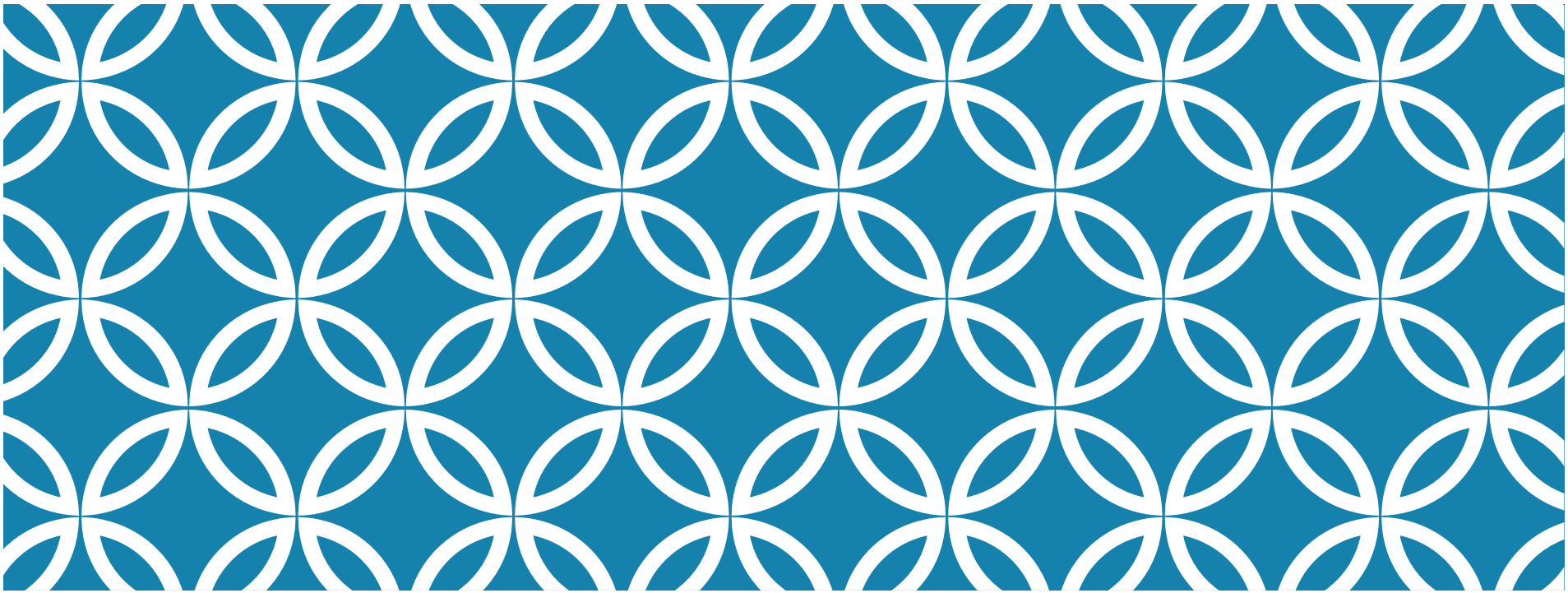


CAP6412 ADVANCED COMPUTER VISION

MUBARAK SHAH
HEC 245

Mubarak Shah
HEC 245

CAP6412

- **Instructor**: Dr. Mubarak Shah
- **Email:** shah@crcv.ucf.edu
- **Office**: HEC 245
- **Phone**: 4078235077
- **Time**: Tuesdays and Thursdays 3:00 to 4:15PM
- **Location:** HEC 117
- **Office Hours:** Tuesday 4:15 to 5:00PM; Thursdays 2:00 to 3:00PM and by appointment
- **Extra Discussion Session**: Wednesdays 4:30 to 5:30, HEC 356
- Pre-requisite: CAp5415
- Course webpage: http://crcv.ucf.edu/courses/CAP6412/Spring2018/

COURSE OBJECTIVES

To expose graduate students to the cutting-edge research. In each class we will discuss one recent research paper related to active areas of current research in particular employing Deep Learning.

STUDENT LEARNING OUTCOMES

Read and understand a research paper.

Write a comprehensive review of the paper.

To identify strong and weak points of the paper.

To come up with own ideas to solve the same problem, which may lead to their first research paper.

To implement known method or work on and successfully complete individual project.

GRADING POLICY

Reports	20%
Presentation	10%
Attendance and Discussion	20%
Projects/Programs	50%

Late Policy

- 0 for late reports
- Projects/Programs
  - 20% off per day
  - up to 4 days

REPORTS (ONE PAGE)

Summary

Good points

Weak points

Questions

Ideas

STATEMENT OF ACADEMIC INTEGRITY

The UCF Golden Rule (http://goldenrule.sdes.ucf.edu/) will be observed in the class. Plagiarism and

Cheating of any kind on an examination, quiz, or assignment will result at least in an "F" for that assignment (and may, depending on the severity of the case, lead to an "F" for the entire course) and may be subject to appropriate referral to the Office of Student Conduct for further action. I will assume for this course that you will adhere to the academic creed of this University and will maintain the highest standards of academic integrity. In other words, don't cheat by giving answers to others or taking them from anyone else. I will also adhere to the highest standards of academic integrity, so please do not ask me to change (or expect me to change) your grade illegitimately or to bend or break rules for one person that will not apply to everyone.

TEXT BOOK

There is no text book for this class. We will discuss recent research papers.

Recommended supplemental textbook:

- Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning. (free electronic version)

RECOMMENDED ONLINE COURSES AND TUTORIALS

http://cs231n.stanford.edu/ CS231n: Convolutional Neural Networks for Visual Recognition

http://web.stanford.edu/class/cs224n/ CS224n: Natural Language Processing with Deep Learning

http://rll.berkeley.edu/deeprlcourse/ CS 294: Deep Reinforcement Learning

http://distill.pub/ Very nice explanations of some DL concepts

https://class.coursera.org/ml003/lecture/preview

https://media.nips.cc/Conferences/2016/Slides/6203-Slides.pdf

https://media.nips.cc/Conferences/2016/Slides/6198-Slides.pdf

https://adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html

https://github.com/adeshpande3?tab=repositories

HOW TO READ A RESEARCH PAPER?

You have to read the paper several times to understand it.

- When you read the paper first time, if you do not understand something do not get stuck, keep reading assuming you will figure out that later.
- When you read it the second time, you will understand much more, and the third time even more ...

HOW TO READ A RESEARCH PAPER?

Try first to get a general idea of the paper

- What problem is being solved?
- What are the main steps?
- How can I implement the method?, even though I do not understand why each step is performed the way it is performed?

Try to relate the method to other methods you know, and conceptually find similarities and differences.

HOW TO READ A RESEARCH PAPER?

In the first reading it may be a good idea to skip the related work.

Do not use dictionary to just look up the meaning of technical terms like

Try to understand each concept in isolation, and then integrate them to understand the whole paper.

COMPUTER VISION CONFERENCES AND JOURNALS

- Conferences
  - International Conference on Computer Vision (ICCV)
  - Computer Vision and Pattern Recognition (CVPR)
  - European Conference on Computer Vision (ECCV)
- Journals
  - IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)
  - International Journal of Computer Vision (IJCV)
  - Computer Visio and Image Understanding (CVIU)

MACHINE LEARNING CONFERENCES

- Neural Information Processing Systems (**NIPS**)
- International Conference on Machine Learning (**ICML**)
- International Conference on Learning Representations (**ICLR**)

DEEP LEARNING HAS BEEN DISRUPTIVE

Very Different Paradigm

Real Learning

Excellent results

Rapid Progress

Computer Vision is impacting other areas

DEEP LEARNING HAS BEEN DISRUPTIVE

Dramatic Increase in Number of

- Publications
- Attendance in Conferences
- Datasets
- Startups
- Academics moving to Industry
- Software platforms/libraries
  - Café, Keras, Tensor Flow, Chainer, ..
- GPUs
- Tutorials, videos, online courses

DEEP LEARNING HAS BEEN DISRUPTIVE

- Computer Vision is impacting other areas
  - Natural Language Understanding
  - Robotics
  - Computer Graphics
  - Audio/Sketches
  - Alpha Go

DEEP LEARNING HAS BEEN DISRUPTIVE

- Learning
- Supervised
- Semi-Supervised
- Weakly-Supervised
- Unsupervised
- Self Supervised
- Reinforcement

MAIN THEMES

GAN: Generative- Adversarial Network

Reinforcement Learning

Transfer Learning/Domain Adaptation

Multi-modal Analysis

End-to-End Learning

Bayesian Deep Learning

PAPERS

SOME OF OUR RECENT RESEARCH

Semantic Segmentation (**ICCV-17**)

Deep Learning Human Mind for Automated Visual Classification (**CVPR-17**)

Generative Adversarial Networks Conditioned by Brain Signals (**ICCV-17**)

T-CNN for Action Detection in Videos (**ICCV-17**)

Improving Facial Attribute Prediction using Semantic Segmentation (**CVPR-17**)

Video Fill In the Blank using LR/RL LSTMs with Spatial-Temporal Attentions (**ICCV-17**)



Center for Research in Computer Vision

UNIVERSITY OF CENTRAL FLORIDA

Semi Supervised Semantic Segmentation Using Generative Adversarial Network

Nasim Souly, Concetto Spampinato and Mubarak Shah

ICCV 2017

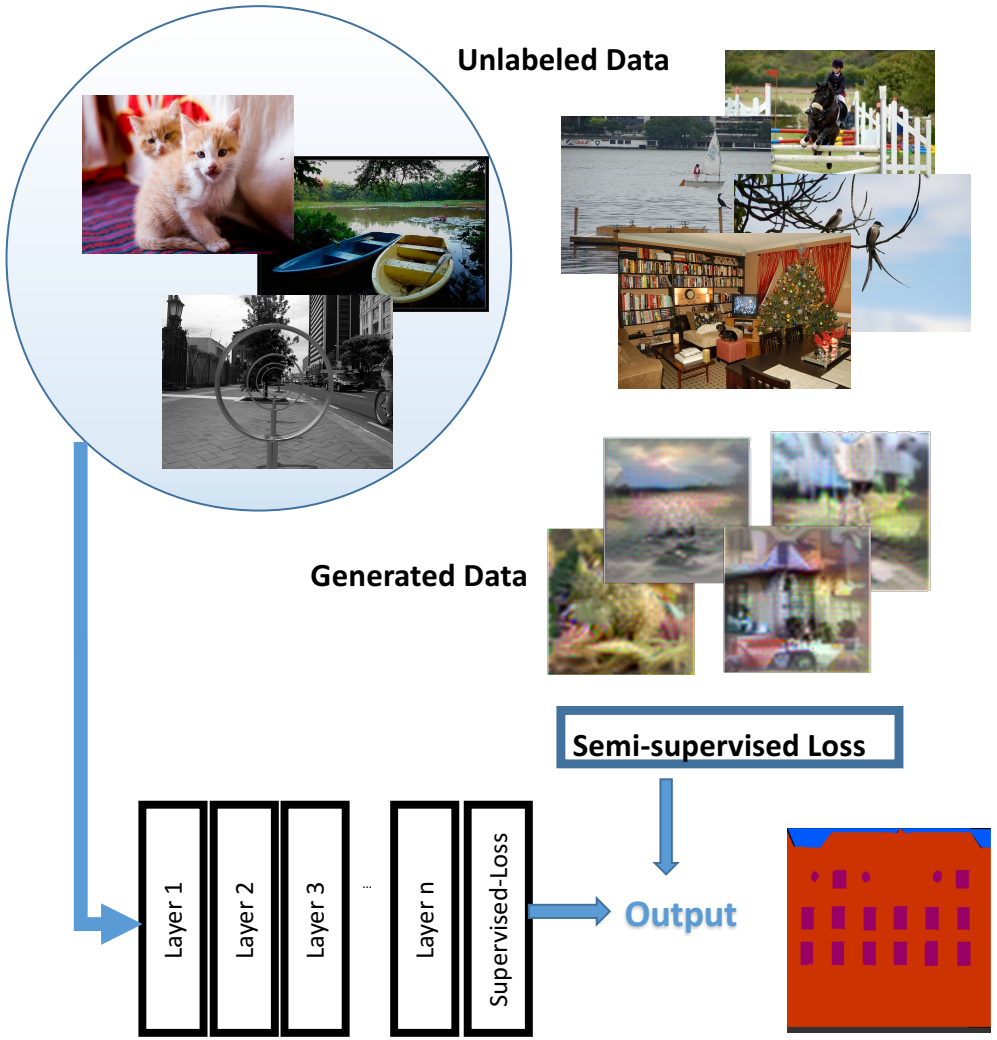
SEMANTIC SEGMENTATION (SCENE LABELLING)

Assigning a semantic label to each pixel of an image.



Motivation

- Lack of enough annotated data
- Plentiful unlabeled data
- Use generative model to improve classifiers



SEMI SUPERVISED LEARNING (SSL)

Halfway between supervised and unsupervised learning

Data points lying on the same feature manifold are more expected to be classified into the same class

Leverage the unlabeled data to find this structure.

Cost function for SSL

$$\boldsymbol{Loss} = \sum_{\boldsymbol{n=1}}^{\boldsymbol{Nl}} \boldsymbol{Loss_l}(\boldsymbol{y_n}, \boldsymbol{x_n}) + \boldsymbol{w} \sum_{\boldsymbol{n=1}}^{\boldsymbol{Nu}} \boldsymbol{Loss_u}(\boldsymbol{x_n})$$

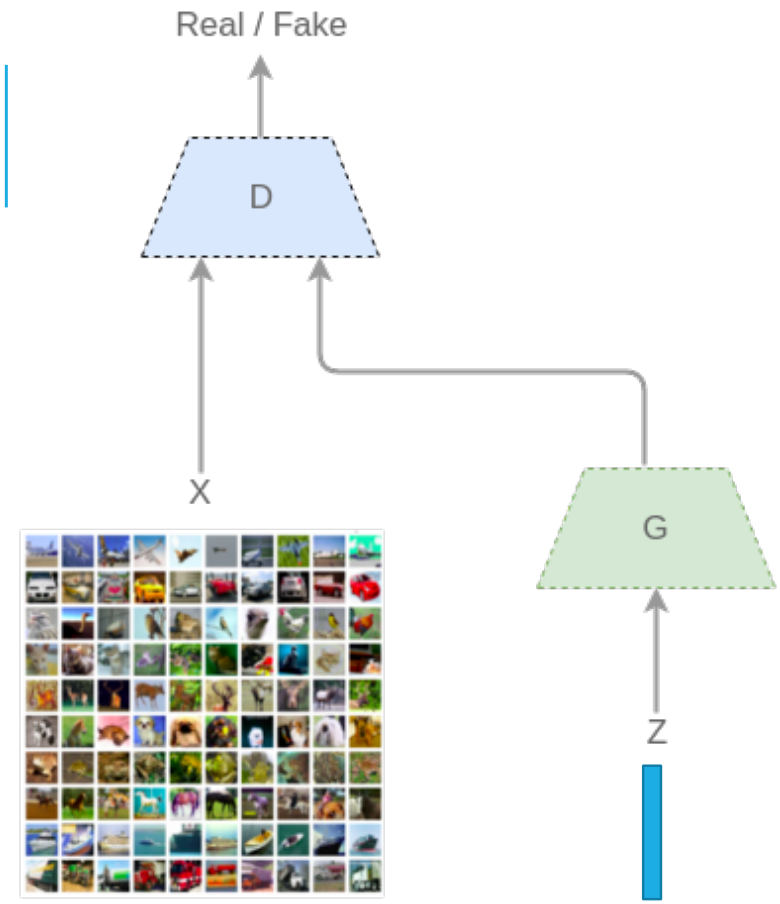
GENERATIVE ADVERSARIAL NETWORK

Enables models to tackle unsupervised learning

The intuitive idea:

- A painter who wants to do art forgery (*G*), (of Picasso)
- Someone is judging paintings (*D*)
- Then *G* produces paintings in an attempt to fool *D*
- *D* starts learning more about Picasso, *G* has a harder time fooling *D*
- *D* gets really good in telling apart what is Picasso and what is not?
- *G* gets really good at forging Picasso paintings

From Kdnuggets http://www.kdnuggets.com



Images from dataset

Noise

Generated Images



GAN

Constant competition between two networks :

- a **generator** (*G*) and
- **discriminator** (*D*).

G starts from some noise, *z*, generate images *G(z)*.

D takes images from the distribution (real) and fake (from *G*) and classifies them: *D(x)* and *D(G(z))*.

$$\min_G \max_D V(D,G) = \mathbb{E}_{x \sim p_{data}(x)}[log(D(x))] + \mathbb{E}_{z \sim p_z(z)}[log(1 - D(G(z))]$$

SEMI SUPERVISED LEARNING USING GANS

Labels are not available for all training images,

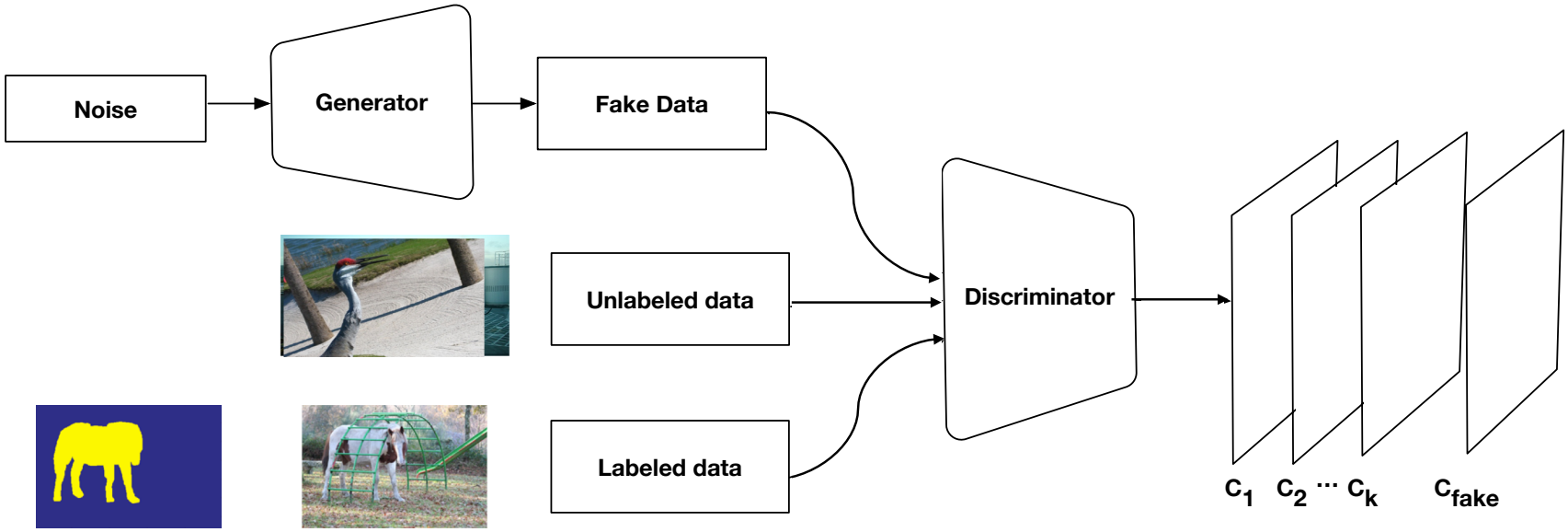
- leverage the unlabeled data by estimating a proper prior.

This prior is used by a classifier to improve.

In GAN :

- Unlabeled data belongs to the same distribution of labeled data
- Generated (fake) data does not.

SEMI SUPERVISED LEARNING USING GANS



QUANTITATIVE RESULTS

StanfordBG

method	pixel accuracy	mean accuracy	mean IU
Standard [15]	73.3	66.5	51.3

- **CamVid**

method	pixel accuracy	mean accuracy	mean IU
SegNet(Basic) [1]	82.2	62.3	43.6

STANFORD BG

Image











QUANTITATIVE RESULTS: PASCAL VOC 2012

Using all fully labeled and unlabeled data in train set.

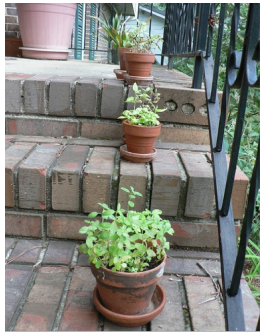
method	pixel accuracy	mean accuracy	mean IU
Fully supervised	90.3	75.9	62.2

Using 30% of fully labeled data and all unlabeled data in train set.

method	pixel accuracy	mean accuracy	mean IU
Fully supervised	83.15	53.1	38.9

QUALITATIVE RESULTS: VOC 2012

Image





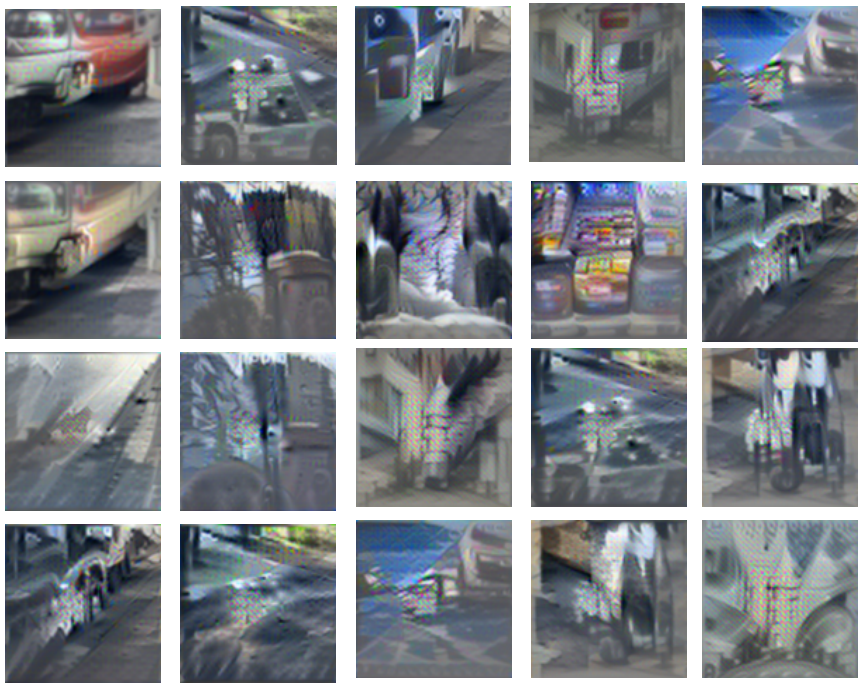


Human
Bicycle
MotorBike
Bottle
Cat
Potted plant
Aeroplane
Bus

GENERATED IMAGES SIFTFLOW



GENERATED IMAGES FROM CAMVID



GENERATED IMAGES

