Organization: **ML4SCI: DeepLense**

Project Title: Updating the DeepLense Pipeline

INTRODUCTION

About Me

Full Name: Diptarko Choudhury

Time Zone: IST (UTC+5:30)

Email: diptarko.choudhury@niser.ac.in

GitHub ID: dc2506

Matrix Handle: dc2506

Whatsapp/Phone: 9593169517

Education

University: National Institute of Science Education and Research (NISER), Bhubaneswar

Program: 5-Year Integrated Masters in Physics, with a Minor in Computer Science

Current Year: Second year (4th semester ongoing)

PERSONAL BACKGROUND

I am an undergraduate 2nd Year student at the National Institute of Science Education and Research (NISER), pursuing my Integrated Masters in Physics. Along with my Integrated Masters, I am also doing a Minor in Computer Science. My main interest lies in the field of computational physics and Machine Learning. I am proficient in C++, Python and Java. I have some experience using Machine Learning Frameworks such as Tensorflow and Pytorch. I enjoy doing Machine Learning and Deep Learning in my free time. I started it as a hobby, but it grew to become a passion with time. I have participated in several Kaggle and private competitions and even secured the First Place in ML4SCI 2021 Higgs Challenge. I am also quite interested in particle and high energy physics since I often like to read about them in my leisure time.

I am also part of many extracurricular clubs in my college and love to take part in events such as debate and extempore. I am also an active member of RTC (Robotech Tech Club) of NISER, where I mentor Computer Vision and Machine Learning projects. I am also working on a pet project of the RTC Club, AMAR(an autonomous rover), where I am currently dealing with Object Detection and Image Segmentation.

Being highly interested in physics and Machine Learning, I came across the ML4SCI 2021 Hackathon. I participated in it, and my interest grew even further to work on a project related to this specific field.

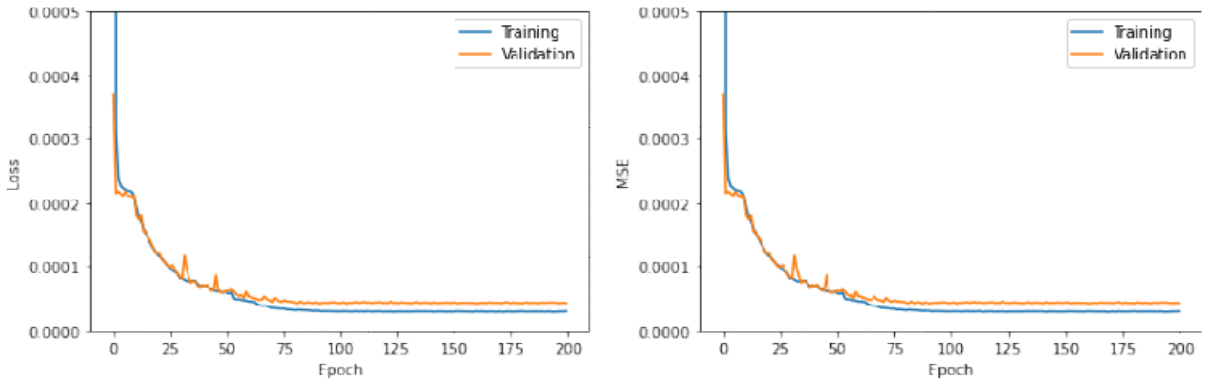
Why this project?

ML4SCI had a great variety in terms of projects to choose from. However, DeepLense was one of those few projects that captured my attention. The main motive behind machine learning or Artificial Intelligence is to replicate the human learning process in computers. Machine Learning tries to make computers have cogitative abilities similar to humans. There are very few examples in machine learning where the machine's ability to do a particular task is greater than that of a general human. The same is true for this project. A simple look at the problem images(simulated) gives us little to no idea about the nature of the sub-structure.

Nevertheless, an algorithm like a CNN can easily beat a human in this task, classify a strong gravitational lensing image, and give us an idea about its substructure. It also quite accurately predicts the mass of the dark matter halo, which we humans can never do by looking at the image. This is what makes me interested in this project, as it can act as a tool for understanding the universe, which otherwise might be pretty tricky. Moreover, there is still much experimentation and work for this project, and there is quite a lot of room left for improvements.

Insights Gained from the Evaluation Tests

The organization wanted me to complete an evaluation test. I have completed the two evaluation tests required for this project and have submitted them as per the instruction. Since the evaluation tests were highly correlated with my current project/proposal, I gained quite some insight. There were two tasks given, classification and regression. EfficientNet architecture was used for both of them. Now let us first discuss the regression problem.

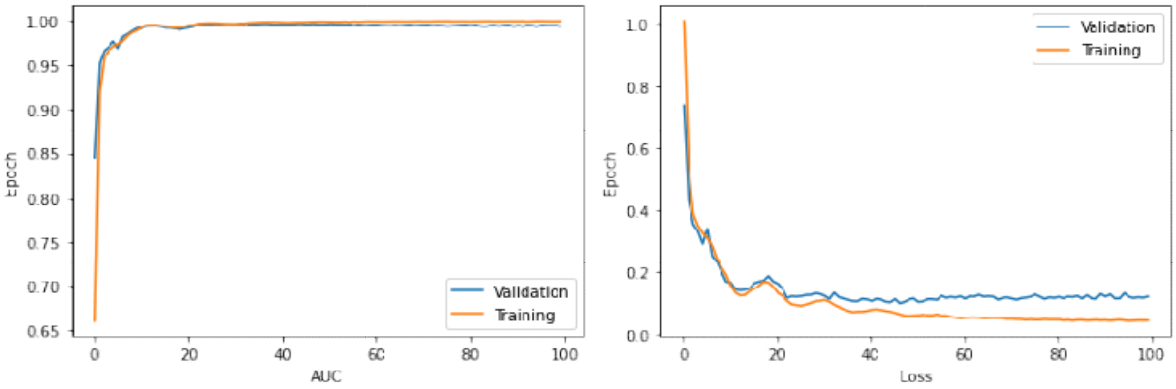


From the above plot, it is clear that the model has quite a good Generalization and needs not be stopped early, showing signs that we can scale up the model without much overfitting. Moreover, when we tried to use the Cyclic Learning Rate policy, we found that the network got trapped in bad local minima. Hence we had to use Reduce Learning Rate on Plateau. Better experimentation and proper research might help us find a much better learning rate policy to get the network to converge better and faster. Since the images are simulated, we will need a better and well-researched image pre-processing technique to get the best results.

For the classification problem, a pretty different scenario is seen. While the regression problem seems to do well with the given data, the classification problem starts to overfit after some epochs. This shows a need for more

advanced regularisation techniques for the classification algorithm, and more data might help.

The Loss and AUC plots of the Classification problem are given below.



The above plots give us a clear idea that the model starts to overfit after some time. Also, a cyclic learning rate policy was employed, which gave better results. This shows that more robust and advanced regularisation techniques must be used for this problem before scaling up the model and cyclic lr policy works well for the classification problem.

The Proposal

Abstract:

We are trying to implement a method to use Deep Learning to infer information from the Strong Gravitational Lensing Images. The main motive is to make Classification and Regression models for the new data generated from the strong lensing simulation using the Lenstronomy python module. This project will focus on building state of the art models with the highest validation for classification and regression tasks for the new dark matter models and then merging them with the pre-existing DeepLense pipeline.

Description:

The project will focus on building state of the art models for regression and classification tasks on the new dark matter models. The images are being simulated using the Lenstronomy package in python. We got perfect results by using Strong Lensing Images using pyauto lense earlier. Now we will try to extend the work for new dark matter models. This project will mainly focus on experimenting with state of the art machine learning models to find out which one works out the best for classification and regression tasks. The models will be built separately, trained and finetuned to get the highest validation scores and merged with the pre-existing DeepLense pipeline.

Deliverables of my project-

- Robust and fast input pipelines
- Efficient and state of the art Image Augmentation and pre-processing Techniques

- Selection of models which is best for the use case.
- Using state of the art training, optimization and regularisation techniques to achieve the highest validation score.
- Develop a robust and user-friendly output pipeline and integrate the new architectures with the pre-existing DeepLense pipeline.

Timeline

CURRENT PROGRESS

I have been working on the Evaluation test of the projects and finished working on them. I gained quite a bit of insight after building the basic architecture and enough experience to build a much more complex and tailor-made architecture to suit my problem the best.

COMMUNITY BONDING

I will set up a proper channel and weekly schedule for communications and discussions in consultation with the mentors. I will use this time to further my understanding of the problem statement and clarify doubts (if any) with my mentors. I try to mainly understand the problem statement during this time and get a shallow idea of the quantities or properties I will be dealing with. After consultation with my mentor, I will also shortlist some papers to read and understand the details of Gravitational Lensing and Dark Matter.

Week 1-2 (2 Weeks)

I will devote the first two weeks to reviewing the literature. I will start to read about various new architectures used in Computer Vision, which can be used to build my model. I will also read about gravitational lensing and get a basic understanding of what I am working with. Also, I will try to get myself exposed to various state of the art techniques(like stochastic depths, Image Augmentation techniques etc.) and their application in the field of computer vision, which might help increase my model's performance. I will also read the research papers given on the website is much more depth to understand what I am working with.

Week 3

I will start building the basic input pipeline for our neural network, which will be robust enough to accommodate any changes in future without getting bottlenecked.

Week 4

This week I will try to understand the nature of the input data and read and understand the simulation techniques and software used to generate the data. I will try to find which type of image pre-processing works the best for the data. I will use a small sample model to benchmark the image processing techniques.
I will try to implement some new Image processing and Augmentation techniques, which we researched in the first two weeks.

Week 5-6 (2 Weeks)

These two weeks will be entirely spent on experimenting and benchmarking various state of the art neural network architectures for the classification model. I will be experimenting with various regularisation techniques (like Stochastic depth, block drop, dropout, normalization etc.) to fight overfitting to efficiently scale the model. By the end of the sixth week, I will finalize the final architecture of our classification model and start training it. I will also decide on the learning rate policy in these two weeks.

Week 7-8 (2 Weeks)

I will spend these two weeks building the regression model. I will repeat all the same things and experiments that I did while building my classification model. I will complete building the regression model and start training it.

Week 9-11 (3 Weeks)

I will do a hyperparameter search for both models and finetune them to get the highest possible validation scores. I will deploy multiple methods such as Bayesian, Evolutionary and Pruning to effectively tune the hyper-parameters and obtain improved scores. I will complete finetuning both the networks by the end of the eleventh week.

Week 12

I will merge all our code this week and add comments and documentation (any remaining) to various parts of our code for it to be easily understandable to the end-user. I will also make an output pipeline for the end-user to use the network. I will merge both the models and their respective pipelines with the already present DeepLense pipeline.

FUTURE DEVELOPMENT

Although our aim in this project will be to achieve the highest result using CNNs still, there will be much room for improvement remains. To further scale up our model, I would like to use Transformer and Graph neural network architectures to get better validation metrics. Also, using more extensive networks calls for higher computational resources; hence to scale my network, I would like to use distributed training strategies to reduce my training time. I would also want to make the model much more deployable in future so that it can evaluate using minimum computational power.