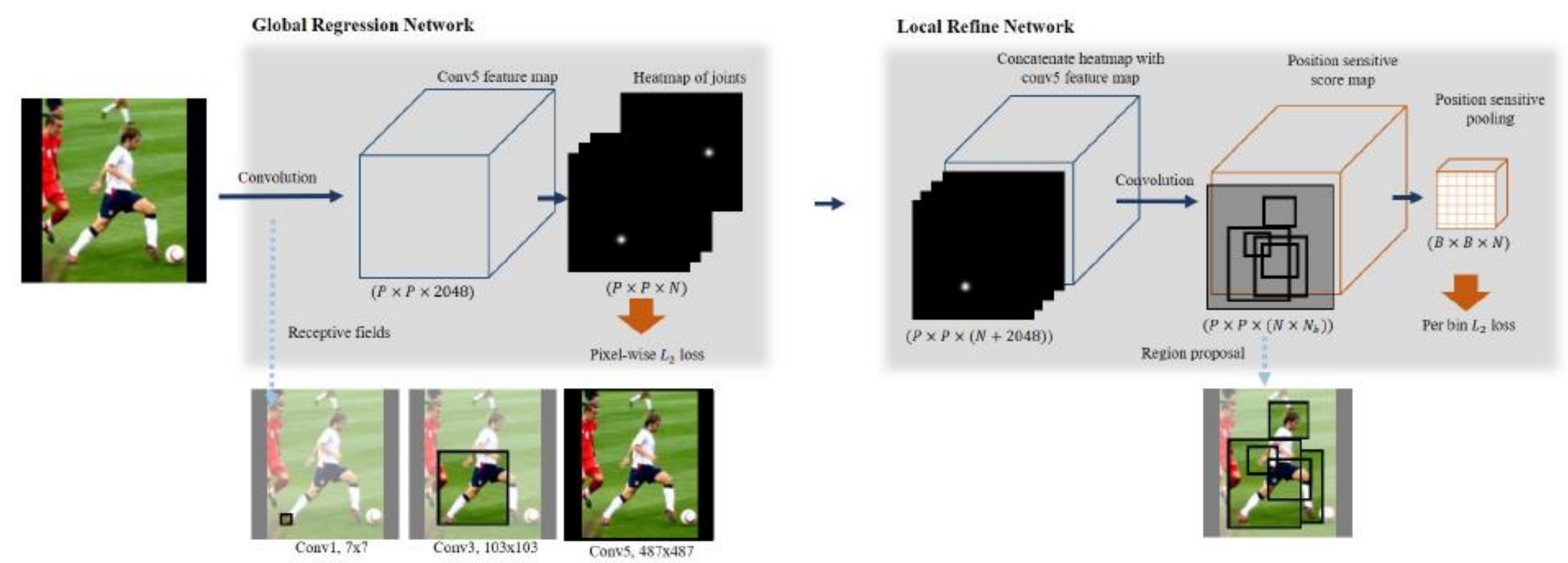


Figure 2. Overall architecture of the proposed method (global-local network). The output of the global network is used as an input to a local network to refine the location using a variety of region proposals. On the left, we show each receptive field of features after the corresponding convolution layer with its size (e.g., 7 × 7 for conv1 layer). On the right, several region proposals are shown.

	Head	Shoulder	Elbow	Wrist	Hip	Knee	Ankle	Total
Local	74.8	63.9	44.7	29.7	66.6	47.9	28.3	50.8
Global	89.3	71.5	58.0	51.0	70.5	66.5	62.5	67.0
Global(14)-local	91.8	76.0	64.7	58.6	76.9	72.9	68.8	72.8
Global(14)-local*	92.3	79.1	69.2	62.9	80.8	76.0	71.5	76.0
Global(28)-local*	96.2	85.4	76.1	71.2	85.7	81.8	76.2	**81.8**
Fan[4] et al. CVPR'15	92.4	75.2	65.3	64.0	75.7	68.3	70.4	73.0
Yang [19] et al. CVPR'16	90.6	78.1	73.8	68.8	74.8	69.9	58.9	73.6

Table 2. PCK-based comparison on LSP. A threshold value was measured at 0.2 (@0.2). The mark * indicates weights from the additional fine-tuning step is used.

Semantic Segmentation

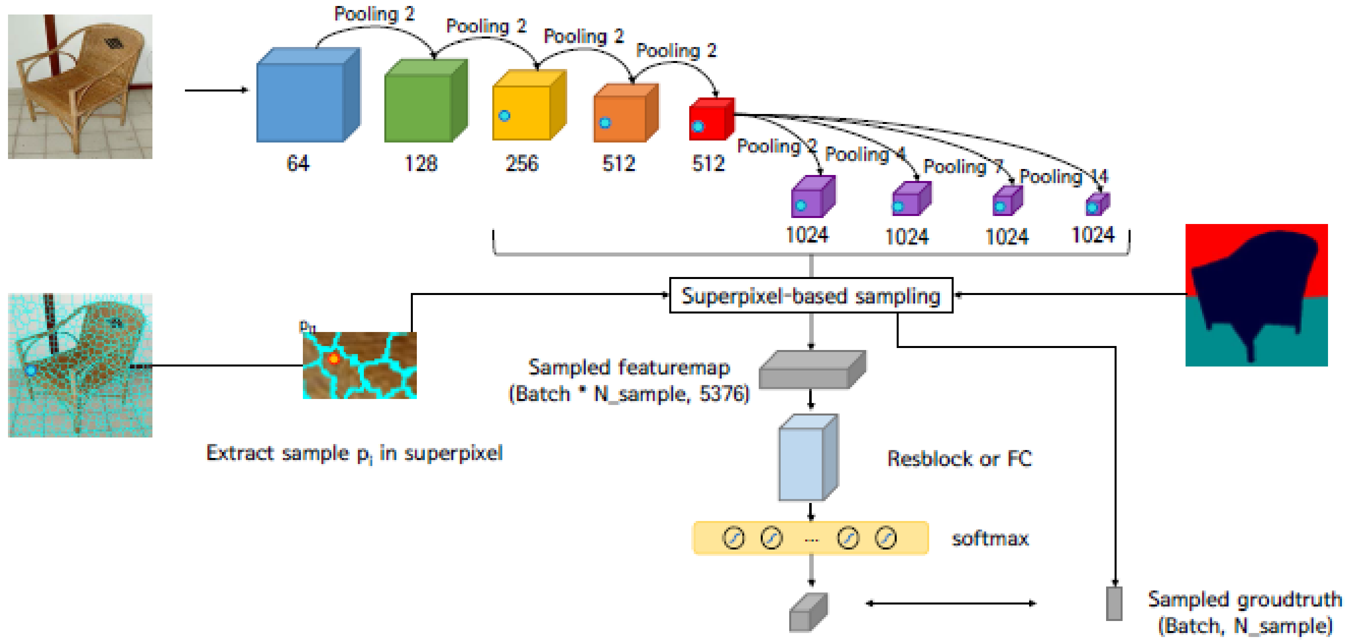
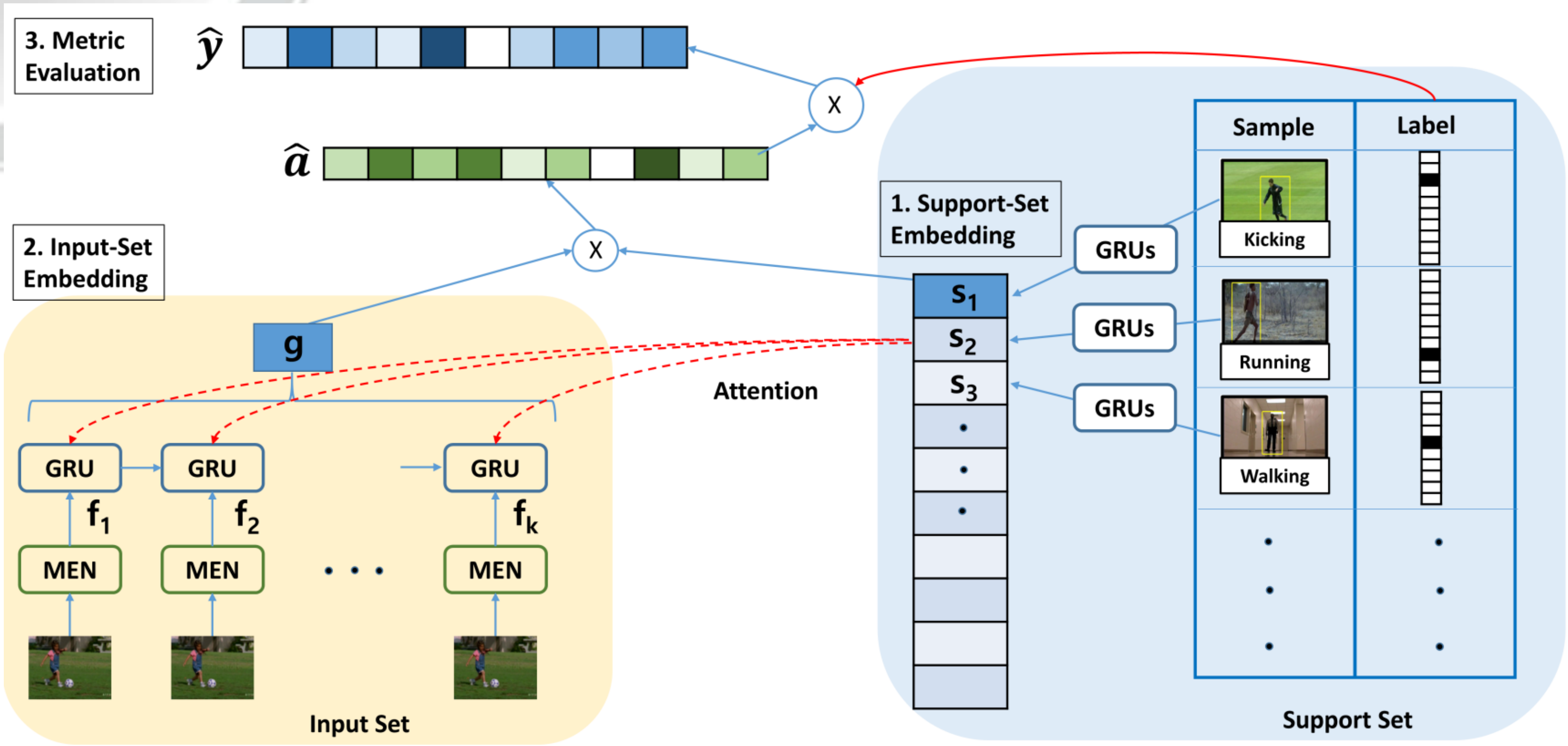


Figure 2: Overall structure of the proposed method (HP-SPS)

(a) input (b) ground truth (c) our result (d) input (e) ground truth (f) our result

Figure 6: Some prediction example of our method on Pascal context dataset

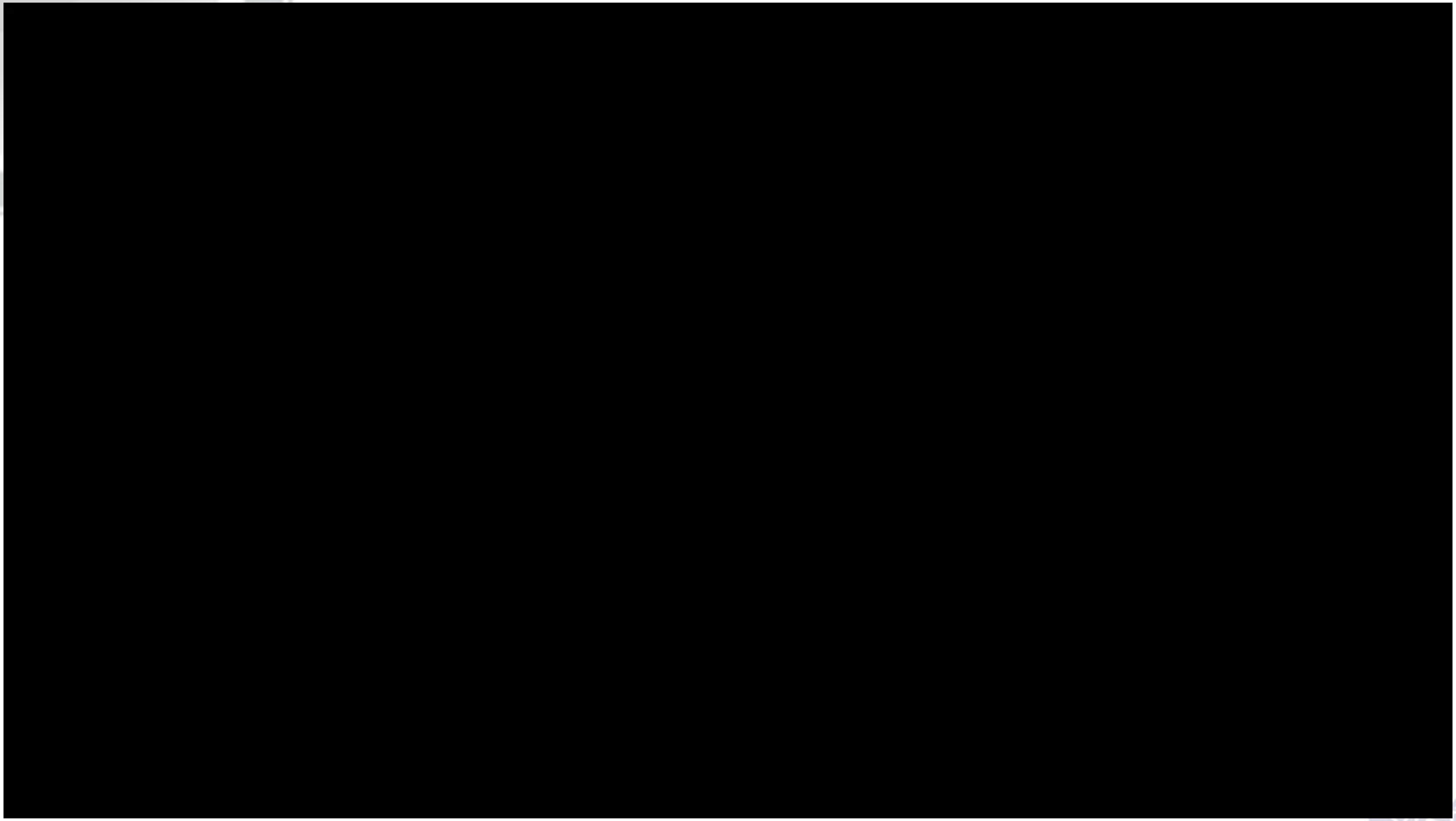
Video action recognition



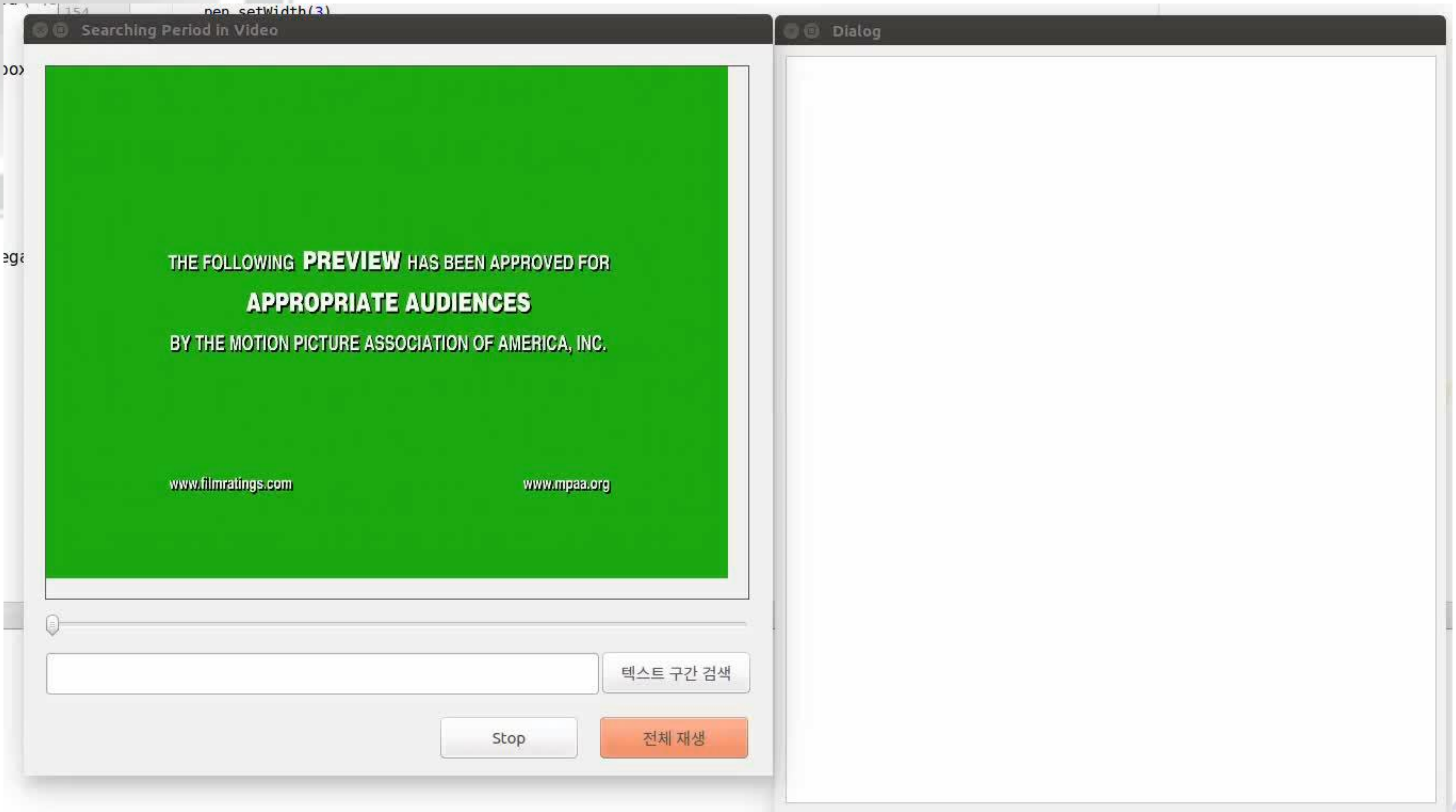
MIPA LABORATORY

서 울 대 학 교
SEOUL NATIONAL UNIVERSITY

Network Pruning



Semantic Search in Video



Reinforcement Learning

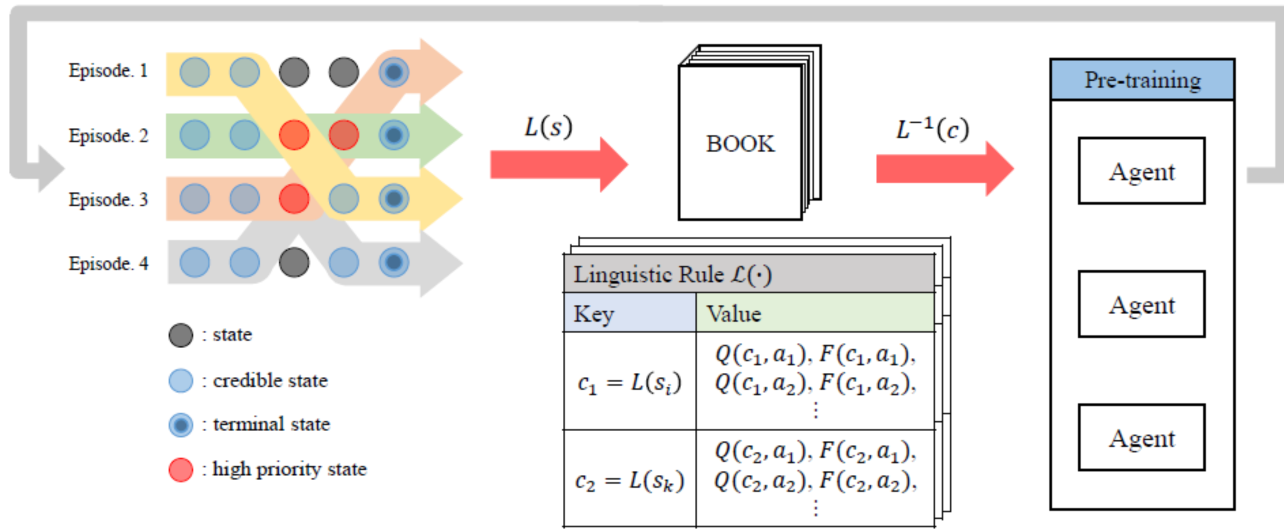
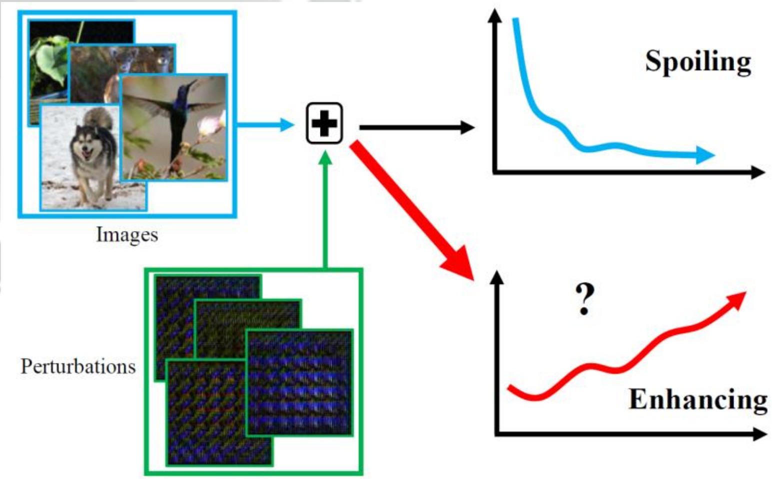
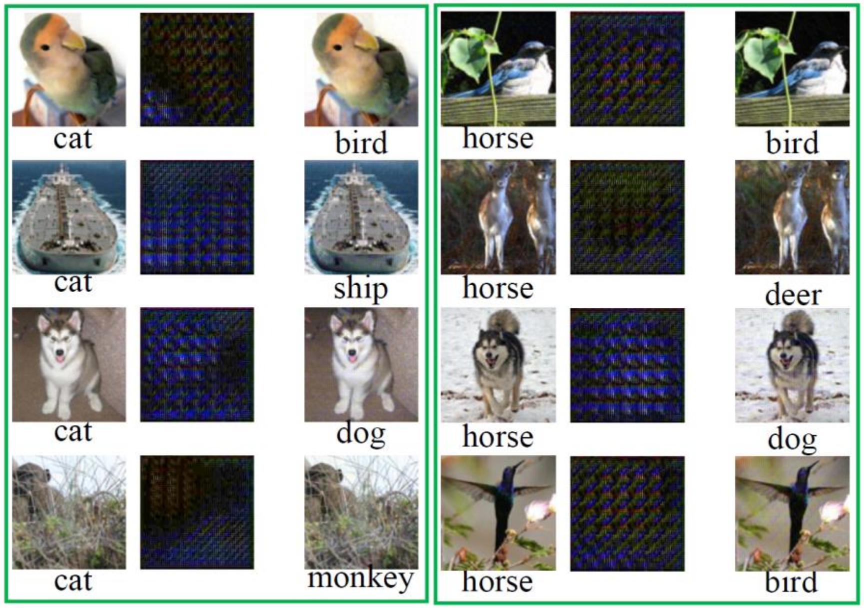
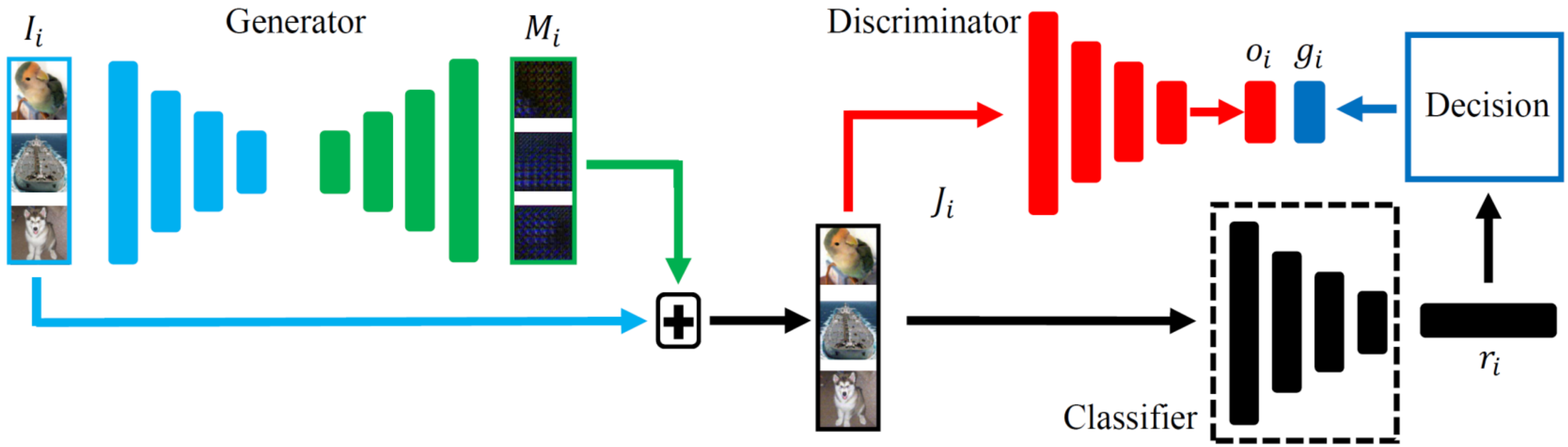


Figure 1: Overall framework of the proposed model. Similar experiences (state-action pairs) from multiple episodes are grouped into a cluster and the credible experiences corresponding to large clusters are written in a BOOK with their Q-values and frequencies Fs. Then, the agent uses this information in training.

Noise GAN


Images
Perturbations
Spoiling
?
Enhancing


cat
bird
cat
ship
cat
dog
cat
monkey
horse
bird
horse
deer
horse
dog
horse
bird


I_i
Generator
M_i
Discriminator
o_i
g_i
Decision
J_i
Classifier
r_i

TQA

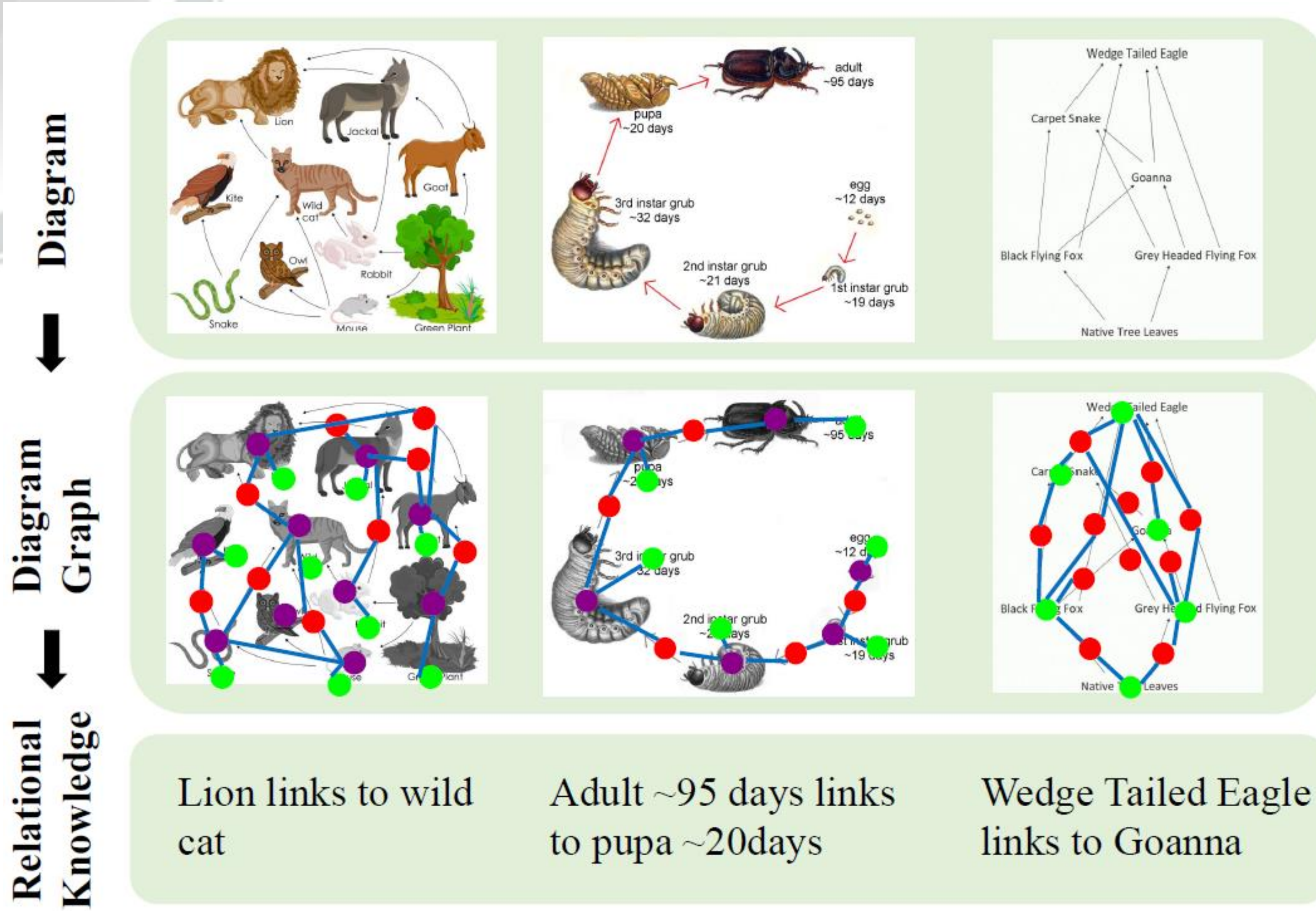


Diagram
Diagram Graph
Relational Knowledge
Lion
Jackal
Goat
Kite
Wild cat
Owl
Rabbit
Snake
Mouse
Green Plant
adult ~95 days
pupa ~20 days
3rd instar grub ~32 days
egg ~12 days
2nd instar grub ~21 days
1st instar grub ~19 days
Wedge Tailed Eagle
Carpet Snake
Goanna
Black Flying Fox
Grey Headed Flying Fox
Native Tree Leaves
Lion links to wild cat
Adult ~95 days links to pupa ~20days
Wedge Tailed Eagle links to Goanna

JPEG Artifact Removal

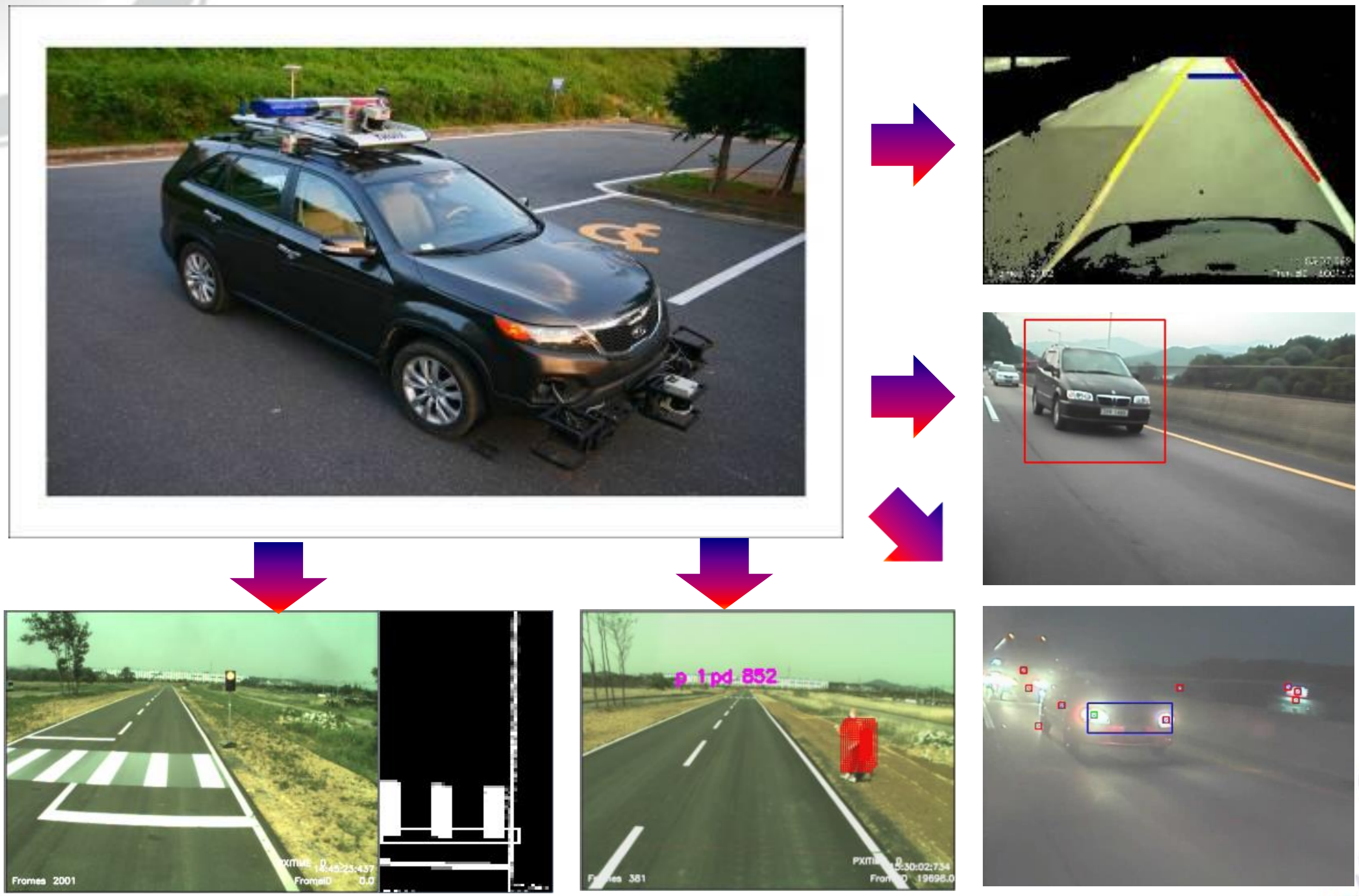
로봇/자동차 응용

MIPA LABORATORY

서울대학교
SEOUL NATIONAL UNIVERSITY

Autonomous Vehicles

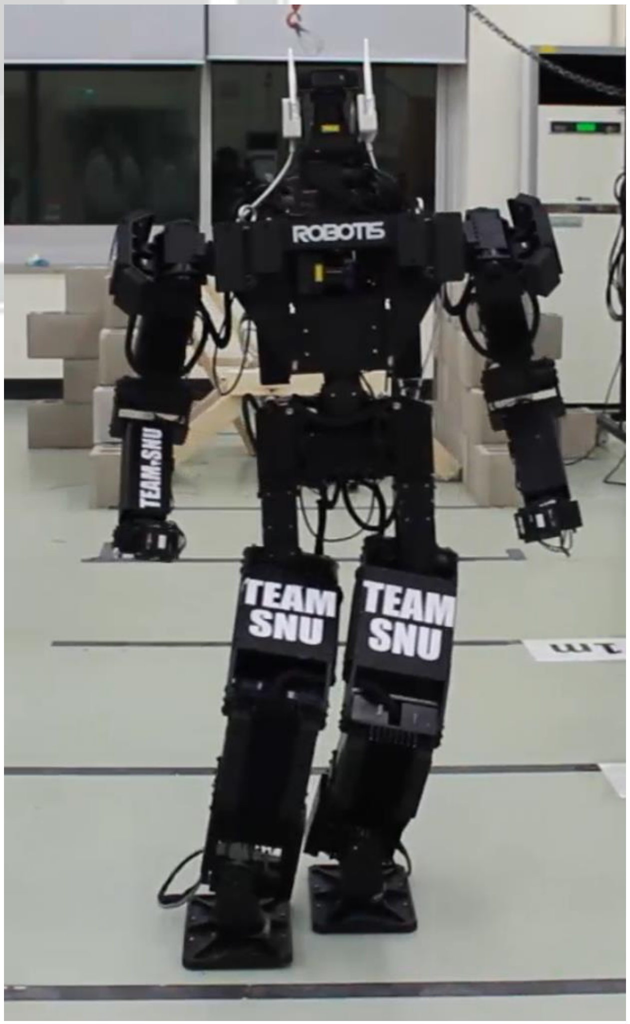
- Lane, Vehicles, Traffic lights, Traffic Signs, Crosswalk, Pedestrian, Obstacles …

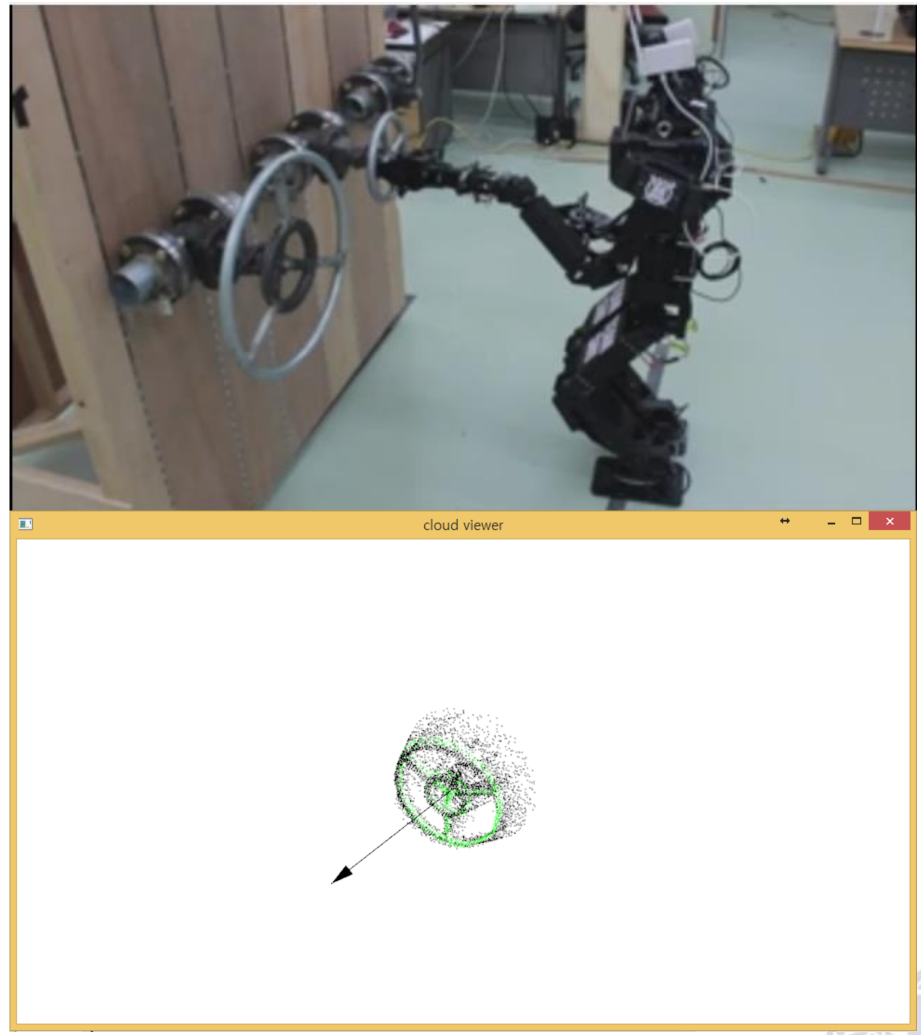


Vehicle Detection

- BSD & Frontal

DARPA Robotics Challenge





SNU in DRC

THANK YOU
QUESTION?

MIPA LABORATORY

서 울 대 학 교
SEOUL NATIONAL UNIVERSITY