# Transformers for Dark Matter Morphology with Strong Gravitational Lensing

Google Summer of Code 2022

## ➔ PERSONAL INFORMATION


- Name: Aritra Das
- University: Jadavpur University
- Branch: Computer Science Engineering
- Batch: 2019-2023
- Location: Kolkata, India
- Github: AritraStark (Aritra Das) (github.com)
- Resume: MyResume.pdf


## ➔ CONTENTS



## ➔ PROJECT DESCRIPTION

Strong gravitational lensing is a promising probe of the substructure of dark matter to better understand its underlying nature. Deep learning methods have the potential to accurately identify images containing substructure, and differentiate WIMP particle dark matter from other well motivated models, including vortex substructure of dark matter condensates and superfluids.

This project will focus on further development of the DeepLense pipeline that combines state-of-the art deep learning models with strong lensing simulations based on lenstronomy. The focus of this project is using transformers (e.g. vision transformers) to augment the performance of DeepLense algorithms (e.g. classification and regression).

## ➔ CONCEPT INTRODUCTION AND APPROACH FOR THE PROJECT

The canonical candidate for dark matter is a weakly interacting massive particle (WIMP). Indeed, extensions of the standard model generally include WIMPs of mass 100 GeV

that accurately reproduce the observed dark matter density. However, WIMPS have thus far evaded detection, both by direct detection and colliders. There are also hints at cracks in the WIMP paradigm, for example, the core versus cusp problem: observations of halos have consistently shown that actual dark matter halos lack cusps like that of the Navarro–Frenk–White (NFW) profile found from simulation. The treatment of substructure searches as a classification problem using deep learning models complements the existing approaches of statistical detection and identification of individual substructures. Convolutional neural networks (CNNs), Equivariant Neural Networks, Autoencoders have been used to distinguish different classes of substructure in lensing images and further processed to find the position, mass, and other properties of individual substructures.

This project aims to include Transformer to extend the DeepLense pipeline for classification and regression problems. Transformer is a novel architecture of neural networks that has shown promise in NLP and produced better results than networks like RNNs and LSTMs. Transformers work in a Encoder-Decoder architecture and uses Masked Multi Headed Attention block to provide attention to a feature of a data and how it is related to several other features. Transformers also facilitate parallel processing to make use of the latest high speed GPUs to train models faster.

Vision Transformers (ViT) from the 2017 paper “An Image is worth 16*16 words” is an implementation of Transformers for image classification and computer vision that has produced better results compared to CNNs(including ResNet) if pre-training data is sufficiently large. In my task I have implemented a ViT that had an ROC AUC score of 0.93 which is comparable to CNN’s results. ViTs have shown promise in many image classification problems and including them in the DeepLense pipeline can help augment its already cutting edge capabilities. ViTs have the promise to train faster on large datasets and use its Masked Multi Headed Attention to better recognise features.

## ➔ PRIMARY GOALS FOR THE PROJECT

- Researching on transformers and Vision transformers architecture.
- Creating and implementing an architecture best suited for DeepLense algorithms like classification and regression.
- Training and testing on various data to check performance and adjust models and architectures accordingly.
- Integrating the Transformer with the DeepLense pipeline and producing well written documentation for all stages.

## ➔ TIMELINE

This project has a mentioned time of 175 hours over 12 weeks. But I am willing to provide as much time required beyond the given timeline.

17 May - 12 June: During the community bonding period I aim to familiarize myself with the mentors and their related works and engage with the community as well. I would also go through the existing DeepLense pipeline and PyAutoLens codebase to

better understand them. I would also brush up on Dark Matter related research and read up on related theories and concepts again.

Week 1 - Week 3: During this period I will research the Transformer and Vision transformer architecture to find a suitable model for the project. Discuss with the mentors about my research and get their inputs and start preparing for the implementation of the model.

Week 3 - Week 6: During this period I will work on coding and training the model on different datasets and monitor the results. Discuss the results with my mentors and get their inputs about further improving the model.

Week 6 - Week 9: During this period I will start integrating the model with the DeepLense pipeline and start the documentation process of the different stages of development and related research.

Week 9 - Week 12: During this period I will finalize my work and work closely with my mentors to review and improve the project to my best capabilities. Prepare for the final submission and do bug and error fixes if required.

Extended Period (20 September - 21 November): During the extended period I plan to monitor the performance of the model integrated in DeepLense and get feedback from the community on it. Keep communicating with the mentors to contribute to the model according to the feedback and be on the lookout for further avenues of contributions.

## ➔ POST GSOC

After the completion of my project I want to stay invested in this field and related research as my goal is to research and contribute something meaningful for advancement in Astrophysics and CS research. My dream is to help in the discovery of the Theory of Everything and if the results of research on dark matter prove to be in accordance with predictions of String theory then it can be a proof of concept and make its candidature for the Theory of Everything stronger. Even if I don't get selected for this project in GSOC I will keep working to find a way to contribute to Astrophysics research with CS through other avenues. And if I do then I will try to make the most of this opportunity to help advance my career towards my goal and be invested in it.

## ➔ REASON FOR MY INTEREST IN THIS PROJECT

I have been a fan of Physics since high school and have sustained that interest into my CS undergrad. I have been keeping up with it and have studied works of Brian Greene, Michio Kaku, Stephen Hawking, Roger Penrose, Kip Thorne, PJE Peebles and many more. I have a good understanding of the Standard Model, QED, QCD, CMB, Relativity which all are important for understanding what Dark Matter is. I can understand the physics behind the problem and can help model my deep learning solution around it. I can combine my knowledge in both fields to produce the optimal solution to this problem. Dark Matter and Dark Energy are two areas where advancement in research can help us

further the gaps in the Standard Model and lead us one step closer towards the Theory of Everything.

## ➔ ABOUT ME

I am a third year student of Computer Science Engineering in Jadavpur University. I am passionate about Physics as well. I have experience in Machine Learning and Deep Learning focusing on computer vision. Recently I have been working on a research project for IEEE with a professor of my university to create a DL model to detect medical accidents and emergencies. Physics was my favorite subject since high school and I have kept up with studying it. I have in-depth knowledge about Standard Model, String Theory, Relativity and many more related astrophysics topics. I aim to contribute to Physics Research using my CS knowledge. I was motivated a lot by the EHT project and the CHIRP algorithm and hope to do something significant like that in my career.

## ➔ RELEVANT PAST EXPERIENCES

- Deep Learning model for Detecting Higgs Boson with the help of TPUs
- Vision Transformer for classification on the CIFAR-100 dataset
- Comparing performance of different CNN models like ResNet, VGG on MNIST dataset
- Deep Learning model to detect anomaly in regular behavior of people in medical accidents

## ➔ REFERENCES


- [2010.11929] An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale (arxiv.org)
- Deep Learning the Morphology of Dark Matter Substructure - IOPscience
- [1706.03762] Attention Is All You Need (arxiv.org)