Google Summer of Code

Proposal - Google Summer of Code 2023

SUPER RESOLUTION FOR STRONG GRAVITATIONAL LENSING

By

Amin Mohamed

ML 4 SCI

CONTENTS

CONTACT INFORMATION

- Name: Amin Mohamed Rabea
- Email: aminmoe707@gmail.com
- Contact No. : +201096289637
- University: Ain Shams University
- Major: Artificial Intelligence
- Country: Egypt
- CV: Link
- Other Links: [LinkedIn] [GitHub] [Kaggle]

ABSTRACT

Strong gravitational lensing is a phenomenon where the gravity of a massive object can distort and magnify light from distant sources, resulting in multiple images that are warped.

This phenomenon is a useful tool for studying the substructure of dark matter, which is a mysterious substance that makes up most of the universe's matter.

However, the quality of lensing images can be limited by instrument constraints and atmospheric conditions, which make it difficult to estimate lensing parameters and understand the nature of dark matter.

To address this challenge, we propose using deep learning techniques, specifically image super-resolution methods, to improve the quality of lensing images. Image super-resolution involves using deep learning models, such as convolutional neural networks, to produce high-quality images from low-quality ones. We will develop and use these models to improve the precision and speed of strong lensing analysis, and explore their potential for other tasks, such as regression and lens finding. Our ultimate goal is to enhance our understanding of dark matter and cosmology using advanced image processing techniques.

PROJECT DESCRIPTION

Our aim for this project is to explore the performance of various deep learning algorithms that can handle the strong gravitational lensing data and how will it impact the performance of other related tasks.

To achieve this, we plan to carry out a number of experiments that will compare the strengths and weaknesses of different kinds of deep learning models, ranging from the simple Convolutional Neural Networks (CNNs) that use convolutional filters to learn features from images, to the more advanced Vision Transformers (ViTs) that use attention mechanisms to learn global representations of images, as well as the state-of-the-art Generative Adversarial Networks (GANs) and Diffusion models that use adversarial and stochastic processes to output high-quality images.

Experiments will also include preprocessing decisions such as:

- Type of normalization
- Type of augmentation

Also, model making and hyperparameter choices such as:

- Type of optimizer and it's specific parameter tuning
- Type of learning rate scheduler and what metric to monitor
- Modification that could be made to the specific network due to some observation during training
- What metrics to monitor and optimize for
- Along with the hyperparameters that belong to each specific model

The choice of the model's type would be very much dependent on the complexity of the data, and the result we could acquire from simpler models with less inference. The decision of trying a more complex model would be made if the monitored metrics indicated that there was a room for improvement.

The choice of the model's type could be divided as:

1. Basic CNN Models
2. GAN Based Models

BASIC CNN MODELS

RESIDUAL DENSE NETWORK (RDN)

RDN utilizes the residual learning concept, where the network learns to predict the difference between the high-resolution and low-resolution images rather than directly predicting the high-resolution image. Additionally, it incorporates dense connections between all the layers in the network, allowing information to flow more efficiently through the network.

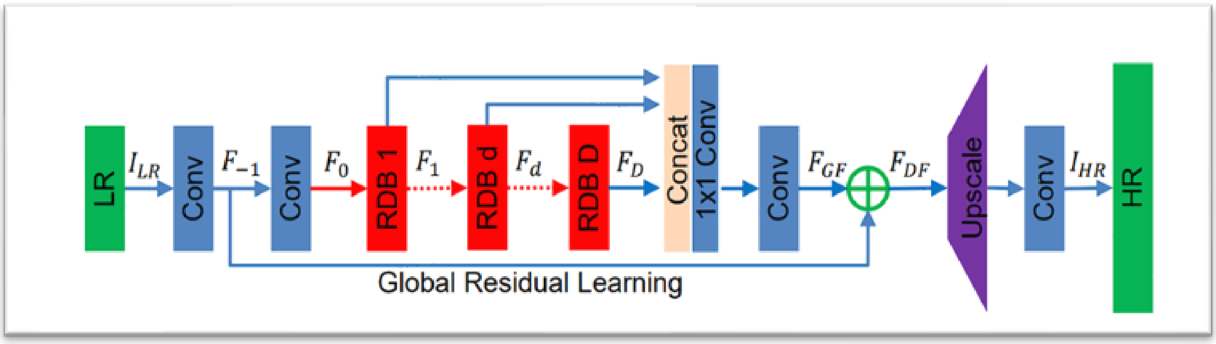

FIGURE 1: RDN NETWORK ARCHITECTURE [1]

The RDN architecture comprises a set of dense blocks, where each block consists of several convolutional layers that extract features from the input image.

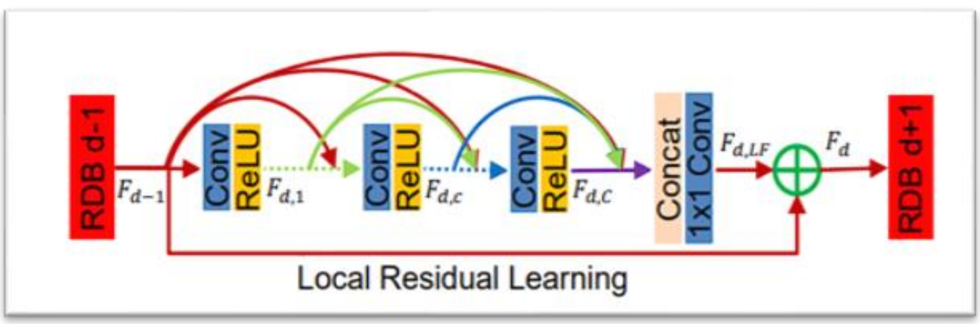

FIGURE 2: RESIDUAL DENSE BLOCK (RDB)[2]

The outputs of all the convolutional layers in a block are concatenated, and the resulting features are fed into the next block. The final output of the network is the predicted high-resolution image.

GENEARTIVE ADVERSERIAL NETWORKS (GANS)

The advantage of using a GAN based architecture for the problem of strong gravitational lensing is that GANs tend to converge to a more perceptually convincing solutions, whereas MSE based architecture tends to be overly smooth due to pixel wise average of possible solutions according to (Ledig et al., 2017).[3]

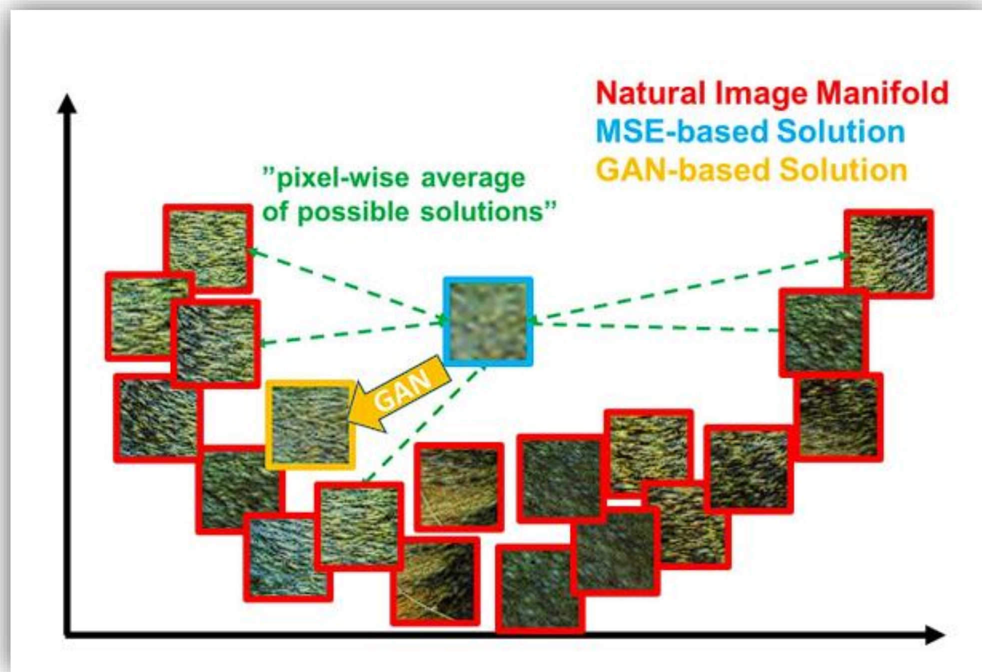

FIGURE 3: MSE-BASED SOLUTION APPEARS TO BE OVERLY SMOOTH, WHILE GAN DRIVES THE RECONSTRUCTION TOWARDS THE NATURAL IMAGE MANIFOLD PRODUCING PERCEPTUALLY MORE CONVINCING SOLUTIONS.

Super Resolution GAN (SRGAN)

SRGAN[4] combines the power of two neural network models: a generator and a discriminator. The generator is trained to generate high-resolution images from low-resolution inputs, while the discriminator is trained to distinguish between the generated high-resolution images and real high-resolution images.

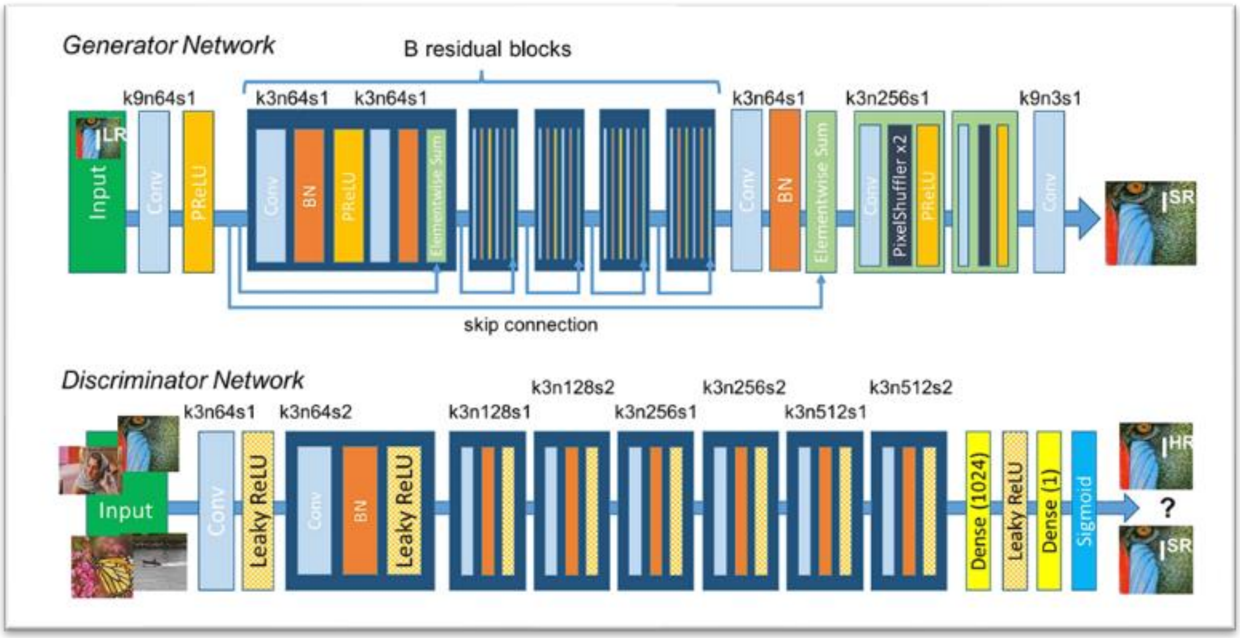

FIGURE 4: ARCHITECTURE OF GENERATOR AND DISCRIMINATOR NETWORK WITH CORRESPONDING KERNEL SIZE (K), NUMBER OF FEATURE MAPS[5]

The generator and the discriminator are trained together in an adversarial manner, where the generator tries to fool the discriminator into thinking that its generated images are real, while the discriminator tries to correctly distinguish between the generated and real images.

SRGAN utilizes a deep residual network as its generator, similar to SRResNet, but adds a perceptual loss term to the training objective function. This loss term namely "perceptual loss" consists of two essential elements namely: content loss and adversarial loss.

The content loss is produced from the Euclidean distance between the generated image's embedding and the ground truth image's embedding. The adversarial loss comes from the adversarial training of the two networks.

Perceptual loss encourages the generator to produce high-resolution images that not only have high similarity with the ground truth high-resolution images but also have high perceptual quality.

Another approach -which I took in the tests- is to train the generator network only which is called SRResNet , and this might yield better performance for specific tasks.

Enhanced Super-Resolution GAN (ESRGAN)

ESRGAN [6] is an improvement over SRGAN, which incorporates several key modifications to the original architecture.

- It uses a deeper generator network with more residual blocks to capture more complex features and improve the quality of the generated high-resolution images.
- It removes Batch Normalization from the network according to (Xintao Wang et al 2018) [7] it has been proven that it increases performance.
- It uses a residual-in-residual dense block (RRDB) instead of a residual block as the basic network building unit. The RRDB can effectively extract rich features from low-resolution images and avoid the degradation problem caused by batch normalization.

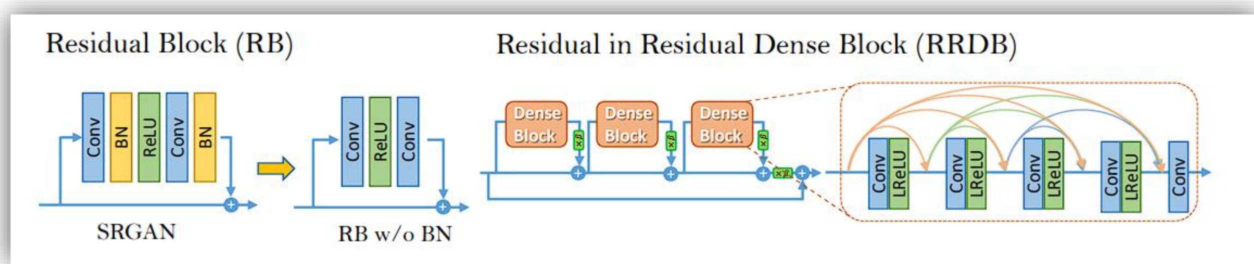

FIGURE 5: LEFT: AUTHORS REMOVED THE BN LAYERS IN RESIDUAL BLOCK IN SRGAN. RIGHT: RRDB BLOCK WAS USED IN THEIR DEEPER MODEL AND β IS THE RESIDUAL SCALING PARAMETER. [8]

- It employs a new residual scaling module that allows the network to better handle scale changes and produce sharper and more detailed images.

- ESRGAN also introduces a new training method, called relativistic GAN training, which further enhances the quality of the generated images by making the discriminator more sensitive to the relative quality of the generated and real images.

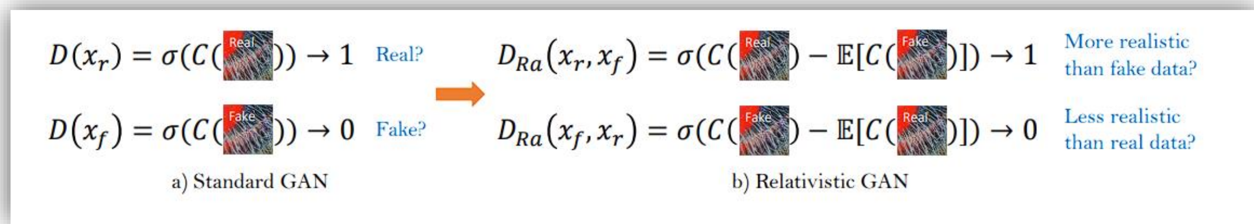

FIGURE 6: DIFFERENCE BETWEEN STANDARD DISCRIMINATOR AND RELATIVISTIC DISCRIMINATOR.[9]

ESRGAN has achieved state-of-the-art performance in several image super-resolution benchmarks, producing high-quality and visually pleasing high-resolution images with rich details and textures.

OTHER SR APPROCHES

As I stated earlier, the experimentation of these techniques is a sequential process that goes from simpler, lighter and more efficient models. We monitor their metrics and assess them accordingly. This process helps us identify the best models that can be used for our research.

If the experiments conducted have hinted that there is room for improvement even after using complex models, we will research more complex - likely SOTA - techniques that utilize ViTs and DDPMs. These techniques are state-of-the-art and can help us achieve better results in our research.

PROJECT TIMELINE

Task	Duration	Description	Start date/End date
Community Bonding Period	4 weeks	Establish communication with mentors, dig more into the project's idea and what is the proper way of implementing it	May 04 – May 28
CNN models implementation	2 weeks	Implement and experiment with CNN-based models..	May 29 – June 12
CNN models tuning and evaluation.	1 week	Try out different preprocessing techniques, data augmentation and hyperparameter choices. Document results for comparisons.	June 12 – June 19
Implement GAN based models	2 weeks	Implement and experiment with different GAN-based models..	June 20, July 3

Task	Duration	Description	Start date/End date
GAN models tuning and evaluation.	1 week	Try out different preprocessing techniques, data augmentation and hyperparameter choices. Document results for comparisons.	July 4, July 14
Comparisons, final Evaluations and Documentation.	2 weeks	Make the final document showing results of monitored metrics, used techniques and conclusions.	July 14, July 28
Investigate leveraging super-resolution models for related tasks	2 week	Explore how the enhanced strong lensing Images could boost the performance of preexisting models.	July 28, August 11
Modifications, Feedback and further experiments	2 weeks	Investigate with my mentors the possible improvements that could be made, and if the use of more complex models is needed. Final Documentation of the project.	August 11, August 28

FEATURE DELIVERABLES

I'd love to be a part of this research organization and contribute to it after this project.

I might be particularly of help in those criteria’s:

- A comprehensive report on the performance of various deep learning algorithms on strong gravitational lensing data, including metrics such as accuracy, precision, recall, F1-score, and ROC curve.
- A comparison of the advantages and disadvantages of different kinds of deep learning models, such as CNNs, ViTs, GANs, and Diffusion models, for strong gravitational lensing tasks, such as detection, classification, parameter estimation, and image reconstruction.
- A discussion of the best practices and challenges for preprocessing, model making, and hyperparameter tuning for strong gravitational lensing data, such as type of normalization, augmentation, optimizer, learning rate scheduler, modification of network

COMMITMENTS IN SUMMER

During the interval that spans from June 8 to June 26, I will be taking my final exams for this academic year, and I will have limited availability to work on this project. Therefore, I will only be able to dedicate a maximum of **25 hours per week** for this period. However, I will make sure to work at my full capacity of **35 hours per week** before and after this interval, in order to catch up with the work and meet the deadlines.

WHY ME?

I have a lot of experience in developing, analyzing and improving models that are based on deep learning, especially those that involve computer vision tasks.

I have taken part in several competitions that are related to deep learning and machine learning and I have managed to score high ranks among the participants.

I am currently majoring in Artificial Intelligence and I have been working with it for two years now.

My graduation project, which is called “Super Resolution and Face Recognition”, is oriented around the techniques of super resolution, which are used to enhance the quality of images.

I’ve extensive experience with deep learning frameworks like PyTorch and TensorFlow, as well as Machine Learning Libraries like scikit-learn.

This gives me relevant experience in dealing with the kind of models that I will be developing for this task, which also involves super resolution.

I have a deep passion and curiosity for this field of work, and I would love to fulfill this passion by working on this project.

WHY THIS PROJECT?

This project is very interesting to me because it involves developing models that use super resolution techniques, which are techniques that are used to enhance the quality of images by increasing their resolution and reducing their noise. These techniques are related to my graduation project and my domain of expertise, which is computer vision.

I have a lot of experience in working with deep learning and computer vision models, and I have participated in several competitions that challenged my skills and knowledge in this field.

I am very passionate and curious about this field of work, and I would love to work on this project to learn more and contribute to the advancement of science and technology.

REFRENCES

[1][2] Zhang, Y., Tian, Y., Kong, Y., Zhong, B., & Fu, Y. (n.d.). Residual Dense Network for Image Super-Resolution. Retrieved March 29, 2023

[3][4][5] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi Twitter, W. (n.d.). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network.

[6][7][8][9] Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Loy, C., Qiao, Y., & Tang, X. (n.d.). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks.