GSoC 2023 Project Proposal

Organization: ML4Sci

Equivariant Neural Networks for Dark Matter Morphology with Strong Gravitational Lensing

MENTORS

Anna Parul, Sergei Gleyzer, Michael Toomey, Pranath Reddy, Saranga Mahanta, Karthik Sachdev

AUTHOR

Prajwal Uday

CONTENTS

1. Contributor Information....................3

2. Project Abstract....................3

3. Technical Details....................3

3.1. Strong Gravitational Lensing....................3

3.2. Equivariance....................4

3.3. Steerable CNNs and Group Theory....................4

4. Related Works....................6

4.1. Thesis on Equivariant Neural Networks....................6

4.2. Evaluation Task....................7

4.3. Previous work at GSoC....................8

5. Goals and Deliverables....................8

5.1. Goals....................8

5.2. Deliverables....................8

6. Timeline and Plan....................9

6.1. Application Review Period (05 April - 04 May)....................9

6.2. Community Bonding (04 May - 28 May)....................9

6.3. Coding Period (29 May - 21 Aug)....................9

6.4. Post GSoC and Future Work....................11

7. About Me....................11

7.1. Technical Experience....................11

7.2. Why ML4Sci and DeepLense?....................12

8. Other Commitments....................12

9. References....................12

1. Contributor Information

- **Name**: Prajwal Uday
- **Email**: uprajwal20@gmail.com, uprajwal.144@gmail.com
- **Github**: prajwal-144
- **Contact No.**: +91-8851383285
- **University/Institute**: Manipal Academy of Higher Education - Manipal Institute of Technology (MIT Manipal), Manipal, India/GMT +5:30
- **LinkedIn**: https://www.linkedin.com/in/prajwal-uday-1b9678229/
- **Major**: B.Tech. Mechanical and Manufacturing Engineering
- **Minor**: Computational Mathematics

2. Project Abstract

The exploration of dark matter substructures has emerged as a promising avenue for tackling the enduring and complex challenge of understanding the true nature of dark matter. To this end, deep learning techniques are effective in identifying substructures within images and distinguishing WIMP particle dark matter from other plausible models such as dark matter condensates, superfluids, and vortex substructures.

DeepLense is a deep learning pipeline designed for particle dark matter searches with strong gravitational lensing. Developed as a part of the umbrella organization, ML4Sci, it incorporates state-of-the-art deep learning models with simulations based on lenstronomy to deliver robust and accurate results.

3. Technical Details

3.1. Strong Gravitational Lensing

One of the most pressing concerns in physics today is the identity of dark matter. Even though several plausible ideas have been offered in recent decades that might reveal dark matter's true identity, the underlying nature of dark matter has remained a mystery. According to Einstein's general theory of relativity, the presence of a significant amount of matter can distort in the space surrounding it. This phenomenon, known as **gravitational lensing**, occurs when a massive object, such as a cluster of galaxies, creates a gravitational field that magnifies and bends the light from distant galaxies that are located behind it but in the same line of sight. The effect is similar to viewing an object through a magnifying glass, which enables scientists to examine the features of ancient galaxies that would otherwise be too distant to observe with current telescopic technology. In summary, gravitational lensing offers a valuable tool for researchers to study the intricate details of distant galaxies.

Thus, strong gravitational lensing provides us with an effective way to study the **substructure** of dark matter. Unfortunately, there aren't ample strong gravitational lens data available, which means that if we want to train a machine learning model to probe into dark matter substructures, simulations need to be made and used.

3.2. Equivariance

In mathematical terms, equivariance is a form of a specific type of symmetry exhibited by functions that map representations from one space with symmetry to another. A function's output maintains the same relationship concerning symmetrical transformations as its input i.e. if the input undergoes a symmetric transformation, the resulting output should also undergo a corresponding transformation. A network is equivariant if the representations it produces transform in a predictable linear manner under transformations of the input. Given a group G that has actions (a group action is a mapping from a group G and a space X to a space X) on two homogenous spaces X_1 and X_2, a G-equivariant neural network is a linear map $\Phi : f(X_1) \rightarrow f'(X_2)$ that has the property:

$$\Phi[\mathbb{T}_g f(x)] \; = \; \mathbb{T}'_g \Phi[f(x)] \; \forall \, f(x) \qquad (1)$$

Where, $\mathbb{T}_g$ and $\mathbb{T}'_g$ are G-function transforms on the two spaces. For more about Groups or GCNNs, refer to section 3.3.

Equation (1) shows that transforming an input x by a transformation g (forming $\mathbb{T}_g x$) and then passing it through the learned map Φ should give the same result as first mapping x through Φ and then transforming the representation. The layer Φ that maps one representation to another should be *structure-preserving* [1].

Current convolutional neural networks are *equivariant to only translation* of their input, which means translation of an image leads to a corresponding translation of the network's feature maps. However classical CNNs fail to exploit a *larger group of symmetries (including rotation and reflection)* that is present in the data. Equivariant neural networks *(e2cnn)* propose a solution by guaranteeing a specified transformation behaviour of their feature spaces under transformations of the input.

3.3. Steerable CNNs and Group Theory

For the required task, we will use E(2)-equivariant convolutions [2] using the framework of Steerable CNNs [3] in PyTorch [4]. The neural network is equivariant under all 2-D isometries E(2) of the image plane $\mathbb{R}^2$, that is under translations, rotations, and reflections. They generalize over rotationally-transformed images by design, hence reducing the amount of intra-class variability that they have to learn. We will implement the code using *e2cnn* which is a PyTorch extension for Equivariant Deep Learning. E2cnn supports planar isometries and is the predecessor of the *escnn* [5] library. Escnn expands to E(n)-equivariant Steerable CNN which supports steerable CNNs equivariant to both 2D and 3D isometries.

G-Steerable CNNs are a more general framework based on Group Convolutional Neural Networks (GCNN) [1] which describes a neural architecture design not only equivariant to translations $\mathbb{R}^2$, but also equivariant to a broader group of interest *G.* GCNNs directly generalize conventional CNNs by replacing the operation of convolution, usually defined over *planar images*, with that of *group-convolution*, i.e., a convolution performed over a group. However, implementation of group convolution requires storing a response for each group element hence making the framework unfavourable to implement networks equivariant to groups with infinite elements. Steerable CNNs, instead of storing the value of a feature map on each group element, stores the *Fourier Transform* of the feature map, up to a finite number of frequencies.

Steerable CNNs consider sets of transformations that have a group structure. The output of a group convolution is not a signal over the input space but rather a function over group G.

Group symmetry is of two types: *Continuous* and *discrete.*
Continuous group symmetry refers to a mathematical concept that describes the equivariance of a system under continuous transformations. It refers to a group of transformations that are continuous, meaning that they can be smoothly varied by continuous parameters, and preserve certain properties of the system.

Discrete group symmetry describes the equivariance of a system under discrete transformations. It refers to a group of transformations that are discrete, meaning that they can only be varied in finite steps by changing discrete parameters, and preserving certain properties of the system
There are various types of group symmetries that we can consider:

- ***Cyclic Group Equivariance*** is equivariant to N rotations. A group of N discrete planar rotations is represented as C_N i.e. the *cyclic* [6] group of order N. Groups are useful to describe the symmetries of space. Cyclic groups are finite groups generated by a single element and they repeat periodically. They are useful for modeling discrete symmetries that involve rotations by multiples of a fixed angle. The group generated by a rotation $\frac{2\pi}{N}$ is represented as a cyclic group of order N.

- ***Dihedral Group Equivariance.*** A *Dihedral* group is a group of N discrete planar rotations and reflections and is represented as D_N of order *2N*. The symmetry mappings of D_N include the identity, rotations by $\pm\frac{\pi}{N}$, and horizontal/vertical reflections.

- ***Special Orthogonal Group SO(2)*** contains continuous planar rotations. Any cyclic group C_N is a finite subgroup of *SO(2)*. It is the group of *2×2* matrices that are both orthogonal (i.e., their transpose is equal to their inverse) and have a determinant of +1. These matrices represent rotations in *2D* space, and their elements can be parameterized by a single angle θ, which determines the amount and direction of the rotation. The action of a rotation $r_\theta \in SO(2)$ by an angle θ can be defined as

$$\forall x \in X = \mathbb{R}^2 \quad r_\theta, x \to r_\theta . x \ = \ \psi(\theta)x \tag{2}$$

Where $\psi(\theta)$ is the rotation matrix.

- ***Orthogonal Group O(2)*** contains all the continuous planar rotations and reflections. The action of a rotation $r_\theta \in O(2)$ is defined as before for SO(2). A reflection *f* reflects the points around the *x-axis* by inverting the sign of the first coordinate of a point. It is also a group of *2×2 orthogonal* real matrices.

Table 1: Overview of the different groups covered in our task

	Order $\|G\|$	$G \leq O(2)$	$(\mathbb{R}^2, +) \rtimes G$
orthogonal	-	$O(2)$	$E(2) \simeq (\mathbb{R}^2, +) \rtimes O(2)$
special orthogonal	-	$SO(2)$	$SE(2) \simeq (\mathbb{R}^2, +) \rtimes SO(2)$
cyclic	N	C_N	$(\mathbb{R}^2, +) \rtimes C_N$
reflection	2	$(\{\pm 1\}, \star) \simeq D_1$	$(\mathbb{R}^2, +) \rtimes (\{\pm 1\}, \star)$
dihedral	2N	$D_N \simeq C_N \rtimes (\{\pm 1\}, \star)$	$(\mathbb{R}^2, +) \rtimes D_N$

4. Related Works

4.1. Thesis on Equivariant Neural Networks

The master thesis Gabriele Cesa, E(2) - Equivariant Steerable CNNs: A General Solution and Implementation of Equivariance to Planar Isometries [7] provides an implementation of different mathematical objects (groups, representations, irreps, direct sum, induced representations, etc.) and many equivariant layers and operations. They extend *PyTorch's* tensors to geometric tensors to be able to describe feature fields. Feature fields are typed features associated with transformation law. The author has shown an implementation of the first convolution layer of a C_8 - equivariant network built. Defining a convolution layer requires choosing a group $G = (\mathbb{R}^2, +) \rtimes H$. Here, the chosen symmetry group is $H = C_8$. Steerable convolution layer is constructed by passing input ad output field types. The author uses point-wise non-linearities like *ReLU*. An example of an input feature is constructed as a torch.Tensor and wrapped in a GeometricTensor. Finally, layers are composed by applying them sequentially on the input tensor.

4.2. Evaluation Task

The jupyter notebooks and results of the evaluation task are summarised in the GitHub repository: prajwal-144/DeepLense_Tests_2023. The overview of the implementation of the task is thoroughly documented in Task_4_Report. In the evaluation task, we have implemented *finite discrete* groups (C_N or D_N) to model symmetries involving rotations by angles and reflections for the purpose of Binary Classification.

The provided dataset is a set of simulated strong gravitational lensing images with and without substructure. We closely monitor and apply transformations to the data. The images are cropped to 148×148 from the original size of 150×150. We do not crop much because by doing so, we may end up losing important pixel information of the images. The images are then padded to 149×149, allowing us to use odd-sized filters with stride 2 while downsampling a feature map. We then upsample the image by a factor of 3, rotate the image, and finally downsample it again. This helps us reduce interpolation artifacts.

My evaluation task work considered building models using finite discrete groups i.e. Cyclic and Dihedral Symmetries. We have used the C_8 group symmetry, which is a Cyclic group equivariant to 8 rotations. The group is generated by a rotation of $\frac{2\pi}{8}$ radians resulting in a group order $\left|C_8\right| = 8$. We have also implemented D_4 for this task whose symmetry mappings include the identity, rotations by $\pm\frac{\pi}{2}$ radians, and horizontal/vertical reflections. They are mirrored around $x = 0$, resulting in a group order $\left|D_4\right| = 8$. For the input data, we use trivial representation [8]and for all subsequent steps in the G-steerable implementation, we adopt the *regular* representation.

Each block in each model is composed of a convolutional layer, a batch normalization layer, and a ReLU activation function. After each pair of layers, we perform channel-wise average pooling and finally we use a fully connected layer with ELU [8]

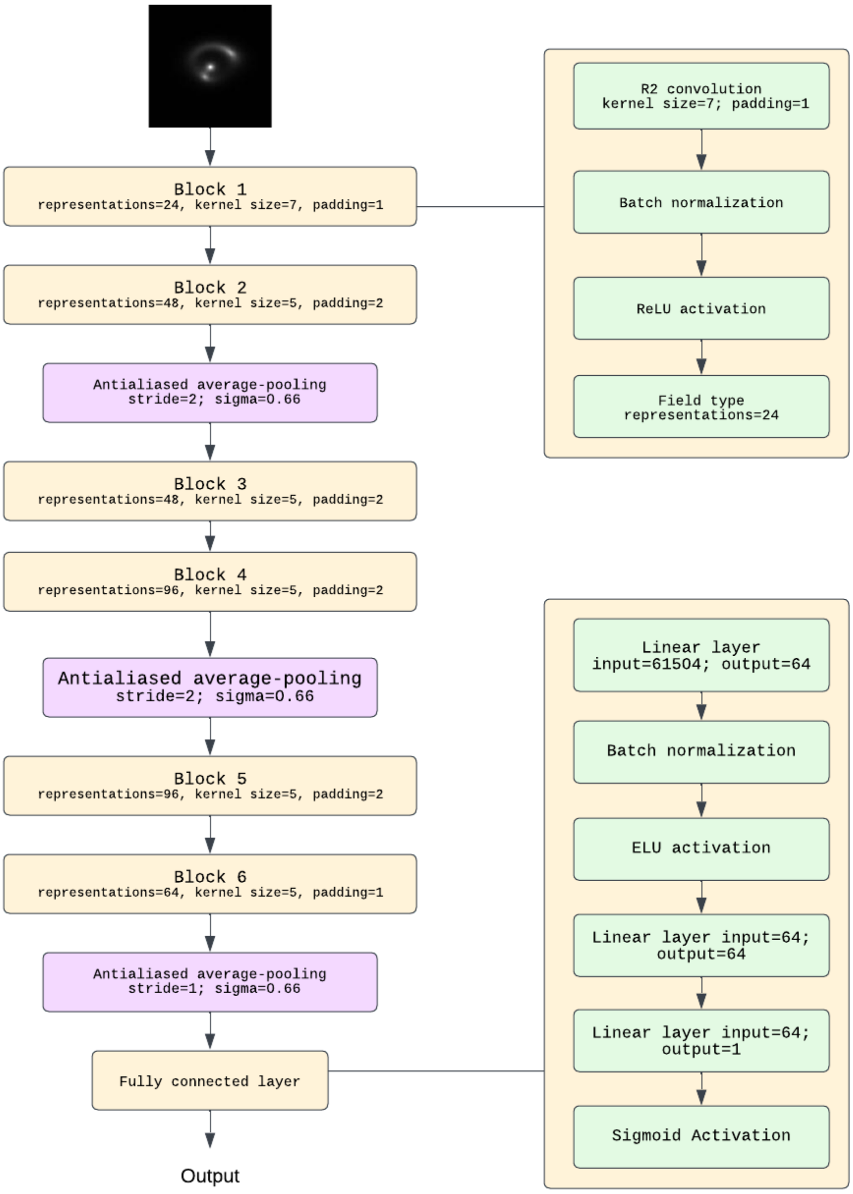

Figure 1: Schematic of the equivariant (ENN) architecture.

non-linearity/activation function and sigmoid function for binary classification in the end. Exponential Linear Unit (ELU) tends to converge loss to zero faster and produce more accurate results. *Fig. 1* shows the architecture of the Equivariant Neural Network used for the evaluation task.

The model design shows very good performance with 99.6% accuracy for the Cyclic Group and 99.4% for Dihedral.

4.3. Previous work at GSoC

The project by Apoorva Vikram Singh contributed to Equivariant Neural Networks for Dark Mater Morphology with Strong Gravitational Lensing in GSoC-2021. It was a follow-up work of the previous attempts to study and determine the morphology of dark matter substructure using deep learning-based approaches. The author made use of the E2CNN library to implement Equivariant Neural Networks. The work majorly focussed on finite discrete groups i.e. Circular (Cyclic) and Dihedral Symmetries. The author worked with C_4 , C_8 , D_4 ,and D_8 symmetries. ResNet-18 model was used for comparison. The subsequent experiments showed very good performance using the ROC AUC score (micro average one-vs-rest) as the evaluation metric.

5. Goals and Deliverables

5.1. Goals

Based on the project description page, this project aims to explore and develop Equivariant Neural Networks to study the substructures in strong gravitational lensing images. The central focus will be on using ENNs for the classification and regression of dark matter particle candidates. The simultaneous focus will be on scientific and mathematical research on ENNs.

5.2. Deliverables

- Implement a Multi-class classification model to classify strong lensing images into three different classes namely, images with no substructure, spherical substructure, and vortex substructure.
- Having previously used the C_8 and D_4 symmetries for binary classification, implement them and other finite cyclic groups C_4 , C_{12} , and C_{16} and dihedral groups D_8 , D_{12} , and D_{16} for multi-class classification.
- Current work focuses majorly on finite discrete groups and symmetries. Though the results are great, we can expand the work by designing networks that use *continuous symmetry* groups like the *orthogonal* symmetry group and the *special orthogonal* symmetry group.
- A report that contains a summary of previous related works, an explanation of the model implementation and how it is trained, and an analysis of their performance compared to other models.
- If time allows, we can expand to unsupervised anomaly detection using Equivariant Neural Networks. This can also be done after the GSoC period is over.

6. Timeline and Plan

6.1. Application Review Period (05 April - 04 May)

During the application period, I will look more into relevant literature useful for the project and brush up on physics-based concepts to get a deep understanding of the broader problem statement. I will also look into the thesis by Gabriele Cesa [7] on E(2) - Equivariant Steerable CNNs which contains accurate mathematical concepts behind the idea of Steerable CNNs. The thesis briefly goes through *Group Theory* and its implementations in order to create Equivariant Neural Networks. This would help me get an overall research idea about the solution we would be using for the more generalized problem which is classifying dark matter particle candidates.

6.2. Community Bonding (04 May - 28 May)

Get feedback from the mentors about the evaluation task and the proposal. Discuss with mentors the goals and deliverables, meeting schedules, and modify them accordingly. After that, discuss with mentors the dataset and preprocessing, the whole pipeline including optimizers, training-testing procedures, and evaluation metrics, and modify the research plan if required accordingly.

Prepare the working environment and repository. Obtain the dataset and clean it. Do all the required pre-processing.

Start by implementing finite discrete symmetries (that were used in the Evaluation task) on the obtained dataset. Perform multi-class classification, analyze and obtain the results using appropriate evaluation metrics (ROC-AUC curve, AUC Score, and Confusion Matrix). This will be done in Jupyter Notebook. Make a short report on the performance, and discuss it with mentors at the first meeting during the "Coding" period.

6.3. Coding Period (29 May - 21 Aug)

- **Week 1 (29 May - 04 Jun)**
  Work on developing a multi-class classification model using finite discrete cyclic symmetries i.e. C_4 , C_{12} , and C_{16} . Evaluate the performance of each of the model symmetries.

- **Week 2 (05 Jun -11 Jun)**
  Optimize the model by tweaking functions and hyperparameters (batch size, number of epochs, learning rate) using the *e2cnn* library and benchmark the performance of each model. Try different pooling methods using the *e2cnn.nn* module to better the performance. Use different dropout layers to prevent overfitting. Evaluate and visualize all the obtained results (loss curves, accuracy curves, ROC curve, and AUC score). Tabulate the average accuracy and AUC score of the best model.

- **Week 3 (12 Jun - 18 Jun)**
  Having implemented the cyclic group symmetries, I will start implementing different Dihedral symmetries for the model. I will make use of D_6 , D_8 , D_{12} , and D_{16} symmetries. In the past works [9] of G-steerable CNNs for astronomical data-based classification, the D_{16} group is the best-performing model.

- **Week 4 (19 Jun - 25 Jun)**
  Optimize the models created using Dihedral symmetries by playing with hyperparameters (batch size, number of epochs, learning rate) and benchmark the performance of each model. Use different dropout layers to prevent overfitting. Evaluate and visualize all the obtained results (loss curves, accuracy curves, ROC curve, and AUC score). Tabulate the average accuracy and AUC score of the best model.

- **Week 5 (26 Jun - 02 Jul)**
  Create a base model which will be used to compare the performance of all the other models. We can choose between many available architectures like ResNet, DenseNet, InceptionNet, and EfficientNet. Use pre-trained models of these architectures and modify/add the classifier head at the end. Compare all the previously implemented models using discrete group symmetries with the base model and obtain the results using the appropriate evaluation metrics. Reduce the overall generalization error. Visualize the Feature space of images and the obtained results.

- **Week 6 (03 Jul - 09 Jul)**
  This week is a buffer week for any unprecedented delays. Complete all the previous tasks if remaining, prepare a blog post, and prepare a report for Phase 1 evaluation (Mid-term evaluation).

- **Week 7 (10 Jul - 16 Jul)**
  Now that I would have implemented models using discrete group symmetries, I will start looking into example usages of continuous group symmetries. Design networks that can use continuous symmetry groups. I will implement a model considering an *orthogonal group O(2)* and a *special orthogonal group SO(2).* Evaluate the performance of the models built.

- **Week 8 (17 Jul - 23 Jul)**
  Optimize the continuous models by tweaking the functions and important hyperparameters in order to get the best performance and reduce the generalization error. Evaluate and visualize all the obtained results (loss curves, accuracy curves, ROC curve, and AUC score). Tabulate the average accuracy and AUC score of the best model.

- **Week 9 (24 Jul - 30 Jul)**

Compare the models built using continuous group symmetries with the base model, obtain and verify the results, and write additional unit tests. Visualize the Feature space of images and the obtained results.

- **Week 10 (31 Jul - 06 Aug)**
  Finalize the complete documentation/report for the final evaluation and clean up all the notebooks used for the project. Add comments and markdowns explaining every step in the notebooks. This will help potential future contributors to begin and expand on the project.

6.4. Post GSoC and Future Work

After completing the final evaluation of GSoC and discussing it with the mentors, I would love to implement any additional features and contribute to open scientific research in ML4Sci.

7. About Me

7.1. Technical Experience

I am currently pursuing (3^{rd} year) a Bachelor of Technology in Mechanical Engineering with a Minor in Computational Mathematics at Manipal Institute of Technology, Manipal. I have been an active member in student projects involving autonomous UAVs. As a part of the team, I also look after research projects based on applications of computer vision and deep learning on UAVs. Currently, I am working as an independent researcher with ASUS Singapore where the primary work is focussed on domain adaptation for ICD Medical Coding. Having previously interned at TATA Advanced Systems Ltd. as a Computer Vision Intern, I possess basic corporate/industry-level experience that is required to put my work into real-life usage.

I have been able to acquire a diverse range of knowledge and skill sets, spanning from computer vision, image processing, data science to machine learning and natural language processing. As a result, I feel confident in my programming abilities, with most of my experience being in Python, and decent experience in C++, C, and MATLAB.

As a person, I strive to give my utmost effort into things that I am passionate about. I enjoy taking up challenges and the endless possibilities of AI encourage me to explore the depths of it. Having spent the past two years researching and implementing Deep learning techniques for various projects, I have found myself to be deeply interested in machine learning, multi-modal ML, and computational mathematics.

I'm always open to new experiences and I hope to be able to continuously learn and apply my knowledge to the fullest.

7.2. Why ML4Sci and DeepLense?

I have always been an enthusiast when it comes to astronomical studies. During my search through various organizations, I found ML4Sci as one such unique organization that is dedicated to leveraging the power of machine learning in the basic sciences. As I explored the various projects, I noticed a few that specifically centered around integrating deep learning techniques into astronomical problem statements. This immediately sparked my interest in these projects. As I continued exploring the evaluation tasks, I got deeply interested in the project Equivariant Neural Networks for Dark Matter Morphology with Strong Gravitational Lensing. During the initial days of research, I came across the thesis on General E(2)-Equivariant Steerable CNNs which provides an implementation of different mathematical objects and many equivariant layers and operations. I found the thesis very intriguing. Having a Mathematics Minor, my interest in the field of artificial intelligence will go well with this GSoC summer project. With my experience in Programming, Data Science, and the impact of my work, I am sure I'll be a great asset to researchers exploring various phenomena in the broader field of Artificial intelligence and Basic Sciences.

8. Other Commitments

- I have my semester-end exams from 22 May - 03 Jun which fall in the "Community Bonding Period" and in the first week of the "Coding Period". Any backlogs during this time period shall be cleared subsequently and the buffer week would be made use of if required to complete the remaining part of the planned task.
- I have a semester break from 04 Jun - 1st week of Aug. Therefore I have nothing else to do this summer, hence I will devote 6-8 hours per day to the project which would sum up to about 42 - 56 hours every week. Once my new semester begins i.e. from the first week of Aug, I will be able to devote 4-5 hours per day which would sum up to 28-35 hours every week.

9. References

[1] Cohen, Taco, and Max Welling. "Group equivariant convolutional networks." *International Conference on Machine Learning*. PMLR, 2016. URL: https://arxiv.org/abs/1602.07576

[2] Weiler, Maurice, and Gabriele Cesa. "General e (2)-equivariant steerable cnns." *Advances in Neural Information Processing Systems* 32 (2019). URL: https://arxiv.org/abs/1911.08251

[3] Cohen, Taco S., and Max Welling. "Steerable cnns." *arXiv preprint arXiv:1612.08498* (2016). URL: https://arxiv.org/abs/1612.08498

[4] Paszke, Adam, et al. "Pytorch: An imperative style, high-performance deep learning library." *Advances in neural information processing systems* 32 (2019). URL: https://papers.nips.cc/paper_files/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html

[5] QUVA Lab, “escnn” [Unpublished], https://quva-lab.github.io/escnn/

[6] Dieleman, Sander, Jeffrey De Fauw, and Koray Kavukcuoglu. "Exploiting cyclic symmetry in convolutional neural networks." *International Conference on Machine Learning*. PMLR, 2016. URL:https://proceedings.mlr.press/v48/dieleman16.html

[7] QUVA Lab - Thesis- E(2) - Equivariant Steerable Cnns A General Solution And Implementation Of Equivariance To Planar Isometries, Gabriele Casa https://gabri95.github.io/Thesis/thesis.pdf

[8] Clevert, Djork-Arné, Thomas Unterthiner, and Sepp Hochreiter. "Fast and accurate deep network learning by exponential linear units (elus)." *arXiv preprint arXiv:1511.07289* (2015). URL: https://arxiv.org/abs/1511.07289

[9] Scaife, A. M., & Porter, F. (2021). Fanaroff–Riley “classification of radio galaxies using group-equivariant convolutional neural networks”. *Monthly Notices of the Royal Astronomical Society*, *503*(2), 2369-2379. URL: https://arxiv.org/abs/2102.08252