Project Proposal: **Domain Adaptation for Improved Lens Finding**

Project Overview:

The goal of this project is to improve the detection of gravitational lenses in wide-area surveys using machine learning techniques. The study will specifically focus on bridging the gap between simulated and actual observational data utilizing domain adaptation approaches. Due to the scarcity of real-world data, simulations will be utilized for training before applying the generated model to real-world observational data. The predicted result is an accurate and effective machine learning framework that can detect lensed systems in observational data, allowing for additional investigation into numerous astrophysical problems, such as examining the substructure in the lensing galaxies' dark matter haloes.

Project Deliverables:

The following deliverables will be provided at the end of the project:

1. A report on the data collection, preprocessing, and analysis process.
2. A machine learning model for lens finding using domain adaptation methods.

3. An evaluation of the model's performance and comparison with existing methods for lens finds.
4. A presentation summarizing the project's findings and contributions.

Project Scope:

The project will involve the following steps:

1. Data Collection: Collect simulated and real observational data from wide-area surveys such as Hyper Suprime-Cam or Dark Energy Survey.
2. Data Preprocessing: Preprocessing the collected data to remove any noise, bias, or artifacts that may affect the lens-finding process.
3. Model Development: Developing a machine learning model for lens finding using domain adaptation methods. The model will be trained on simulated data and validated on real observational data.
4. Model Testing: Testing the developed model on a subset of real observational data that has not been used for training or validation.
5. Performance Evaluation: Evaluating the performance of the developed model and comparing it with existing methods for lens finds.
6. Result Analysis: Analyzing the results to identify the strengths and weaknesses of the developed model and suggest areas for further improvement.

Methodology:

The project will follow the following methodology:

1. Data collection and preparation: We will collect data from wide-area surveys, such as Hyper Suprime-Cam or Dark Energy Survey, and simulated data for lens finds. We will clean and preprocess the data, and format it for use in the machine learning model.
2. Model development and training: We will develop and train a machine-learning model for lens finding using simulated data as the source domain. We will use convolutional neural networks (CNNs) and other deep learning techniques to build the model. We will select appropriate hyperparameters and optimize the performance of the model on the simulated data.
3. Domain adaptation techniques: We will apply domain adaptation techniques, such as adversarial training or domain-specific normalization, to the trained model to improve its accuracy on real observational data. We will test different techniques to find the most effective approach for the given data.
4. Testing and validation: We will test the trained and adapted model on a separate dataset of real observational data to validate its accuracy and performance. We will compare the results of the model with known lensed systems and identify new lensed systems that were previously unknown or undetected.
5. Documentation and presentation: We will document the entire process, including the data preparation, model development and training, domain adaptation techniques used, testing and validation, and the results obtained. We

will prepare a presentation or report to present the findings to stakeholders.

Timeline:

The project will take approximately 175/350 hours to complete. The following is a rough timeline for each stage:

1. Data preparation: 25-50 hours
2. Model development and training: 75-150 hours
3. Domain adaptation techniques: 25-50 hours
4. Testing and validation: 25-50 hours
5. Documentation and presentation: 25-50 hours

Expected outcomes:

The expected outcomes of the project are:

1. Improved accuracy of the machine learning model in identifying lensed systems in real observational data.

Conclusion:

In essence, the goal of this project is to create and evaluate methods for detecting lensed systems in wide-area survey data utilizing domain adaptation approaches using simulated data as the source and real observational photos as the objective. The project's output is a precise and efficient machine learning framework capable of identifying lensed systems in observational data, which will aid future studies into numerous astrophysical topics.