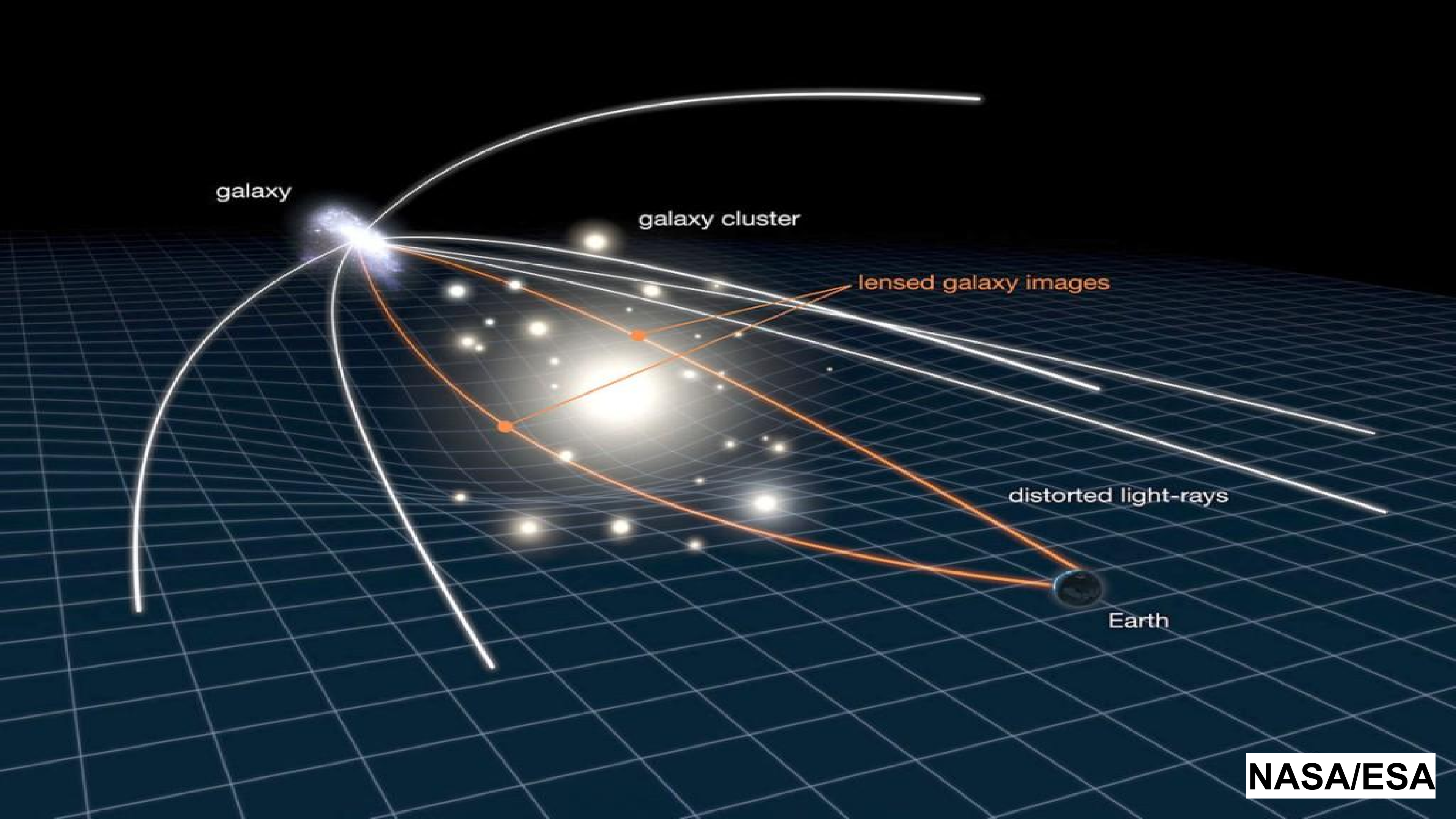
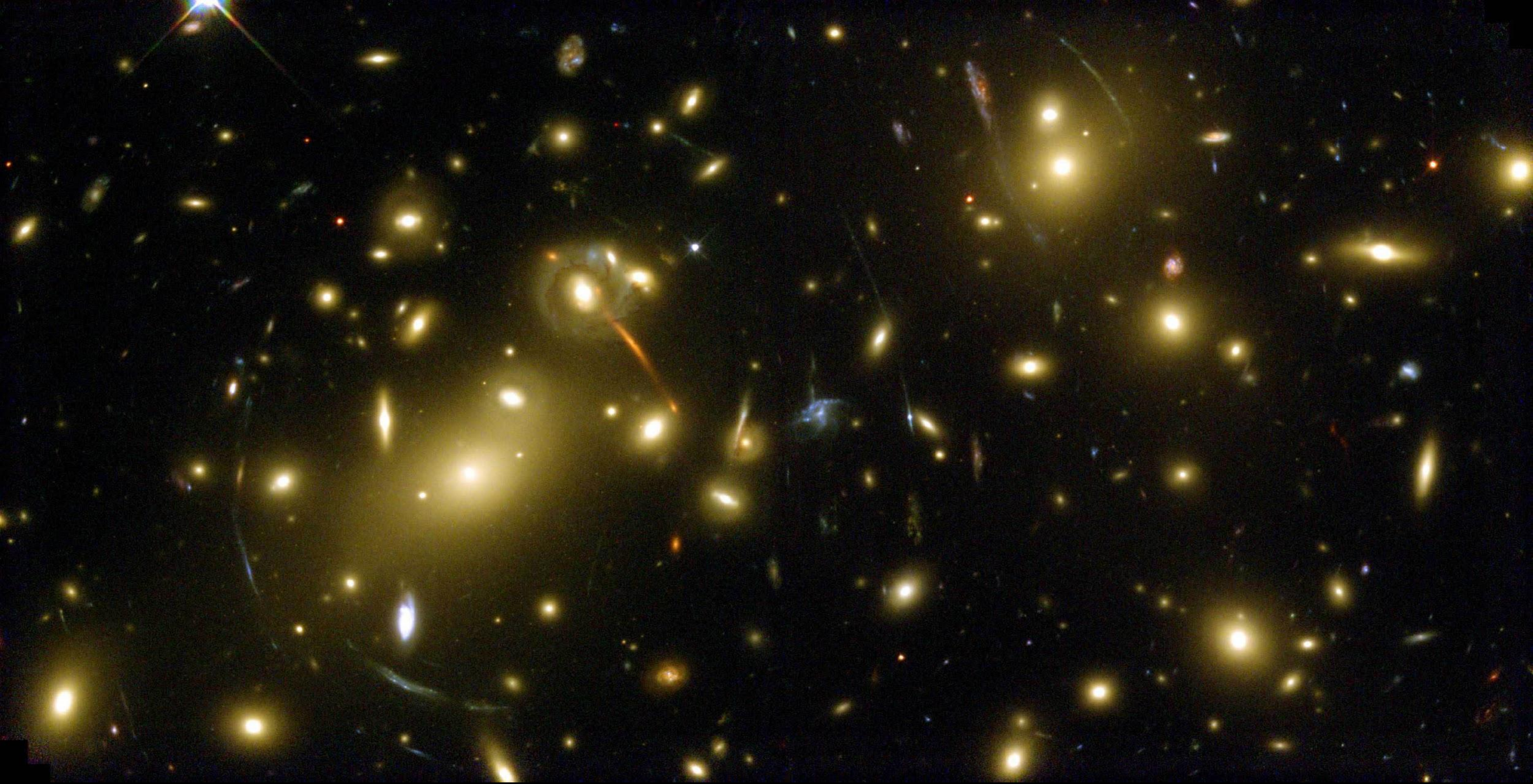
Automated Analysis Of Strong Lenses in the Era of Big Data

Nan Li (NAOC)
on Behave of the CSST-SLWG


galaxy
galaxy cluster
lensed galaxy images
distorted light-rays
Earth
NASA/ESA



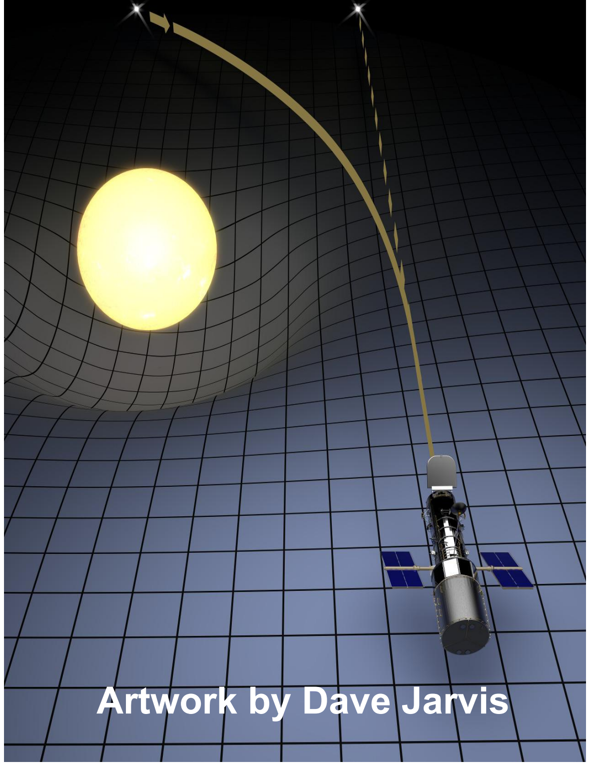
http://www.roe.ac.uk/~heymans/website_images/abell2218.jpg

NASA/ESA


Cosmic Smiling Faces
NASA/ESA

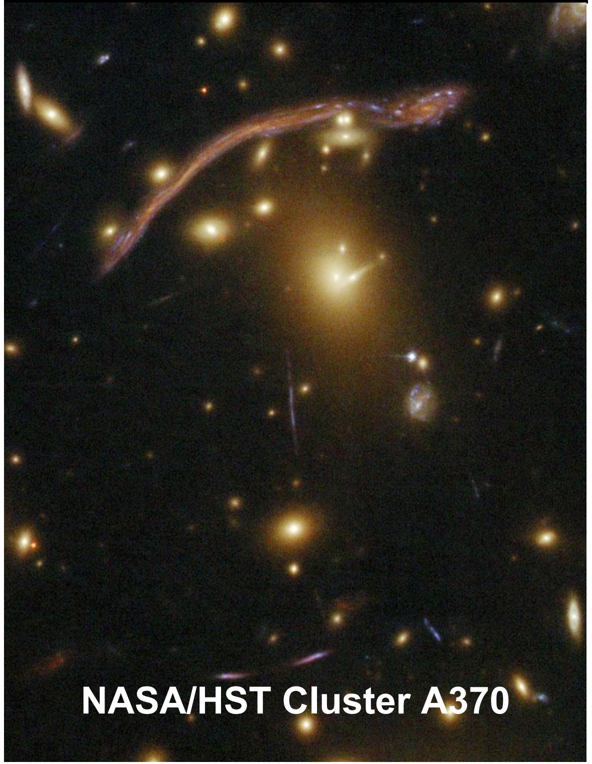
Why Gravitational Lensing Matters?

- Test the theory of GR
- Geometry of the space



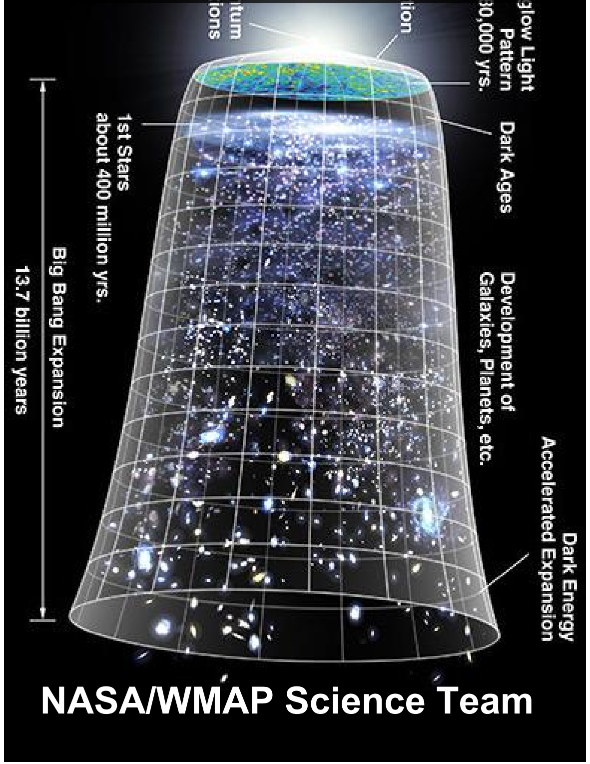
Artwork by Dave Jarvis

- Mass distribution of lenses
- Dark matter substructures



NASA/HST Cluster A370

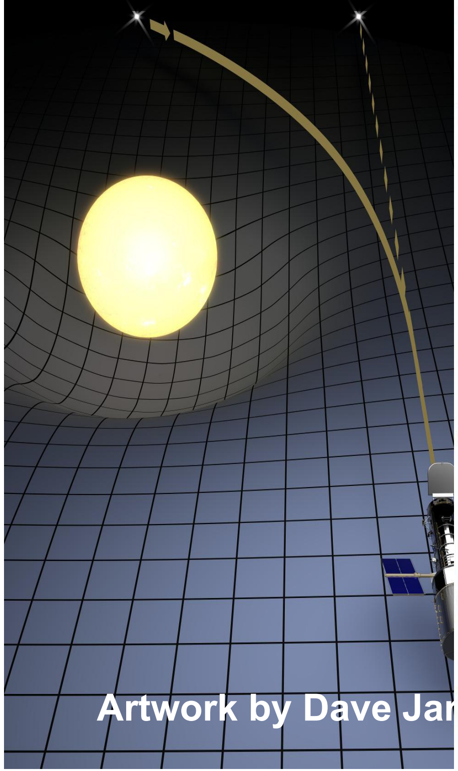
- Structure Formation History
- Cosmological Parameters



NASA/WMAP Science Team

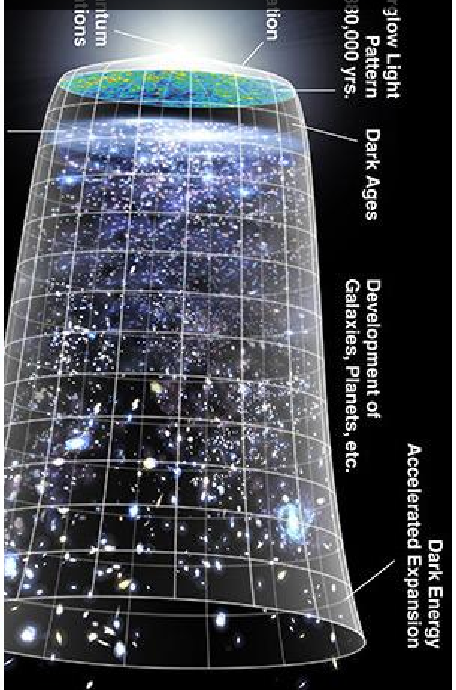
Why Gravi rs?

- Test the theory of GR
- Geometry of the space



Artwork by Dave Jar

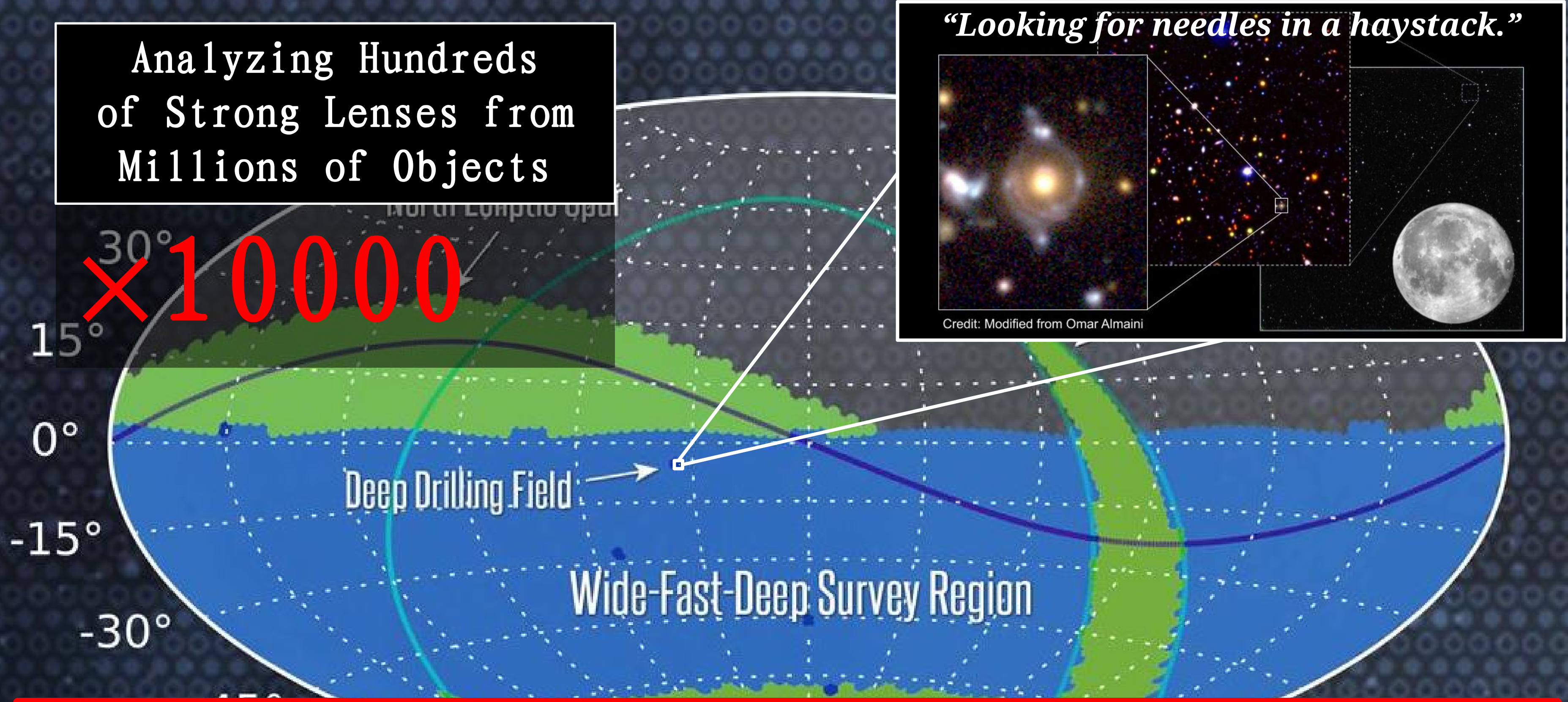
ture Formation History
ological Parameters



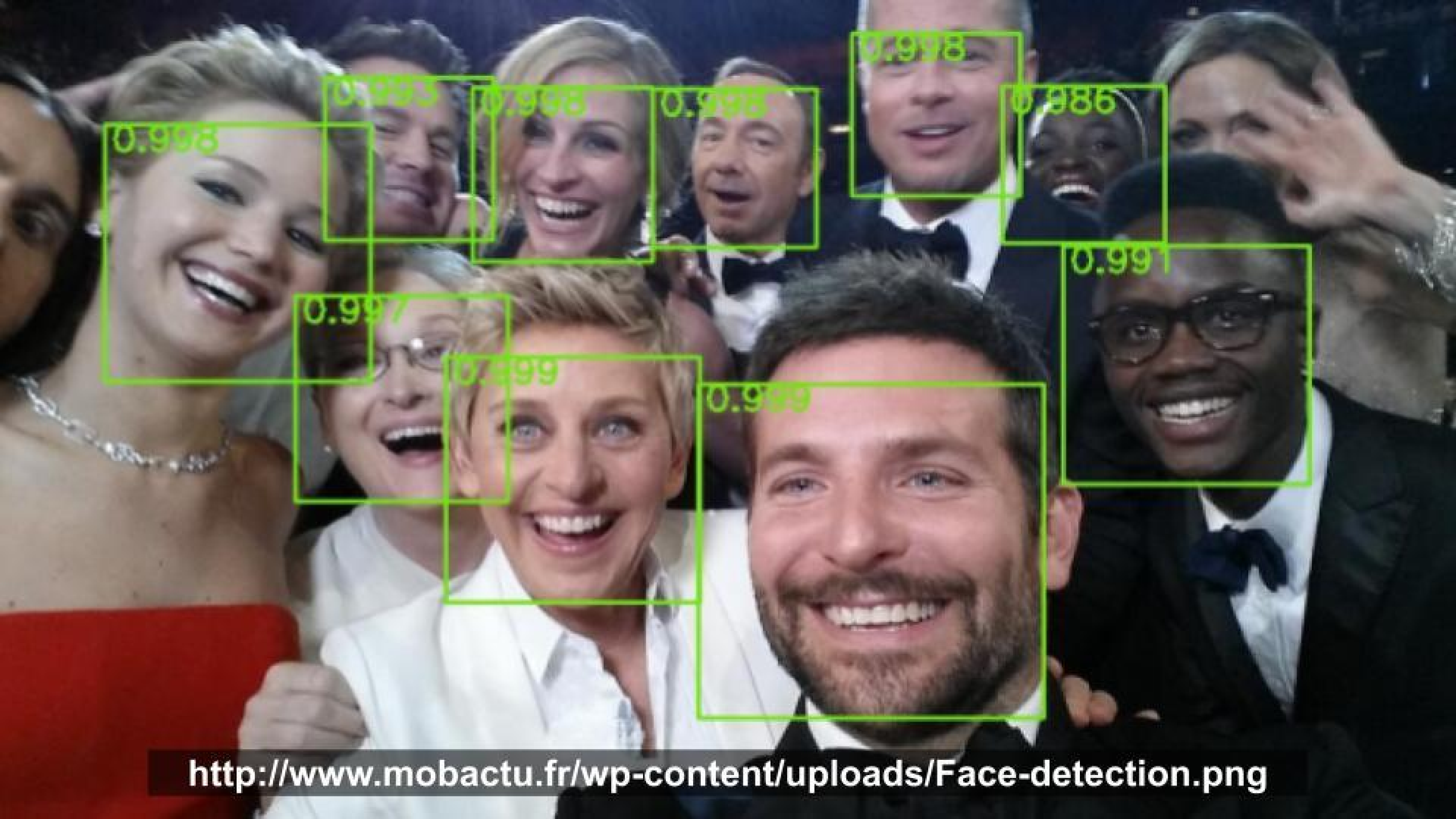
/WMAP Science Team

Analyzing Hundreds of Strong Lenses from Millions of Objects

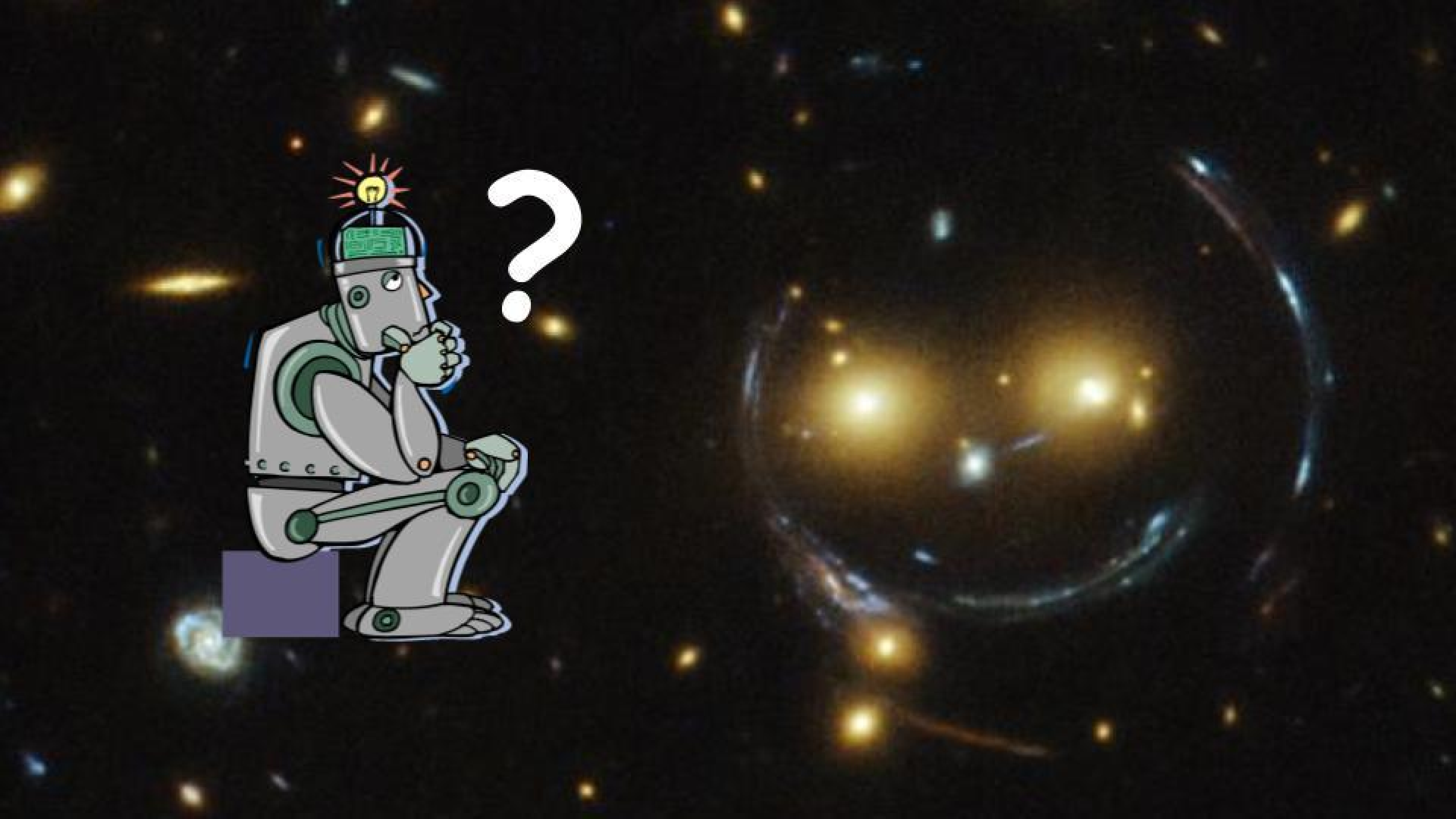
×10000



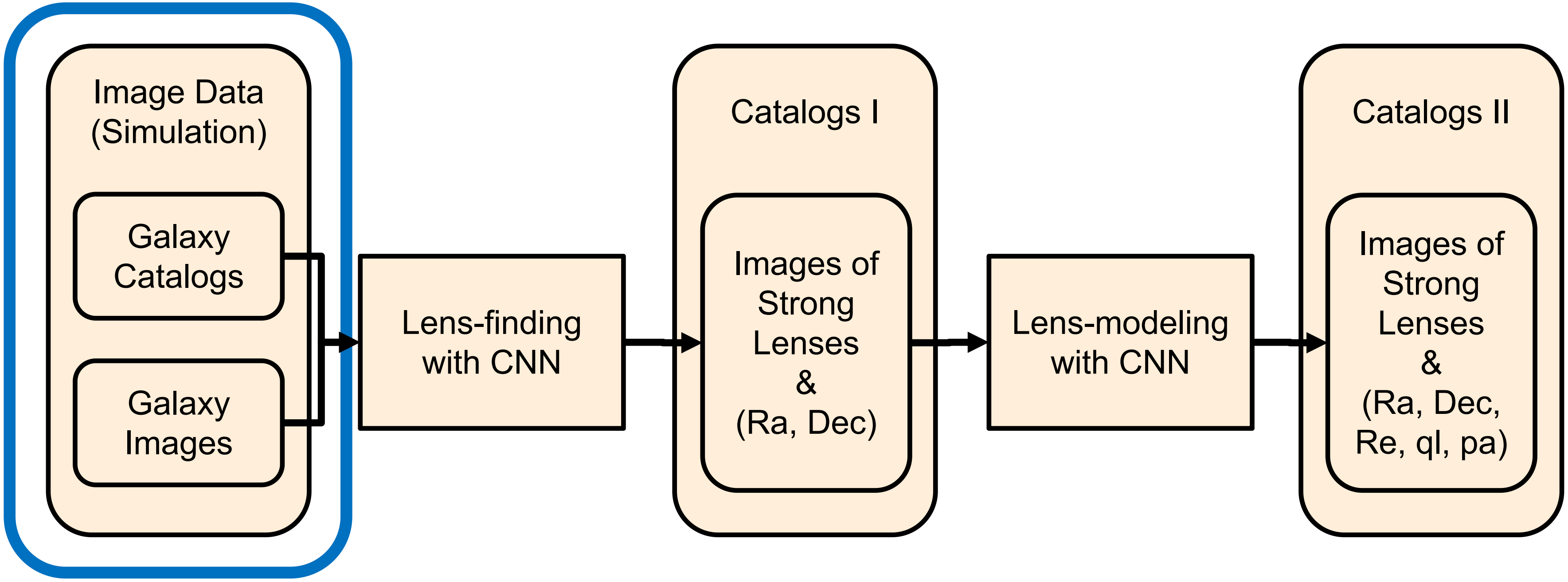
How to identify and model tens of thousands of strong lenses from tens of billions of astronomical objects in the big data era ? ? ?



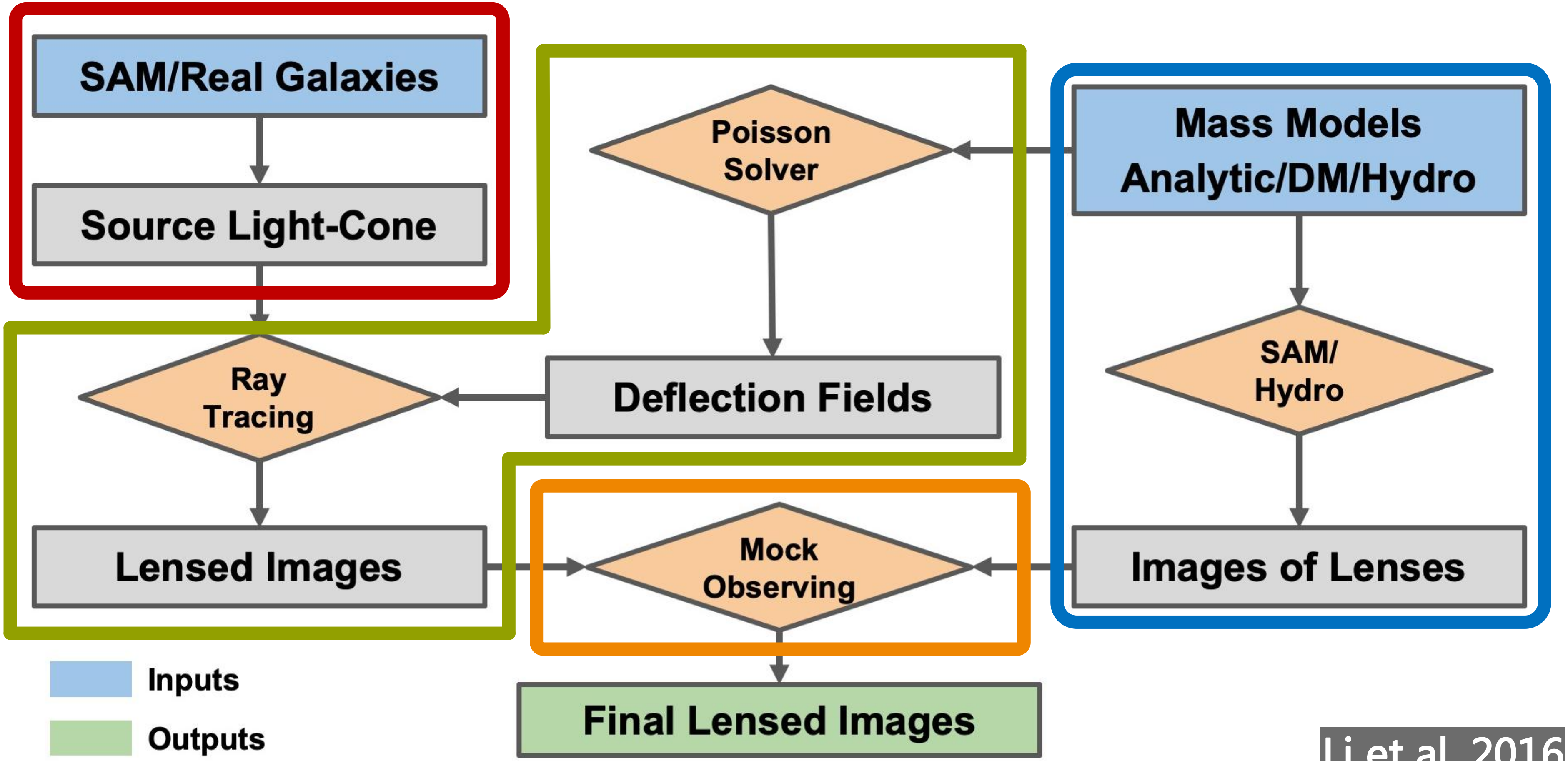
http://www.mobactu.fr/wp-content/uploads/Face-detection.png



How to identify and model tens of thousands of strong lenses from tens of billions of astronomical objects in the big data era ? ? ?

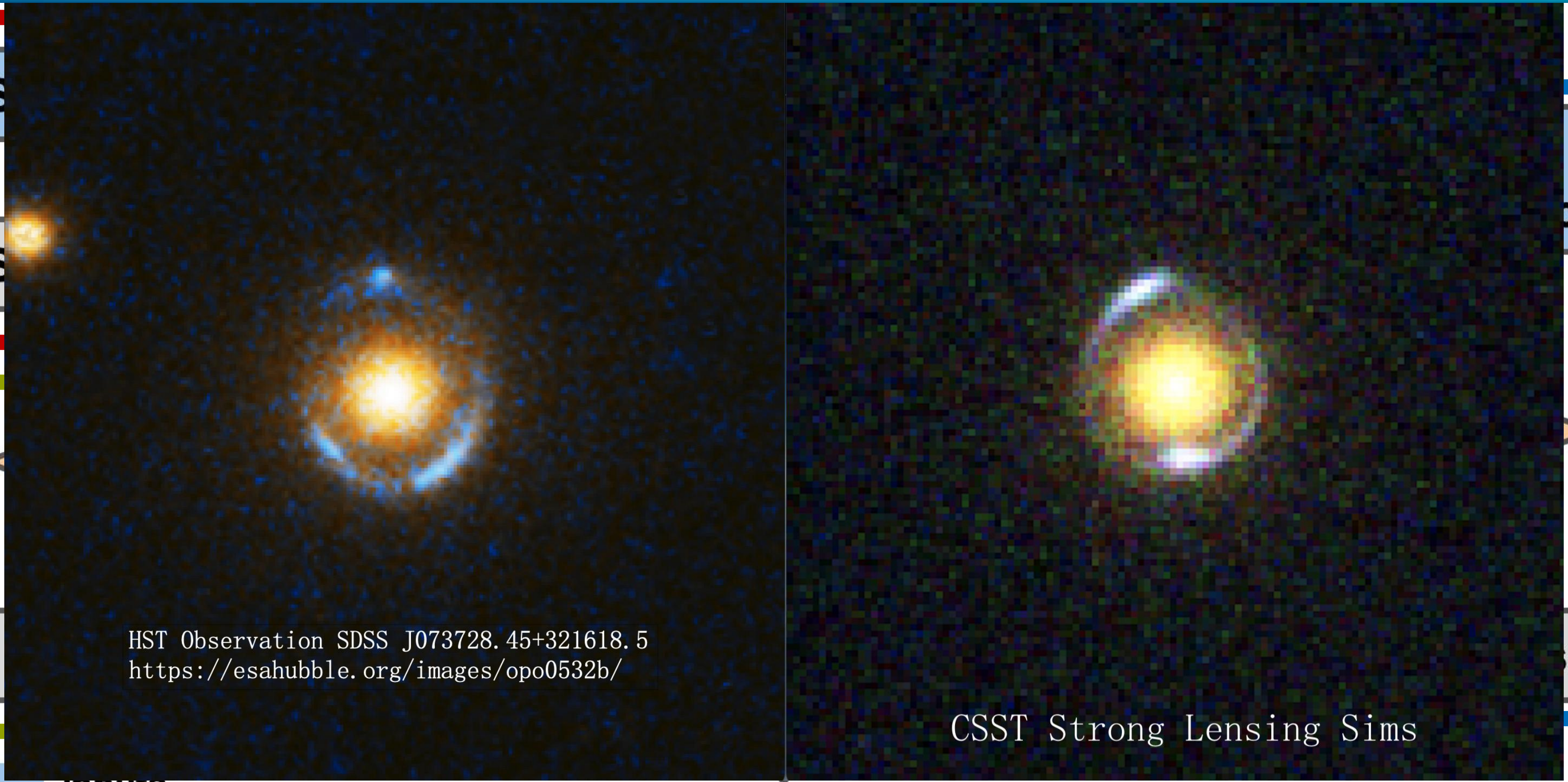


Simulations of Gravitational Lenses (PICS)



Li et al. 2016

Simulations of Gravitational Lenses (PICS)

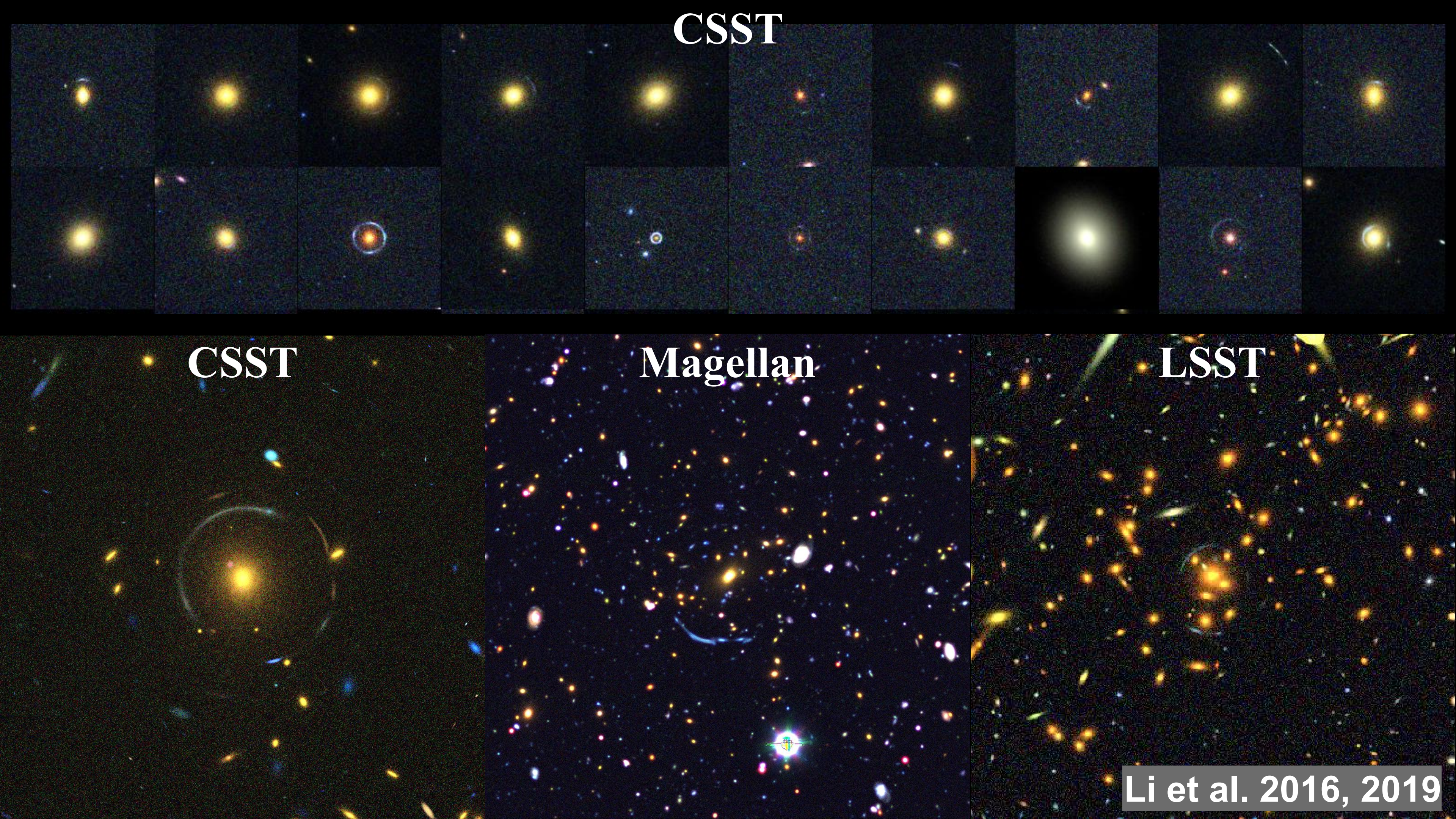


Inputs

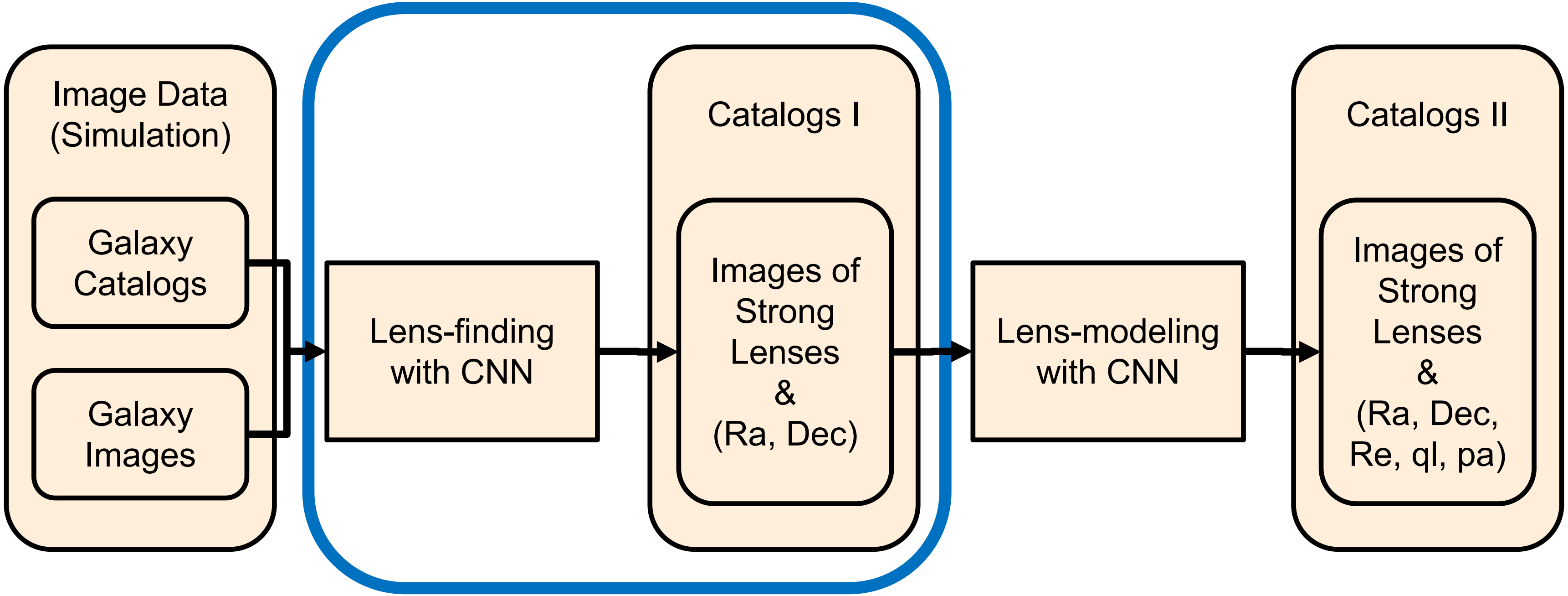
Outputs

Final Lensed Images

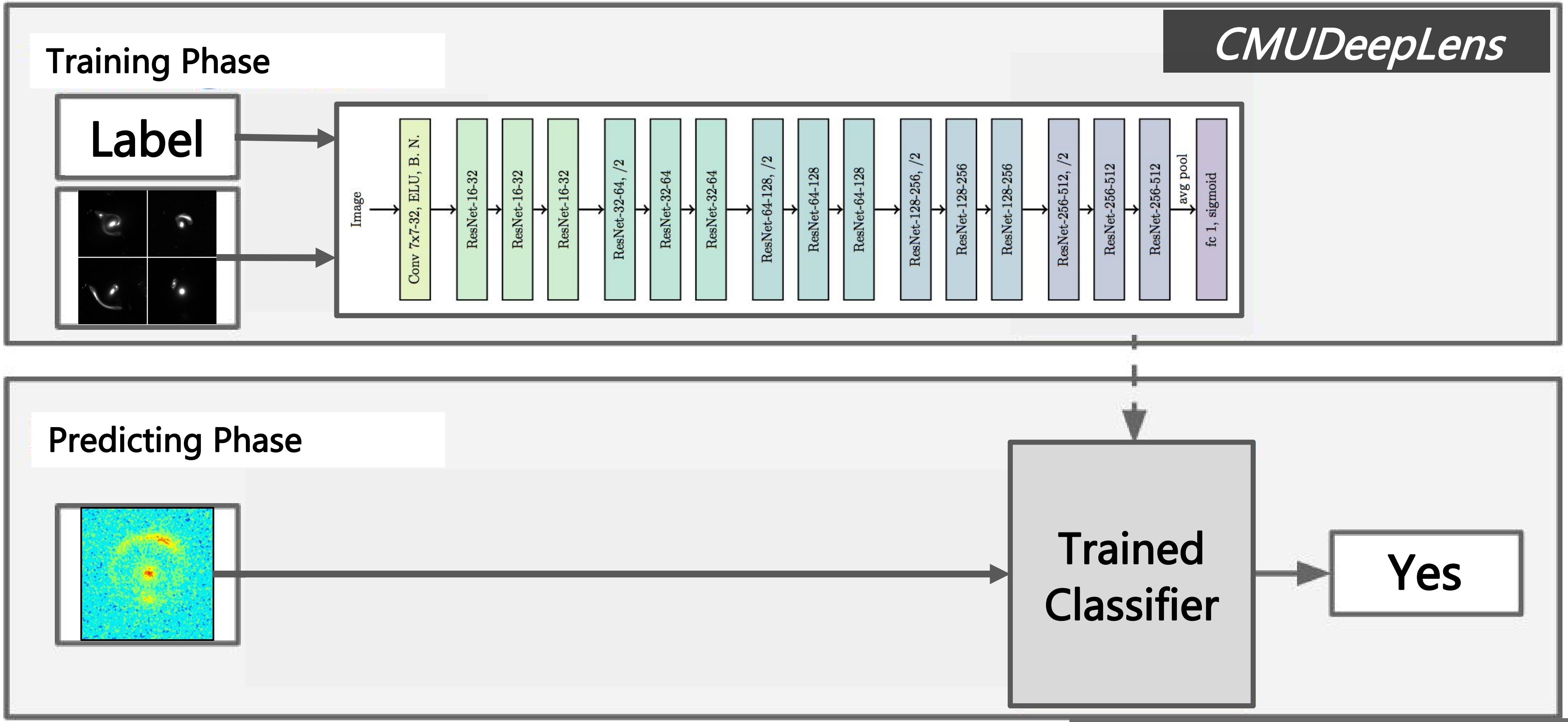
Li et al. 2016


CSST
CSST
Magellan
LSST
Li et al. 2016, 2019

How to identify and model tens of thousands of strong lenses from tens of billions of astronomical objects in the big data era ? ? ?



Identifying Strong Lenses with Deep Learning



Lanusse, Ma, Li et al. 2017

Sorted by area under the ROC curve

CMUDeepLens 98%

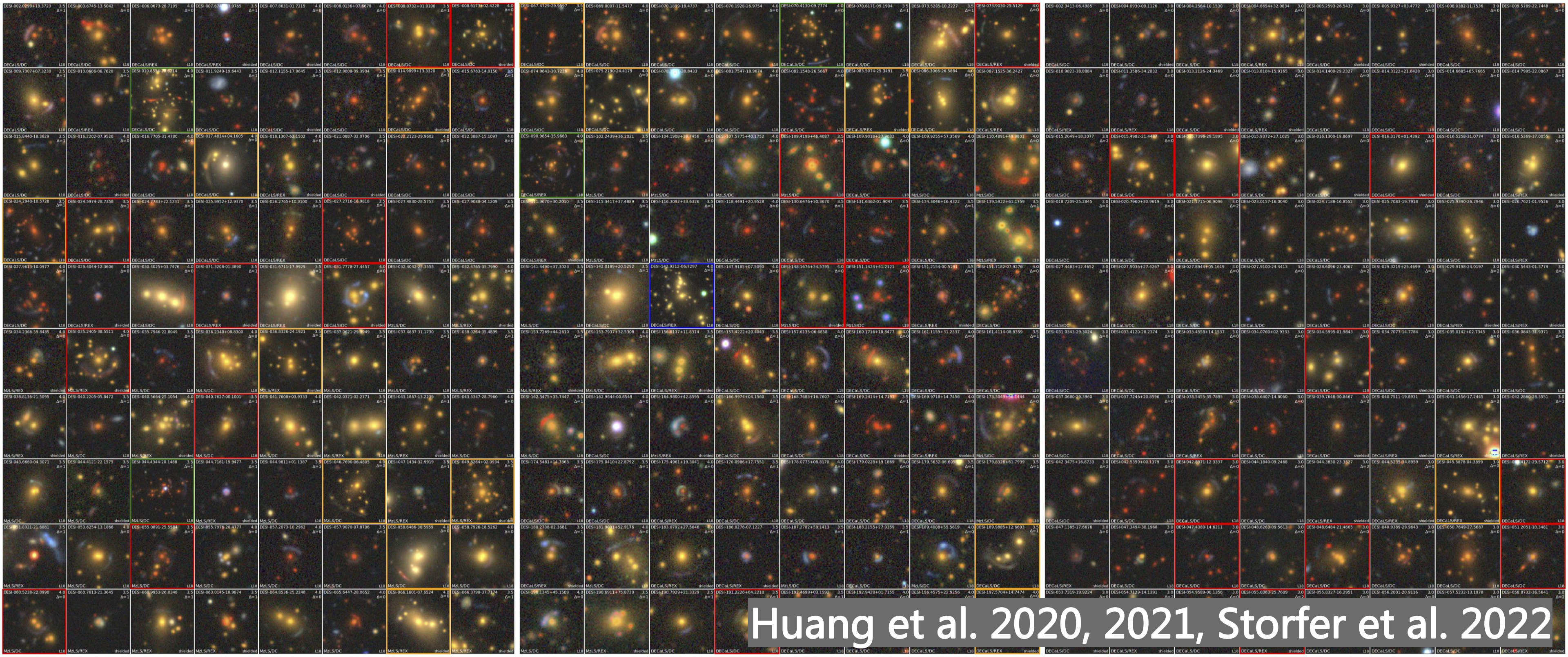
Human Inspection 89%

Traditional Methods 83%

##				TPR0	TPR10	description_short	author.1
## 14				3e-02	0.4534297041	CNN	Francois Lanusse
## 10				0e-02	0.1027314963	CNN	Quanbin Ma
## 20	LASTRO EPFL (111)	Ground-Based	0.9749255	7.493794e-02	0.1131977256	CNN	Mario Geiger
## 3	cas_convnet_mean	Ground-Based	0.9634215	2.022629e-02	0.0761790327	CNN	Colin Jacobs
## 22	Ground	Ground-Based	0.9557059	0.000000e+00	0.0071018193	CNN	Emmanuel Bertin
## 23	Ground	Ground-Based	0.9557059	0.000000e+00	0.0071018193	CNN	Emmanuel Bertin
## 24	Ground_fixed	Ground-Based	0.9557059	0.000000e+00	0.0071018193	CNN	Emmanuel Bertin
## 25	Ground_fixed	Ground-Based	0.9557059	0.000000e+00	0.0071018193	CNN	Emmanuel Bertin
## 9	Philippa Hartley2	Ground-Based	0.9310191	2.237273e-01	0.3453159911	SVM / Gabor	Philippa Hartley
## 7	Philippa Hartley	Ground-Based	0.9293543	2.123763e-01	0.3316908714	SVM / Gabor	Philippa Hartley
## 27	Manchester-NA2	Ground-Based	0.8913778	2.803645e-04	0.0075297887		
## 28	Manchester-NA2	Ground-Based	0.8913778	2.803645e-04	0.0075297887		
## 29	Manchester-NA2-Submission2	Ground-Based	0.8913778	2.803645e-04	0.0075297887		
## 30	Manchester-NA2-Submission2	Ground-Based	0.8913778	2.803645e-04	0.0075297887		
## 4				5e-03	0.0186123524	edges/gradiants and Logistic Reg.	Camille Avestruz
## 13				5e-05	0.0003810517	CNN / SVM	Clecio Roque De Bom
## 31				2e-04	0.0021145867	SExtractor	Alessandro Sonnenfeld
## 16	LASTRO EPFL (13b)	Space-Based	0.9325338	4.773626e-03	0.0779692201	CNN	Mario Geiger
## 8	resnet_5d0aad0	Space-Based	0.9225303	2.206807e-01	0.2904204271	CNN	Francois Lanusse
## 15	GAMOCLASS	Space-Based	0.9210117	7.416406e-02	0.3570444584	DL / CNN	Marc Huertas-Company
## 6	CMU-DeepLens-Resnet-Voting	Space-Based	0.9145407	0.000000e+00	0.0082046692	CNN	Quanbin Ma
## 1	space	Space-Based	0.9143197	6.755404e-04	0.0127852282	CNN	Emmanuel Bertin
## 19	res_bottleneck_87b7e8a	Space-Based	0.9068996	7.506005e-05	0.0038030424	CNN	Eric Ma
## 32	CNN_kapteyn	Space-Based	0.8179482	1.000625e-04	0.0002001251	CNN	Enrico Petrillo
## 21	CAST-SB	Space-Based	0.8128851	6.909326e-02	0.1186942145	CNN	Clecio Roque De Bom
## 5	Manchester1	Space-Based	0.8101726	7.354597e-03	0.1739837398	Human Inspection	Neal Jackson
## 18	Philippa Hartley2	Space-Based	0.8092423	2.859788e-02	0.0812650120	SVM / Gabor	Philippa Hartley
## 17	Philippa Hartley	Space-Based	0.8012731	2.934848e-02	0.0717323859	SVM / Gabor	Philippa Hartley
## 12	Attempt2	Space-Based	0.7626792	0.000000e+00	0.0008265498	CNN / wavelets	Andrew Davies
## 11	YattaLensLite	Space-Based	0.7622929	0.000000e+00	0.0003502802	Arcs / SExtractor	Alessandro Sonnenfeld
## 26	All-now	Space-Based	0.7346352	4.900040e-02	0.0659031545	edges/gradiants and Logistic Reg.	Camille Avestruz
## 2	GAHEC IRAP 1	Space-Based	0.6580909	1.127113e-03	0.0090920476	arc	

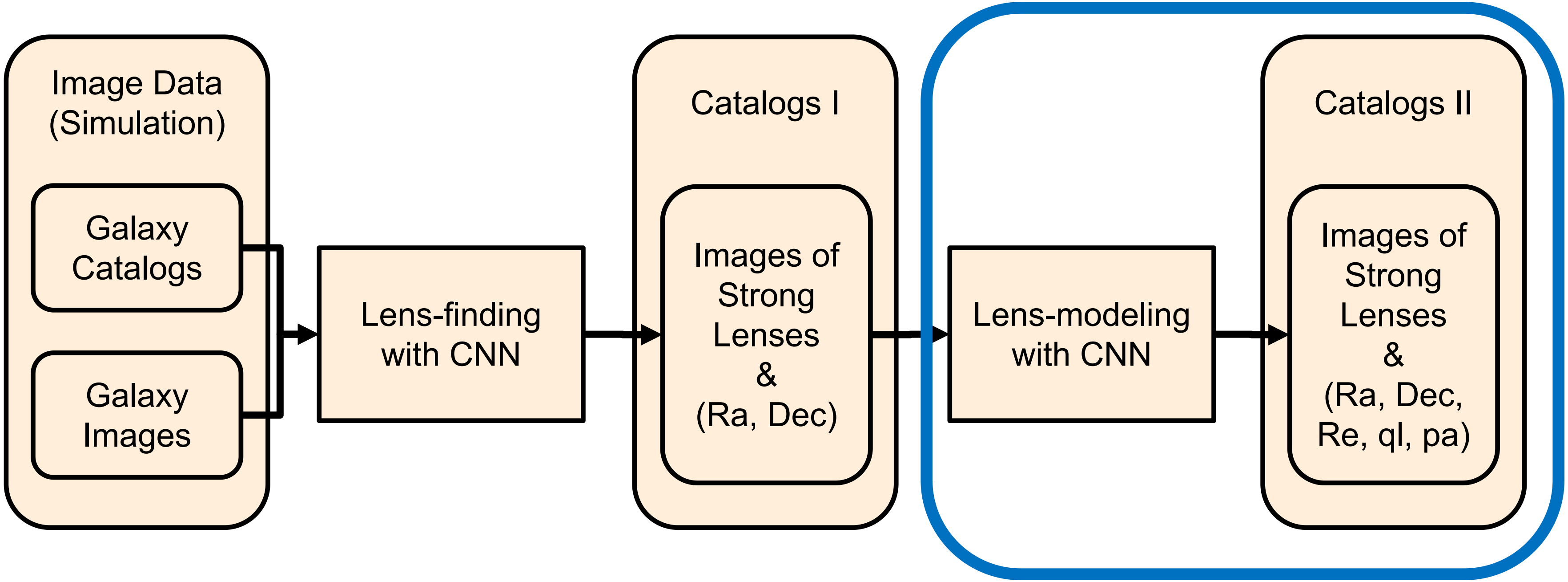
Metcalf et al. 2019

CMUDeepLens is used to search strong Lenses in DESI Legacy Imaging Surveys, found **> 3000 high-quality** candidates, **~ 10 X Current Strong Lens Sample**



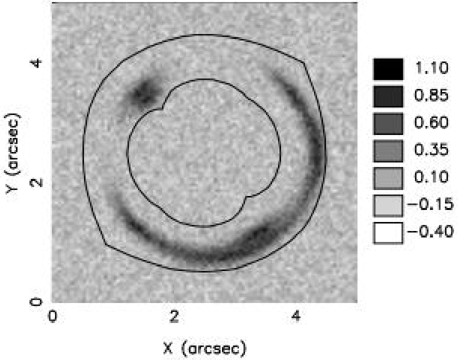
Huang et al. 2020, 2021, Storfer et al. 2022

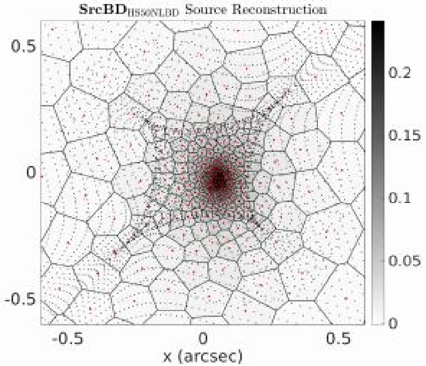
How to identify and model tens of thousands of strong lenses from tens of billions of astronomical objects in the big data era ? ? ?



Modeling Strong Lenses with CNN

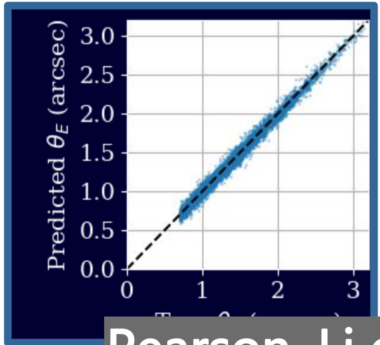
Traditional MCMC： Slow、Low Automatization

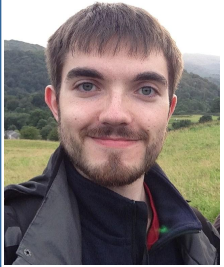




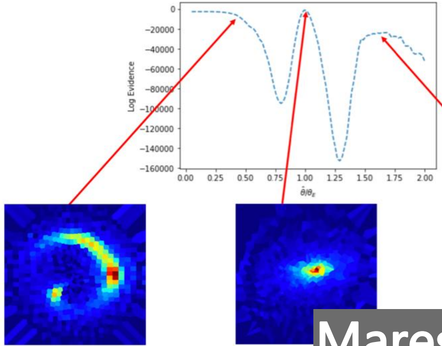
1 system /person /week

CNN-based Method： 10^6 Faster、Fully Automatic





Pearson, Li et al. 2019

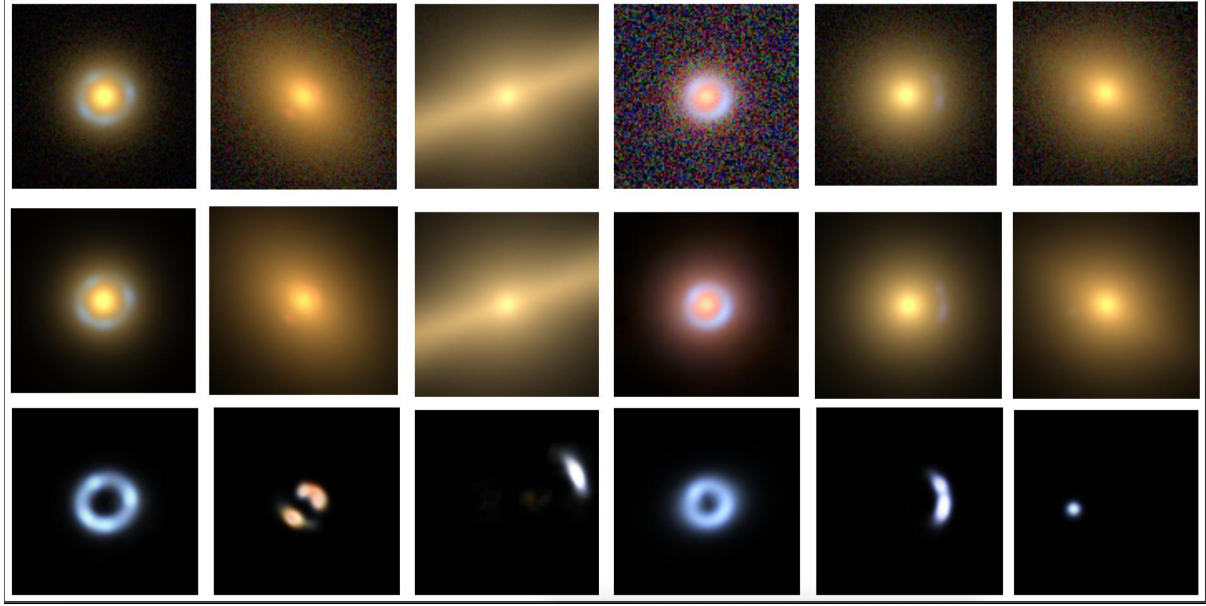


Maresca, Dye, Li 2021

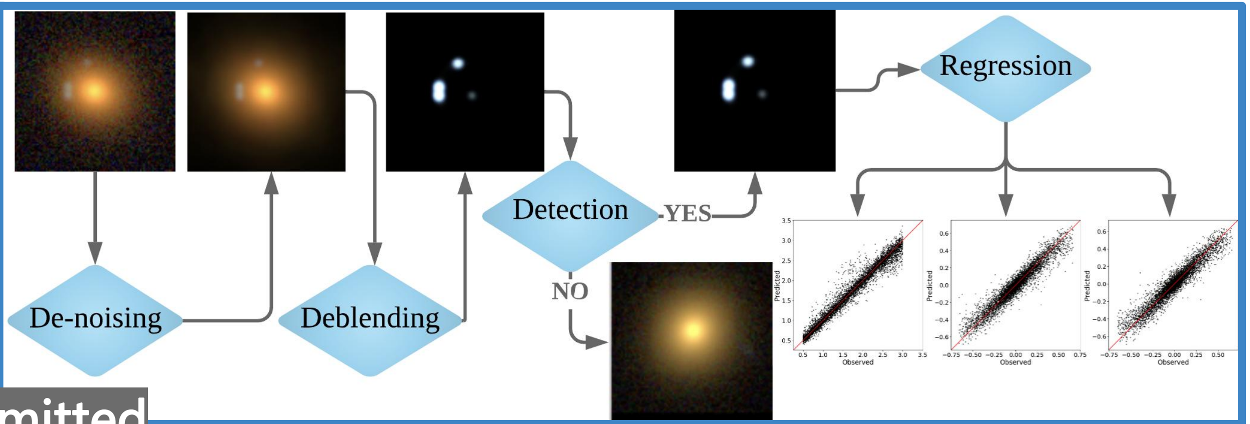
Images

Denoise

Deblend

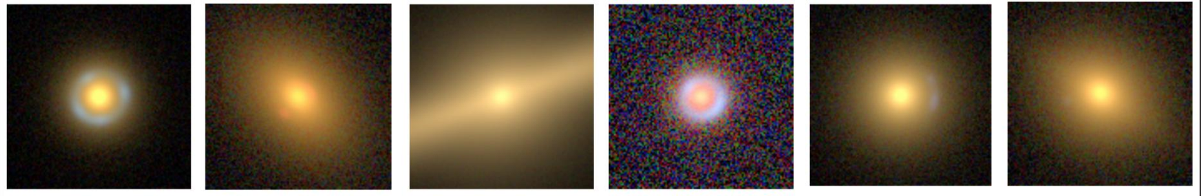


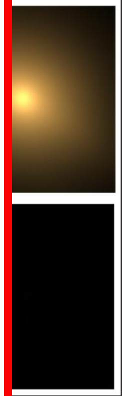
An End to End Tool for Analyzing Galaxy Scale Strong Lenses



Madireddy et al. submitted

Images

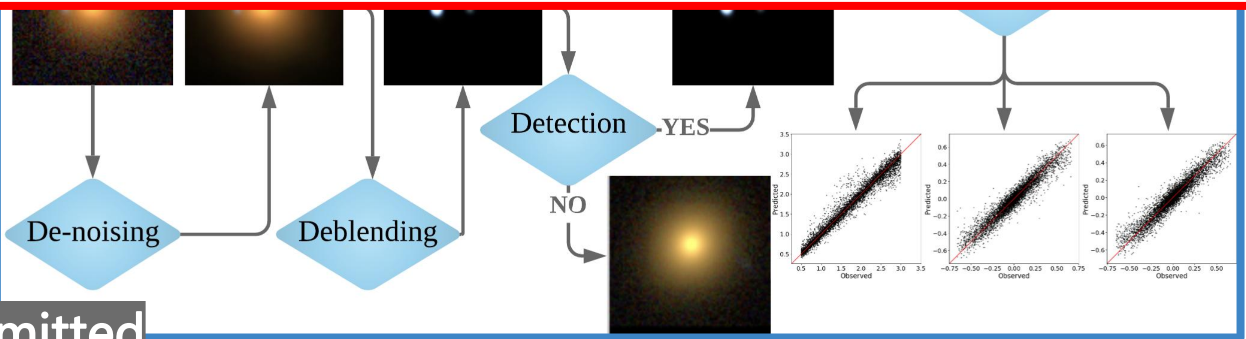




The CSST-SLWG is intensely forging this pipeline for CSST at the Moment.

Welcome to join us!

End Tool for Analyzing Galaxy Scale Strong Lenses



Madireddy et al. submitted

Summary

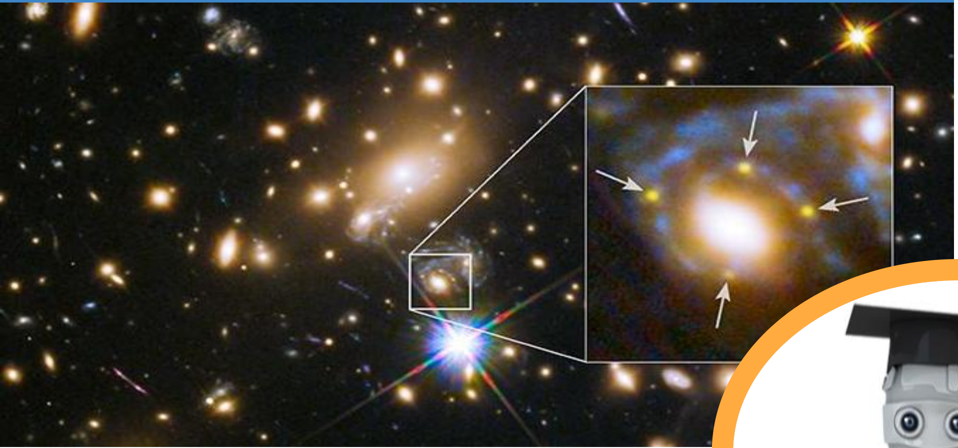
1. Gravitational lensing is useful, but lens-finding and modeling is challenging in the Big Data Era.

2. Deep learning works better than traditional methods and human-inspection in the detection of SGLs.

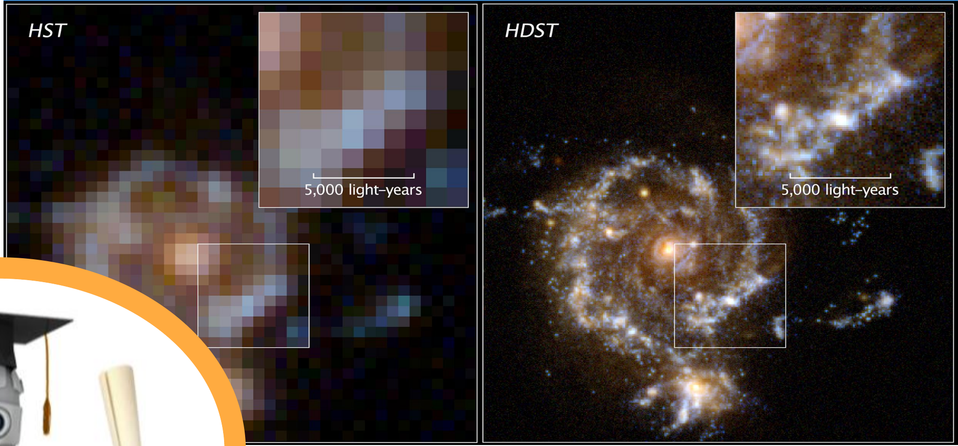
3. Deep learning also can implement strong-lens-modeling efficiently and automatically.

4. An End2End Strong lensing Pipeline for LSST is ready. The CSST-SLWG is intensely forging it for CSST.

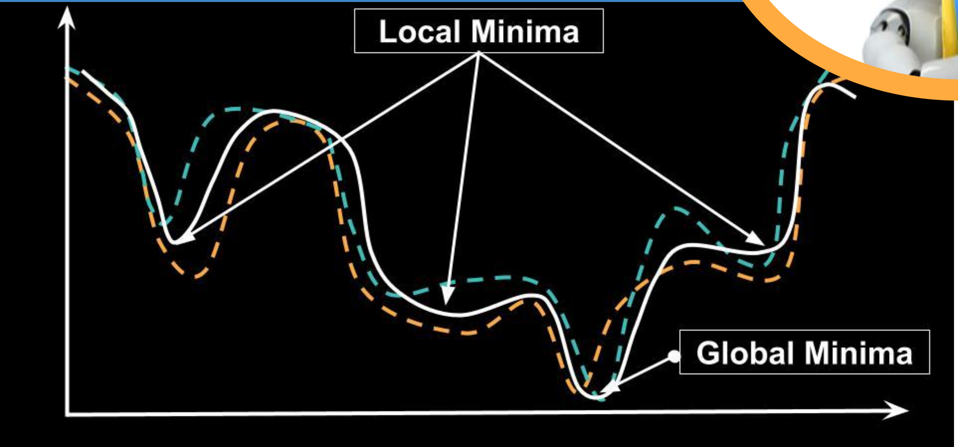
Detecting



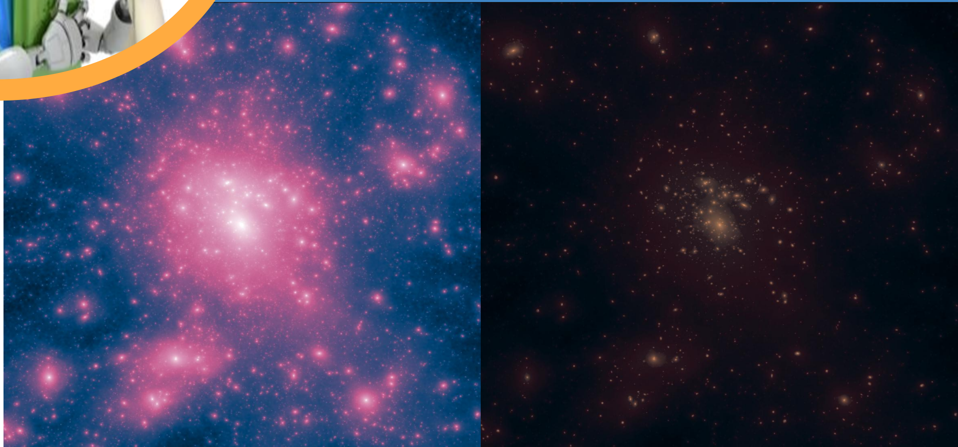
Predicting



Fitting



Modeling



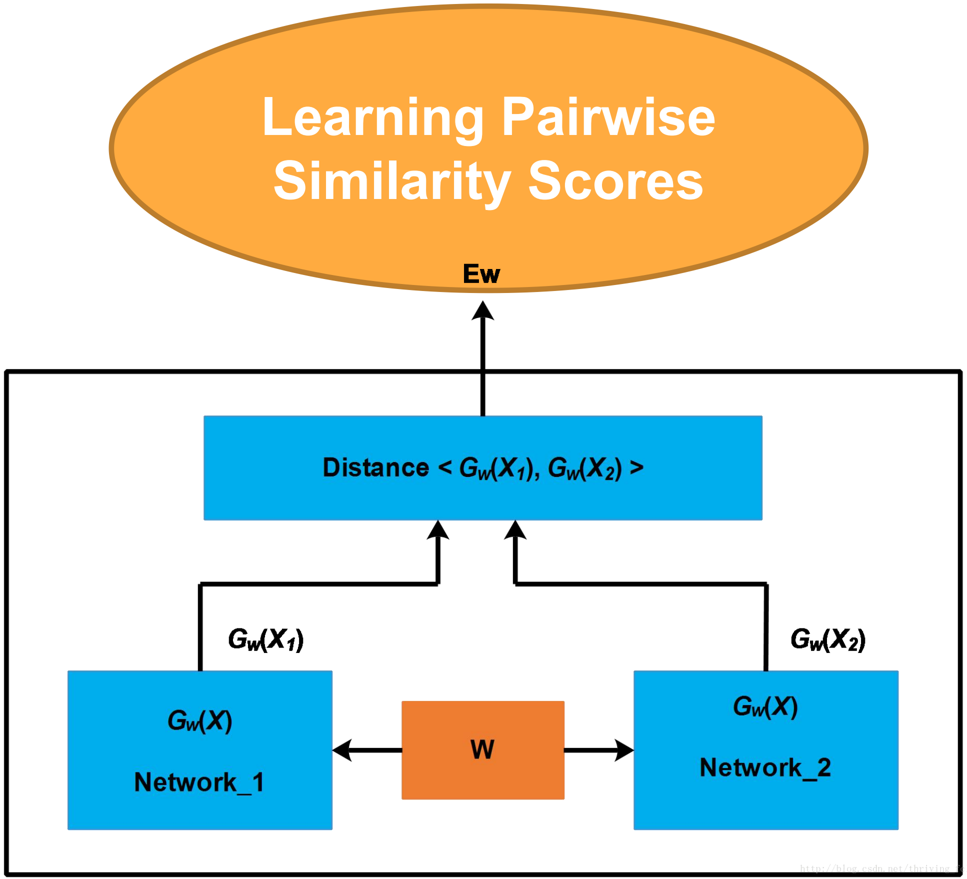
Backup Slides

Few-shots Learning and Siamese Network

Basic Idea of Few-Shot Learning

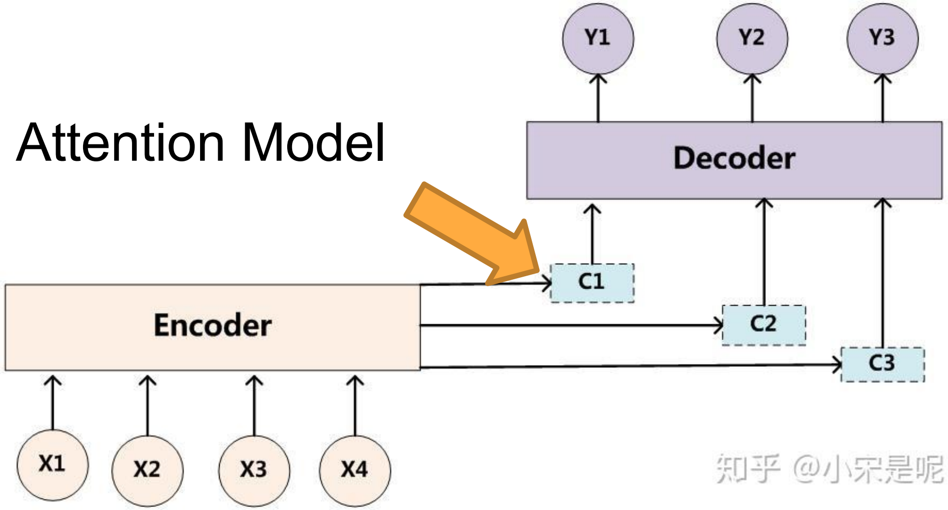
- Train a Siamese network on large-scale training set.
- Given a support set of k-way n-shot.
  - k-way means k classes.
  - n-shot means every class has n samples.
  - The training set does not contain the k classes.
- Given a query, predict its class.
  - Use the Siamese network to compute similarity or distance.

Positive Samples	Negative Samples



Encoder-Decoder and Attention Model

Attention Model

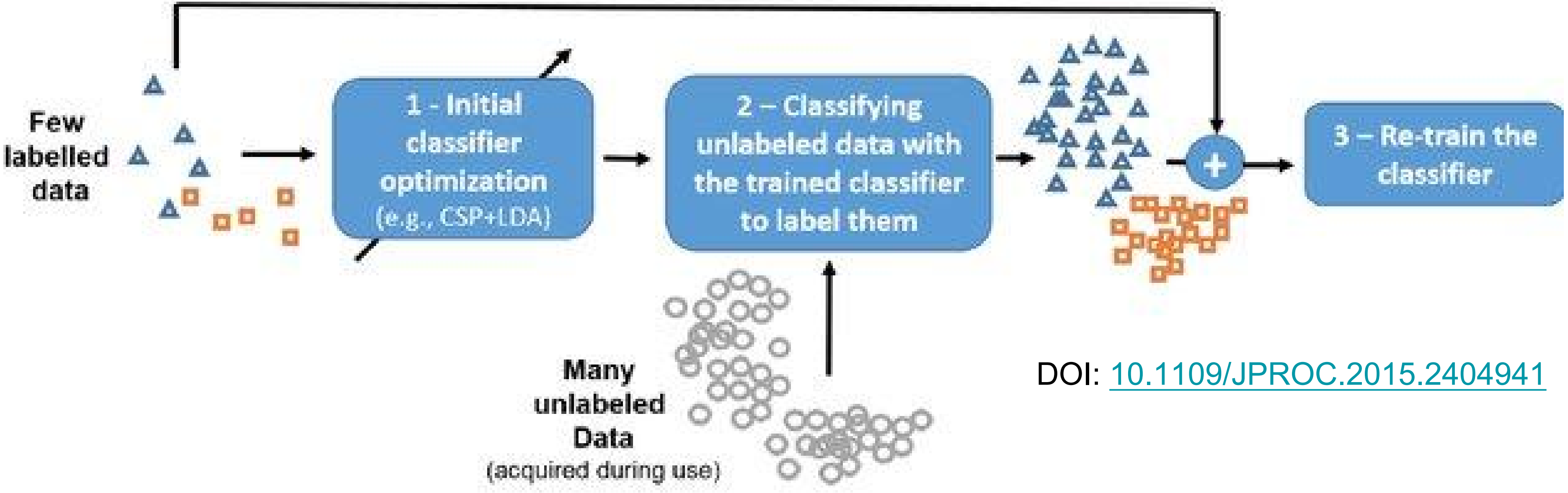


Mapping X to Y with Weights

A woman is throwing a frisbee in a park.

A little girl sitting on a bed with a teddy bear.

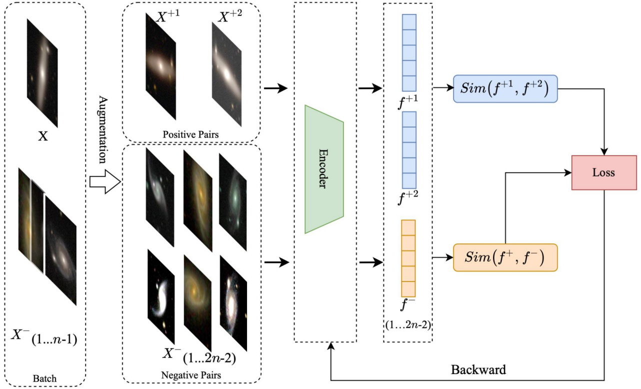
Semi-supervised Learning



DOI: 10.1109/JPROC.2015.2404941

- A model is first trained on the few available labelled training data.
- The model is then used to classify and label the many unlabeled data available.
- The newly labelled data are combined with the originally available labelled ones to retrain the model with many more data, and thus hopefully to obtain a better model.

Unsupervised Deep Contrastive Learning



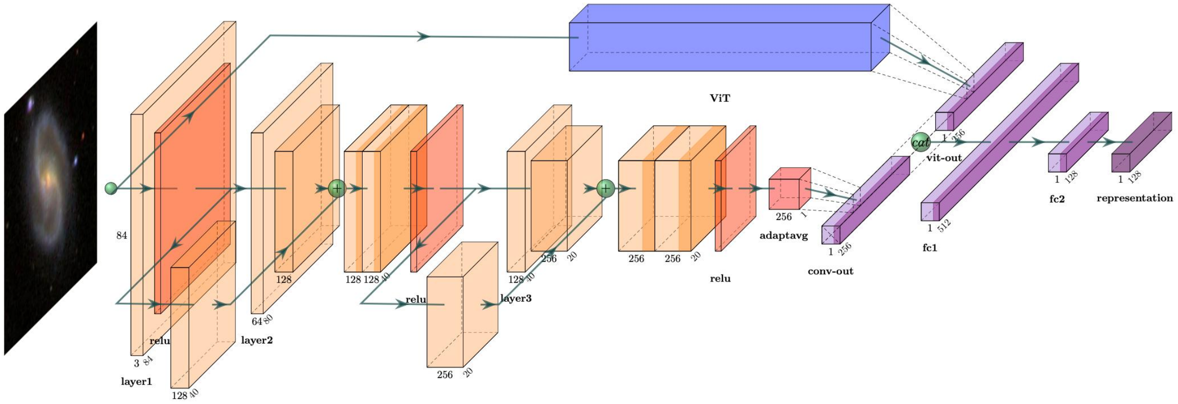


Figure 3. The overall architecture of the proposed network, which learns representation with multi-hierarchy features by *Encoder* and backward the loss calculated from the similarities between the positive pair and negative pairs.

g with ViT to learn

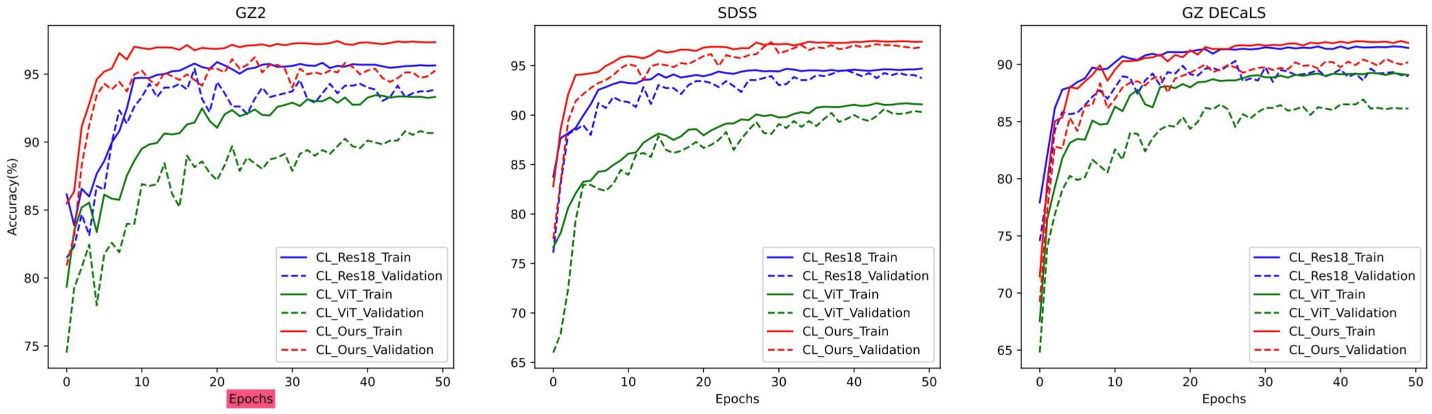


Figure 6. Performance comparison of the accuracy of training and validation along with the epochs. CL_Res18, CL_ViT and CL_Ours represents the CL methods using ResNet-18, ViT and our proposed model as encoder.