Proposal: Transformers for Dark Matter Morphology with Strong Gravitational Lensing

Personal Details

Name: Tanmay Ambadkar
Email: tanmay.ambadkar@gmail.com / tanmay.ambadkar@psu.edu
University: The Pennsylvania State University
Year of study: 1st Year
Major: M.S in Computer Science and Engineering
GitHub: http://github.com/TanmayAmbadkar
Linkedin: https://linkedin.com/in/tanmay-ambadkar
Resume: Tanmay_s Resume.pdf
Timezone: EST/EDT

About Me

I am Tanmay Ambadkar. I am pursuing my MS in Computer Science and Engineering at PennState. I am currently in the 2nd semester of my degree. I am interested in Deep reinforcement learning and computer vision using deep learning.

I am working on a few projects that involve inverse image reconstruction problems like colorization. In addition, I am working on Deep RL-based computer vision problems and composition in reinforcement learning. I work with Pytorch and TensorFlow, with PyTorch being my main library because it is easier to define and manipulate networks and their inputs. I have used TensorFlow for internship projects which have been deployed into production.

I have worked with various deep-learning algorithms for image processing and have taken several courses that have helped me understand how they work. I am keen on working with ML4Sci for the DeepLense projects because I want to apply my deep learning knowledge to a new domain.

The time I can dedicate to the project.

I can dedicate 30 hours a week to the entire project.

Preferred communication medium

I am reachable via email, Google Meet, and Slack.

Why ML4Sci and DeepLense?

I am very interested in applying my deep learning knowledge to various domains. Dark matter has always fascinated me, and the proposed projects have intrigued me.
I would like to understand the work being done at the laboratories and how conventional deep-learning methods can support their research. In addition, super-resolution has always interested me. Deep-learning-based super-resolution has to understand the semantics of the low-res image to then upscale it. I have tried to implement super-resolution algorithms to see which performs the best on a small dataset and how well it has generalized. This is why I would like to work on the Super-resolution project at DeepLense.

Experience with Deep learning

I started working with deep learning in 2019, with my first framework being TensorFlow. I implemented basic projects in computer vision using CNNs for classification. Reinforcement learning interested me, which motivated me to learn Pytorch, as it is much easier to define the training loop for reinforcement learning.

I have taken several courses in deep learning that introduced me to the applications of neural nets. I have worked with residual architectures, self-supervised learning, and GANs in my classes. I have implemented all these algorithms in PyTorch.

My research is at the intersection of deep reinforcement learning and computer vision. My publication - "Deep reinforcement learning approach to predict head movement in 360◦ videos," has been implemented in PyTorch. Using a custom Deep RL framework, I created a multi-frame convolutional network to predict head movement in the videos. In addition, I worked on image colorization, creating a simple network in PyTorch that leverages end–to–end training to show that we don't need a multi-stage very-deep network with millions of parameters to learn semantic transfer of color. The work has been submitted to IEEE SSP 2023.

In addition, I have dabbled with Natural language processing by working with sentence transformers for a project titled Paragraph boundary detection. We have created a custom CNN to process BERT embeddings for binary classification.

Courses at PennState have helped me explore more complex topics in deep learning, like image captioning, image-text retrieval, visual question answering, and neurosymbolic AI, which I have implemented in PyTorch. I am currently studying how to implement robust classifiers in PyTorch.

My internship at Siemens required production-ready code; thus, I developed all networks in TensorFlow, which is easier to deploy. Part of my research at Siemens has been published at IDSTA 2022.

Some relevant projects

The following are some of my relevant deep learning projects that are part of Open source repositories:
1) Deep reinforcement learning approach to predict head movement in 360◦ videos - TanmayAmbadkar/DRL-FOV (github.com)
2) ImageColorNet - Perception consistent image colorization - TanmayAmbadkar/ImageColorNet-Residual-Colorization (github.com)
3) Microsoft Generative Image-to-text transformer - TanmayAmbadkar/GenerativeImage2Text: GIT: A Generative Image-to-text Transformer for Vision and Language (github.com)

Past Open Source experience

I have contributed to open source projects during my undergraduate. My major open-source projects are TanmayAmbadkar/CertificateGenerator (github.com) and TanmayAmbadkar/gymkhana (github.com). Apart from this, all of my research projects are open-source for replication of results and improvement.

The proposal

I wish to apply for "**Transformers for Dark Matter Morphology with Strong Gravitational Lensing**." (350 hours)

Summary

Transformers were introduced in the paper "Attention is all you need" for NLP tasks. They use multiple attention heads to calculate attention scores for tokenized text for different downstream tasks. It is pre-trained on masked language modeling, after which it is finetuned for various tasks like classification and question answering. Following that, "An Image is Worth 16x16 Words" introduced the first Vision transformer. It divides an image into 16x16 patches which are then processed by the attention heads. It is similarly used for various downstream tasks. Multi-modal attention models use transformers for various vision-language tasks like retrieval, captioning, visual-question answering, moment retrieval, etc.

Transformers work extremely well using self-supervised tasks for pretraining and thus are part of every state-of-art model in recent deep learning literature. Transformers can also be fine-tuned for out-of-domain tasks and small-scale datasets. Using transformers for strong lensing images is quite interesting, as the data used differs greatly from what the models are trained on. Thus, I experimented with training a small vision transformer from scratch, which yielded promising results, thus, a promising research direction. With this project, I aim to explore the use of transformers for various tasks using strong lensing images like classification, anomaly detection, regression, and super-resolution.

Milestones

I have created short-term goals that I can follow

1) Pre-training - The transformer will be pre-trained using various tasks like masked-image-modeling. This will also set up a base implementation of the project
2) Classification - Using the pre-trained transformer, the model will be finetuned for classification and trained from scratch to see the difference in performance, training time required, and effect of using a small labeled dataset versus using a large dataset.
3) Regression - Similarly, the model will be finetuned or trained from scratch for regression, and the metrics will be noted.
4) Anomaly detection - Unsupervised anomaly detection is difficult because we must learn the data distribution and then create a statistic to identify out-of-distribution data.

5) Super-resolution - Image super-resolution is one of the tasks in the test. Multiple papers explore super-resolution using transformers, which can directly be adapted for this task.

Note: Anomaly detection and super-resolution are more compute and time-intensive tasks, thus will be given more time in the coding phase.

Current implementation

Huggingface provides a very useful API for using transformers for various tasks. I used the hugging face implementation of ViT for Test V to achieve promising results. I have used 3 attention heads and 3 hidden layers, significantly less than the 12+ heads used for other datasets. The model also takes less time to train (6 minutes on A4000 GPU), which shows that it is efficient and effective for this domain.

I have also trained a model which uses the pytorch VisionTransformer as a base model for the anomaly detection task. This model is also very efficient, using only 4 heads and 4 hidden layers. The use of pytorch VisionTransformer shows that the transformer architecture can be adapted for any downstream task.

Deliverables

1) Tasks - Transformers trained on multiple tasks and datasets, providing a comprehensive evaluation of their use.
2) Library - A library of different transformer implementations, which can be extended for other tasks

Timeline

The goals in the timeline for the coding period is part of the milestones

Task	Duration
Proposal submission	April 4
● Exploring ML4Sci and DeepLense ● Learn in-depth about Strong lensing imaging ● Learn about dark matter properties	April 5 - May 4
Community Bonding [1]	May 4 - May 28

• Learning about the community • Get to know mentors • Learn about other student projects • Get familiar with the review process	
Coding Period starts	May 29 - August 28
Week 1 : Researching about various pre-training tasks and their implementations	May 29 - June 4
Week 2 & 3: Creating multiple pre-training models and adding them to the library	June 5 - June 18
Week 4 & 5: Implement multiple transformers for fine-tuning from pre-trained models and models from scratch for classification	June 18 - July 1
Week 6 & 7: Implement multiple transformers for fine-tuning from pre-trained models and models from scratch for regression	July 1 - July 15
Week 7: Evaluations	July 10 - July 14
Week 8, 9 & 10: Implement transformers for anomaly detection by finetuning and training from scratch	July 16 - August 2
Week 10, 11 & 12: Implement transformers for super-resolution	Aug 2 - August 21
Week 13: Final submission and cleaning up work[2]	August 21 - August 28
Submissions and Final Evaluations	August 28 onwards

[1] I have an exam on May 5, so I will need May 4 and 5 to prepare for the exam.
[2] After working on this project, I would like to report the findings as a research paper.
[3] Since there are 4 weeks of community bonding, I can utilize the time to start researching the datasets and their properties and learn more about transformers and how they can be modified for efficiency and efficacy.

If I cannot complete the target due to unforeseen circumstances (health, family problems), I will put in extra work to ensure the future targets are met. The timeline has been set so that I have

2-3 extra days for every milestone, considering I work 30 hours a week. I do not have any other work/research commitments, so I can dedicate more time if there are delays.

Target

My goal with this project is to understand the domain of strong image lensing and dark matter, as I have not explored this as part of my deep learning research. I want to apply my knowledge to scientific domains that have not used deep learning. Even if I do not get selected, I would like to be in touch with the mentors and contribute towards developing transformers for strong image lensing, enhance my knowledge, and be a part of this contribution to science and open source.

DeepLense test solutions - TanmayAmbadkar/deeplense-tests (github.com)