

WebVision 2017

Visual Understanding by
Learning from Web Data



web
webvision

Workshop Organizers

General Chairs



J. Berent



A. Gupta



R. Sukthankar



L. Van Gool

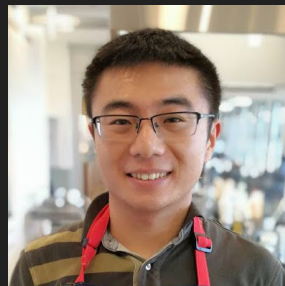
Program Chairs



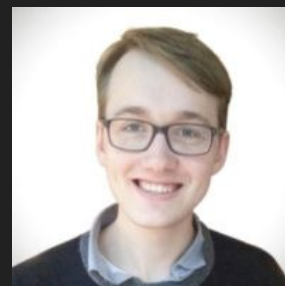
Wen Li



Limin Wang



Wei Li



E. Agustsson

Thanks to Workshop Sponsors & Collaborators



Dataset Collection & Challenge Hosting



Sponsor for Challenge and Award
Collaborator in Challenge Organization



Collaborator in Challenge Organization

Program Schedule

8:30	Opening Remarks
8:40	Invited Talk, Chen Sun (Google Research)
9:30	Challenge Overview
10:00	Coffee Break
10:20	Participant Presentation (Malong AI Research)
10:40	Participant Presentation (SHTU_SIST)
11:00	Poster Session

14:00	Invited Talk, Lamberto Ballan (Stanford & U. Padova)
14:50	Participant Presentation (VISTA)
15:10	Coffee Break
15:30	Participant Presentation (CRCV)
15:50	Invited Talk, Olga Russakovsky (Princeton)
16:40	Award Session
16:55	Closing Remarks

Deep Learning Revolution

Revolutionizing almost all fields of computer vision

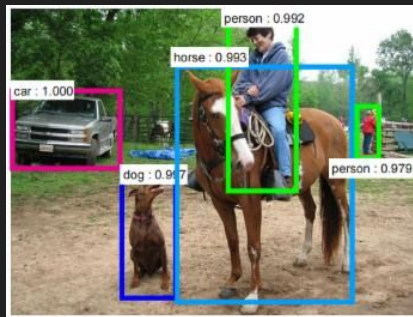
LeNet

AlexNet

GoogLeNet

ResNet

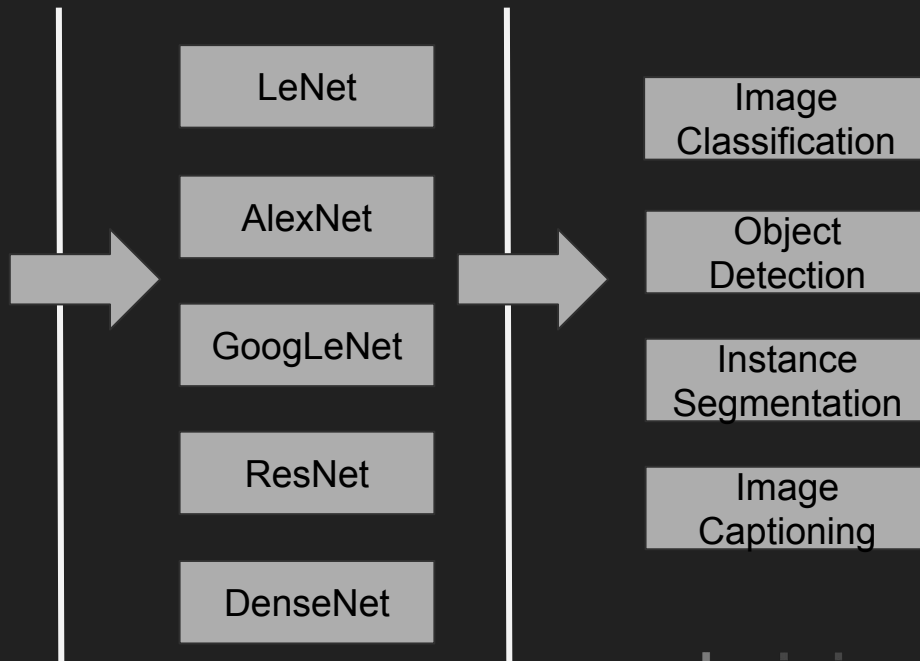
DenseNet



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

Deep Learning Revolution

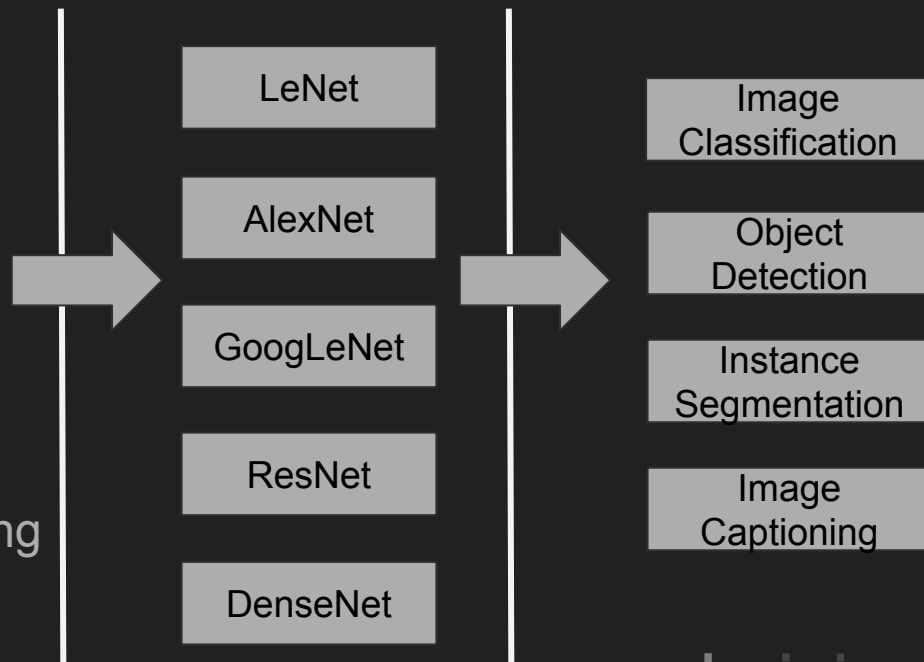
Powered by *human annotated* big data



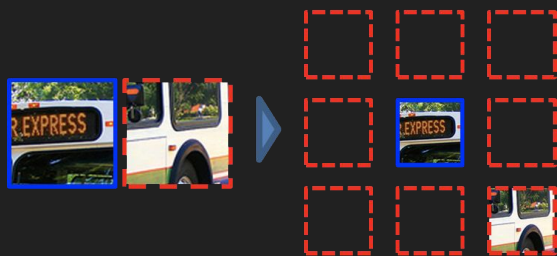
Deep Learning Revolution -- The Hope

Big Data
w/o human
annotation

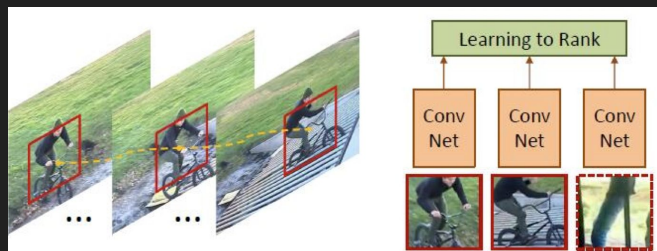
Can we get equivalent performance using
{self, weakly, un}supervised methods?



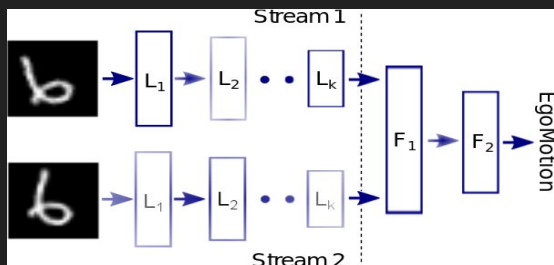
Self-supervised/Unsupervised Learning (Examples)



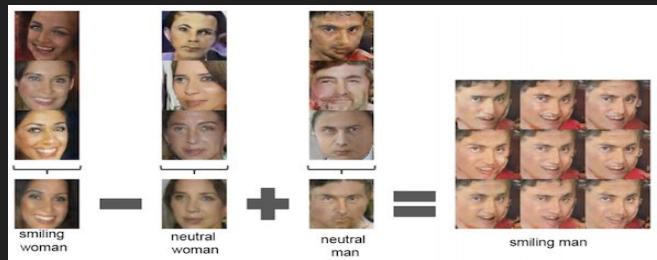
Doersch'15



Wang'15



Agrawal'15



Raford'16

Supervision using noisy & weak web signals



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
- Domain Adaptation Issue
- Meta Information

**Lots of work but hard to compare methods & quantify progress in the field.
Need for a common dataset and challenge**

Workshop Contribution

WebVision benchmark dataset

- 1000 categories
- Flickr & Google
- 2.4M images
- 50,000 validation images
- 50,000 test images

WebVision Challenge

- WebVision Image Classification Track
- PASCAL VOC Transfer Learning 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

8:30	Opening Remarks
8:40	Invited Talk, Chen Sun (Google Research)
9:30	Challenge Overview
10:00	Coffee Break
10:20	Participant Presentation (Malong AI Research)
10:40	Participant Presentation (SHTU_SIST)
11:00	Poster Session

14:00	Invited Talk, Lamberto Ballan (Stanford & U. Padova)
14:50	Participant Presentation (VISTA)
15:10	Coffee Break
15:30	Participant Presentation (CRCV)
15:50	Invited Talk, Olga Russakovsky (Princeton)
16:40	Award Session
16:55	Closing Remarks