# Workshop Organizers

**General Chairs**



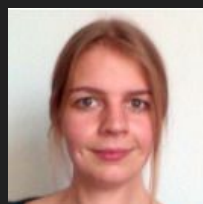J. Berent  A. Gupta  R. Sukthankar  L. Van Gool
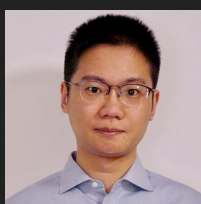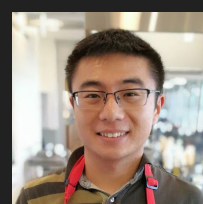
**Program Chairs**



Wen Li  Hilde Kuehne  Suman Saha  Qin Wang  Limin Wang  Wei Li

# Thanks to Workshop Sponsors & Collaborators

Dataset Collection & Challenge Hosting

Sponsor for Challenge and Award
Collaborator in Challenge Organization

Collaborator in Challenge Organization

# Program Schedule

| | |
|---|---|
| 9:00 | Opening Remarks |
| 9:10 | Dataset/Challenge Overview |
| 9:30 | Participant Presentation by Huawei |
| 9:40 | Participant Presentation by Tencent |
| 9:50 | Participant Presentation by Pcitech |
| 10:00 | Live Q&A Session |

| | |
|---|---|
| 10:15 | Paper Session (ID 1-3) |
| 10:30 | Live Q&A Session |
| 10:36 | Paper Session (ID 4-6) |
| 10:51 | Live Q&A Session |
| 11:00 | Award Session & Closing Remarks |

webvision

# Deep Learning Revolution

Revolutionizing almost all fields of computer vision

LeNet

AlexNet

GoogLeNet

ResNet

DenseNet



motor scooter — leopard

| motor scooter | leopard |
| go-kart | jaguar |
| moped | cheetah |
| bumper car | snow leopard |
| golfcart | Egyptian cat |



car : 1.000
person : 0.992
horse : 0.993
dog : 0.9?7
person : 0.979
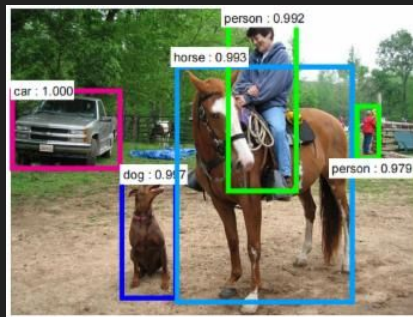

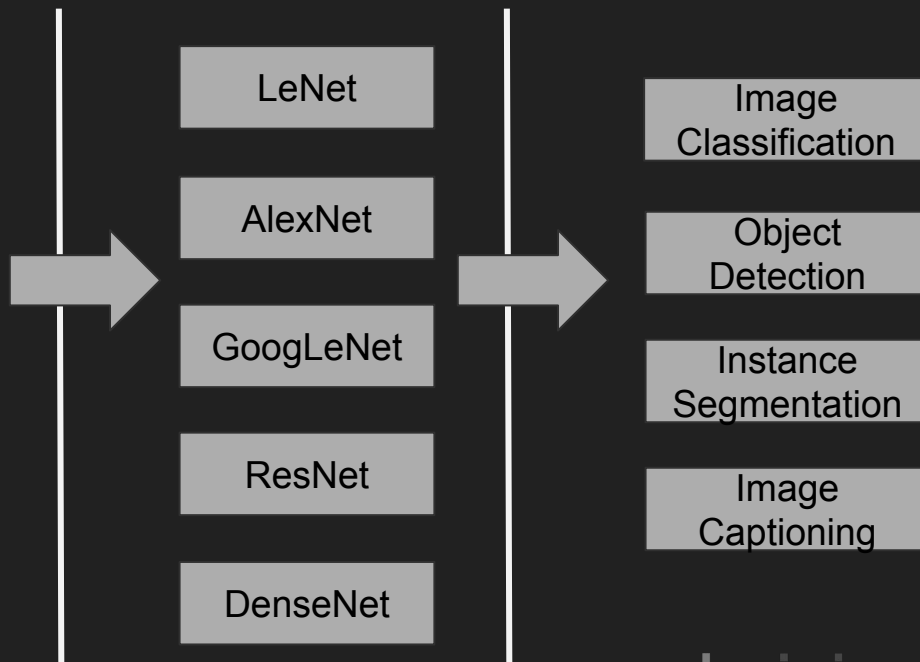
The man at bat readies to swing at the pitch while the umpire looks on.

# Deep Learning Revolution

Powered by *human annotated* big data



| LeNet |
| AlexNet |
| GoogLeNet |
| ResNet |
| DenseNet |

| Image Classification |
| Object Detection |
| Instance Segmentation |
| Image Captioning |

webvision
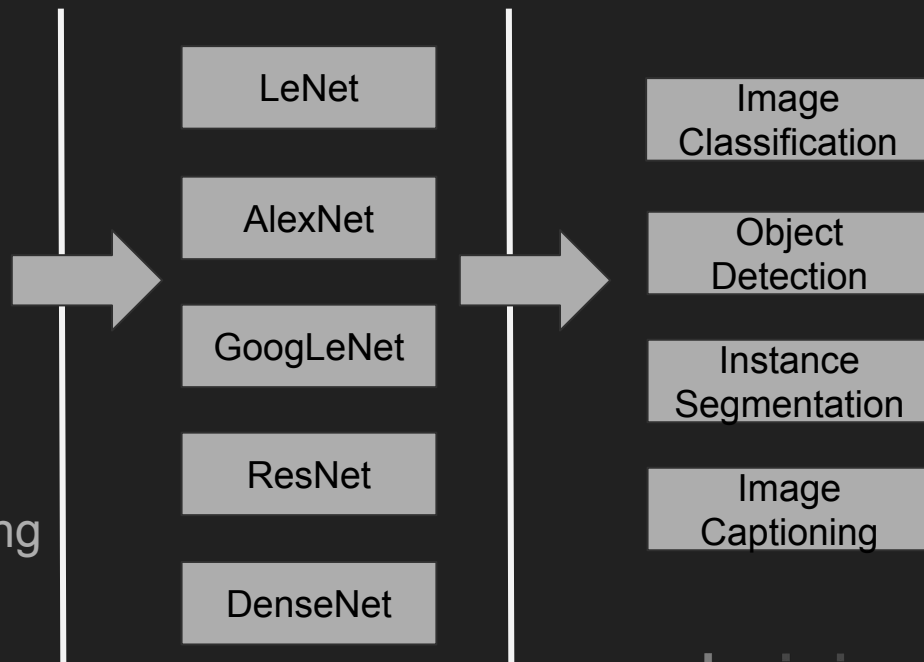
# Deep Learning Revolution -- Our Hope

Big Data

w/o human annotation

Can we get equivalent performance using {self, weakly, un}supervised methods?

| LeNet |
| AlexNet |
| GoogLeNet |
| ResNet |
| DenseNet |

| Image Classification |
| Object Detection |
| Instance Segmentation |
| Image Captioning |

webvision

# Deep Learning Revolution -- Previous Years

WebVision 2017

w/o human annotation

Yes!

LeNet

AlexNet

GoogLeNet

ResNet

DenseNet

Image Classification

Object Detection

Instance Segmentation

Image Captioning
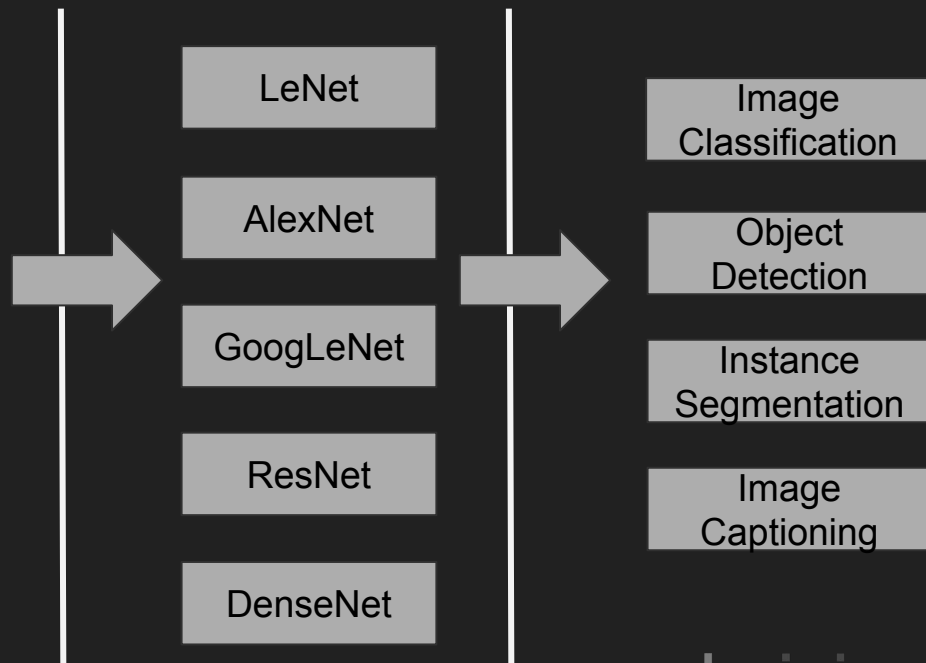
webvision

# Deep Learning Revolution -- Previous Years

WebVision
2018/2019
w/o human annotation

A Bigger Dataset

LeNet

AlexNet

GoogLeNet

ResNet

DenseNet

Image Classification

Object Detection

Instance Segmentation

Image Captioning
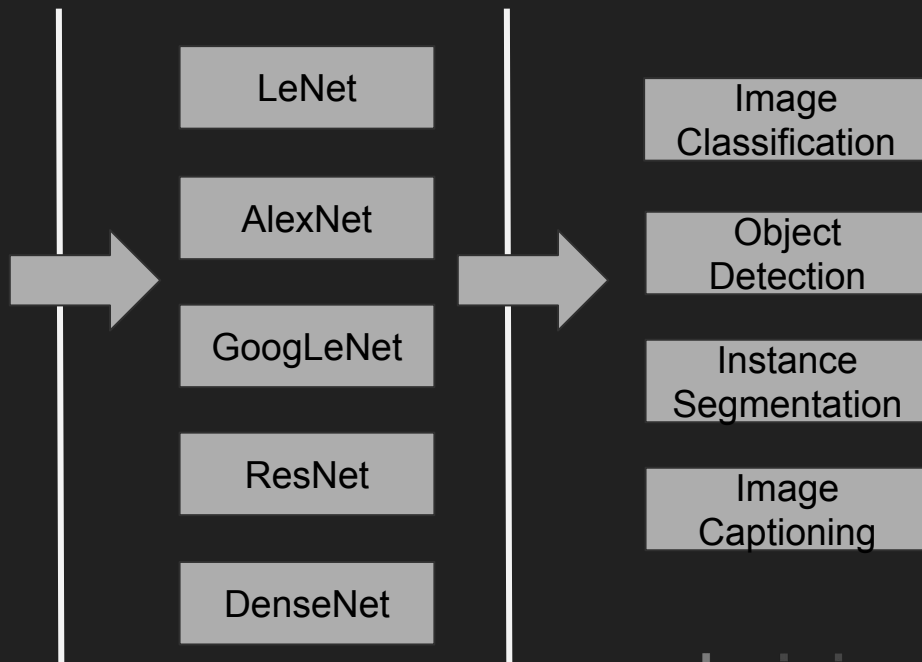
webvision

# Deep Learning Revolution -- This Years

WebVision 2020

w/o human annotation

The Same Big Dataset

| LeNet |
| AlexNet |
| GoogLeNet |
| ResNet |
| DenseNet |

| Image Classification |
| Object Detection |
| Instance Segmentation |
| Image Captioning |

webvision
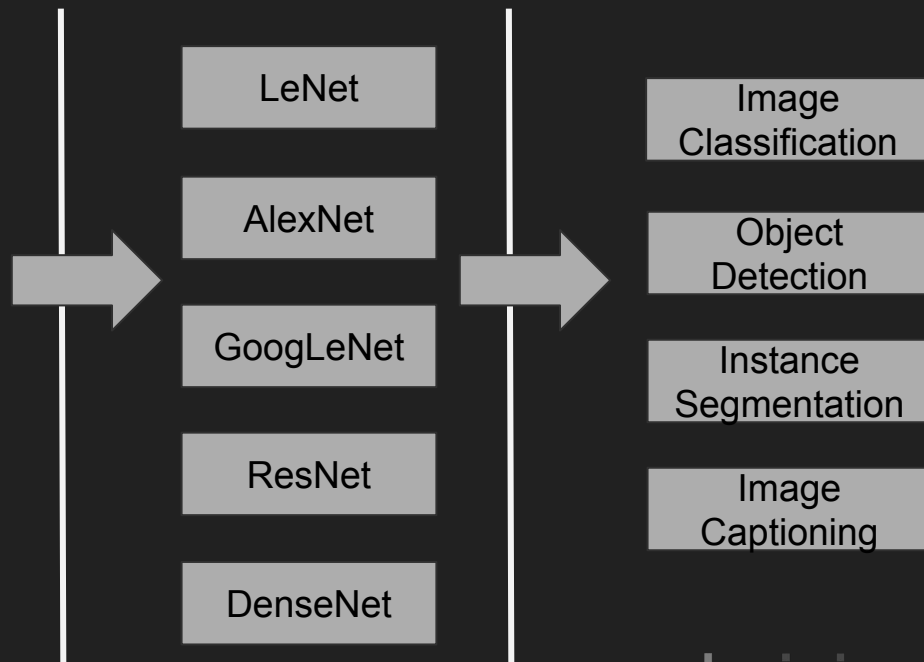
# Supervision using noisy & weak web signals



Internet

*keyword based search*

Training Data

Classifier

**No human annotation is used**

# Learning from Web Data

## Advantages

➢ No human annotation is needed for images
➢ Coarse semantic annotation generated from search engine or social signals
➢ Large number of images and classes
➢ High diversity (multiple sources)

## Challenges

➢ Noisy Labels
➢ Use of meta-information
➢ Domain adaptation issue

# Learning from Web Data

Recent Advances

1. Z. Wei et al. Learning Visual Emotion Representations From Web Data. In CVPR 2020
2. Y. Tu et al. Learning From Web Data With Self-Organizing Memory Module. In CVPR 2020.
3. D. Mahajan et al. Exploring the Limits of Weakly Supervised Pretraining. In arxiv, 2018.
4. C. Sun et al. Revisiting Unreasonable Effectiveness of Data in Deep Learning Era. In ICCV 2017.
5. Y. Li et al. Learning from noisy labels with distillation. In ICCV 2017.
6. A. Veit et al. Learning From Noisy Large-Scale Datasets With Minimal Supervision. In CVPR 2017.
7. A. Joulin et al.  Learning Visual Features from Large Weakly Supervised Data. In ECCV 2016.
8. S. Azadi et al. Auxiliary image regularization for deep cnns with noisy labels. In ICLR 2016.
9. X. Chen and A. Gupta. Webly supervised learning of convolutional networks. In ICCV 2015.
10. T. Xiao et al. Learning from Massive Noisy Labeled Data for Image Classification. In CVPR 2015.
11. S. Sukhbaatar et al. Training convolutional networks with noisy labels. In ICLR 2015.
(and many more…)

**Lots of work but hard to compare methods & quantify progress in the field.**

**Need for a common dataset and challenge.**

# Workshop Contributions

WebVision 2.0 dataset

- 5,000 categories
- 16M internet images
- 290K validation images
- 290K test images

WebVision Challenge

- WebVision Image Classification Track

# Our Vision for WebVision

- **Understand** deep learning from web data by enabling direct comparisons to methods that trained on ImageNet data.
- **Facilitate** research on handling the challenges of learning from web data, e.g., label noise, class imbalance, meta-information
- **Unite** the research community to solve those challenges

# Program Schedule

| | | | | |
|---|---|---|---|---|
| 9:00 | Opening Remarks | | 10:15 | Paper Session (ID 1-3) |
| **9:10** | **Dataset/Challenge Overview** | | 10:30 | Live Q&A Session |
| 9:30 | Participant Presentation by Huawei | | 10:36 | Paper Session (ID 4-6) |
| 9:40 | Participant Presentation by Tencent | | 10:51 | Live Q&A Session |
| 9:50 | Participant Presentation by Pcitech | | 11:00 | Award Session & Closing Remarks |
| 10:00 | Live Q&A Session | | | |

webvision
webvision