Sunnybrook ML Journal Club

Topic: Generative Adversarial Networks (GANs)
March 19, 2018

David Burns

Orthopaedic Surgery Resident
PhD(c) with Orthopaedic Biomechanics Lab, IBBME
University of Toronto

Supervisor: C. Whyne

Paper

SYNTHETIC DATA AUGMENTATION USING GAN FOR IMPROVED LIVER LESION CLASSIFICATION

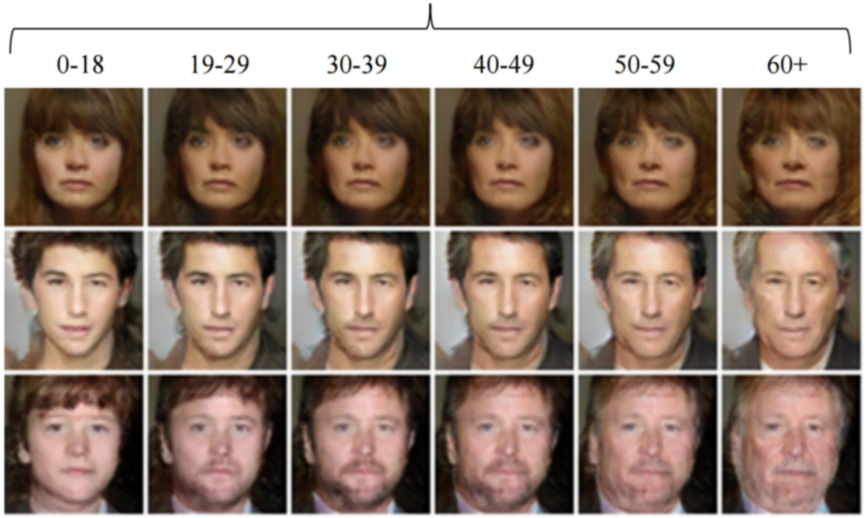
Maayan Frid-Adar[1] *Eyal Klang*[2] *Michal Amitai*[2] *Jacob Goldberger*[3] *Hayit Greenspan*[1]

[1]Department of Biomedical Engineering, Tel Aviv University, Tel Aviv, Israel.
[2]Department of Diagnostic Imaging, The Chaim Sheba Medical Center, Tel-Hashomer, Israel.
[3]Faculty of Engineering, Bar-Ilan University, Ramat-Gan, Israel.

Why Generative Models

- Generative models are poorer discriminators than discriminative models
- Synthetic data!
  - "What I cannot create, I do not understand" - Richard Feynman

Face Aging



Style Transfer

Background: GANs

- Introduced by Ian Goodfellow et al in 2014 [2]
- **Motivation in 2014**
- Successful deep learning applications
  - Discriminative models
  - Supervised learning framework
  - Mapping rich high dimensional input (e.g. image) to class label
- Deep generative models
  - Model estimation techniques for deep learning (backprop) depend on piece-wise linear units (with a well defined gradient)
  - Unclear how to leverage this in a generative context

Generative Models

Probabilistic Discriminative models

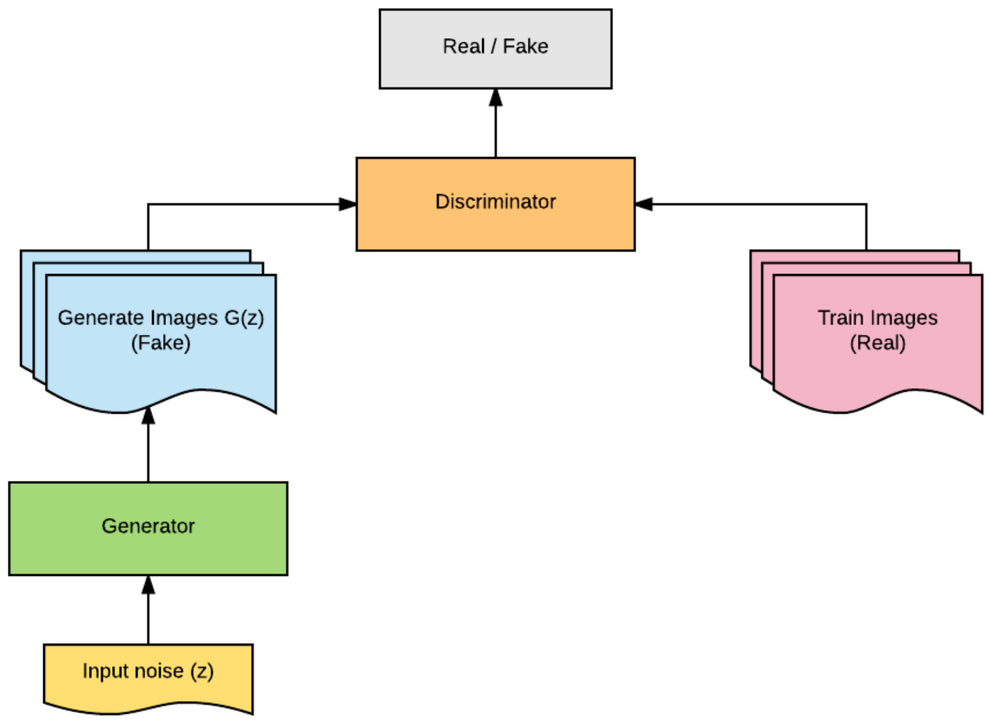
$$p\left(y \middle| \mathbf{x};\boldsymbol{\theta}\right)$$

Generative Models

$$p\left(\mathbf{x},y;\boldsymbol{\theta}\right)$$

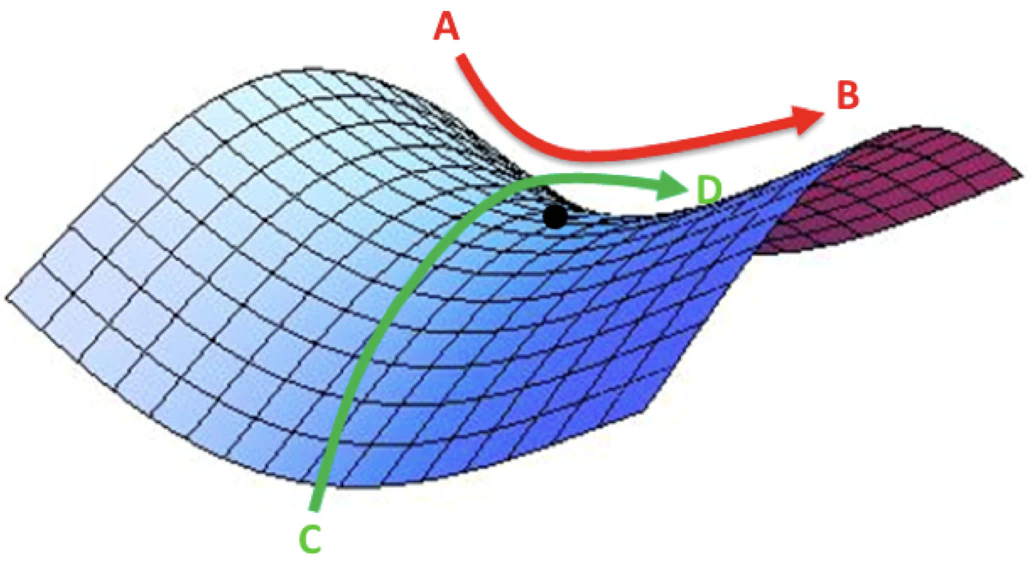
GAN Training Architecture

- Generative machine
  - Does not model likelihood p(**x,** y) directly
  - But can generate samples from the distribution
- Generator maps probabilistic input to data samples
- Discriminator discriminates between real and generated samples



GAN Minimax Optimization

- Single value function that one agent maximizes and one agent minimizes
- Global optimum is a saddle point
  - Minimum with respect to generator parameters
  - Maximum with respect to discriminator parameters
- Discriminator maxes
  - Probability of assigning correct label to both training samples (**x**) and generates samples G(**z**)
- Generator minimizes
  - Probability that discriminator correctly labels generated samples



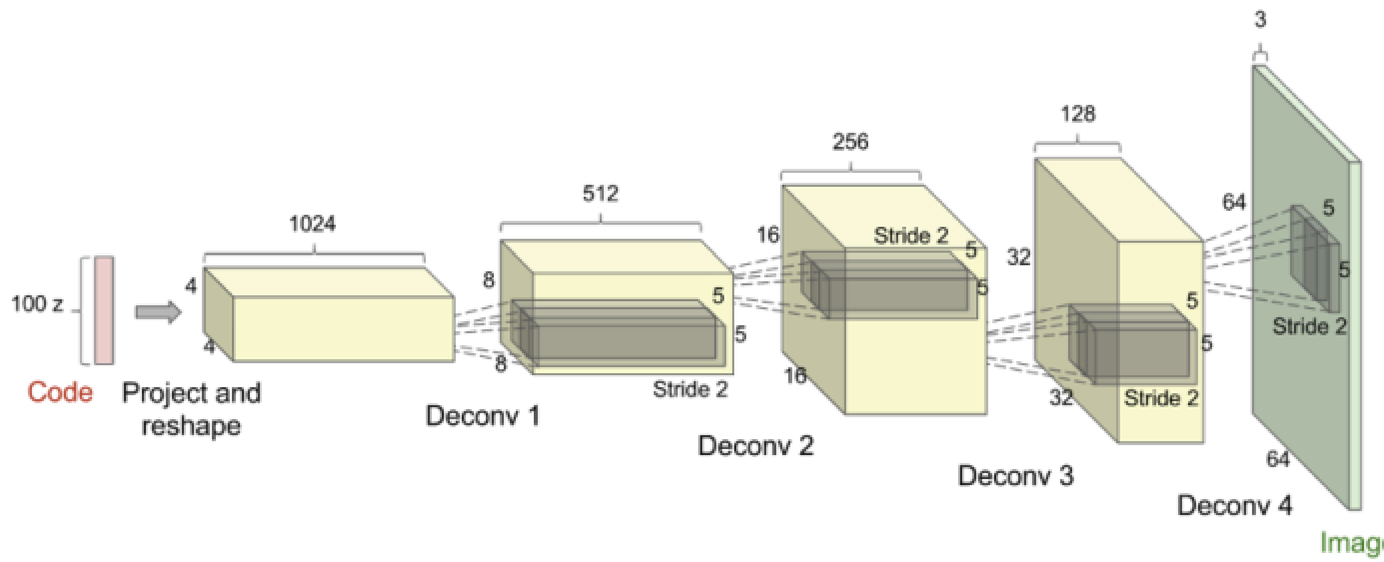
$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

GAN Optimization: In Practice

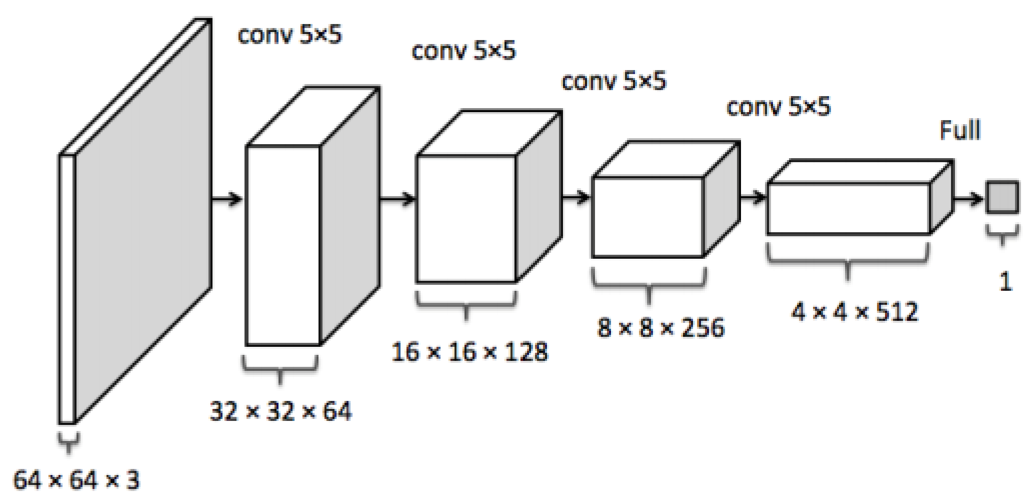
- Simultaneous application of mini-batch SGD
- k steps optimizing D, one step optimizing G
- Training stops when discriminator is unable to distinguish real and fakes: p(G) = p(x)
- Lots of tricks needed to help with convergence / stability:
  - https://github.com/soumith/ganhacks

Early GAN: DCGAN (Radford 2015)

Generator

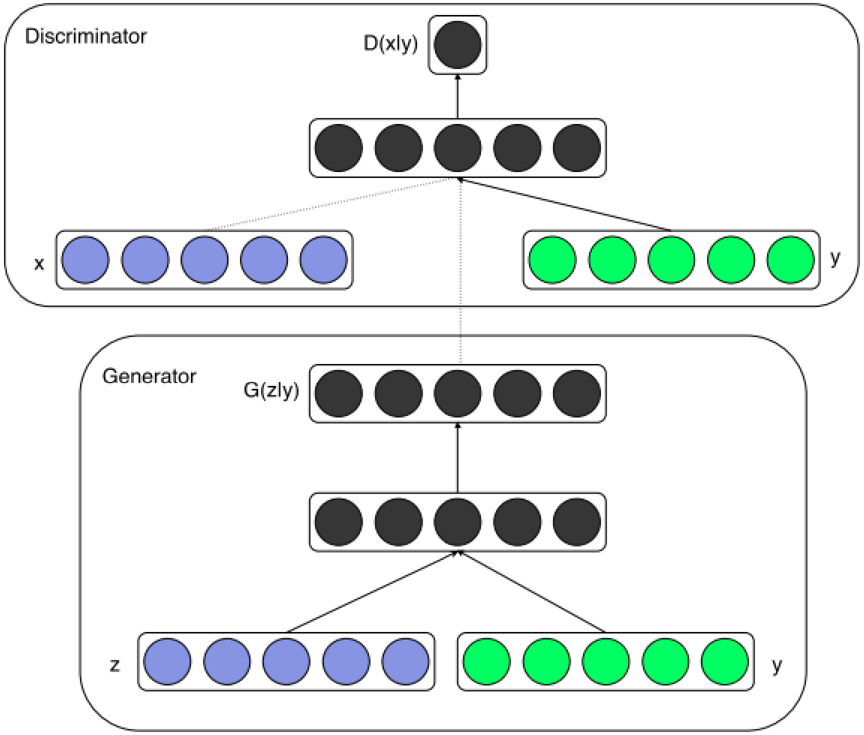


Discriminator



Common GAN Extensions

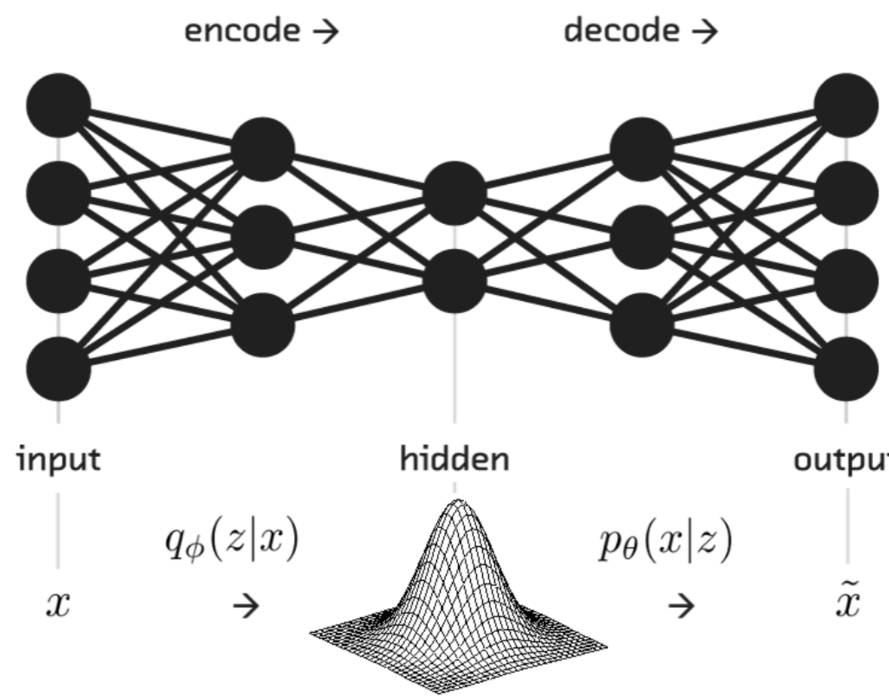
- Conditioned GAN (cGAN)
  - p(**x** | **c**) obtained by adding **c** as input to G and D
  - Text-to-image, image-to-image,
- Inference nets
  - Predict **z** given **x** or G(**z**)
- Semi-supervised learning
  - Features from inference or discriminator nets used to boost performance of classifiers



cGAN

Other Important Deep Generative Models

- Deep Generative auto-encoders
  - Variational Auto Encoders (VAEs)
  - Deep AutoRegressive Networks
    - Google Deep Mind
- Adversarial Auto Encoders
  - Goodfellow Group, Open AI



$$\widetilde{\mathcal{L}}^B(\boldsymbol{\theta}, \phi; \mathbf{x}^{(i)}) = -D_{KL}(q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x}^{(i)})||p_{\boldsymbol{\theta}}(\mathbf{z})) + \frac{1}{L}\sum_{l=1}^{L}(\log p_{\boldsymbol{\theta}}(\mathbf{x}^{(i)}|\mathbf{z}^{(i,l)}))$$

GAN vs VAE

- GAN
  - Optimized for generative task, better sample fidelity
  - Challenging optimization where D and G updates must be synchronized
  - No direct way to evaluate model quality, compare models (no likelihood)
- VAE
  - Easy optimization
  - Tractable likelihood
  - Sample images tend to be lower quality

Paper

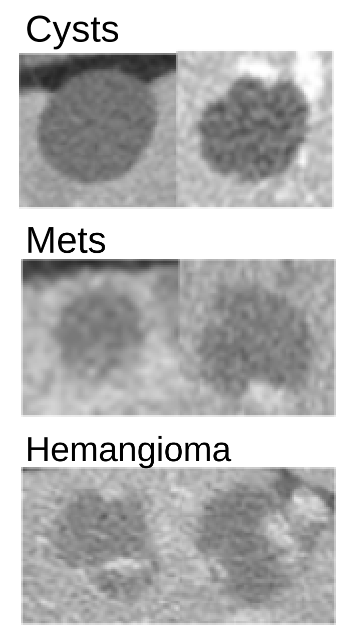
SYNTHETIC DATA AUGMENTATION USING GAN FOR IMPROVED LIVER LESION CLASSIFICATION

Maayan Frid-Adar[1] *Eyal Klang*[2] *Michal Amitai*[2] *Jacob Goldberger*[3] *Hayit Greenspan*[1]

[1]Department of Biomedical Engineering, Tel Aviv University, Tel Aviv, Israel.
[2]Department of Diagnostic Imaging, The Chaim Sheba Medical Center, Tel-Hashomer, Israel.
[3]Faculty of Engineering, Bar-Ilan University, Ramat-Gan, Israel.

Data Augmentation

- Goal
  - Improve CNN liver lesion classifier with synthetic data
  - Expensive dataset
- Data
  - 182 CT scans from 2009-2014
  - 53 cysts, 64 mets, 65 hemangiomas
  - 2D ROIs cropped by radiologist to 64x64



Experiment

- Data augmentation with classical transforms
  - Rigid (translation, rotation, reflections)
  - Scaling
  - No shear
  - 30,000 “Aug” samples created
- Synthetic data created using GAN
  - GAN trained with real and classically augmented data
  - “Synth” samples created

Experiment

- Effect of data augmentation on classifier performance measured over range of augmentation samples / approach
  - Real
  - Real+”Aug”
  - Real+”Synth

Classifier

Conv/Pool
Conv/Pool
Conv/Pool
Dense
Dense
Softmax

GAN Architecture (DCGAN)

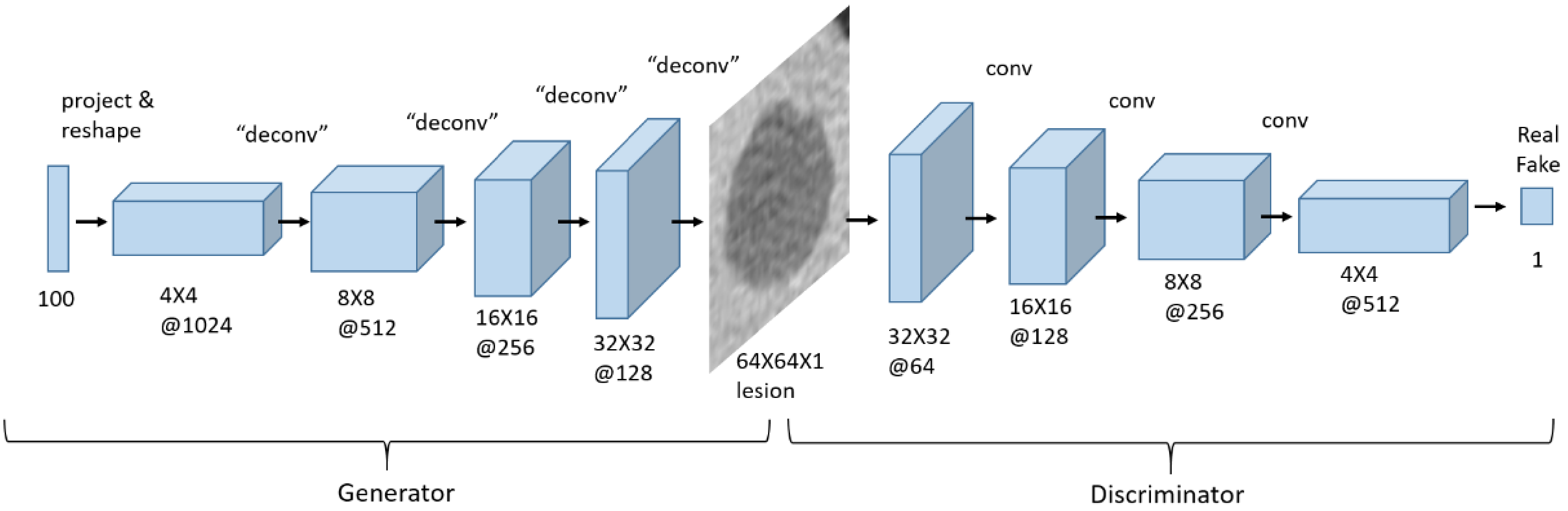
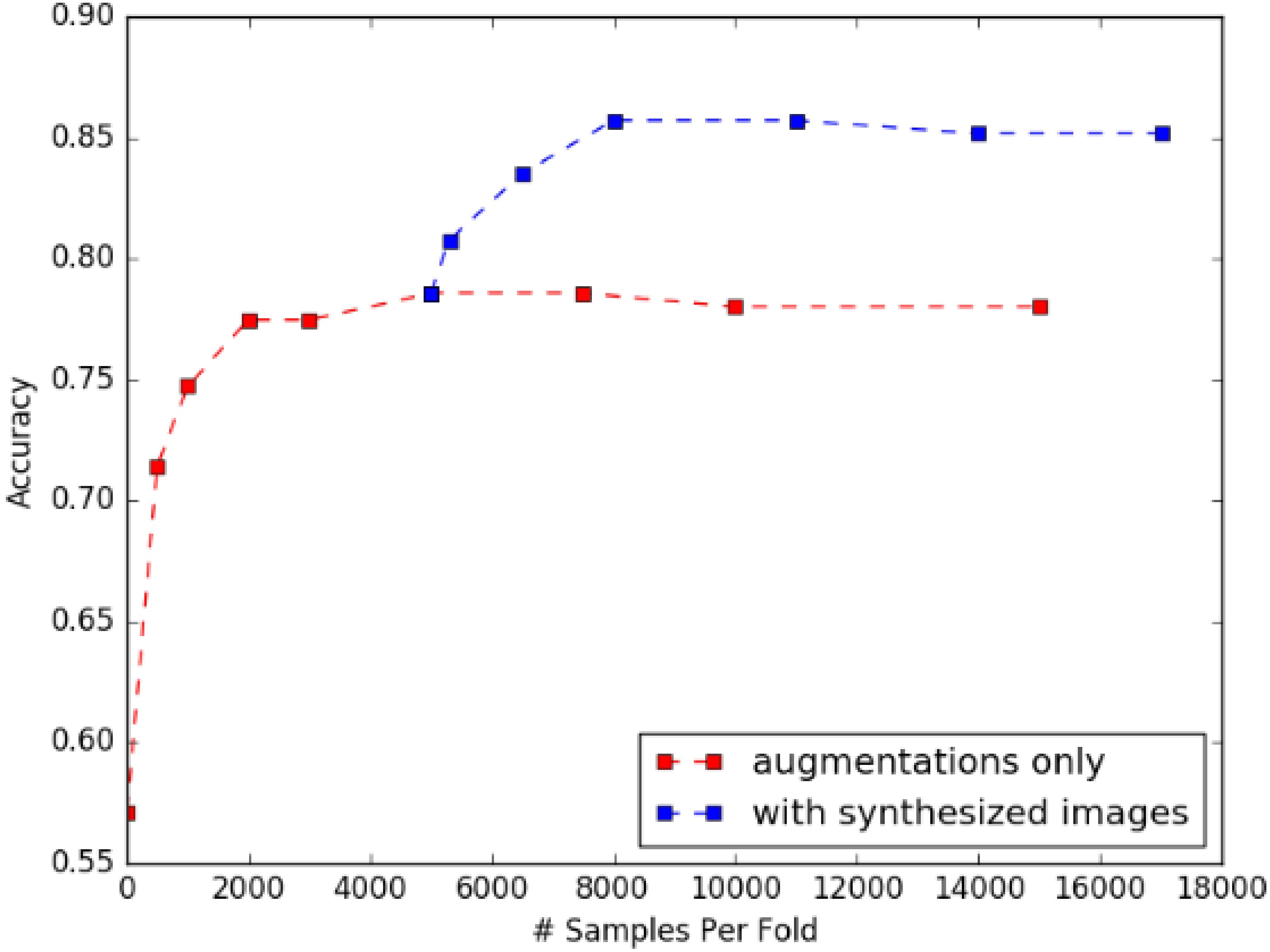


Fig. 2: Deep Convolutional GAN Architecture (generator+descriminator).

Implemented with tensorflow

Note open source tf implementations of DCGAN available online:
https://github.com/carpedm20/DCGAN-tensorflow

Results



0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
Accuracy
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
Samples Per Fold
augmentations only
with synthesized images

Radiologist Assessment of Synthesized Data

- Two radiologists attempted to classify synthetic and real ROIs
- Accuracy was similar but poor on both real and synthetic images (70%, 78%)

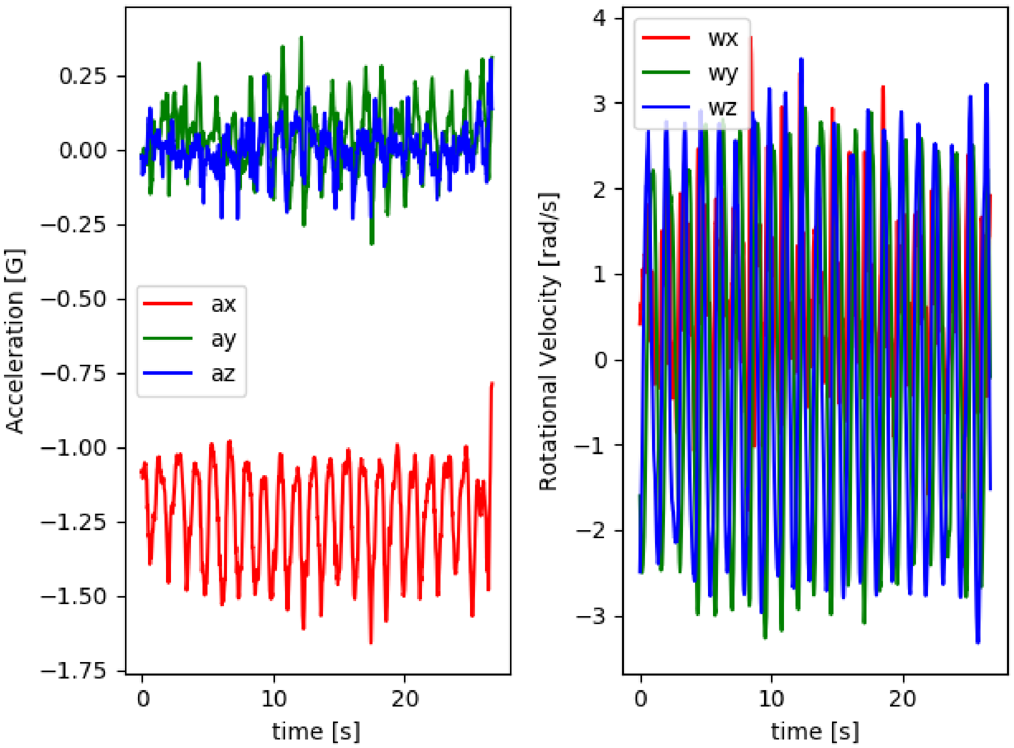
GANs in Medical Imaging

- CT Denoising
- Segmentation
- Anomaly detection
- Medical image synthesis
- Image reconstruction, compressed sensing
- Image classification

https://github.com/xinario/awesome-gan-for-medical-imaging

My Interest in GANs

- Shoulder physiotherapy exercise evaluation from inertial sensor data
- cGAN with upper extremity kinematic model
  - Correct class imbalance
  - Increase training samples



Key References

1) Frid-Adar et al. Synthetic Data Augmentation Using Gan for Improved Liver Lesion Classification. ArXiv:1801.02385. Jan 2018.
2) Goodfellow et al. Generative Adversarial Nets. NIPS proceedings 2014.
3) Radford et al. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (DCGAN). ArXiv:1511.06434. Nov 2015.
4) Kingma and Welling. Auto-Encoding Variational Bayes. ArXiv:1312.6114. Dec 2013.
5) Gregor et al. Deep AutoRegressive Networks.   ArXiv:1310.8499. Oct 2013.
6) Makhzani et al. Adversarial Autoencoders. ArXiv:1511.05644. Nov 2015.