ICCV23

cvpaper.challenge

ICCV 2023 Workshop on

# Representation Learning with Very Limited Images

The potential of self-, synthetic- and formula-supervision

LIMIT Organization Team

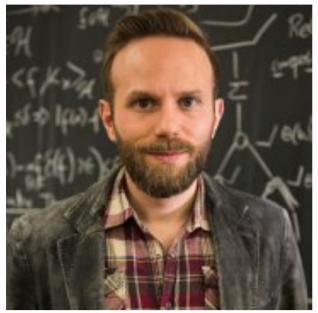
# Organizers

Hirokatsu Kataoka
AIST/LINE

Rio Yokota
Tokyo Tech/AIST

Nakamasa Inoue
Tokyo Tech/AIST

Dan Hendrycks
Center for AI Safety



Xavier Boix
MIT

Yue Qiu
AIST

Connor Anderson
BYU

Ryo Nakamura
Fukuoka Univ./AIST

Ryosuke Yamada
Univ. of Tsukuba/AIST

Risa Shinoda
Kyoto Univ./AIST

# Program committee

| | | |
|---|---|---|
| Aaqib | Saeed | Eindhoven University of Technology |
| Akira | Sakai | Fujitsu Laboratories Ltd. |
| Amir | Rahimi | Massachusetts Institute of Technology |
| Anirban | Sarkar | Massachusetts Institute of Technology |
| Atsushi | Hashimoto | OMRON SINIC X Corp. |
| Avi | Cooper | Yale University, Massachusetts Institute of Technology, Fujitsu, Weizmann Institute of Science |
| Balu | Adsumilli | YouTube/Google |
| Chia-Hsien | Shih | University of Illinois at Urbana-Champaign |
| Daiki | Ikami | Tokyo University of Agriculture and Technology |
| Ece | Ozkan | MIT |
| Erika | Mori | Keio University |
| Fumiya | Matsuzawa | University of Tsukuba |
| Go | Ohtani | Keio University |
| Guoqing | Hao | University of Tsukuba |
| Guy | Ben-Yosef | GE Research |
| Hannah | Pinson | Eindhoven University of Technology |
| Hao-Wei | Yeh | The University of Tokyo |
| Hiroaki | Aizawa | Hiroshima University |
| Hong-You | Chen | The Ohio State University |
| Hyunjung | Shim | KAIST |
| Ian | Mason | MIT |
| Ikuro | Sato | Tokyo Institute of Technology / Denso IT Laboratory |

**Many thanks!**

# Program committee

| | | |
|---|---|---|
| Jin | Yamanaka | SiMa.ai |
| Jingjing | Chen | Fudan University |
| Kasper | Vinken | Harvard |
| Kensho | Hara | National Institute of Advanced Industrial Science and Technology (AIST) |
| Kimberly | Villalobos Carballo | MIT |
| Kodai | Nakashima | CyberAgent, Univ. of Tsukuba, AIST |
| Kuniaki | Saito | Boston University |
| Lile | Cai | Institute for Infocomm Research |
| Manel | Baradad Jurjo | MIT |
| Masatoshi | Tateno | Institute of Industrial Science, The University of Tokyo |
| Mayu | Otani | CyberAgent |
| Moyuru | Yamada | Fujitsu Limited |
| Naoshi | Kaneko | Tokyo Denki University |
| Naoto | Inoue | CyberAgent |
| Naoya | Chiba | Tohoku University |
| Nishant | Rai | Stanford University |
| Pavan | Madhusudana | University of Texas at Austin |
| Qianru | Sun | National University of Singapore |
| Raphael | Achddou | Telecom Paris |
| Rei | Kawakami | Tokyo Institute of Technology |
| Ryosuke | Furuta | The University of Tokyo |
| Ryota | Suzuki | Saitama University |

**Many thanks!**

# Program committee

| | | |
|---|---|---|
| Ryu | Tadokoro | Tohoku University |
| Said | Ladjal | Telecom Paris |
| Sanjana | Srivastava | Stanford University |
| Seitaro | Shinagawa | Nara Institute of Science and Technology |
| Shobhita | Sundaram | MIT |
| Shota | Nakamura | Tokyo Institute of Technology |
| Shuya | Takahashi | Tokyo Denki University |
| Spandan | Madan | Harvard University |
| Takuma | Yagi | National Institute of Advanced Industrial Science and Technology |
| Taro | Sunagawa | Fujitsu Laboratories Ltd. |
| Tatsuya | Yokota | Nagoya Institute of Technology |
| Teppei | Suzuki | Denso IT Laboratory |
| Tommi | Kerola | Preferred Networks, Inc. |
| Tomoyuki | Suzuki | CyberAgent, Inc. |
| Tongzhou | Wang | MIT |
| Vikash | Sehwag | Princeton University |
| Xavier | Boix | MIT |
| Yanjun | Sun | Keio University |
| Yilin | Wang | Google Inc. |
| Youssef | Dawoud | Friedrich-Alexander-Universitat Erlangen-Nurnberg |
| Yu | Liu | Dalian University of Technology |
| Yunhao | Cao | Nanjing University |
| Yusuke | Mukuta | The University of Tokyo |

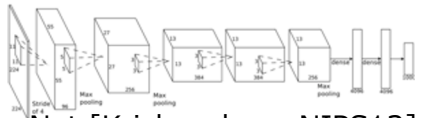
**Many thanks!**

# Paper statistics

- 44 paper submissions
- 41 valid submissions / 3 desk rejects
- 21 accepts (47.7% acceptance rate)

# Motivation

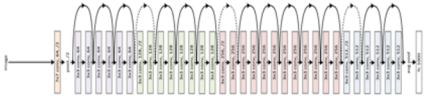
## Models



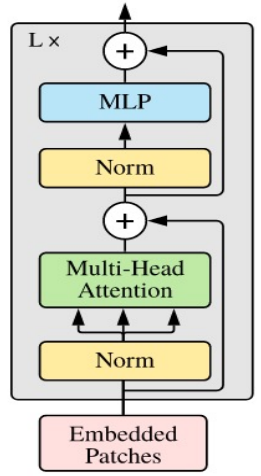
AlexNet [Krizhevsky+, NIPS12]

VGGNet [Simonyan+, ICLR15]

GoogLeNet [Szegedy+, CVPR15]

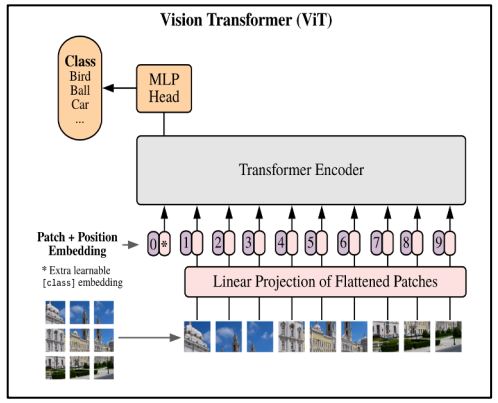


ResNet [He+, CVPR16]



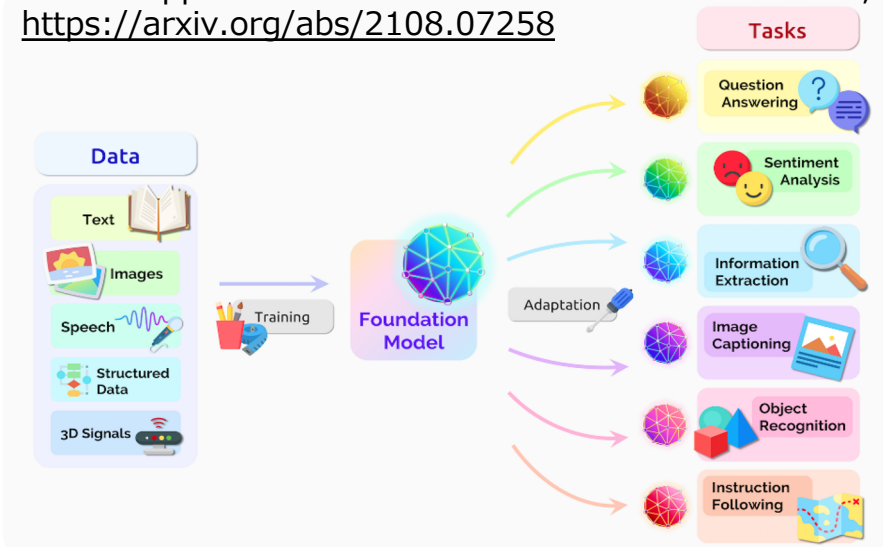

[Vaswani+, NIPS17]

Vision Transformer (ViT)



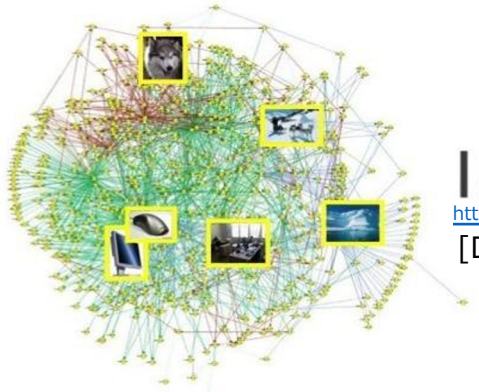

[Dosovitskiy+, ICLR21]

Foundation Models
On the Opporttunities and Risks of Foundation Models, https://arxiv.org/abs/2108.07258

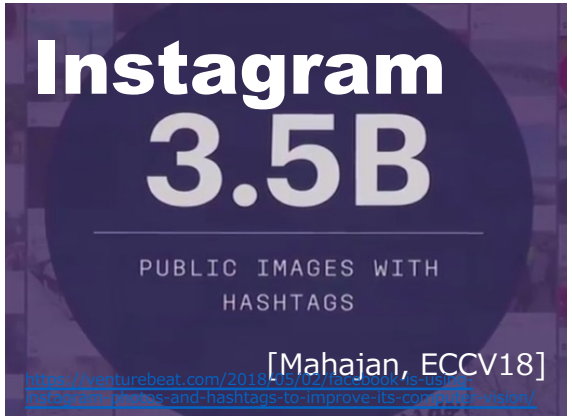



[Bommasani+, arXiv22]

## Datasets



IMAGENET
http://www.image-net.org/
[Deng+, CVPR09]




[Mahajan, ECCV18]
https://venturebeat.com/2018/05/02/facebook-is-using-instagram-photos-and-hashtags-to-improve-its-computer-vision/

JFT-300M/3B
[Sun+, ICCV17][Zhai+, CVPR22]

LAION
https://laion.ai/

Models and datasets are getting larger

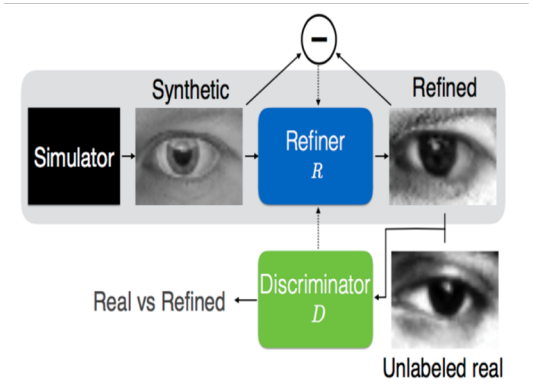
# Critical issues


A larger-curated dataset tends to occur labor/ethical concerns

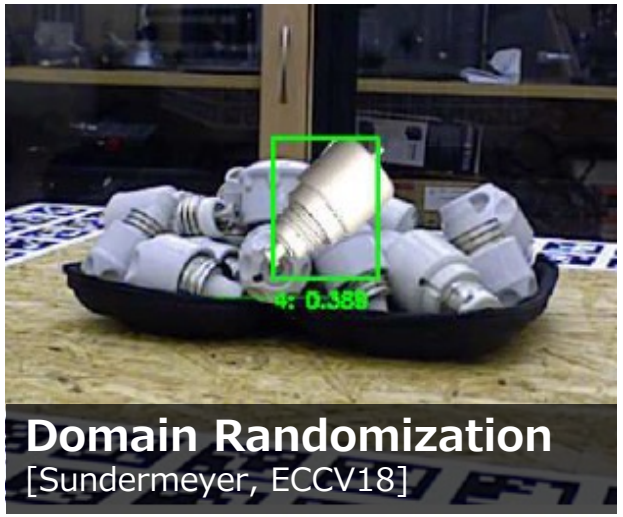
Less data, higher performance is preferable!

# Very limited images (data)?

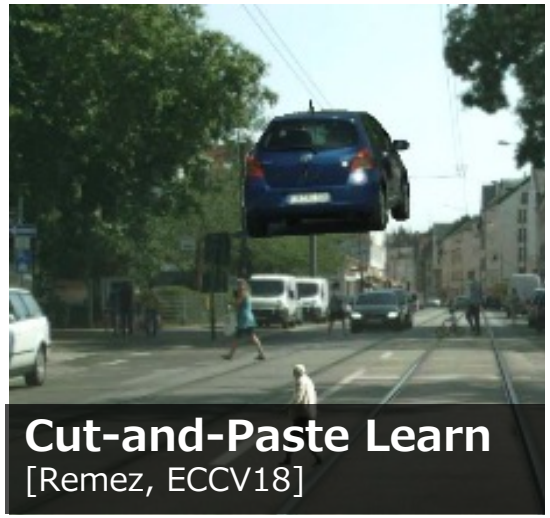
## Our community has awesome solutions



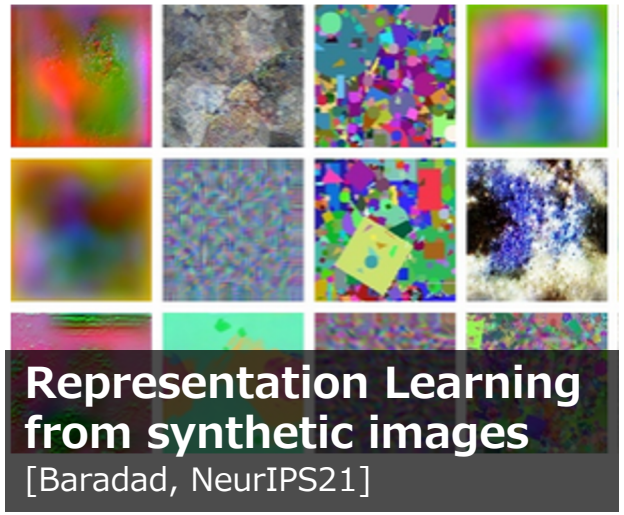

**Sim-to-Real**
[Shrivastava, CVPR17]



**Domain Randomization**
[Sundermeyer, ECCV18]



**Cut-and-Paste Learn**
[Remez, ECCV18]



**Representation Learning from synthetic images**
[Baradad, NeurIPS21]


**Fractal Database**
[Kataoka, IJCV22]

**1-image SSL**
[Asano, ICLR20, ICLR23]

Keyword: Synthetic Data, Adversarial Learning, Data Augmentation, Domain Randomization, Formula-driven Supervised Learning

# Today’s schedule

13:30 - 13:40 Welcome (10 min)
13:40 - 14:20 Invited Talk 1 (Christian Rupprecht; 40 min)
14:20 - 14:30 Short Break (10 min)
14:30 - 15:30 Oral Session (10 min x 6 Full Papers)
15:30 - 15:50 Coffee Break (20 min)
15:50 - 16:30 Invited Talk 2 (Manel Baradad; 40 min)
16:30 - 16:40 Spotlight Session (30 sec x 15 posters; 7.5 min)
16:40 - 16:45 Closing (5 min)
16:45 - 17:00 Short Break (10 min)
17:00 - 18:00 Poster Session (60 min)
* In-person only

# Invited talk 1: Christian Rupprecht (Univ. of Oxford)

**Title: Unsupervised Learning from Limited Data**

Abstract: While current large models are trained on millions or even billions of images, in this talk, we will discuss how unsupervised learning can be performed on a limited number of samples. A special focus of this talk will lie on representation learning, but we will also explore specific applications such as 3D reconstruction, object detection and tracking. Overall, several strategies have shown promise in this area: naturally image augmentations play a strong role in combatting data scarcity and imposing priors on images. Additionally, synthetic data can often be generated with either very simple methods or through pre-trained large-scale models that have already captured the diversity of the real world and allow the distillation of information into downstream applications.

# Invited talk 2: Manel Baradad Jurjo (MIT)

**Title: Learning to see by looking at noise**

Abstract: Current vision systems are trained on huge datasets, and these datasets come with costs: curation is expensive, they inherit human biases, and there are concerns over privacy and usage rights. To counter these costs, interest has surged in learning from cheaper data sources, such as unlabeled images. In this talk, I will present two of our recent approaches for training neural networks without data. In the first one, we train neural networks using a small set of curated programs that generate simple noise-like images, which have properties present in natural images. On the second, we follow the opposite direction: instead of curating a small set of programs, we collect a big dataset of 21k image-generation programs, which we do not manually tune, making it easier to scale up and increase performance. Both yield competitive results on different vision downstream tasks, showing the potential of these approaches to circumvent the usage of realistic data.