Tokyo Institute of Technology
School of Computing
Department of Computer Science

Introduction of Kanezaki Laboratory

Assoc. Prof. Asako Kanezaki

Short bio.

- Mar 2008 Graduated from Department of Mechanical Information Engineering, Faculty of Engineering, The University of Tokyo
- Mar 2010 Completed master's course, Graduate School of Information Science and Technology, The University of Tokyo
  - 2010-2011 (a half year) visiting research at Technische Universität München
- Mar 2013 Ph.D. (Information Science and Technology), The University of Tokyo,
- Apr 2013 Full-time employee, Research and Development Center, Toshiba Corporation
- Dec 2013 Assistant Professor, Graduate School of Information Science and Technology, The University of Tokyo
  - 2015.8-9 visiting research at Microsoft Research Redmond
- Apr 2016 Researcher -> Senior Researcher, Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology (AIST)
- Apr 2020 Associate Professor, School of Computing, Tokyo Institute of Technology

Computer Vision

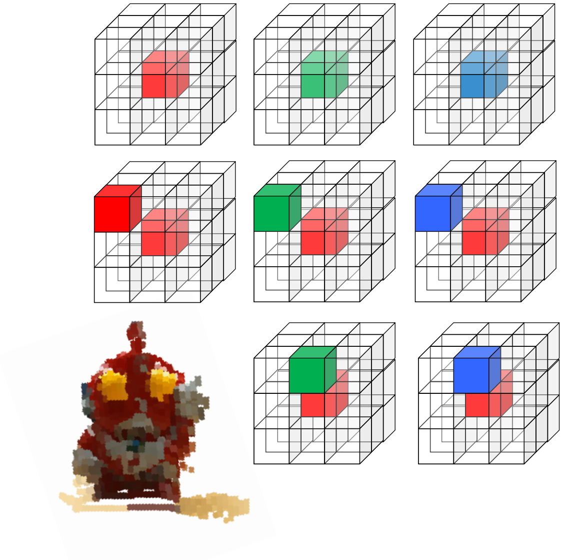
Machine Learning

Robotics

Research

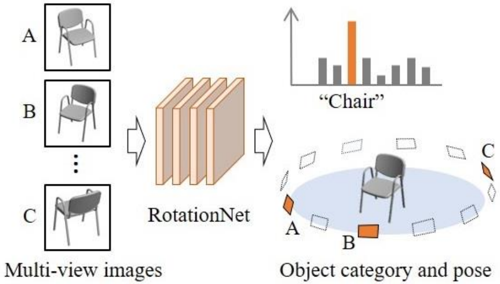
Master – Ph.D. student

3D features & Object recognition



AIST（last 4 years)

3D object recognition



Deep Learning

Unsupervised Image Segmentation

Robot Navigation

Computer Vision

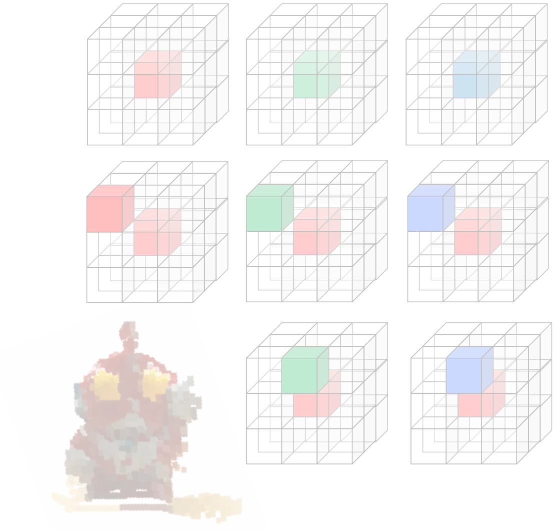
Machine Learning

Robotics

Research

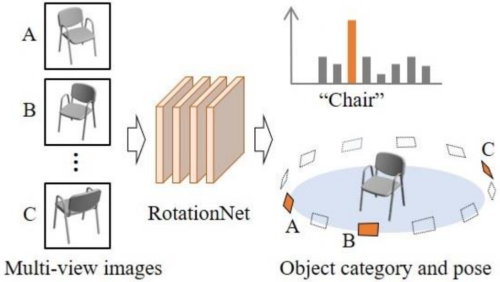
Master – Ph.D. student

3D features & Object recognition



AIST (last 4 years)

3D object recognition



Deep Learning

Unsupervised Image Segmentation

Robot Navigation

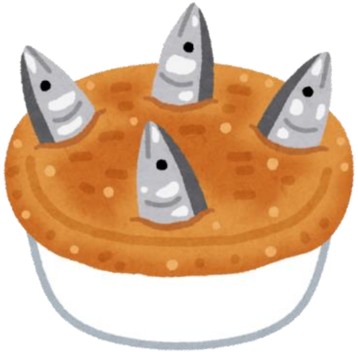
CVPR 2018, accepted to TPAMI

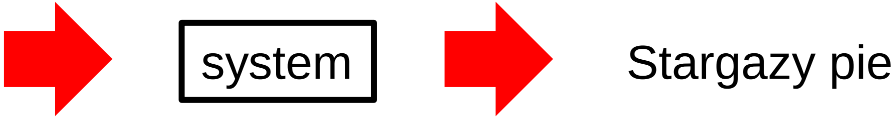
RotationNet: Joint Object Categorization and Pose Estimation Using Multiviews from Unsupervised Viewpoints

Asako Kanezaki, Yasuyuki Matsushita, and Yoshifumi Nishida.
National Institute of Advanced Industrial Science and Technology (AIST)

3D object recognition?

- Takes 3D data as input and outputs the probability of object categories (object classification)



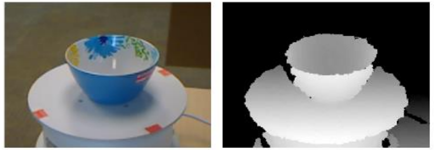


Cf.) object detection, object retrieval, parts segmentation

3D object recognition?

Approaches:

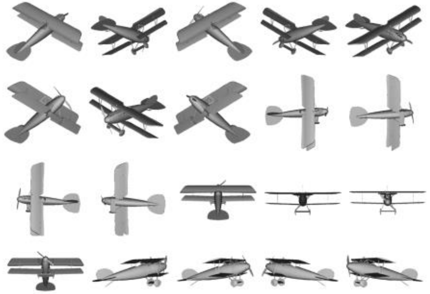
RGBD based



Point Cloud based

Voxel based

Multi-view based



Here

RotationNet is the best!

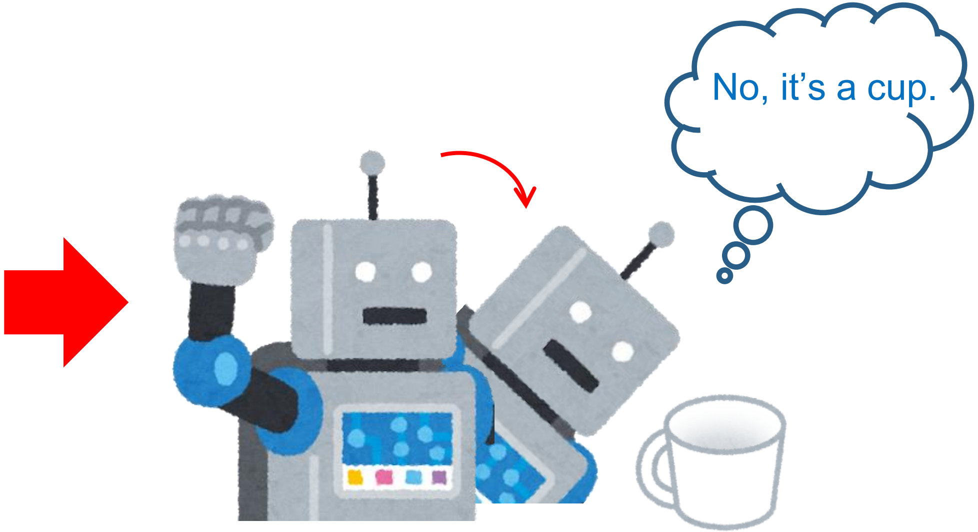
- State-of-the-art scores on the ModelNet dataset http://modelnet.cs.princeton.edu/
  - ModelNet40: 40 categories
  - ModelNet10: 10 categories
  - 2020/4/25 leaderboard ⇒

ModelNet40	ModelNet10
First : RotationNet Multi-view based	First : RotationNet Multi-view based
Second : iMHL Multi-view based	Second : SPNet Multi-view based

Algorithm	ModelNet40 Classification (Accuracy)	ModelNet40 Retrieval (mAP)	ModelNet10 Classification (Accuracy)	ModelNet10 Retrieval (mAP)
RS-CNN[63]	93.6%	-	-	-
LP-3DCNN[62]	92.1%	-	94.4%	-
LDGCNN[61]	92.9%	-	-	-
Primitive-GAN[60]	86.4%	-	92.2%	-
3DCapsule [59]	92.7%	-	94.7%	-
3D2SeqViews [58]	93.40%	90.76%	94.71%	92.12%
OrthographicNet [57]	-	-	88.56%	86.85%
Ma et al. [56]	91.05%	84.34%	95.29%	93.19%
MLVCNN [55]	94.16%	92.84%	-	-
iMHL [54]	97.16%	-	-	-
HGNN [53]	96.6%	-	-	-
SPNet [52]	92.63%	85.21%	97.25%	94.20%
MHBN [51]	94.7	-	95.0	-
VIPGAN [50]	91.98	89.23	94.05	90.69
Point2Sequence [49]	92.60	-	95.30	-
Triplet-Center Loss [48]	-	88.0%	-	-
PVNet[47]	93.2%	89.5%	-	-
GVCNN[46]	93.1%	85.7%	-	-
MLH-MV[45]	93.11%		94.80%	
MVCNN-New[44]	95.0%			
SeqViews2SeqLabels[43]	93.40%	89.09%	94.82%	91.43%
G3DNet[42]	91.13%		93.1%	
VSL [41]	84.5%		91.0%	
3D-CapsNets[40]	82.73%	70.1%	93.08%	88.44%
KCNet[39]	91.0%		94.4%	
FoldingNet[38]	88.4%		94.4%	
binVoxNetPlus[37]	85.47%		92.32%	
DeepSets[36]	90.3%			
3D-DescriptorNet[35]			92.4%	
SO-Net[34]	93.4%		95.7%	
Minto et al.[33]	89.3%		93.6%	
RotationNet[32]	97.37%		98.46%	
LonchaNet[31]			94.37	
Achlioptas et al. [30]	84.5%		95.4%	
PANORAMA-ENN [29]	95.56%	86.34%	96.85%	93.28%
3D-A-Nets [28]	90.5%	80.1%		
Soltani et al. [27]	82.10%			
Arvind et al. [26]	86.50%			
LonchaNet [25]			94.37%	
3DmFV-Net [24]	91.6%		95.2%	
Zanuttigh and Minto [23]	87.8%		91.5%	
Wang et al. [22]	93.8%			
ECC [21]	83.2%		90.0%	
PANORAMA-NN [20]	90.7%	83.5%	91.1%	87.4%
MVCNN-MultiRes [19]	91.4%			
FPNN [18]	88.4%			
PointNet[17]	89.2%			
Klokov and Lempitsky[16]	91.8%		94.0%	
LightNet[15]	88.93%		93.94%	
Xu and Todorovic[14]	81.26%		88.00%	
Geometry Image [13]	83.9%	51.3%	88.4%	74.9%
Set-convolution [11]	90%			
PointNet [12]			77.6%	
3D-GAN [10]	83.3%		91.0%	
VRN Ensemble [9]	95.54%		97.14%	
ORION [8]			93.8%	
FusionNet [7]	90.8%		93.11%	
Pairwise [6]	90.7%		92.8%	
MVCNN [3]	90.1%	79.5%		
GIFT [5]	83.10%	81.94%	92.35%	91.12%
VoxNet [2]	83%		92%	
DeepPano [4]	77.63%	76.81%	85.45%	84.18%
3DShapeNets [1]	77%	49.2%	83.5%	68.3%

Motivation

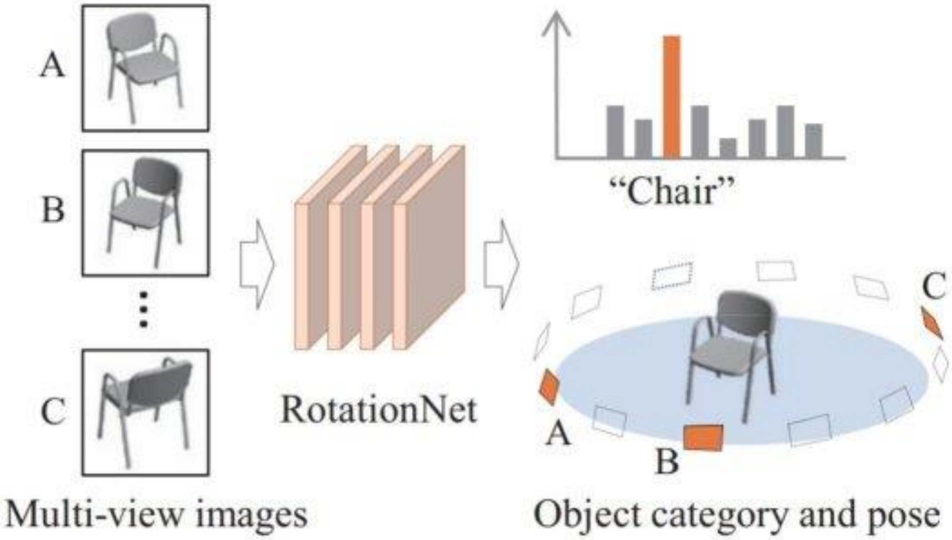
- Object recognition by robots
  “Move and see” to achieve better performance



【a single image input】
Not always captured from a best view to recognize an object.

RotationNet (Overview)

- Takes multi-view images as input and predicts object category and pose
  - Applicable to real-time applications
  - Improved accuracy by rotating in a direction that is easy to recognize



Video

https://kanezaki.github.io/rotationnet/

Competition & Exhibition

- SHREC2017 - 3D Shape Retrieval Contest 2017

  Participated in 2 out of 7 tracks and **won the first prize for both of them!**

[Subject 1]

Competition to find a model similar to a query shape from a large three-dimensional CAD object database

[Subject 2]

Competition to find a CAD model similar to a real query shape data captured by an RGB-D sensor

Demonstration of Hololens application at CEATEC2017: 152,066 participants

Invited Talks on 3D Recognition

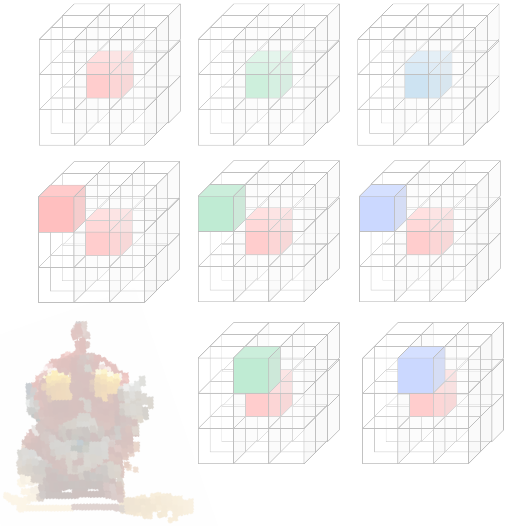
- **「機械学習を用いた3次元データ認識について」**，第3回情報処理学会セミナー AIと歩む未来(2)：画像・映像処理の最前線，千代田区，2019.9
- **「深層学習を用いた三次元物体認識」**，確率場と深層学習に関する第2回CRESTシンポジウム，早稲田大学，2018.10
- **「3次元物体認識技術」**，第112回ロボット工学セミナー ロボットのための画像処理技術，2018.5
- **「色距離センサを用いた点群処理と三次元物体認識に関する研究紹介」**，画像応用技術専門委員会2016年度第３回研究会，2016.9
- **「Kinect等の色距離センサを用いた点群処理と3D物体認識—ベーシックな手法と最新動向・ソフトウェアの紹介—」**, 第22回 画像センシングシンポジウム（SSII2016）チュートリアル講演, みなとみらい, 2016.6
- **「RGBD画像処理と三次元物体認識の研究紹介」**，公益社団法人 精密工学会 大規模環境の３次元計測と認識・モデル化技術専門委員会 第20回定例研究会, 本郷, 2015.12
- **「三次元画像からの物体検出と三次元特徴量」**, 第9回 次世代コンピュータ支援診断ソフトウェア臨床使用・評価プラットフォーム研究会，本郷，2015.6

etc.

Research

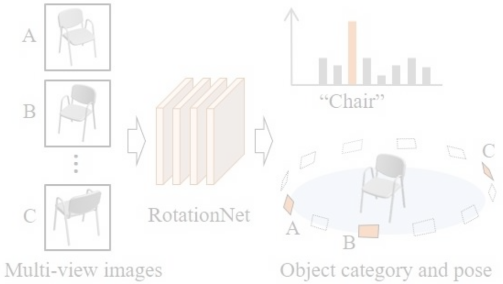
Master – Ph.D. student

3D features & Object recognition



AIST (last 4 years)

3D object recognition



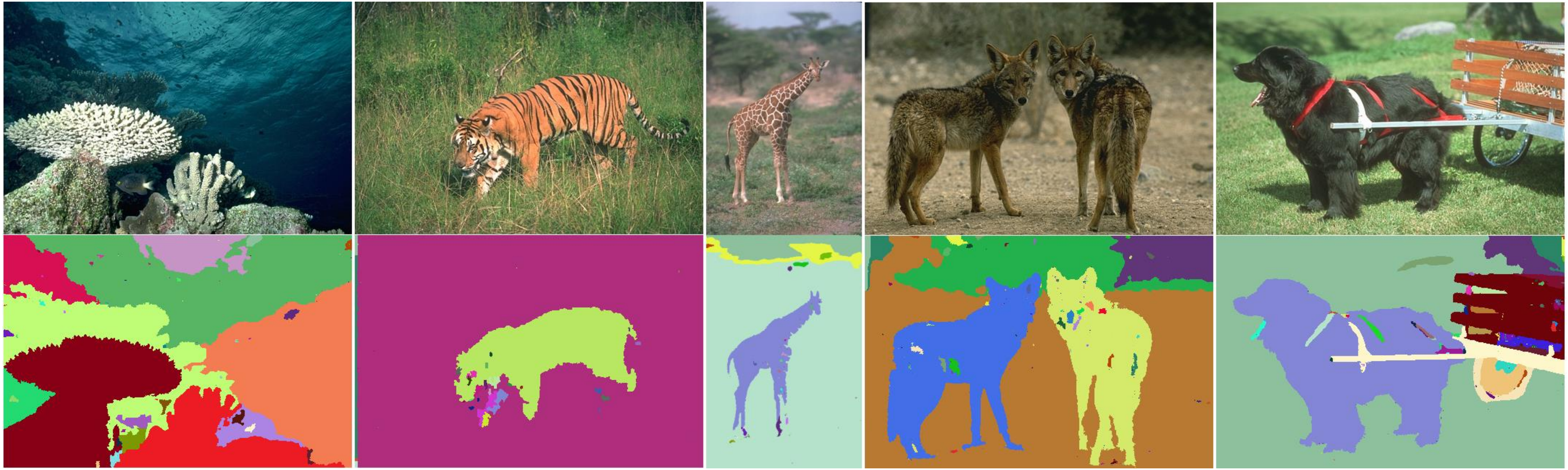
Deep Learning

Unsupervised Image Segmentation

Robot Navigation

https://github.com/kanezaki/pytorch-unsupervised-segmentation

Unsupervised Image Segmentation using Deep Learning



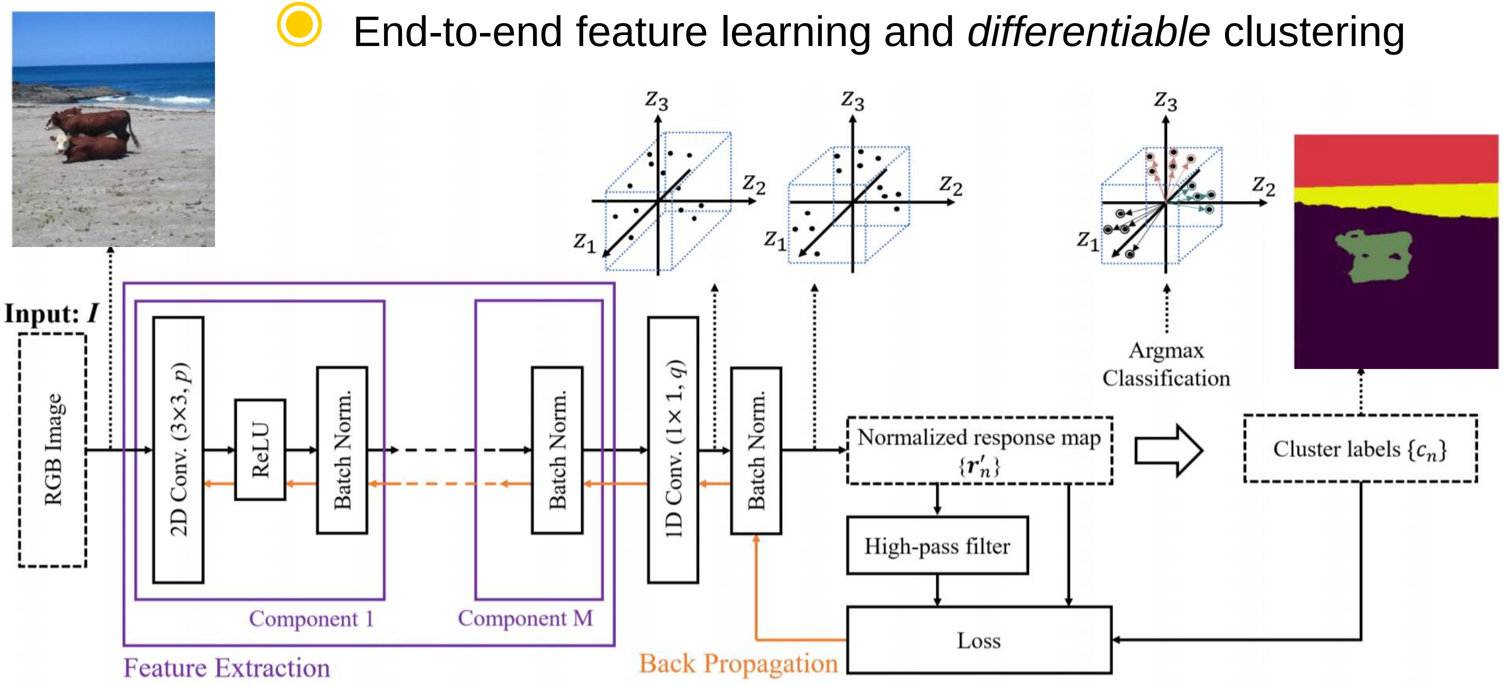
※Each area is shown in the same random color.

Asako Kanezaki. “Unsupervised Image Segmentation by Backpropagation.” *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp.1543-1547, 2018.

https://github.com/kanezaki/pytorch-unsupervised-segmentation

Unsupervised Image Segmentation using Deep Learning

- End-to-end feature learning and *differentiable* clustering

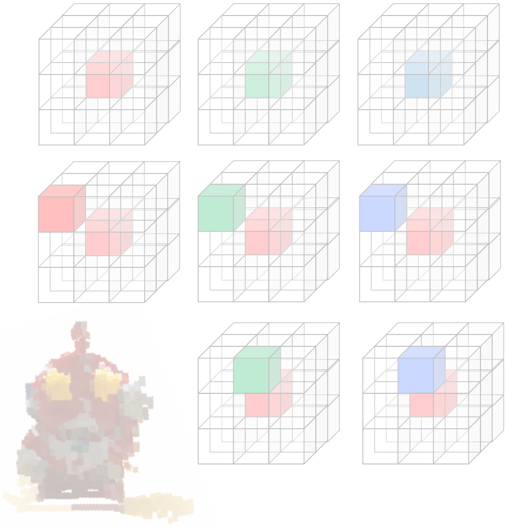


Asako Kanezaki. “Unsupervised Image Segmentation by Backpropagation.” *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp.1543-1547, 2018.

Research

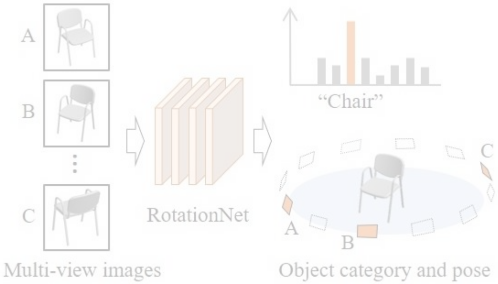
Master – Ph.D. student

3D features & Object recognition



AIST (last 4 years)

3D object recognition



Deep Learning

Unsupervised Image Segmentation

Robot Navigation

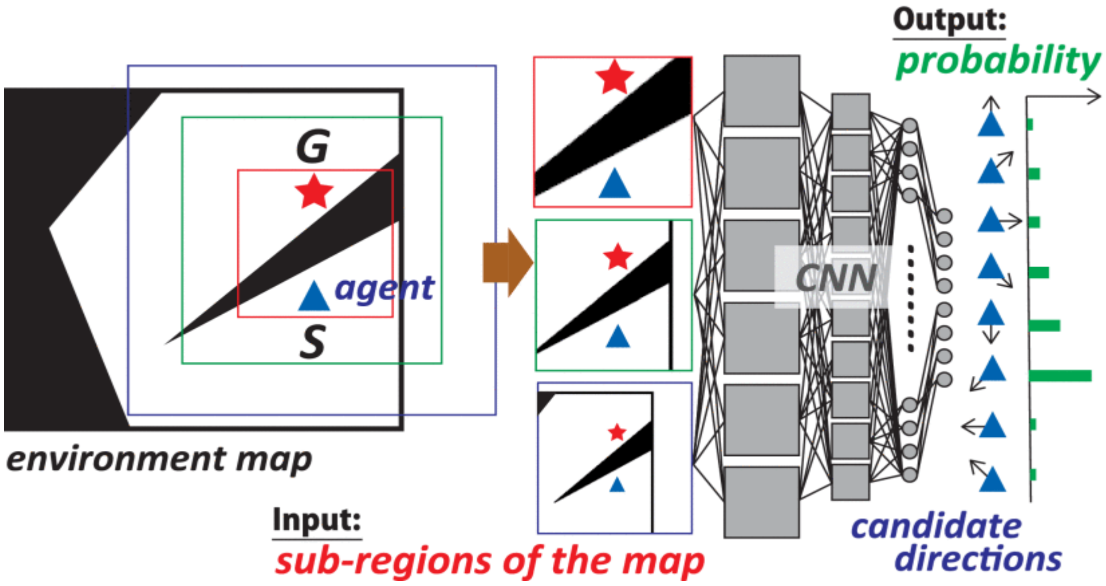
Robot navigation using deep learning

Video
https://kanezaki.github.io/goselo/

A. Kanezaki+, "GOSELO: Goal-Directed Obstacle and Self-Location Map for Robot Navigation using Reactive Neural Networks." *IEEE Robotics and Automation Letters (RA-L),* Vol.3, Issue 2, pp.696-703, 2018. (**presented in ICRA'18**)

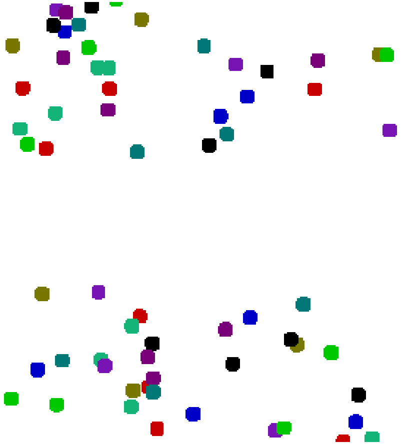
Robot navigation using deep learning

Key Idea: By creating a goal-directed map representation, we can learn the relationship between visual patterns of the surrounding environment and movement patterns without being constrained by the shape of a particular environment!



Follow-up studies（1/2）

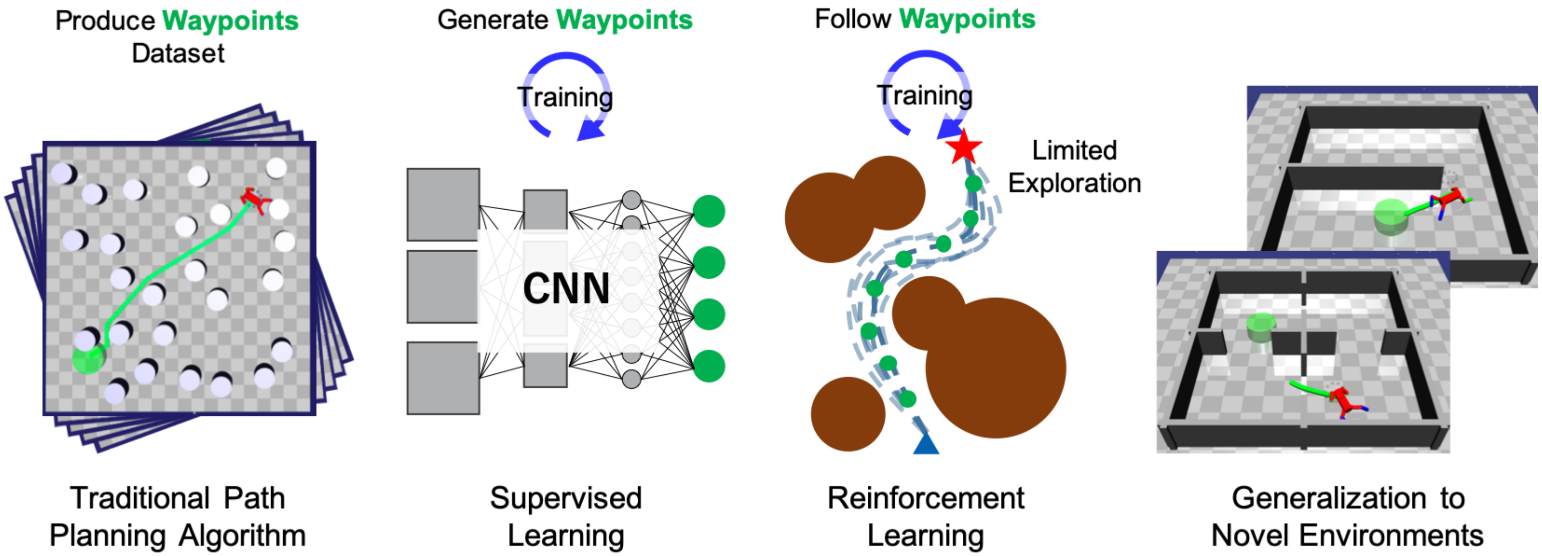
- Yoko Sasaki, Syusuke Matsuo, Asako Kanezaki, and Hiroshi Takemura. “A3C Based Motion Learning for an Autonomous Mobile Robot in Crowds.” *IEEE International Conference on System Man and Cybernetics (SMC2019)*, pp.1046-1052, 2019.
- 佐々木洋子，松尾修佑，金崎朝子，竹村裕. 歩行者観測履歴を用いた深層強化学習による車輪ロボットの雑踏切り抜け動作生成. 日本機械学会ロボティクス・メカトロニクス講演会, 2019.
- 渋谷 薫, 金崎 朝子, 大西 正輝. 深層学習による画像識別問題に帰着した人の流れのシミュレーション. Meeting on Image Recognition and Understandings (MIRU), 2018.



Learned policy simulation

Follow-up studies（2/2）

Kei Ota, Yoko Sasaki, Devesh K. Jha, Yusuke Yoshiyasu, Asako Kanezaki. “Efficient Exploration in Constrained Environments with Goal-Oriented Reference Path”, *arXiv,* 2020.



Student internships at AIST

- Robot navigation by reinforcement learning 理科大

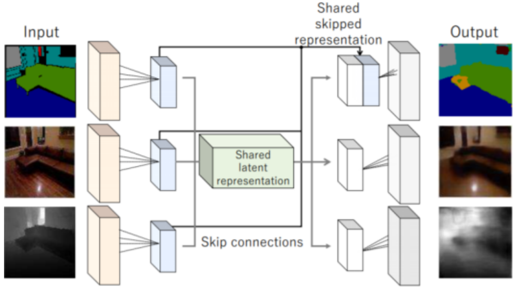
Yoko Sasaki, Syusuke Matsuo, Asako Kanezaki, and Hiroshi Takemura. **“A3C Based Motion Learning for an Autonomous Mobile Robot in Crowds.”** *IEEE International Conference on System Man and Cybernetics (SMC2019)*, pp.1046-1052, 2019.

- Product recognition using deep learning 首都大

加納諒也，和田一義，金崎朝子，冨沢哲雄，谷川民生. 深層学習による物体認識のためのマーカの検討 ～第一報：AlexNetによる認識率の比較～. 日本機械学会ロボティクス・メカトロニクス講演会, 2019.

- Deep Multi-modal learning 阪大

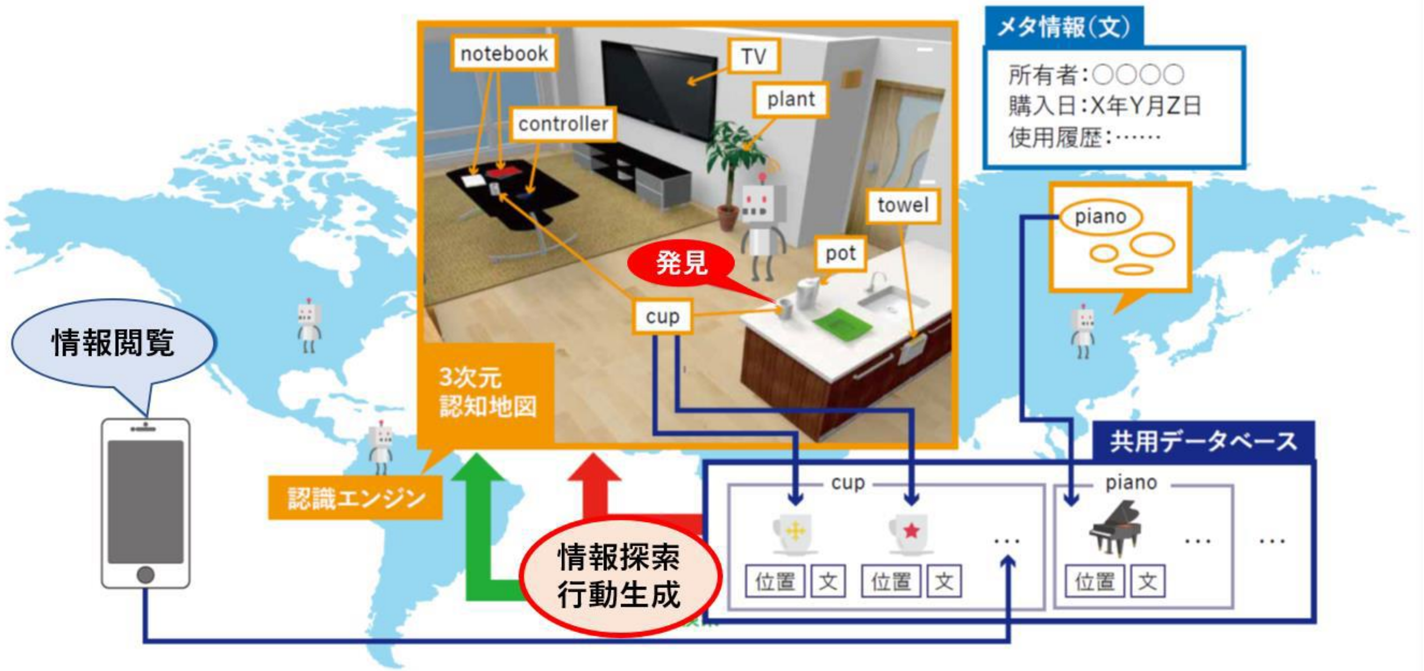
Ryohei Kuga, Asako Kanezaki, Masaki Samejima, Yusuke Sugano, and Yasuyuki Matsushita. **“Multi-modal U-Nets for Multi-task Scene Understanding.”** *IEEE ICCV Workshop on Multi-Sensor Fusion for Dynamic Scene Understanding*, pp.403-411, 2017.



- Unsupervised image segmentation by deep learning 東工大
- Real-time object detection by RotationNet 東工大

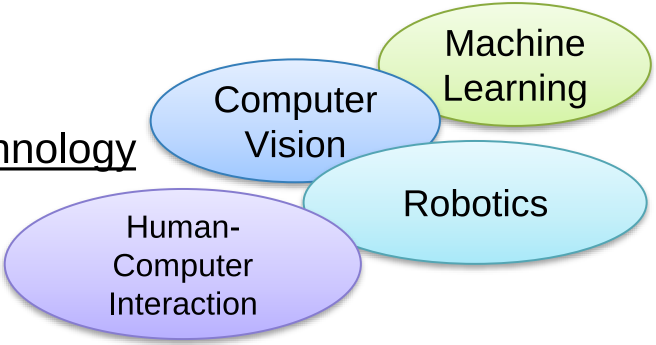
Research plan *IoP (Internet of Perception)?*

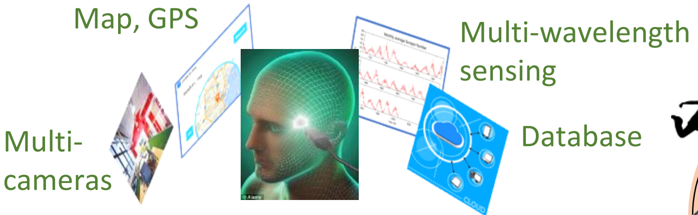
- A robot system for active information search and database construction



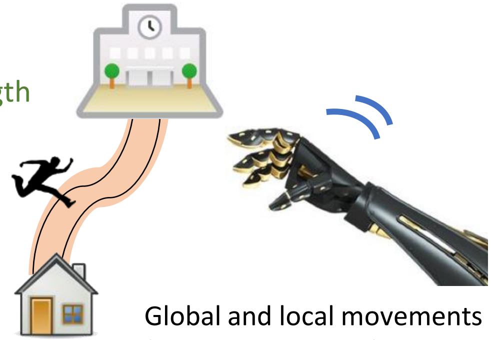
Research topics (examples)

- Action planning for autonomous mobile robots using reinforcement learning
- Automatic generation of daily life search engines
- Information collection planning for interactive robots
- Robot intelligence with superhuman recognition technology
- Global and local action planning for robots

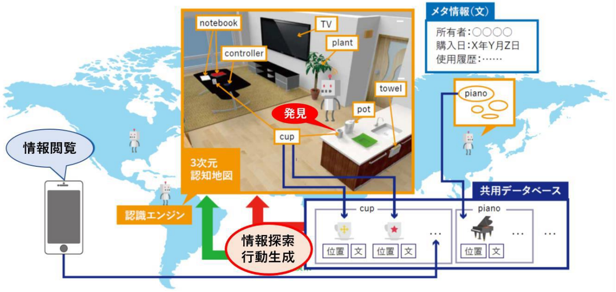




Construction of robot intelligence with superhuman recognition technology

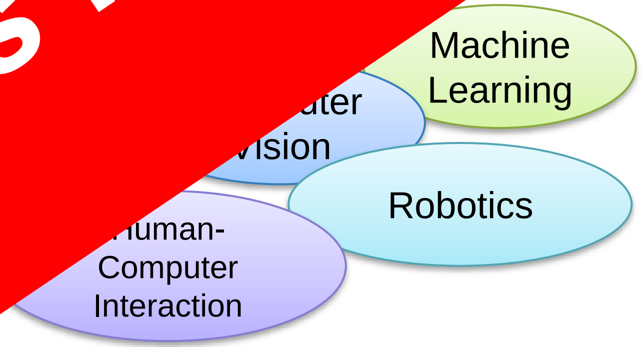


Global and local movements (Manipulation, etc.)



Research topics (examples)

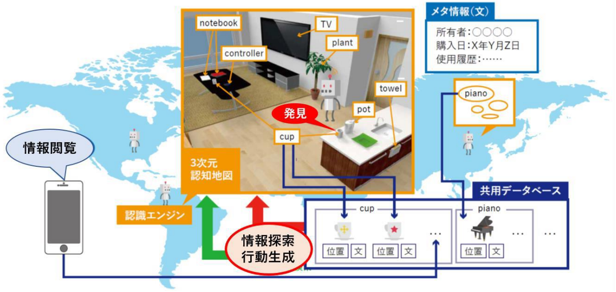
- Action planning for autonomous m
- Automatic generation of daily
- Information collection pl
- Robot intelligence
- Global and lo



Anything else is fine.

Map, GPS

Global and local movements
(Manipulation, etc.)



Why Kanezaki Laboratory? 😉

- New lab. (You will be first members!)
- Can try new things as you wish.
- The infrastructure will be in place.
  - Use of computer clusters such as TSUBAME, AIST-ABCI
  - Managing source code and tips on GitHub
  - Sharing of library know-how such as deep learning and robot simulation
- Ookayama Campus
  - West Building 8 is conveniently located.

Tokyo Tech

Questions?

kanezaki [at] c.titech.ac.jp @kanejaki