K.K.WAGH INSTITUTE OF ENGINEERING EDUCATION & RESEARCH

ASTRONOMICAL IMAGE COLOURISATION AND SUPER-RESOLUTION USING GANs

Group ID: 23
Internal Guide: Prof. Dr. S. M. Kamalapur

K K Wagh

TEAM MEMBERS

Division./ Roll No.	Exam Seat No.	Name of the student
17	B150134261	Shreyas Kalvankar
18	B150134296	Hrushikesh Pandit
19	B150134299	Pranav Parwate
20	B150134303	Atharva Patil

Content

- Problem Definition
- Requirement Specification
- Literature Survey
- Motivation of the Project
- Objectives
- Block diagram of Project/Architecture
- Methodology
- Mathematical Model
- Experimental Set Up
- Performance Parameter
- Efficiency Issues
- References

Problem Definition

- The problem can be divided into two sub-problems:
  - Create an efficient model to colorize grayscale images
  - Take a colorized image and upscale it n times the original size

Keywords: GAN, Neural Network, NodeJS, puppeteer, Convolutional Neural Network, Upscaling, Colorization.

Requirement Specifications

- The following table showcases the minimum hardware requirements:

Sr. No.	Parameter	Minimum Requirement	Justification
1	GPU type	NVIDIA CuDA enabled GPU	Training the model
2	GPU memory	>6 GB	Batch training

- The following are the software requirements:
  - Operating System: Windows/Linux
  - IDE: Jupyter Notebook
  - Programming Languages: python3, javascript
  - Frameworks: Node.js, Tensorflow, Sci-kit learn, plotting libraries, OpenCV

Table 1: Hardware Requirements

Literature Review

Publication and Year	Technology	Summary
TSAI, R. (1984)	Multiframe image restoration and registration	Applied and evaluated the ScSR method for improvement of image quality of magnified MR images (T1-weighted, T2-weighted, FLAIR, and DWI images) in16-bit DICOM format
Tom and Katsaggelos (1996)	Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images	Solution is provided to the problem of obtaining a high resolution image from several low resolution images that have been subsampled and displaced by different amounts of sub-pixel shifts
Welsh, T., Ashikhmin, M., and Mueller, K. (2002)	Transferring color to greyscale images	Introduced a general technique for colorizing greyscale images by transferring color between a source, color image and a destination, greyscale image
Levin, A., Lischinski, D., and Weiss, Y. (2004)	Colorization using optimization	Used quadratic cost function and were able to generate high quality colorizations.

Literature Review

Yatziv, L. and Sapiro, G. (2006)	Fast image and video colorization using chrominance blending	High Quality colorization results are obtsined at a fraction of the complexity and computational cost using concepts of luminance-weighted chrominance blending and fast intrinsic distance computations
Qu, Y., Wong, T.-T., and Heng, P.-A. (2006)	Manga colorization	Proposed a novel colorization technique that propagates color over regions exhibiting pattern-continuity as well as intensity-continuity
Tola, E., Lepetit, V., and Fua, P. (2008)	A fast local descriptor for dense matching	Introduced a novel local image descriptor designed for dense wide-baseline matching purposes
Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014)	Generative Adversarial Networks	Proposed a novel approach of implementing Generative Adversarial Networks using two Neural Networks, viz Generator and Discriminator Networks.

Literature Review

Mirza, M. and Osindero, S. (2014)	Conditional generative adversarial nets	Introduced the conditional version of generative adversarial nets, which can be constructed by simply feeding the data, y, to condition on to both the generator and discriminator
He, K., Zhang, X., Ren, S., and Sun, J. (2015)	Deep residual learning for image recognition	Presented 152 layer using residual learning framework for image recognition and an adaptive edge detection based colorization algorithm and its applications.
Long, J., Shelhamer, E., and Darrell, T. (2015)	Fully convolutional networks for semantic segmentation	Showed that convolutional networks by themselves, trained end-to-end, pixels-to-pixels, improve on the previous best result in semantic segmentation.
Simonyan, K. and Zisserman, A. (2015)	Very deep convolutional networks for large-scale image recognition	Investigated the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting
Cheng, Z., Yang, Q., and Sheng, B. (2016)	Deep Colorization	The paper presented a fully-automatic colorization method using deep neural networks

Literature Review

Dahl, R. (2016)	Automatic Colorization	automatically produce multiple colorized versions of a grayscale image
Radford, A., Metz, L., and Chintala, S. (2016)	Unsupervised representation learning with deep convolutional generative adversarial networks	Introduced a class of CNNs called deep convolutional generative adversarial networks (DCGANs), that have certain architectural constraints, and demonstrate that they are a strong candidate for unsupervised learning
Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., and Shi, W. (2017)	Super Resolution using GAN	Photorealistic single image super-resolution using a generative adversarial network.
Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2018)	Image-to-image translation with conditional adversarial networks	Pix2Pix is a Conditional-GAN with images as the conditions for colorization.

Motivation of the Project

- Application of GANs has successfully improved the performance and computers are getting better and better at predicting accurate missing pixel values and upscaling images many folds the original size.
- All this computation power can be used for astronomical research by processing large data archives.

Motivation of the Project

- A large number of images lie dormant in most of the space survey data archives which never go through any kind of processing and are low resolution and black & white.
- These images could be processed automatically by an algorithm that will colorize and super-resolve the images which can make it easier for astronomers to visually inspect the images.

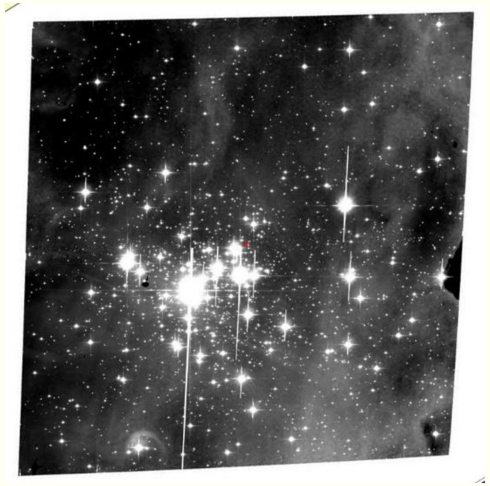

Fig 1a: Black & White image from space archive

Fig 1b: Subsequent colored version of Fig 1a

Objectives

- Auto-Colorization
- Upscaling/super-resolution
- The models may be combined to form a single model that will take a low resolution, grayscale image as its input and produce a high resolution, colorized image as its output.

Block diagram of Project

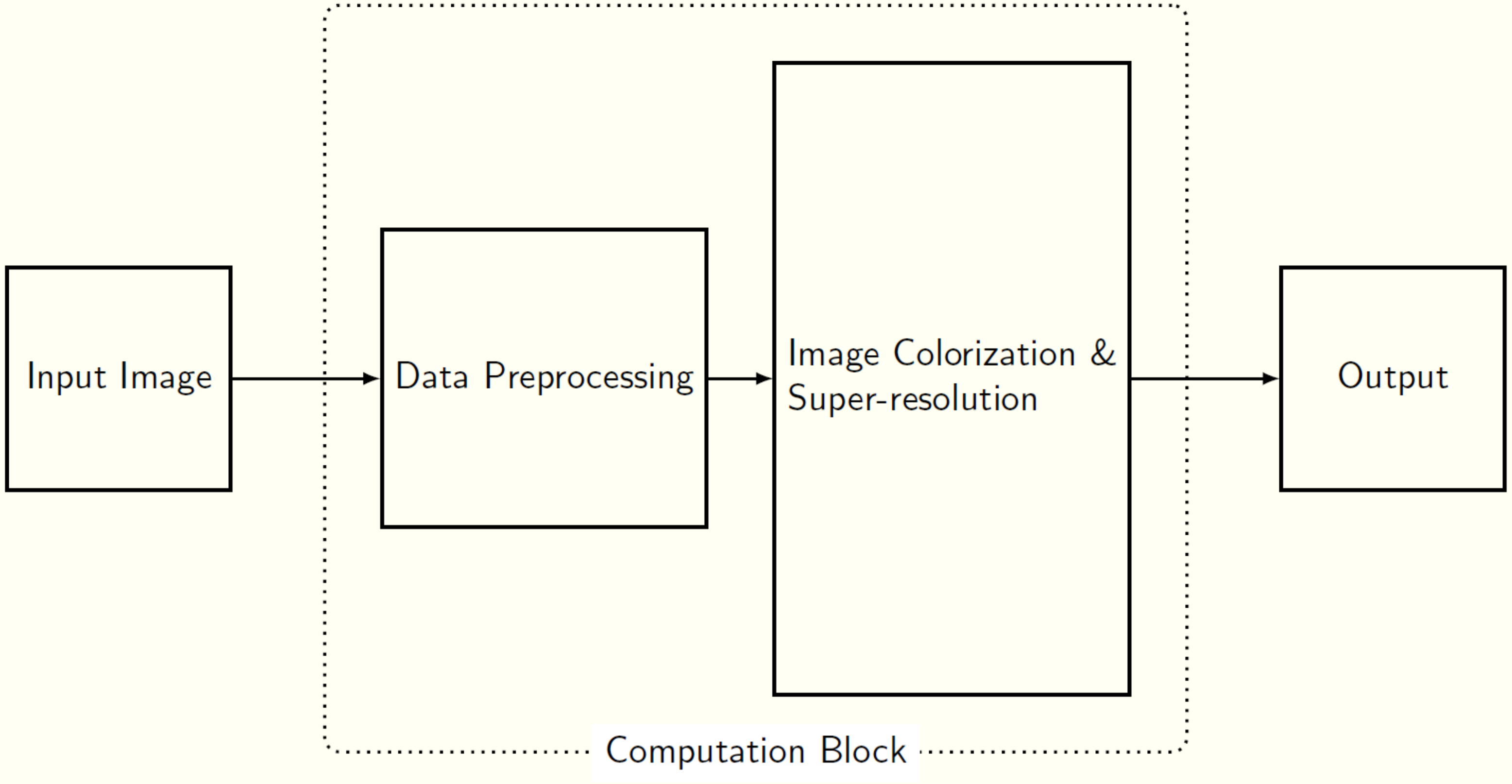

Figure 2: Basic Block diagram

Block diagram of Project

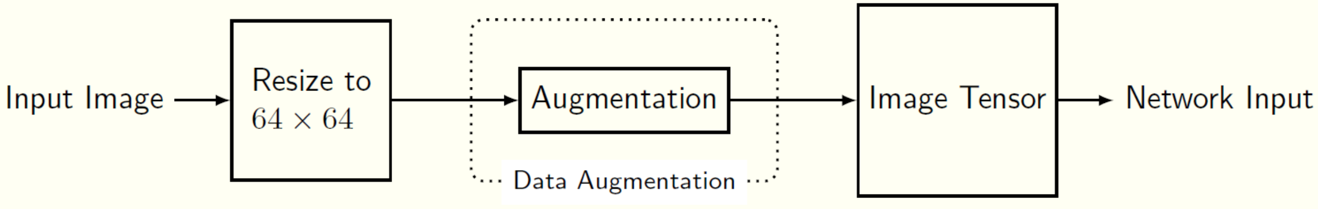

Figure 3: Data Preprocessing

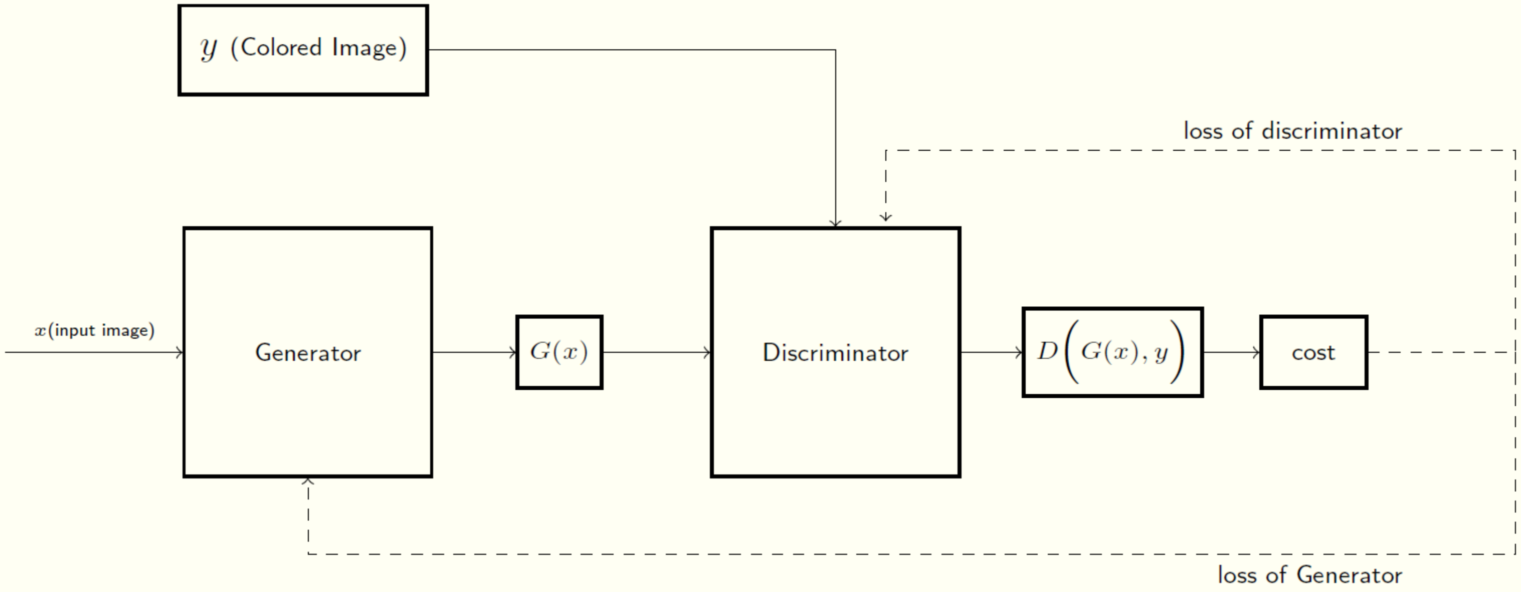

Figure 4: Basic GAN architecture

Block diagram of Project

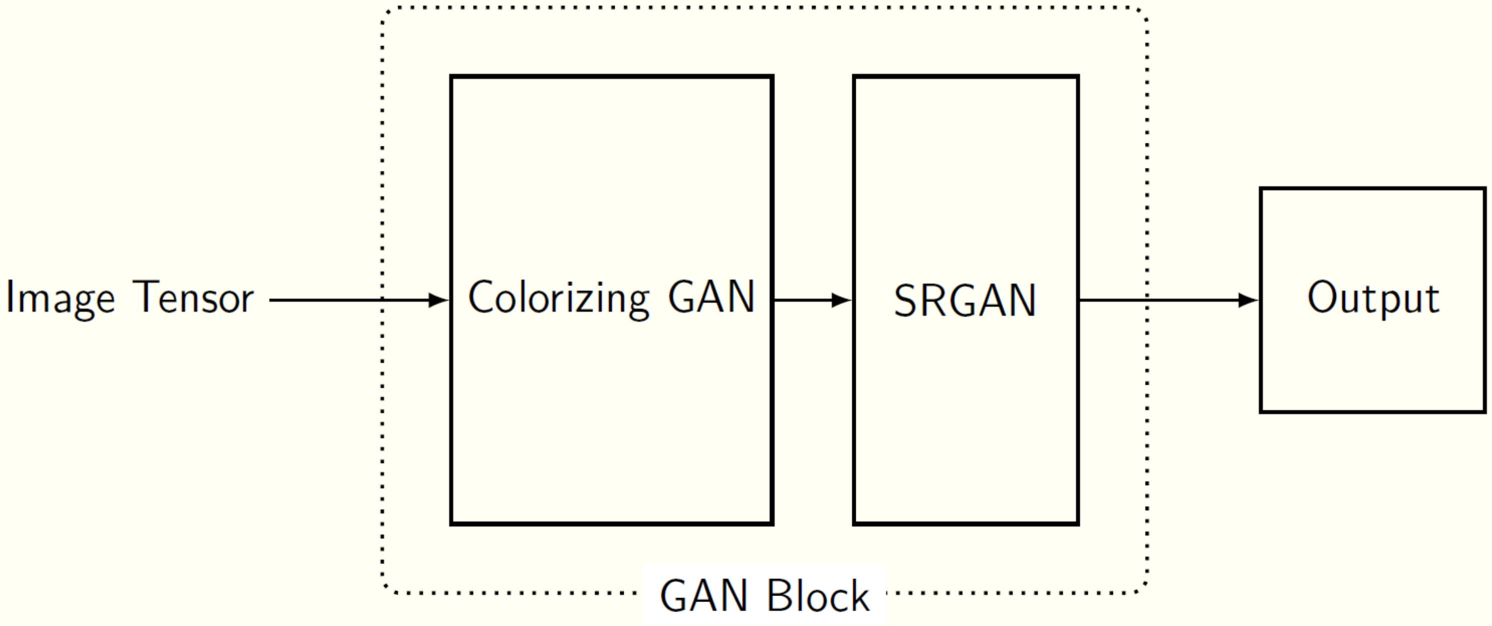

Figure 5: GAN Block Summary

Block diagram of Project

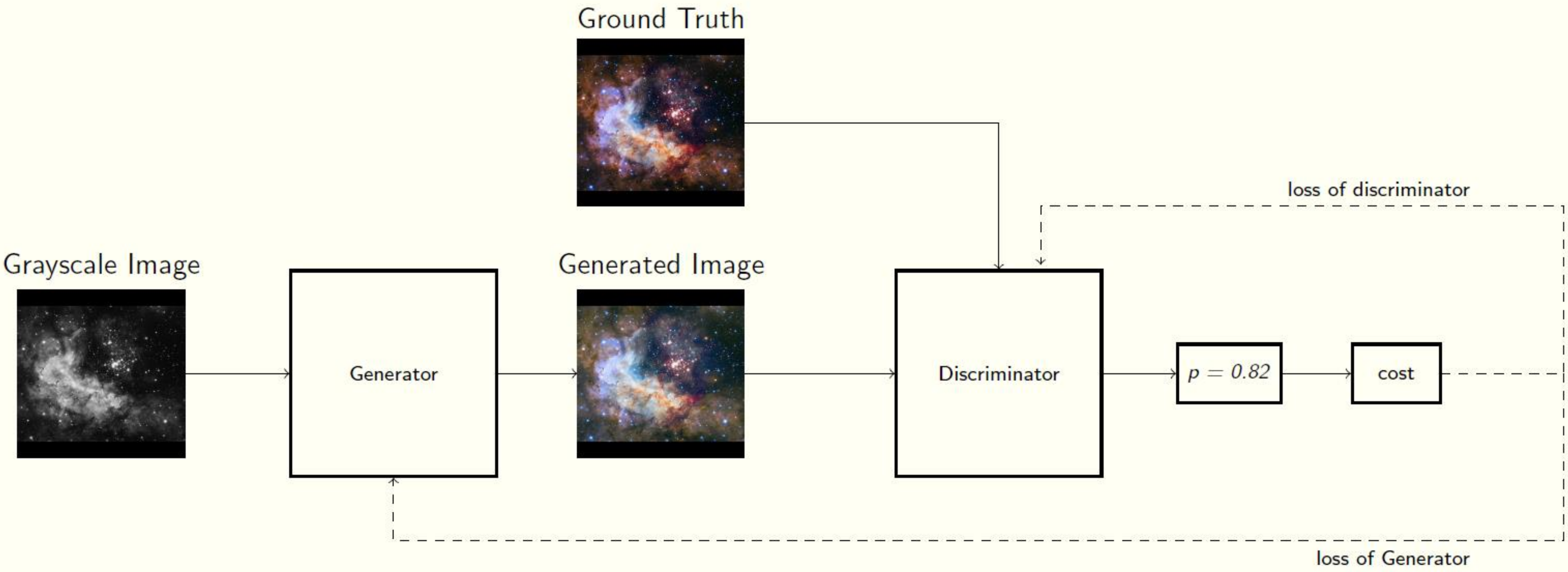

Figure 6: Image conversion by GAN

Block diagram of Project

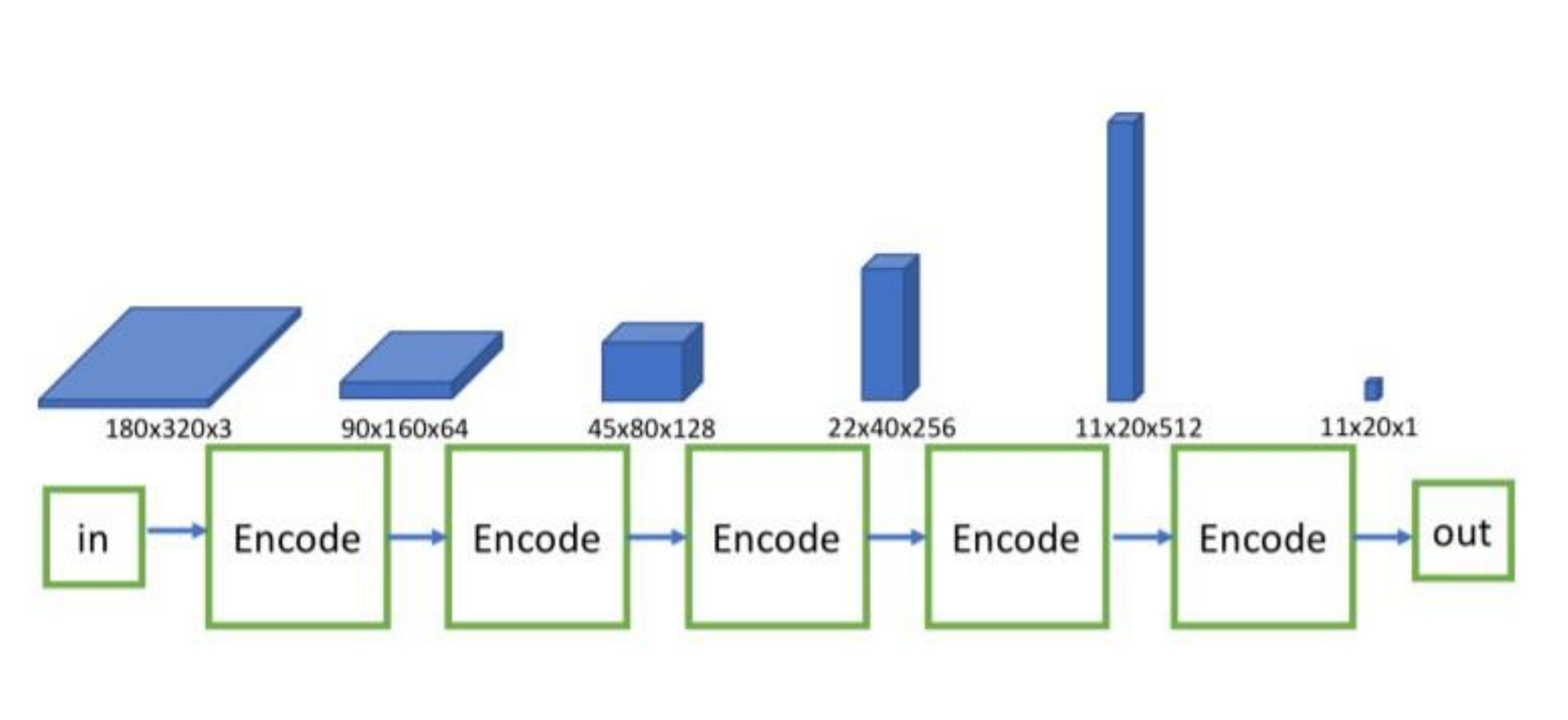

Figure 7: Discriminator (Ronneberger et al., 2015)

Block diagram of Project

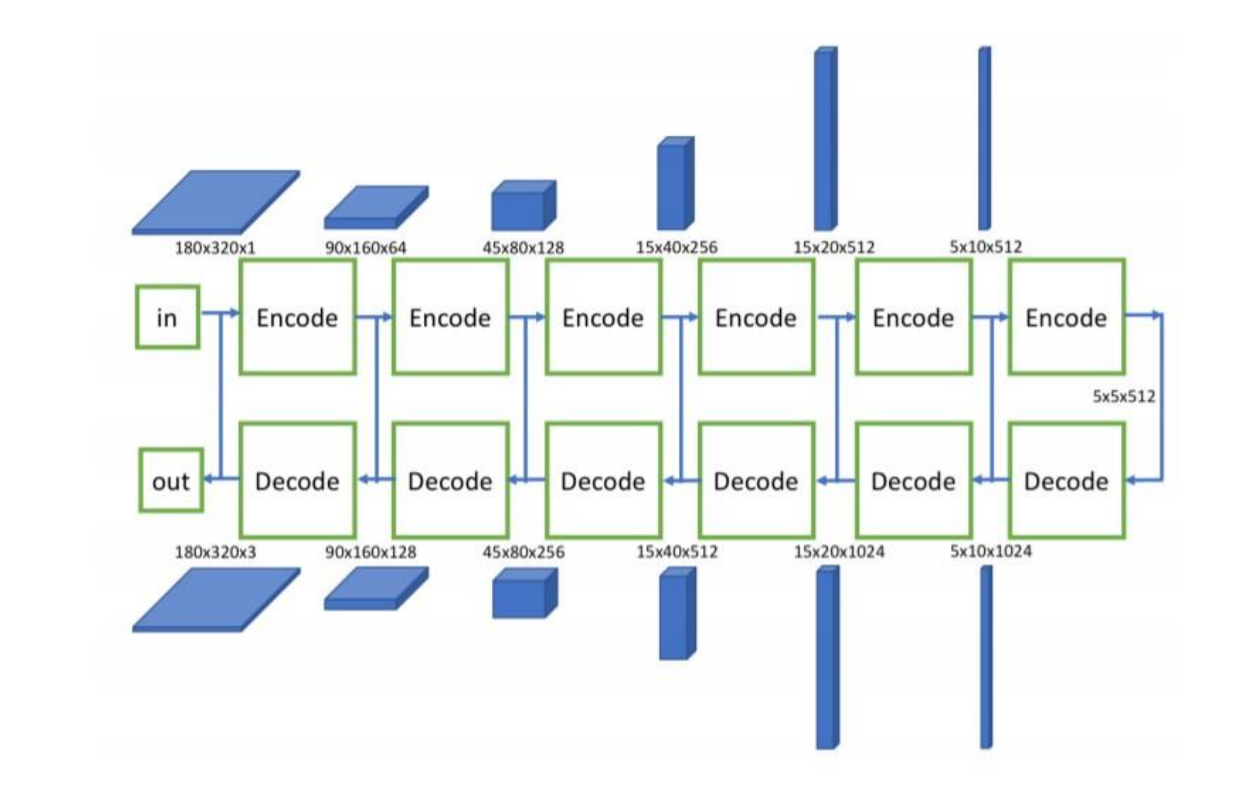

Figure 8: Encoder Decoder Generator (Ronneberger et al., 2015)

Block diagram of Project

K K Wagh

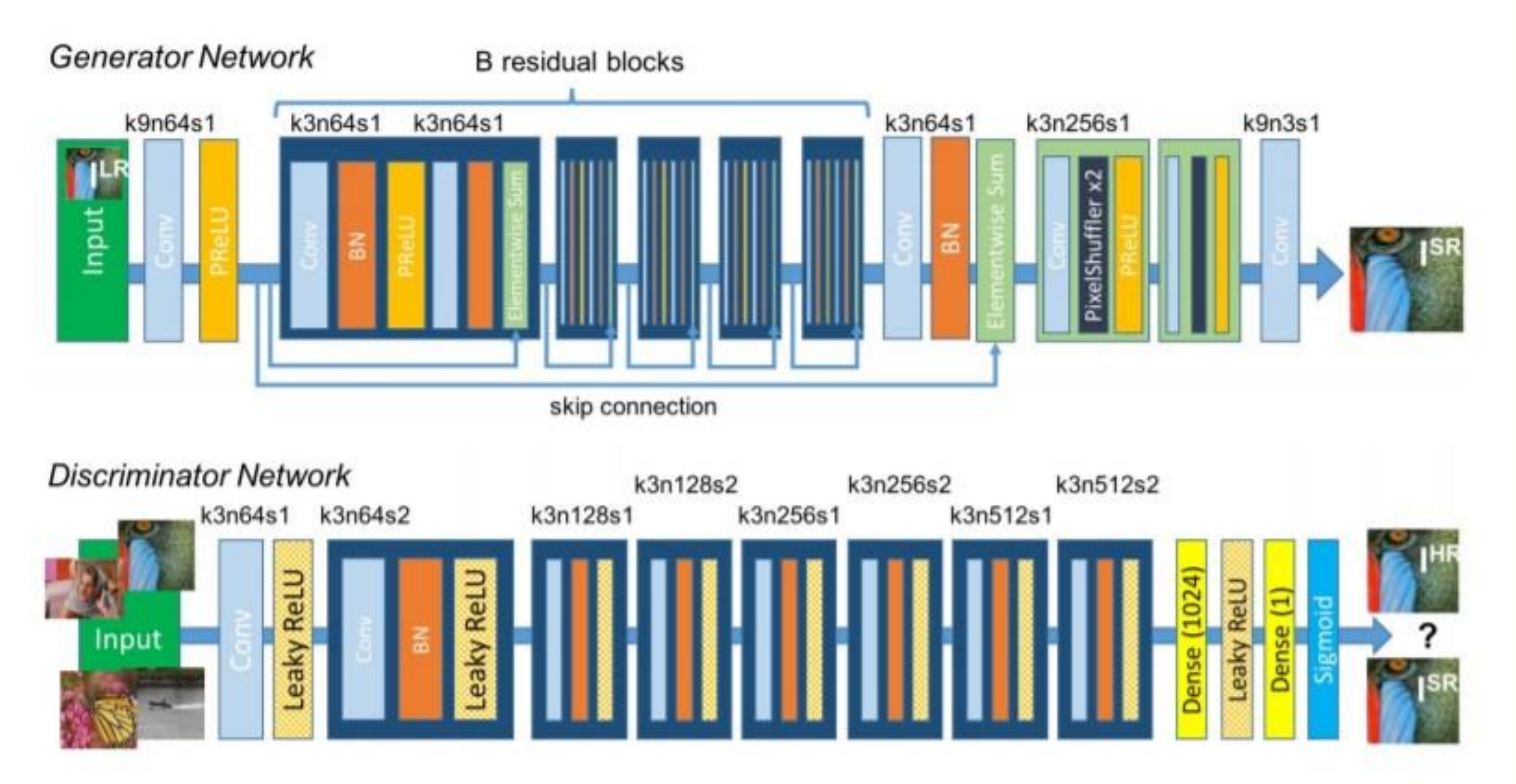

Figure 9: Ledig SRGAN Architecture (Ledig et al., 2017)

Methodology

- Data gathering and processing:
  - Data Scraping
  - Data Cleaning
- Model Building
- Model Training
- Cost Optimization and tuning
- Performance Evaluation and Documentation

Image Colorization

- Image Colorization convolutional neural networks with residual encoders using the VGG16 architecture will be used.
- Generative Adversarial Networks use a minimax loss which is different than the L2 loss as it will choose a color to fill an area rather than averaging. This is similar to a classification based approach.

Fig 10a: Black & White image

Fig 10a: Colorized output

Image Upscaling

- SR-GAN works well with for single image super-resolution as it also uses an intelligent content loss function that uses pre-trained VGG-net layers.

Fig 11a: Blurred image

Fig 11a: Upscaled output

Mathematical Model

- A generative network, G, is supposed to learn the underlying distribution of a latent space, Y.
- The Discriminator network D takes in both the fabricated outputs generated by G and real inputs from the underlying distribution Y.
- The network produces a probability of the image belonging to the real or fabricated space.

Mathematical Model

K K Wagh

Let $x \in X$ be a low resolution/grayscale image and $y \in Y$ be it's underlying distribution from the latent space Y.

$$G(x) = \hat{y}$$

The discriminative network D is fed the fabricated mapping $x \rightarrow \hat{y}$ and the underlying distribution of x i.e. $y \in Y$.

$$D\big(G(x), y\big) = p$$

where $p \in (0, 1)$ is the probability that the image is fabricated or real.

Let the generator be parameterized by θ_g and the discriminator be parameterized by θ_d.

The minimax objective function can be defined as:

$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x,y \sim p_{data}} \log D_{\theta_d}(x, y) + E_{x \sim p_{data}} \log(1 - D_{\theta_d}(x, G_{\theta_g}(x))\right]$$

Mathematical Model

K K Wagh

Where, G_{θ_g} is the output of the generator and D_{θ_d} is the output of the discriminator.

Also, we consider $L1$ difference between input x and output y in generator.

On each iteration, the discriminator would maximize θ_d according to the above expression and generator would minimize θ_g in the following way:

$$\min_{\theta_g} \left[-\log(D_{\theta_d}(x, G_{\theta_g}(x))) + \lambda \| G_{\theta_g}(x) - y \|_1 \right]$$

Experimental Setup

- We aim to implement the neural network models in Tensorflow using Jupyter Notebook and python
- As deep learning models require huge computational power for training, we plan to use Google Colab which provides a Tesla K80 GPU with memory ranging between 8GB to 16GB
- The dataset has been scraped off the Hubble Heritage project and Hubble Legacy Archive
- The processing on the dataset will be done using OpenCV and other image libraries in python and will be fed into the network

Performance Parameters

- To evaluate the performance of the coloring model quantitatively, we propose averaging the L1 and L2 distance (per pixel-channel) between the generated images and the ground truth images
- Another evaluation method is to calculate the Perceptual loss. It is critical for the performance of the Generator network
- The perceptual loss is defined as the weighted sum of the content loss and the adversarial loss component

Efficiency Issues

- The data gathered had to be scraped off websites such as the Hubble Legacy archive and Hubble main website
- This yielded in more images than were useful. So we focused on a particular section of the sky where we could get the images of galaxy M101
- This still yielded in about 400,000 images which had to be manually filtered
- Even with all the images available, the network training will require huge computational resources to perform efficiently
- The network parameters exceed the available training data and will require augmentation to avoid overfitting
- A quantitative evaluation of a GAN is considerably difficult even with the availability of the ground truth images

References

- Cheng, Z., Yang, Q. and Sheng, B. (2016). Deep colorization.
- Dahl, R. (2016). Automatic colorization.
- Dong, C., Loy, C. C., He, K. and Tang, X. (2014). Learning a deep convolutional network for image super-resolution, *in* D. Fleet, T. Pajdla, B. Schiele and T. Tuytelaars (eds), *Computer Vision – ECCV 2014*, Springer International Publishing, Cham, pp. 184-199.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial networks.
- He, K., Zhang, X., Ren, S. and Sun, J. (2015). Deep residual learning for image recognition.
- Huang, Y.-C., Tung, Y.-S., Chen, J.-C., Wang, S.-W. and Wu, J.-L. (2005). An adaptive edge detection based colorization algorithm and its applications, pp. 351–354.
- Isola, P., Zhu, J.-Y., Zhou, T. and Efros, A. A. (2018). Image-to-image translation with conditional adversarial networks.
- Jianchao Yang, Wright, J., Huang, T. and Yi Ma (2008). Image super-resolution as sparse representation of raw image patches, *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8.

References

- Kim, J., Lee, J. K. and Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1646–1654.
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. and Shi, W. (2017). Photorealistic single image super-resolution using a generative adversarial network.
- Levin, A., Lischinski, D. and Weiss, Y. (2004). Colorization using optimization, *ACM SIGGRAPH 2004 Papers*, pp. 689–694.
- Long, J., Shelhamer, E. and Darrell, T. (2015). Fully convolutional networks for semantic segmentation, *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431–3440.
- Mirza, M. and Osindero, S. (2014). Conditional generative adversarial nets.
- Pressman, R. S. (1992). Software Engineering (3rd Ed.): A Practitioner's Approach, McGraw-Hill, Inc., New York, NY, USA.
- Qu, Y.,Wong, T.-T. and Heng, P.-A. (2006). Manga colorization, ACM Transactions on Graphics (TOG) 25(3): 1214–1220.
- Radford, A., Metz, L. and Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks.
- Shi, W., Caballero, J., Husz´ar, F., Totz, J., Aitken, A. P., Bishop, R., Rueckert, D. and Wang, Z. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1874–1883.

References

- Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition.
- Tola, E., Lepetit, V. and Fua, P. (2008). A fast local descriptor for dense matching, *Proc. CVPR* .
- Tom and Katsaggelos (1996). Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images, *in* Anon (ed.), *IEEE International Conference on Image Processing*, Vol. 2, IEEE, pp. 539–542. Proceedings of the 1995 IEEE International Conference on Image Processing. Part 3 (of 3) ; Conference date: 23-10-1995 Through 26-10-1995.
- TSAI, R. (1984). Multiframe image restoration and registration, *Advance Computer Visual and Image Processing* 1: 317–339.
- Welsh, T., Ashikhmin, M. and Mueller, K. (2002). Transferring color to greyscale images, *ACM Trans. Graph*. 21: 277–280.
- Yatziv, L. and Sapiro, G. (2006). Fast image and video colorization using chrominance blending, *IEEE Transactions on Image Processing* 15(5): 1120–1129.
- Zhu, J.-Y., Krahenbuhl, P., Shechtman, E. and Efros, A. A. (2018). Generative visual manipulation on the natural image manifold.

THANK YOU !!

K K Wagh

ANNEXURE

Annexure : Performance Parameters

We define the content loss as the L2 distance between the feature representations of the reconstructed image $G_{\theta_g}(I^{LR})$ and the reference image I^{HR}

$$l^{SR}_{VGG_{i,j}} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_g}(I^{LR}))_{x,y})^2$$

where $W_{i,j}$ and $H_{i,j}$ represent the dimensions of the respective feature maps within VGG19 network. The adversarial generative loss l^{SR}_{Gen} is defined on the probabilities of the discriminator $D_{\theta_d}(G_{\theta_g}(I^{LR}))$ over all the training samples as:

$$l^{SR}_{Gen} = \sum_{n=1}^{N} -\log D_{\theta_d}(G_{\theta_g}(I^{LR}))$$

Annexure: Performance Parameters

$D_{\theta_d}(G_{\theta_g}(I^{LR}))$ is the probability that the reconstructed image $G_{\theta_g}(I^{LR}))$ is a natural HR image. For beter gradient behavior, we minimize $-\log D_{\theta_d}(G_{\theta_g}(I^{LR}))$ instead of $\log\left[1-D_{\theta_d}(G_{\theta_g}(I^{LR}))\right]$.