<u>PROPOSAL FOR DEEPLENSE- ML4SCI, GSoC '23</u>

Ansh Kanotra
BITS Pilani, Hyderabad Campus

Transformers for Dark Matter Morphology with Strong Gravitational Lensing

ABSTRACT

Through this project, the aim is to apply Vision Transformer structures onto gravitational lensing data obtained from various facilities & simulated and further improve the DeepLense pipeline that automatically classifies lensing images, which will help in multiple scientific applications like finding the most distant galaxies, and identifying the true nature of dark matter.

INTRODUCTION

CONCEPTUAL FRAMEWORK

Gravitational Lensing is a phenomenon in which the path of light is bent by the gravitational field of a massive object. It can cause distant objects to appear distorted, magnified, or even multiple images of the same object to appear in the sky.

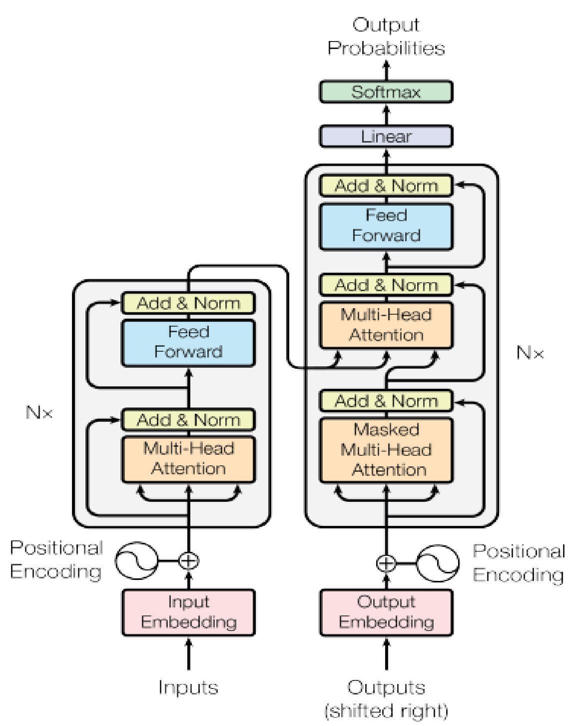

Figure 1: The Transformer - model architecture.

Transformers are a type of deep learning model that were introduced in 2017 by Vaswani et al. They are primarily used for natural language processing tasks such as language translation, language modeling, and text classification. The name "Transformer" comes from the fact that the model operates on the input sequence and transforms it into an output sequence. The main innovation of the Transformer model is the self-attention mechanism, which allows the model to weigh the importance of different parts of the input sequence when generating the output sequence. **Self-attention** allows the model to capture long-range dependencies between different parts of the input sequence, which was previously difficult to achieve with other models. The Transformer model consists of an encoder and a decoder. The **encoder** takes in the input sequence and produces a set of hidden representations, while the **decoder** takes in the output sequence and uses the hidden representations to generate the final output. Both the encoder and decoder consist of multiple layers, each of which contains a multi-head self-attention mechanism and a feedforward neural network. During training, the Transformer model learns to optimize a loss function that measures the difference between the predicted output and the actual output. The model is trained using backpropagation and gradient descent, just like other deep learning models. Overall, the Transformer model has become one of the most popular and effective models in natural language processing, and has been used in a wide variety of applications.

Vision Transformers are also a type of deep learning transformer architecture for image classification that use self-attention mechanisms to capture global context information. They have shown promising results on several benchmarks and are highly parallelizable.

AIM

The aim of my project will be to apply the latest architectures of vision transformers onto the dataset provided in order to improve the accuracy of classification of gravitational lensing images, quite similar to the test provided in **DeepLense**, but increasing the number of **ViT** models, and tuning hyperparameters to improve the pipeline for dark matter morphology. Additionally, I would also like to research about how we can use **Bayesian Learning Approaches** on this problem.

RELATED WORK

As someone deeply interested in both Astrophysics and Machine Learning, I have already worked on a project in a similar domain, namely, [Finding, Classifying & Analyzing ExoPlanets] in which we used **Artificial Neural Networks** to classify whether any given space object is an exoplanet or not. I have also done some other simpler ML Projects like making a [Logistic Regression Model] from scratch

Moreover, both the common test and specific test provided me with much needed experience in the field of **Computer Vision**, introducing me to the fascinating concepts of Transformers and how they work (Self Attention, etc), and also helped me find out how to handle different types of files like tgz.
The Common Test was the first time I used **CNNs and LSTM** in a project and the specific task was very fun to explore as I had to research a lot on how **Transformers** actually worked. I have provided the sources of my research in the GitHub Link(DeepLense Test) to the tests as well.

APPROACH & METHODOLOGY

Since this is a **175**/350 Hour Project, I will demonstrate here what the workflow of the project will be but I am open to suggestions by my mentors as well.

20/40 Hours - *Research*
An Extensive Research on the Topic (Reading Research Papers & Learning about Newer Architectures)

7.5/10 Hours - *Dataset*
Finding, Simulating & Preprocessing the Dataset (using inbuilt & specific techniques)

130/280 Hours - *Model Tuning*
Using the processed data to train different & newer models of ViT Architecture, like Swin, CSwin, DPT, LeViT and even trying to deploy a transformer architecture from scratch.
Around 20 hours will be given to each Model.

17.5/20 Hours - *Documentation*
Documentation of Results and Research done in the form of Slides, Poster and a Research Paper, also hosted on GitHub.

I would also like to research Bayesian Learning Methods for my problem during the research phase, and if something concrete is found, I will try making a model for the same.

Please note that I have divided time according to both, a 175 hour project and a 350 hour project, and therefore I am open to discussion with my mentors for this, and also the division of time. For now, I will be going with the **175 hour project**.

Thank you for this opportunity :)

BIBLIOGRAPHY

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

2. Ellis, R. S. (2010). Gravitational lensing: a unique probe of dark matter and dark energy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *368*(1914), 967-987.

3. Graham, B., El-Nouby, A., Touvron, H., Stock, P., Joulin, A., Jégou, H., & Douze, M. (2021). Levit: a vision transformer in convnet's clothing for faster inference. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 12259-12269).

4. Alexander, S., Gleyzer, S., Parul, H., Reddy, P., Toomey, M. W., Usai, E., & Von Klar, R. (2020). Decoding dark matter substructure without supervision. *arXiv preprint arXiv:2008.12731*.

5. Alexander, S., Gleyzer, S., McDonough, E., Toomey, M. W., & Usai, E. (2020). Deep learning the morphology of dark matter substructure. *The Astrophysical Journal*, *893*(1), 15.