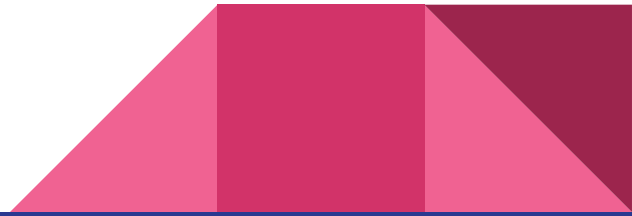
REU Update 2

Adam Vest

Refresher

- Problem definition
- Discussed several adversarial models:
  - GAN
  - AAE
  - ALI
  - AGE
- Additional Considerations
  - Datasets
  - Loss functions
  - Evaluation Metrics



Week in Review

- Read several papers on image-to-image translation, adversarial models
  - CycleGAN
- Designed initial architecture
- Made some decisions concerning datasets, investigated possible loss functions to boost performance
- Implemented current state-of-the-art SR model SRGAN

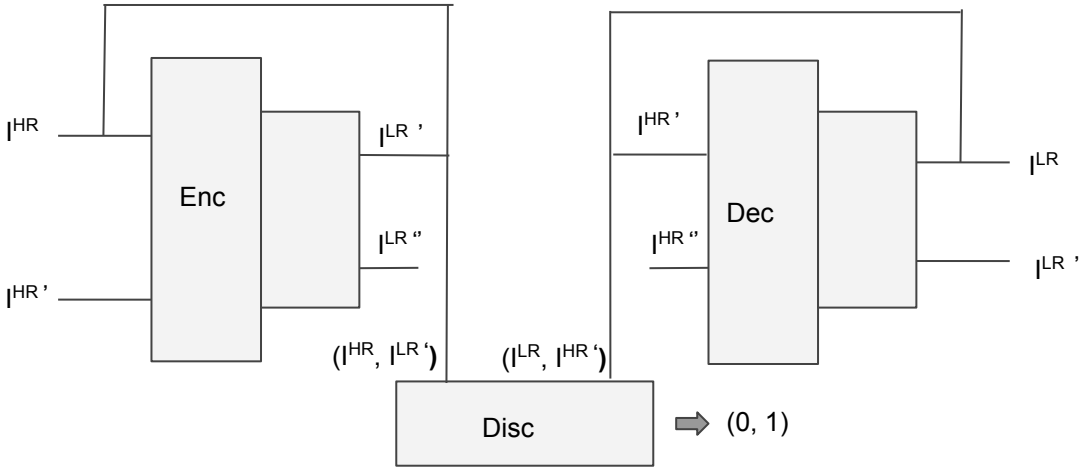


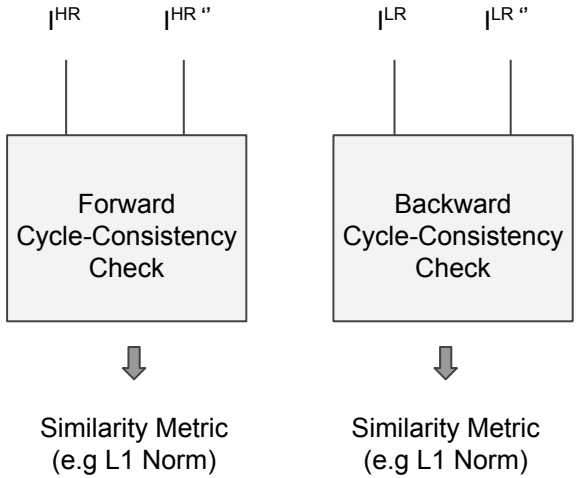
CycleGAN

- Model described in Zhu/Park et al. “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks”
- Utilizes two generators, two discriminators to map images from domain X to domain Y
- Each domain has its own generator/discriminator pair, where e.g. $G_x(y) = x \sim X$ and vice versa (i.e., each generator maps an image in one domain to an image in another)
- Cycle consistency check serves as a kind of reconstruction loss and forces the generators to be inverses of each other
    - Uses L1 Norm
    - Cycle consistency check ensures that $G_x(G_y(x)) = x'$ is approx. x

Proposed Model

- Combination of CycleGAN and ALI





Model Notes

- Loss functions used would be the adversarial loss, cycle consistency loss, and perhaps some other content loss (MSE, VGG, L1, etc.)
- Decoder could have the same structure as the SRGAN generator, encoder would have the “inverse” of this structure
- Might have difficulty using a single discriminator to handle a tuple of images, could break into two discriminators as in CycleGAN
- Can we integrate Patch-based discriminators and WassesteinGAN?

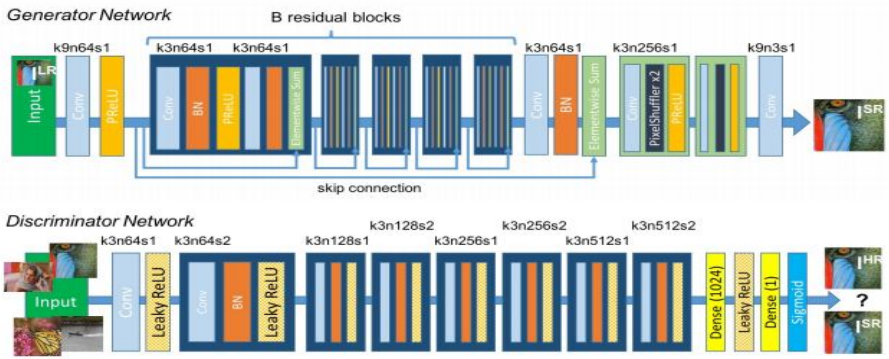
Miscellaneous

- Basic training pipeline:
  - Subset of Imagenet (~350,000 images used in SRGAN)
  - Take several 128x128 random crops of each image, downsize those subimages using bicubic interpolation to 32x32
- Evaluation Metrics still an open question (likely PSNR + some qualitative analysis for ease/speed)



SRGAN Implementation

- Implemented the full SRGAN model from scratch based on Twitter paper
- Some differences:
  - Use 128x128 image crops instead of 96x96
  - Trained from scratch instead of using Twitter's SRResNet to initialize generator weights
  - Had to modify a couple of hyperparameters (adversarial weight loss and discriminator learning rate) due to the above change



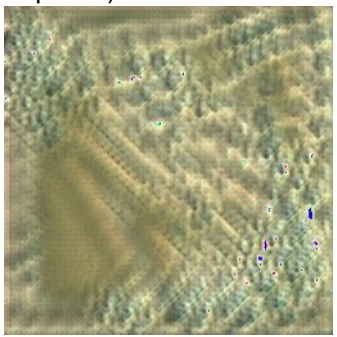
SRGAN Implementation Cont...

- Trained on BSD100
- Ran for 3 epochs using VGG loss, 1 epoch using MSE loss
- Results:

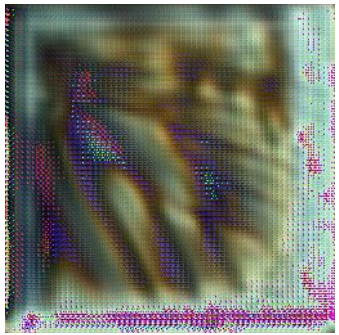
Ground Truth HR



SR (Vgg Loss, 3 epochs)

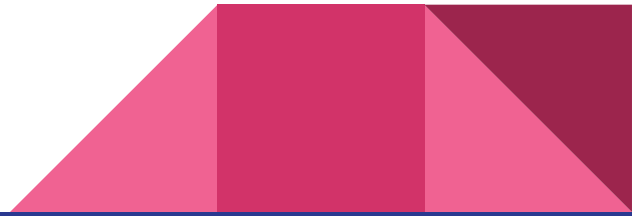


SR (MSE Loss, 1 epoch)



Goals for the Next Week

- Implement and evaluate model prototype
- Download ImageNet?



Thank you! Questions?