Super Resolution WGAN Model for Low Resolution Images & GIFs

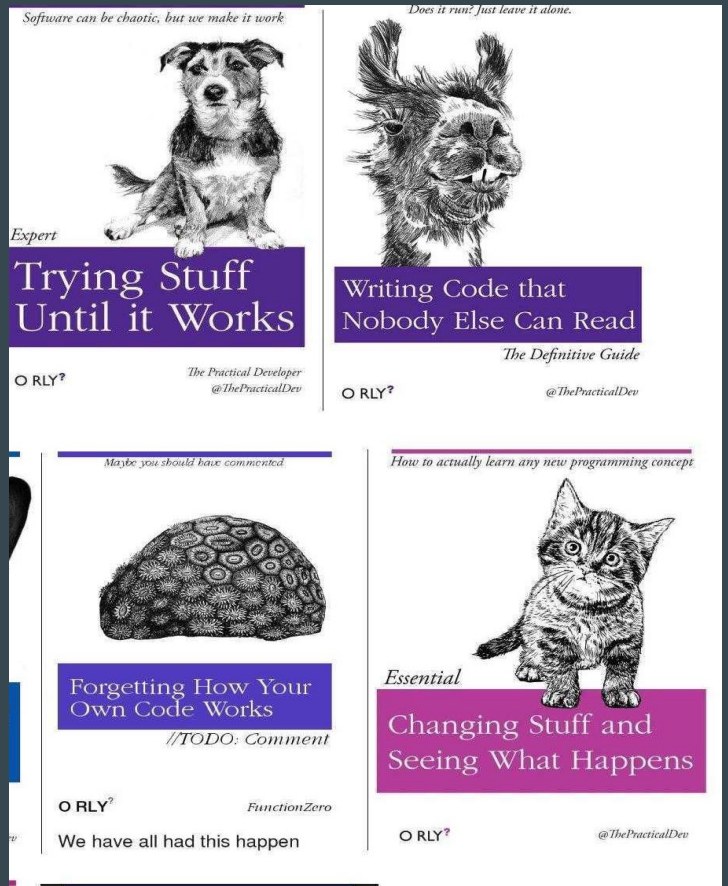
Honglin Wu
Sang Uk Park
Shilai Yang
Yash Bafna

CSCI 2470

Everyone be like:

...

Def not us:



Introduction and Aim of the Project

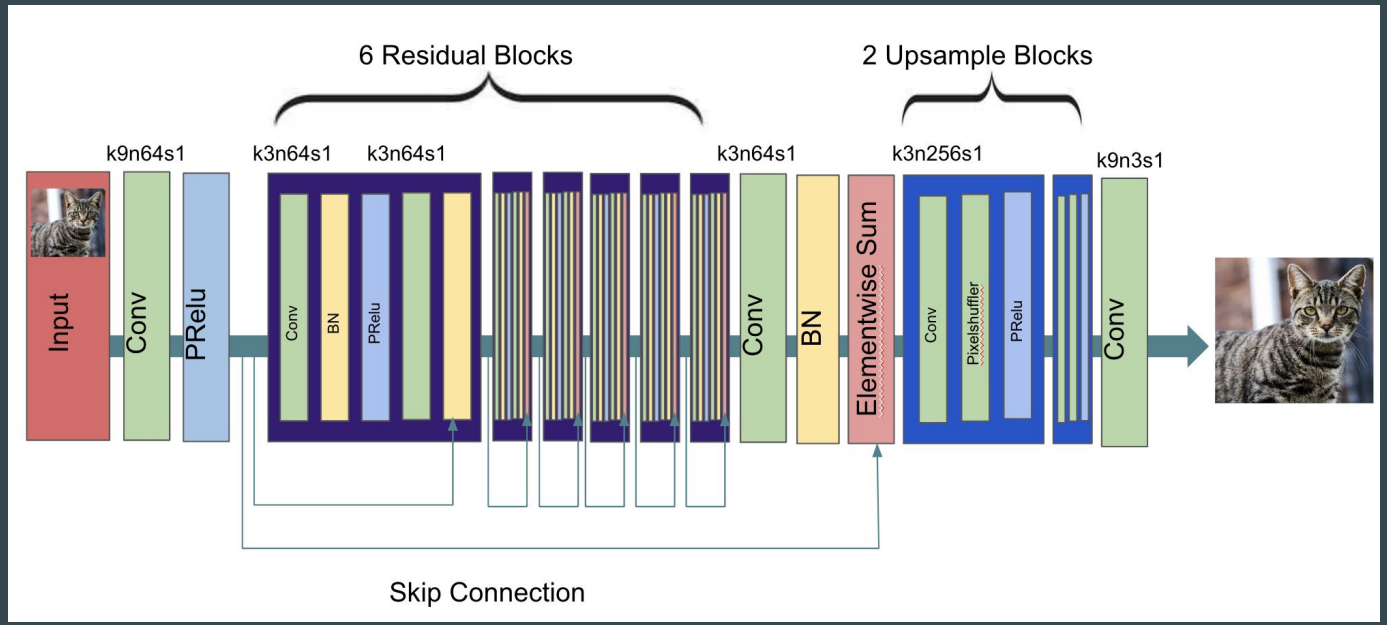
- Leverage SRGAN & WGAN to increase the resolution of an image by a factor of four or more
- What is SRGAN (Super Resolution GAN)?
  - A variant of GAN used to generate the super resolution of a single image
- What is WGAN - GP (Wasserstein Distance with Gradient Penalty)
  - A variant of GAN with a modified loss function (Wasserstein Distance v.s. Jensen-Shannon Divergence) and a gradient penalty term (Lipschitz Constraint) to avoid mode collapse and promote stability

Dataset & Preprocessing

- We used the DIV2K dataset
- It is a popular super-resolution dataset of different scenes which contains 800 images for training and 100 images for validation
- All the images were of 2k resolution
- Preprocessing involved cropping all the images to a specific size
- A low resolution version of the images by was obtained down scaling the images by a factor of 4

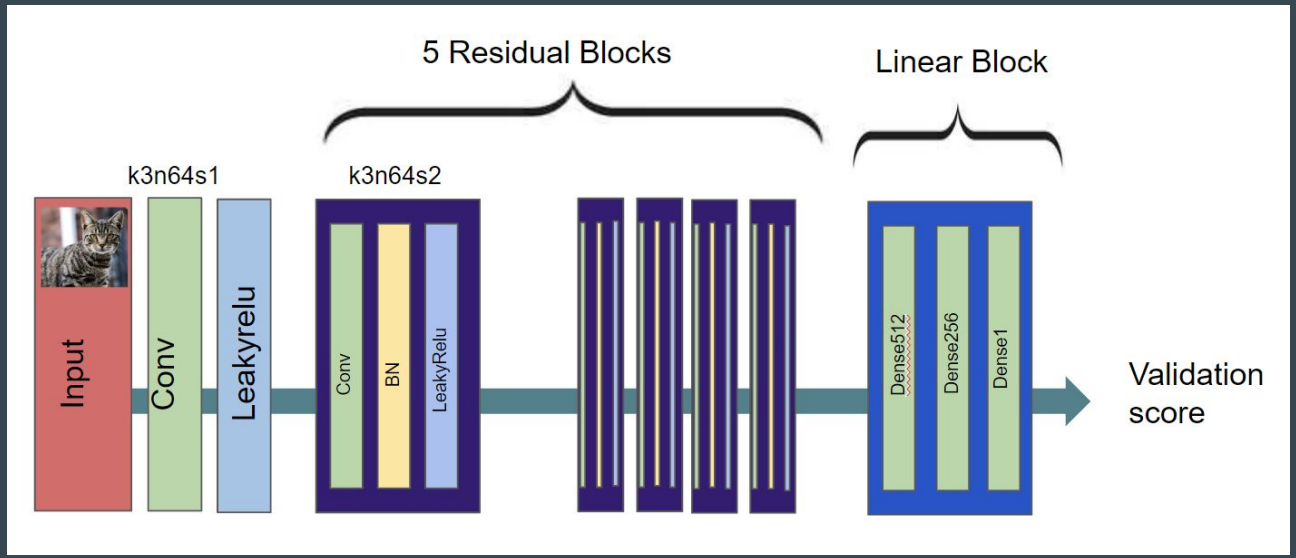
Generator

Architecture of the Generator: (PReLU + Residual Blocks + SubPixel Convs)



Discriminator

Architecture of the Discriminator: (LeakyReLU + Convs + Linear Block)



Loss Function

Given G(Low-Res) = Super-Res

Discriminator Loss:

E(D(Super-Res)) - E(D(High-Res)) + L * Gradient-Penalty

Lipschitz Constant

~ 0.20

Generator Loss:

E(-D(Super-Res)) + Reconstruction_Loss + Content_Loss

~0.0875 ~ 0.1163

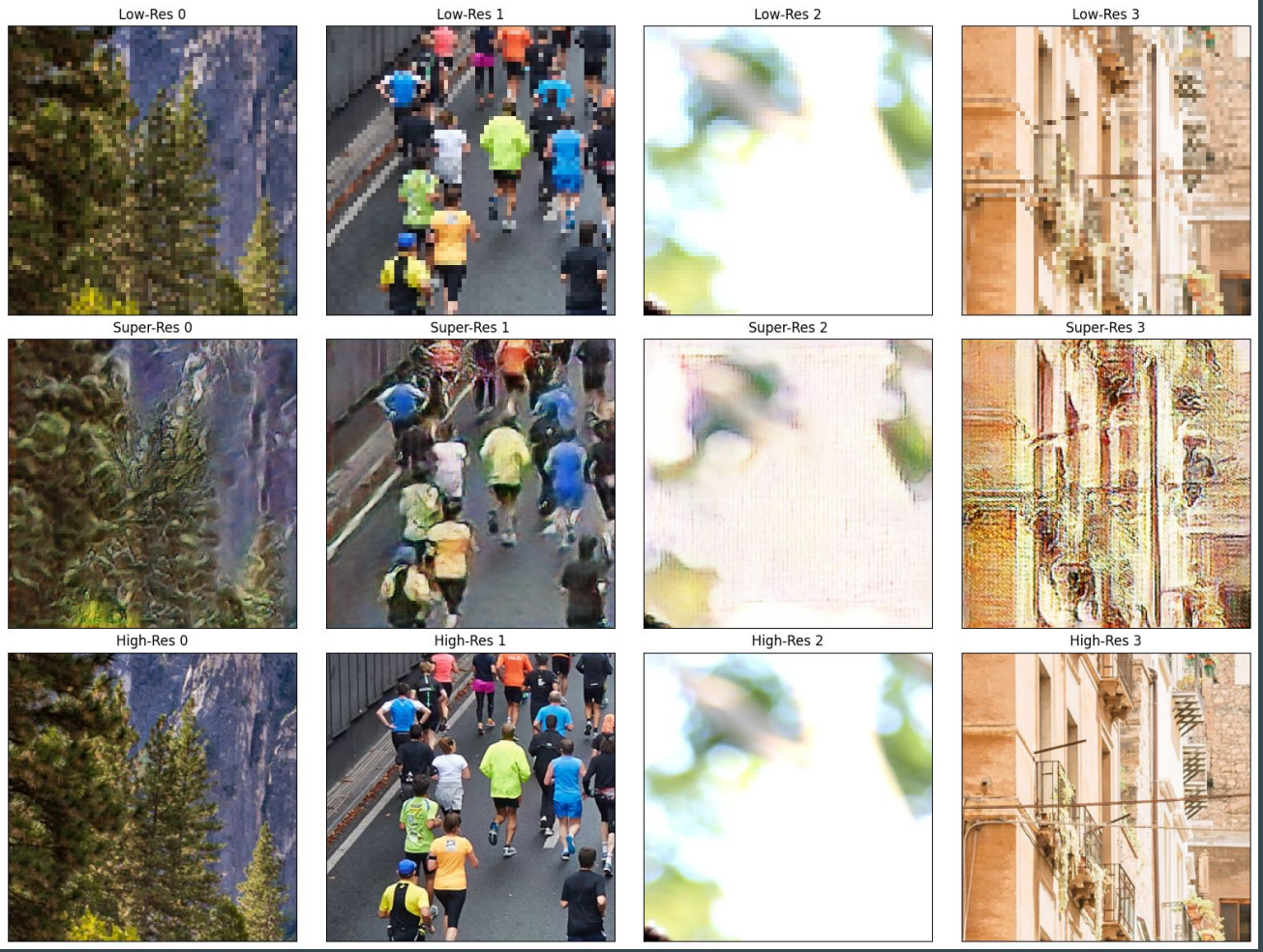
Evaluation Metrics : Peak Signal-to-Noise Ratio (PSNR)

PSNR :

A metric used to measure the quality of an image signal. It compares the peak signal power to the noise power (MSE differences) in pixel values, expressed in decibels (dB).

$$\mathbf{PSNR = 10 \cdot \log (MAX^2 / MSE)}$$

Results



PSNR: 20.02493 17.960169 24.903938 14.021114

Results

PSNR:	15.887534	25.741234	24.10814	18.081806

Future Improvements

- Video Support
- Adding a feature for the user to be able to choose by what factor to increase the image (of different sizes)
- Adding a feature where the model selects the faces and other important areas of the image specifically to increase the resolution
- Incorporating Optimal Transport Theory in generating Super Resolution Images (code for Sinkhorn Algorithm included)

Ethics

- Privacy
  - E.g.) Facial Recognition via low resolution CCTV Footage
- Manipulation and misrepresentation of the original image
  - E.g.) Medical Imaging
  - E.g.) Misidentification in Facial Recognition

Thank you!