ML Journal Club SKA & JBCA

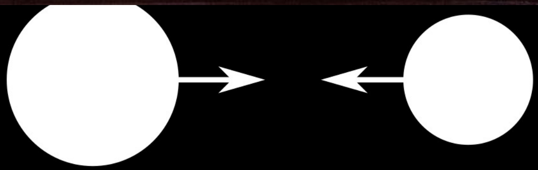
Support vector machine classification of strong gravitational lenses
Hartley et al 2017
and
The Strong Gravitational Lens Finding Challenge
Metcalf et al 2019

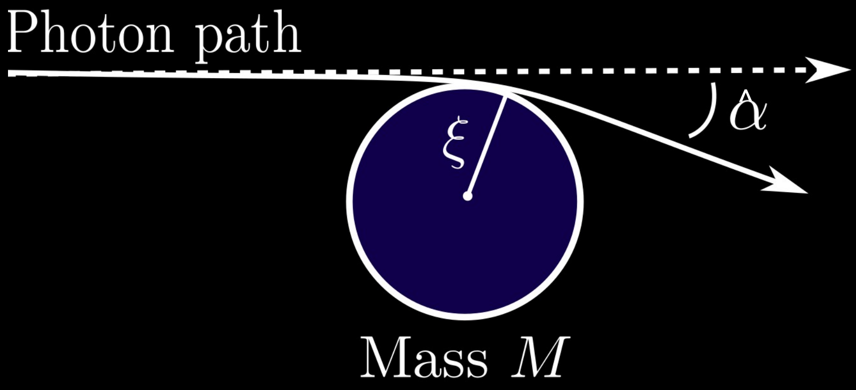
Overview

- **Lensing overview**
  - Lensing history
  - Types of gravitational lensing
- **Science from strong lenses**
  - Current applications of lenses
  - Euclid sky survey: more lenses
- **An automatic lens finder**
  - Support vector machine approach
- **Application of the lens finder**
  - Lens finding challenge
  - Kilo Degree Survey
- **The future….**

1704: Newton suspects gravitational deflection of light

"Do not bodies act upon light at a distance, and by their action bend its Rays; and is not this action strongest at the least distance ?", *Opticks*

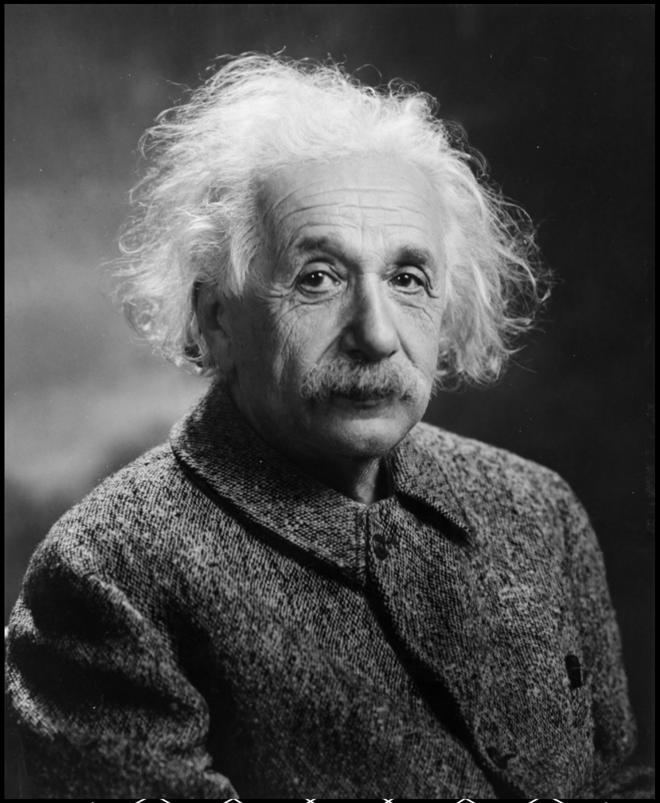


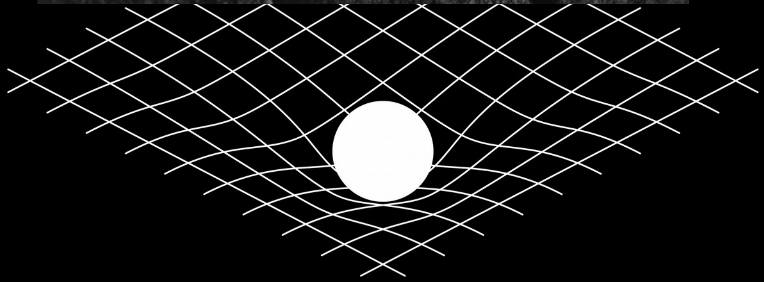


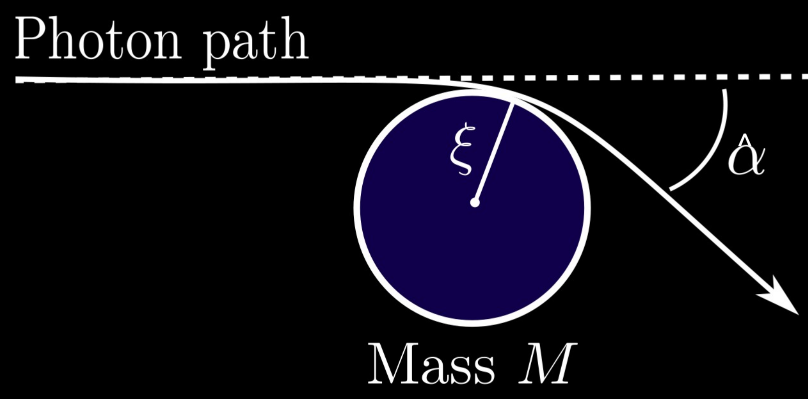
$$\hat{\alpha} = \frac{2GM}{c^2 \xi}$$

J. Soldner, 1804

1915: General relativity predicts twice the deflection

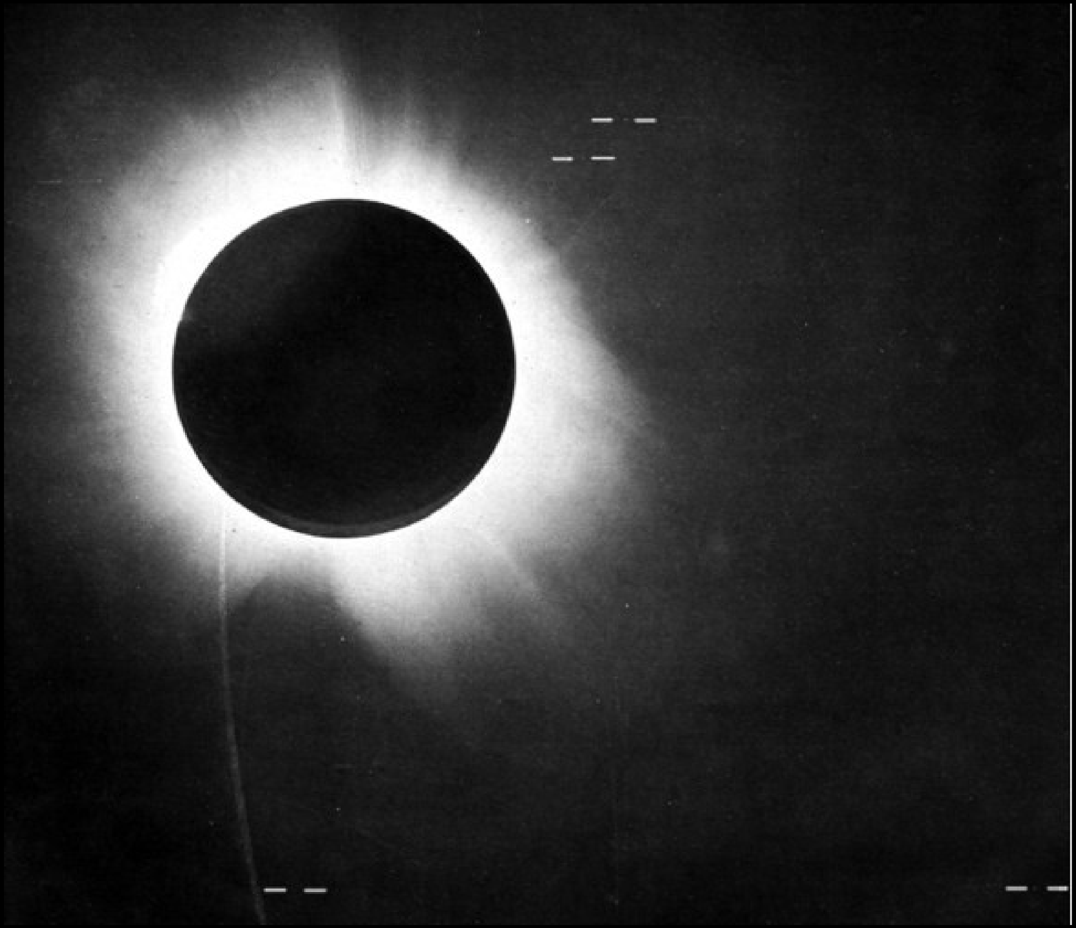






$$\hat{\alpha} = \frac{4GM}{c^2 \xi}$$

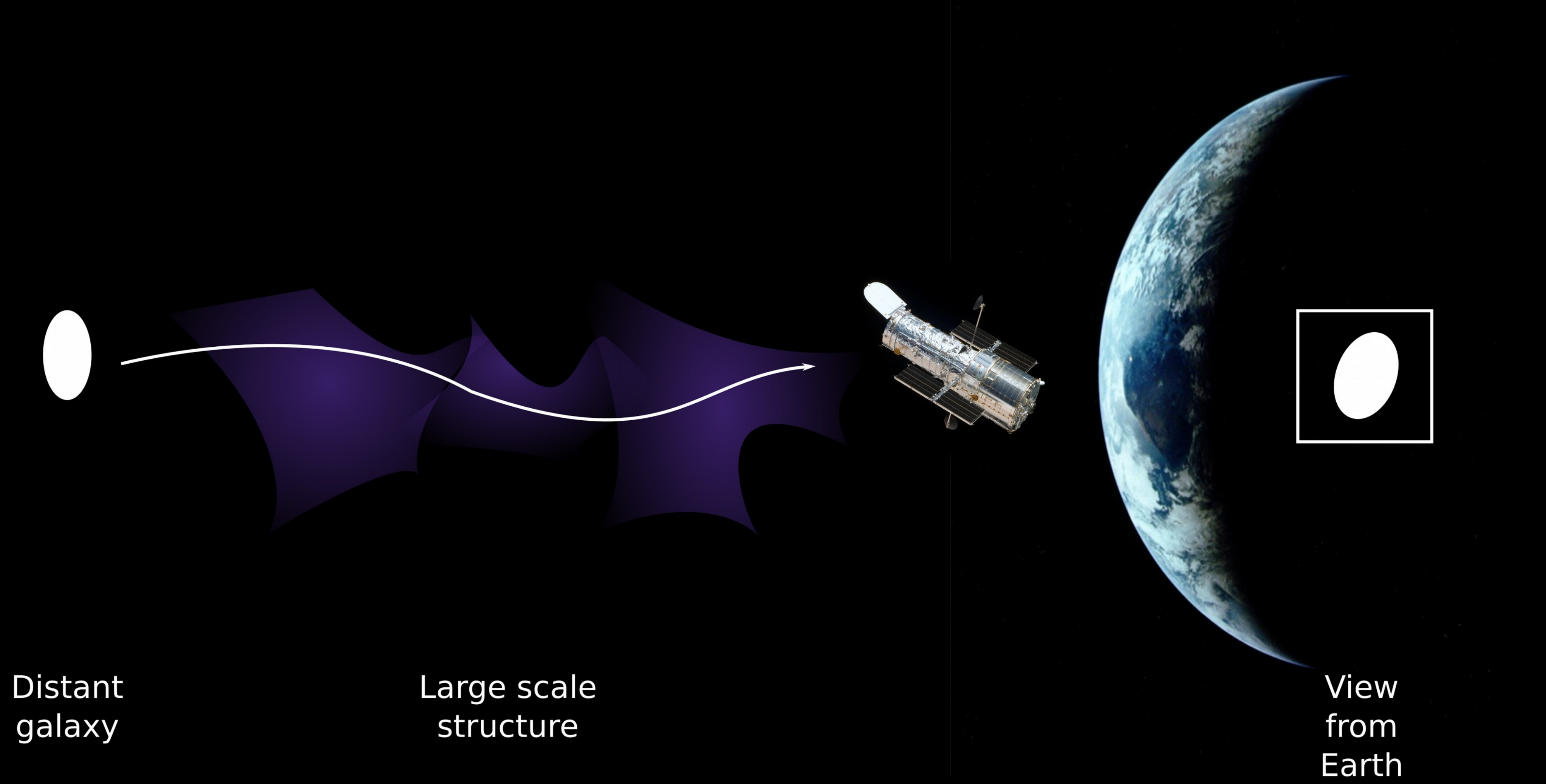
1919: Lensing effect observed by Arthur Eddington
GR confirmed



Solar eclipse of 1919, shifted star locations marked

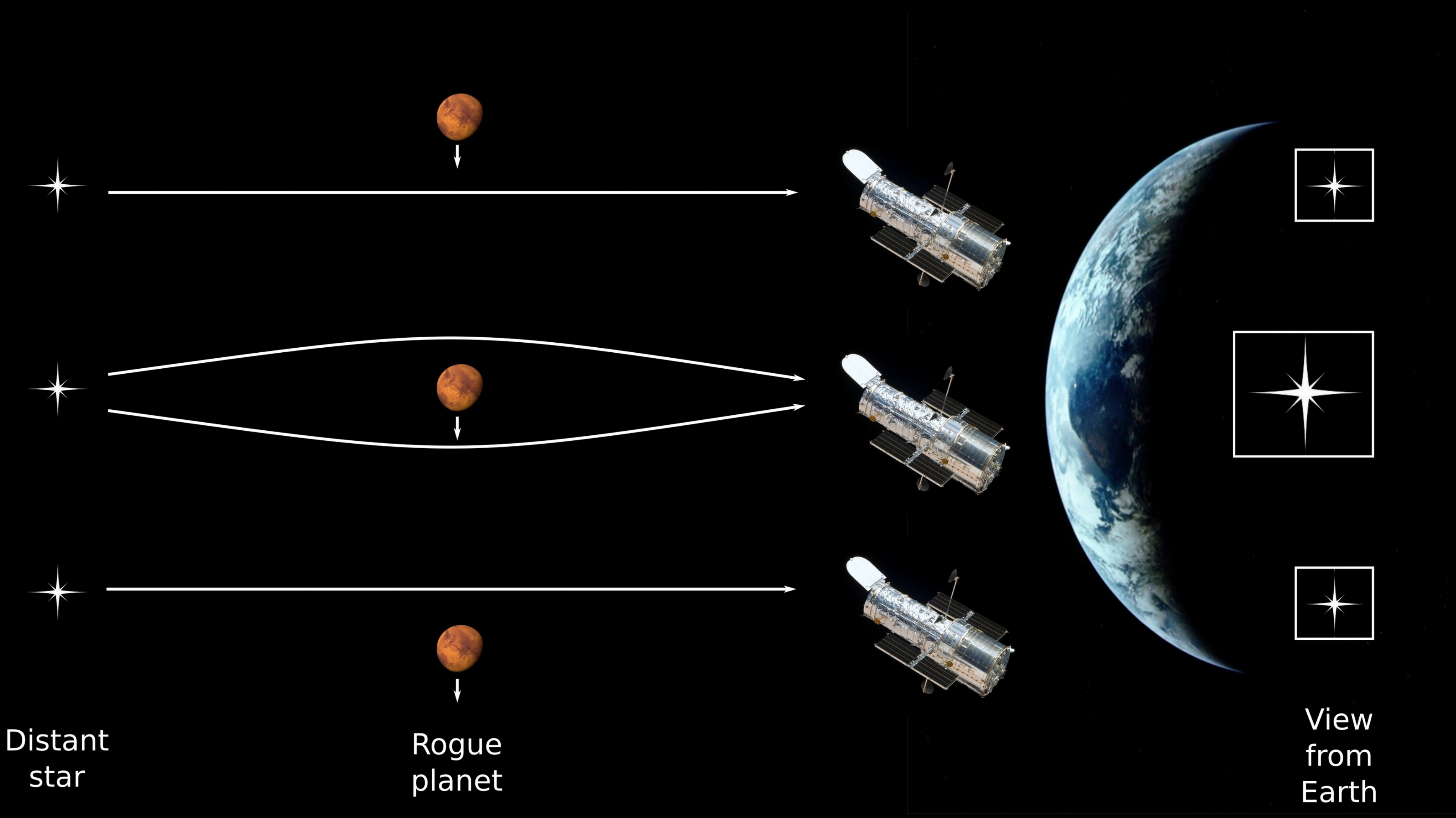
Weak lensing

A statistical measurement of **cosmic shear**



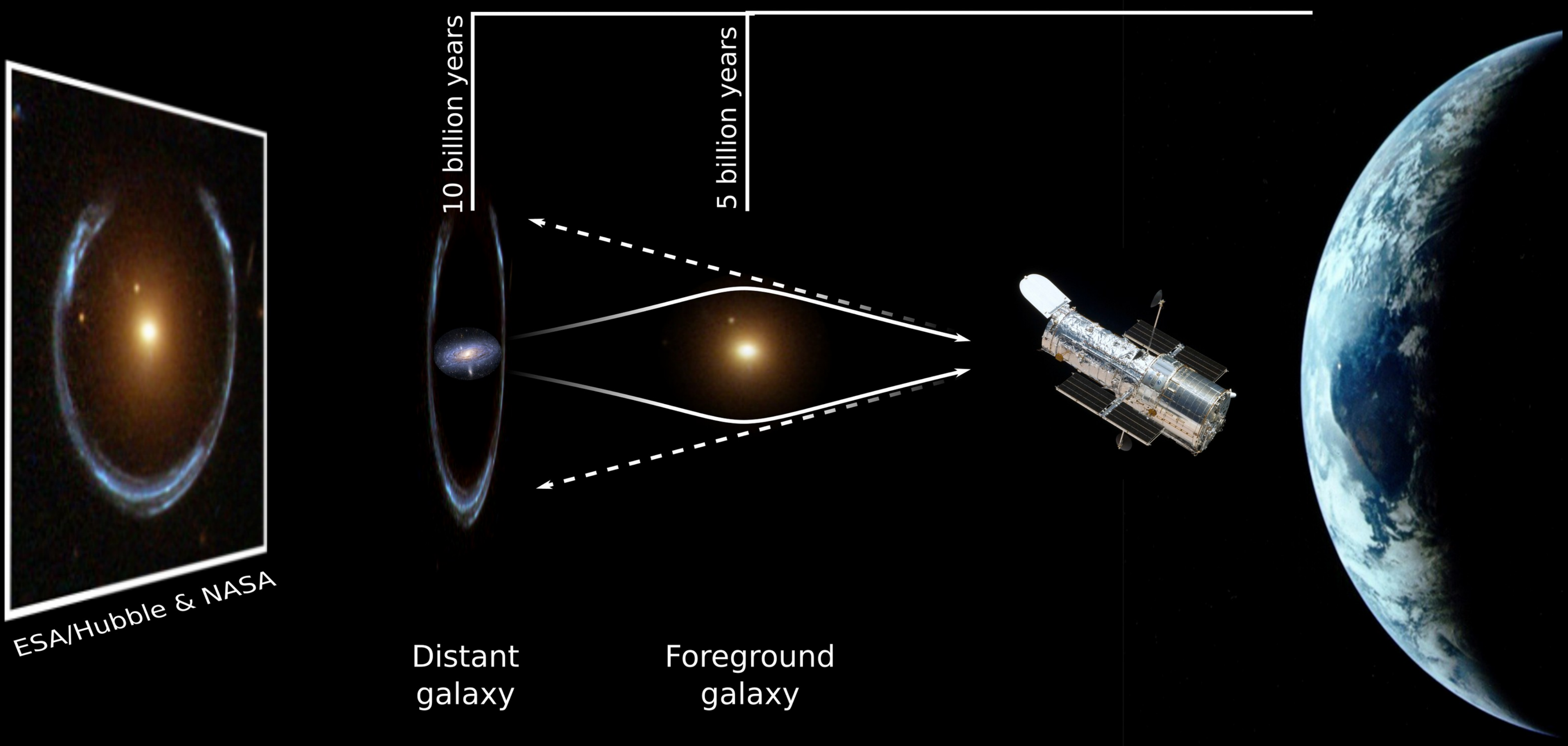
Microlensing

Compact projected mass exceeds **critical density**

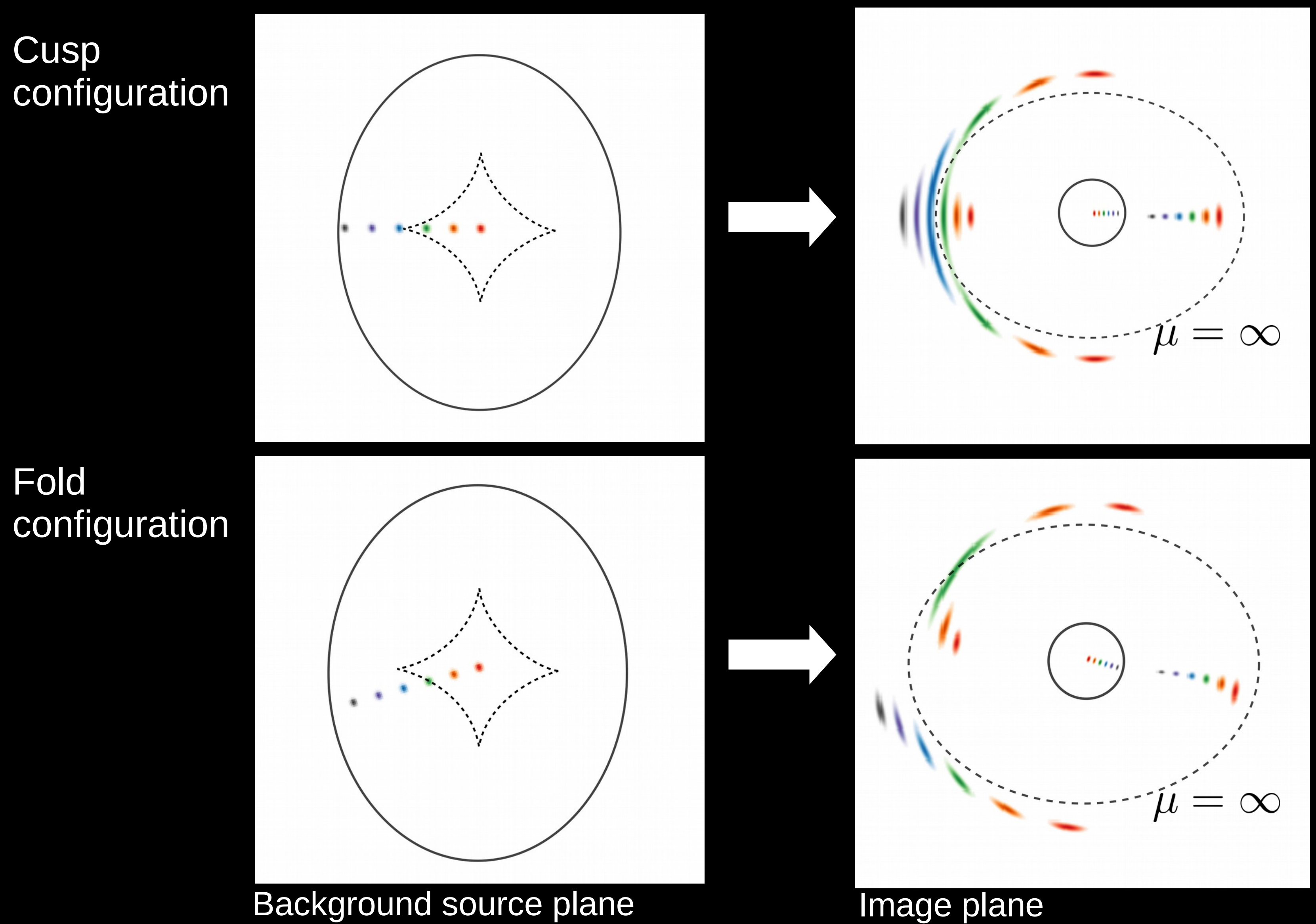


Strong lensing

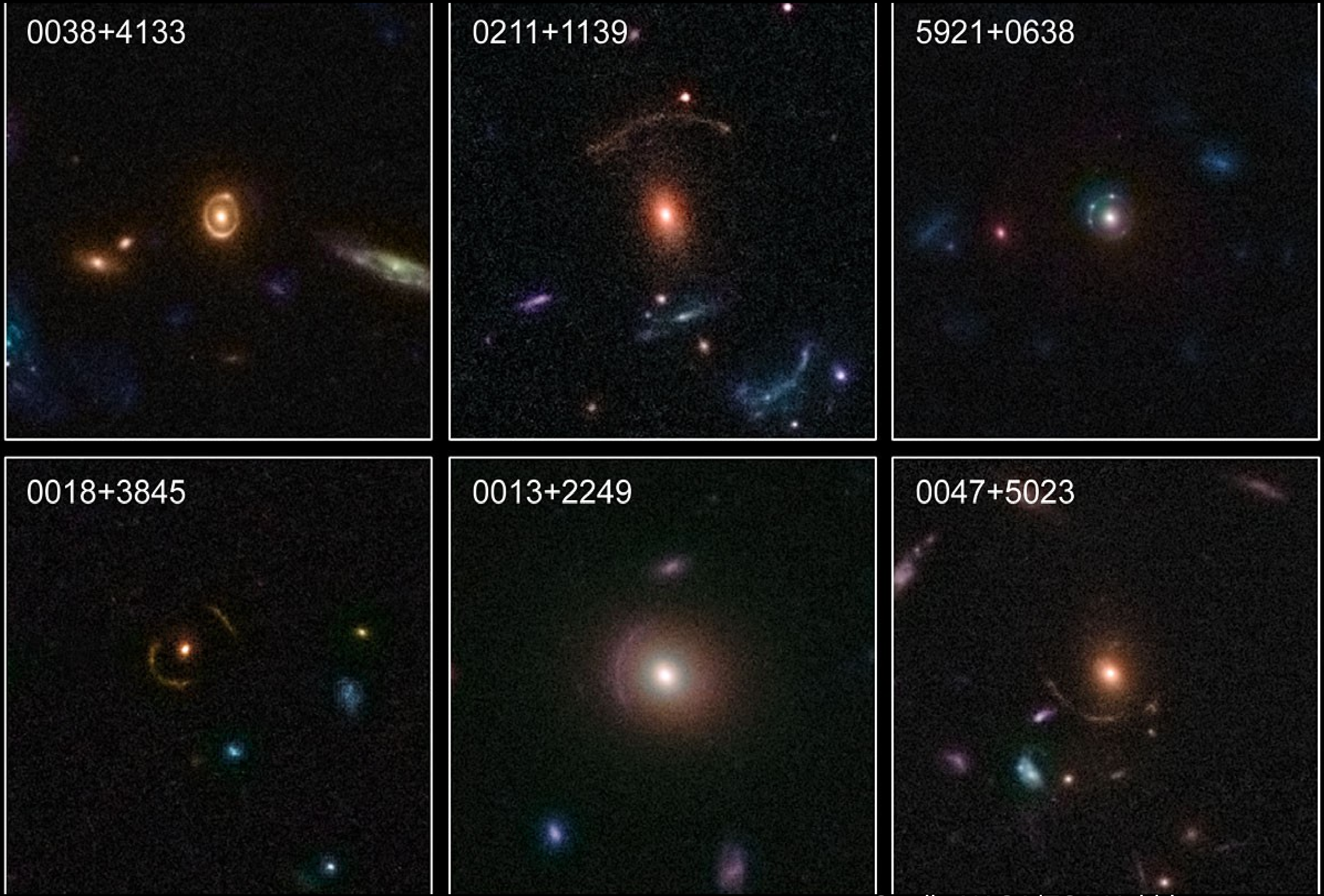
Extended projected lens mass exceeds **critical density**



Strong lensing configurations



Strong lensing configurations



Credit: NASA/ESA Hubble

Overview

- **Lensing overview**
  - Lensing history
  - Types of gravitational lensing
- **Science from strong lenses**
  - Current applications of lenses
  - Euclid sky survey: more lenses
- **An automatic lens finder**
  - Support vector machine approach
- **Application of the lens finder**
  - Lens finding challenge
  - Kilo Degree Survey
- **The future….**

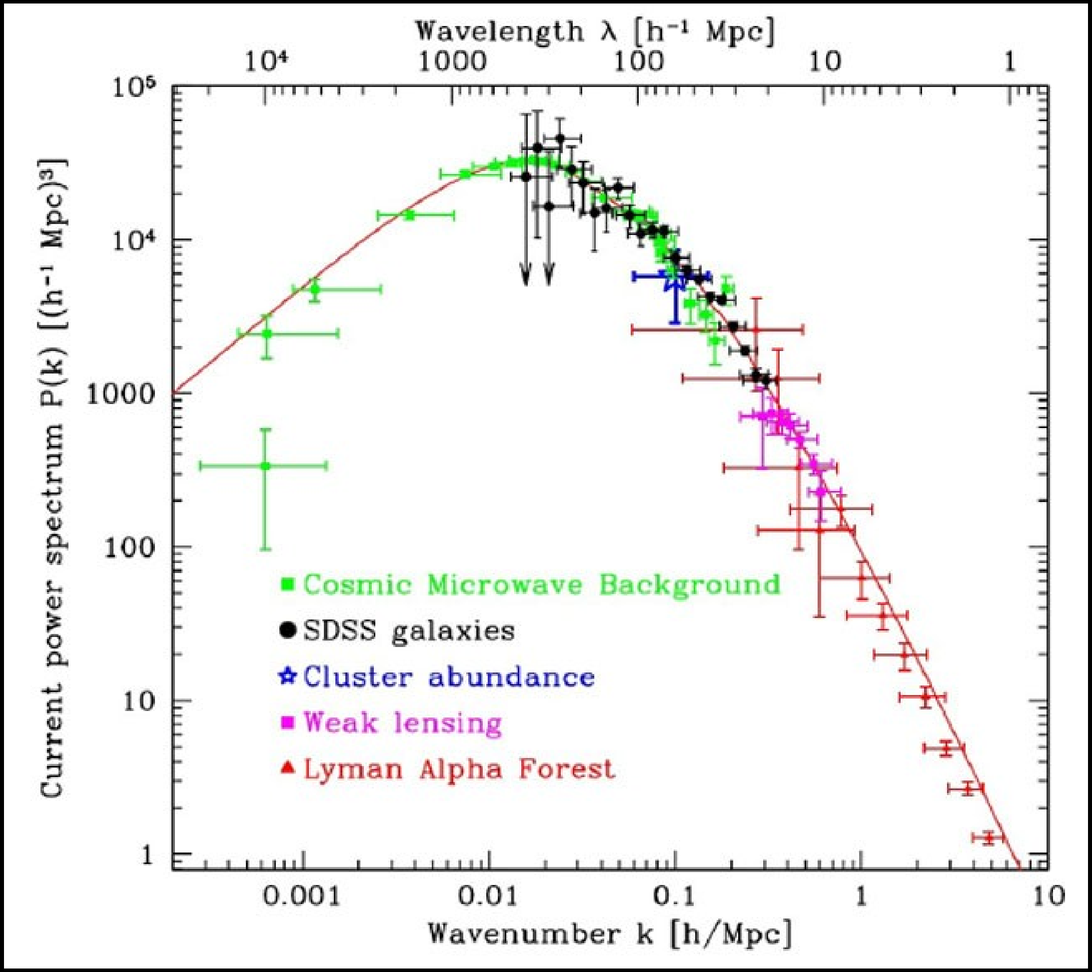
Science from lenses: dark matter structure

- Sub-halos of dark matter produce anomalies in lensed images

Millennium simulation of dark matter halos

Moore at al. 1999

Dark matter power spectrum



Smaller scales of matter →

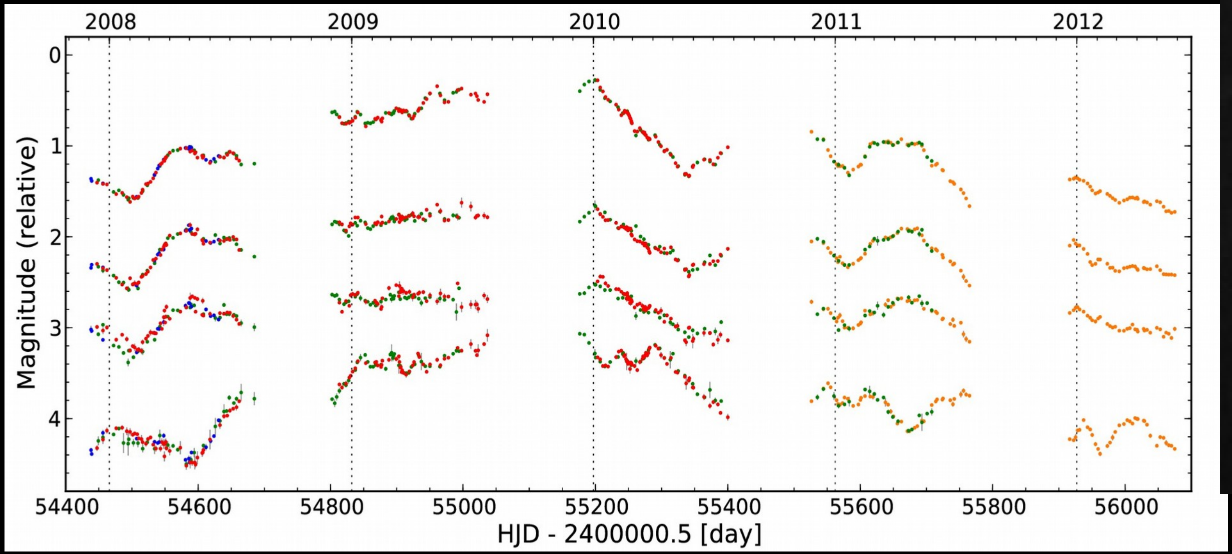
Science from lenses: the Hubble parameter

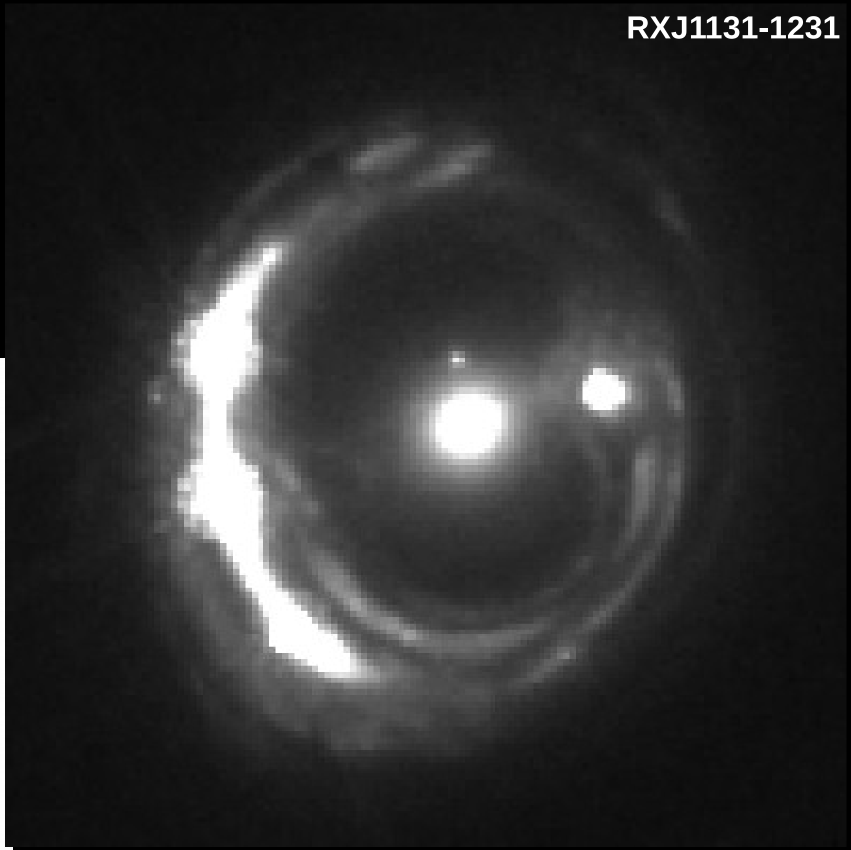
Measure **time delay** between images

Measure and model lensing galaxy

Infer time delay distance

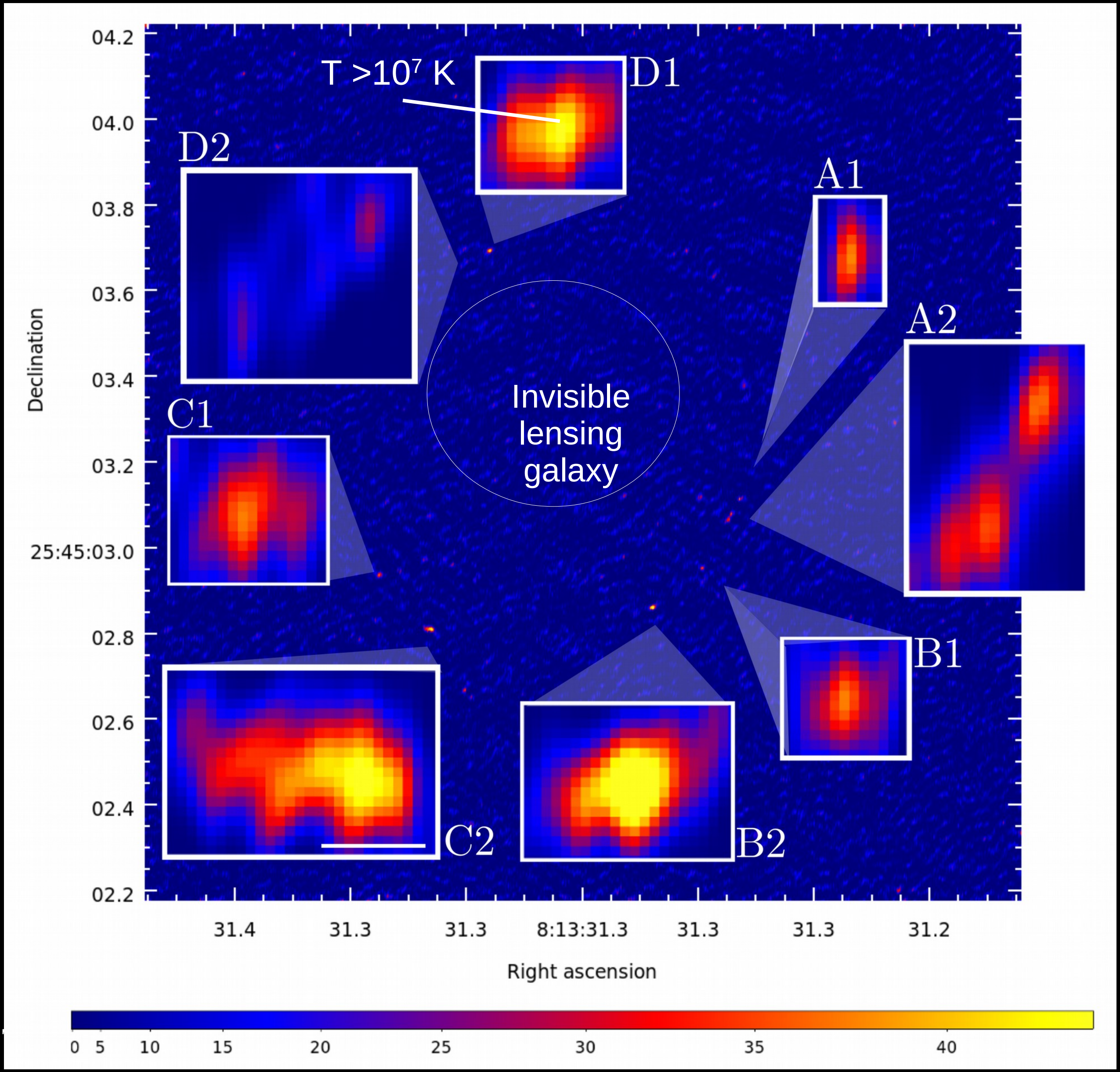
Convert into cosmological parameters





Suyu, Treu et al. 2014

Science from lenses: cosmic telescopes



VLBI observations of a strongly lensed **radio quiet quasar** at z=1.51

Reveal **twin jets** located either side of the quasar core

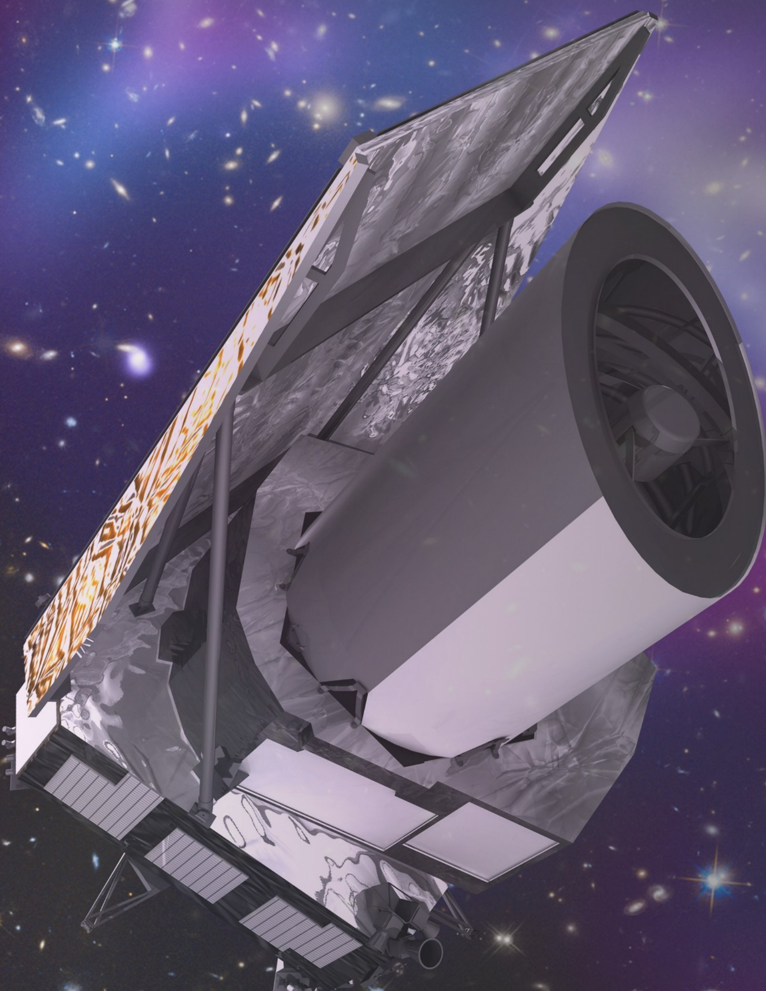
Lensing results in x100 magnification of this **sub-microJy** source at **sub-parcec** resolution

Hartley et al., 2019

Philippa Hartley

Euclid mission: Strong Lens Science Working Group

\- ***white paper in prep.***



Current sample

~1000 strong lenses

Expectations

~10 000 000 000 sources

~300 000 galaxy-galaxy lenses

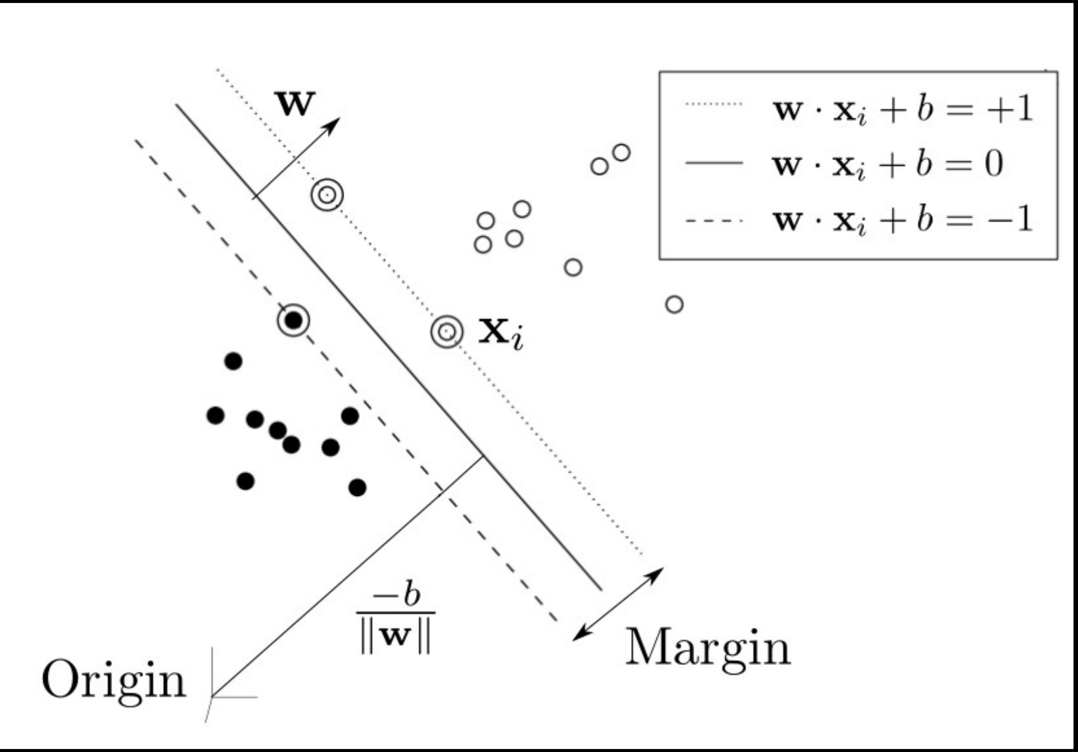
~3000 cluster lenses

Where are they?

Overview

- **Lensing overview**
  - Lensing history
  - Types of gravitational lensing
- **Science from strong lenses**
  - Current applications of lenses
  - Euclid sky survey: more lenses
- **An automatic lens finder**
  - Support vector machine approach
- **Application of the lens finder**
  - Lens finding challenge
  - Kilo Degree Survey
- **The future….**

Support vector machines



Find optimal hyperplane separating two classes of data

Support vector machines

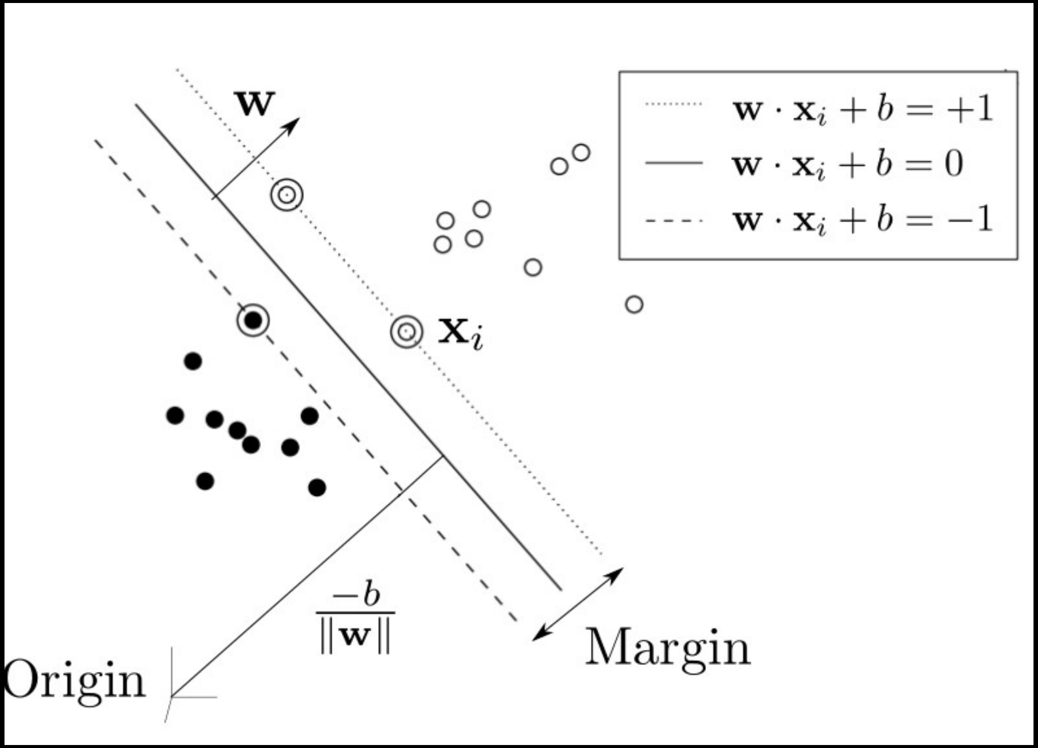
$$\{\mathbf{x}_i, y_i\}, \quad i = 1, \cdots, l, \quad y_i \in \{-1, 1\}, \quad \mathbf{x}_i \in \mathbf{R}^d$$

$$\mathbf{x}_i \cdot \mathbf{w} + b \geq +1 \quad \text{for } y_i = +1$$
$$\mathbf{x}_i \cdot \mathbf{w} + b \leq -1 \quad \text{for } y_i = -1$$

$$L_P \equiv \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{l} \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{w} + b) + \sum_{i=1}^{l} \alpha_i$$

$$\mathbf{w} = \sum_i \alpha_i y_i \mathbf{x}_i \qquad \sum_i \alpha_i y_i = 0.$$

$$L_D = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j$$



Vapnik et al. 1979, Cortes & Vapnik 1995

Find optimal hyperplane separating two classes of data

Optimisation depends only on dot products of support vectors, found on the edge of each class

Support vector machines

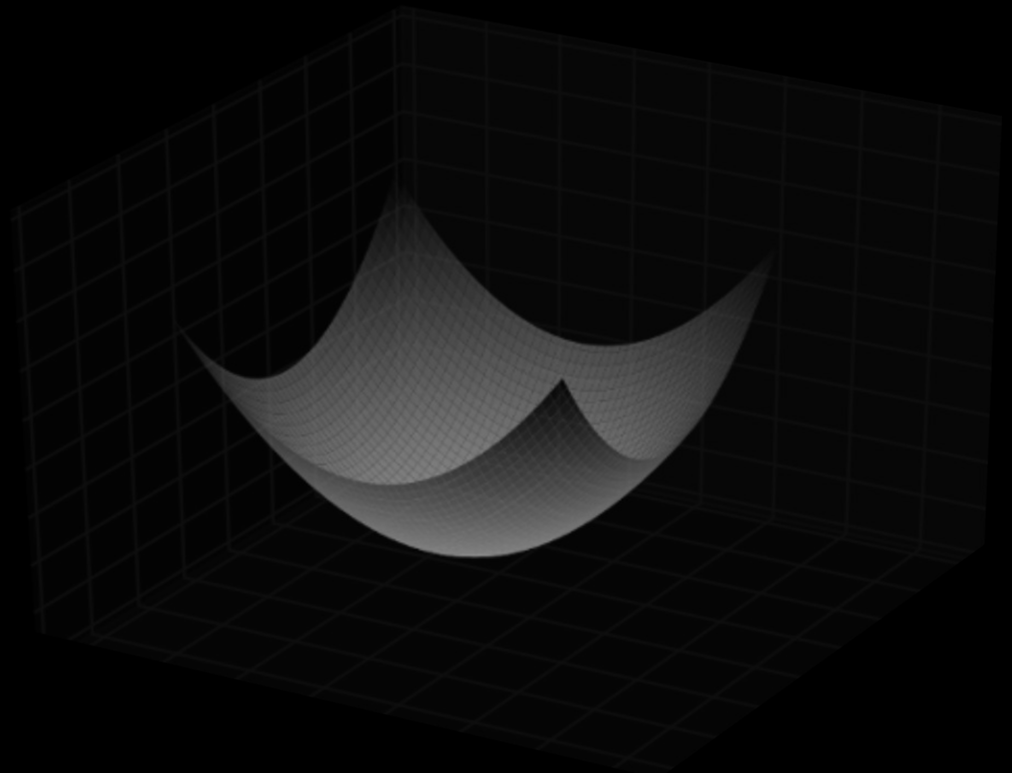
$$\{\mathbf{x}_i, y_i\}, \quad i = 1, \cdots, l, \quad y_i \in \{-1, 1\}, \quad \mathbf{x}_i \in \mathbf{R}^d$$

$$\mathbf{x}_i \cdot \mathbf{w} + b \geq +1 \quad \text{for } y_i = +1$$
$$\mathbf{x}_i \cdot \mathbf{w} + b \leq -1 \quad \text{for } y_i = -1$$

$$L_P \equiv \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{l} \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{w} + b) + \sum_{i=1}^{l} \alpha_i$$

$$\mathbf{w} = \sum_i \alpha_i y_i \mathbf{x}_i \qquad \sum_i \alpha_i y_i = 0.$$

$$L_D = \sum_i \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j$$



Vapnik et al. 1979, Cortes & Vapnik 1995

- Function is convex: every local solution is a global one – no local minima

Support vector machines

$$\{\mathbf{x}_i, y_i\}, \quad i = 1, \cdots, l, \quad y_i \in \{-1, 1\}, \quad \mathbf{x}_i \in \mathbf{R}^d$$

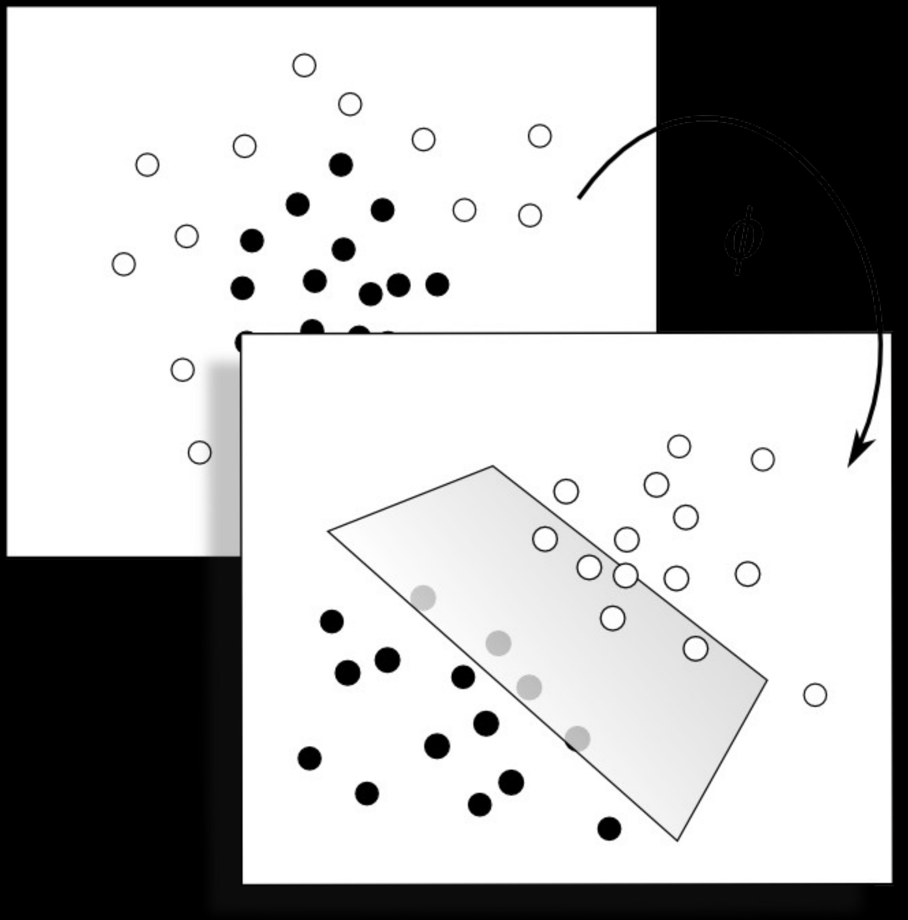
$$\mathbf{x}_i \cdot \mathbf{w} + b \geq +1 \quad \text{for } y_i = +1$$
$$\mathbf{x}_i \cdot \mathbf{w} + b \leq -1 \quad \text{for } y_i = -1$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j).$$

$$\mathbf{w} = \sum_i \alpha_i y_i \mathbf{x}_i \qquad \sum_i \alpha_i y_i = 0.$$

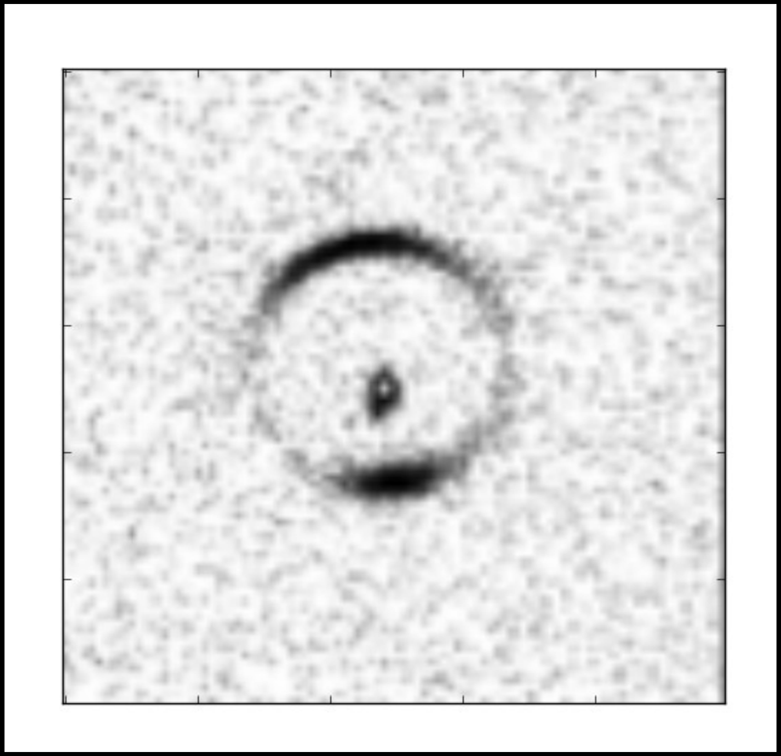
$$L_D = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j$$

Boser et al. 1992

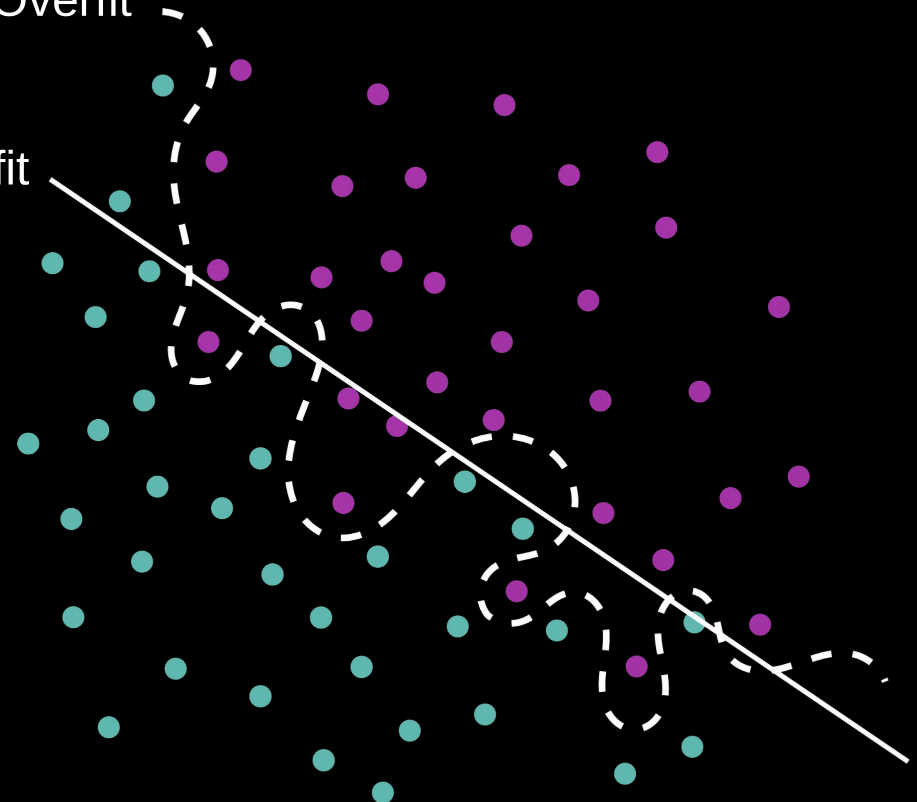


- Coordinate transformation can deal with non-linear separation
- Unknown kernel function replaces dot product

Feature extraction

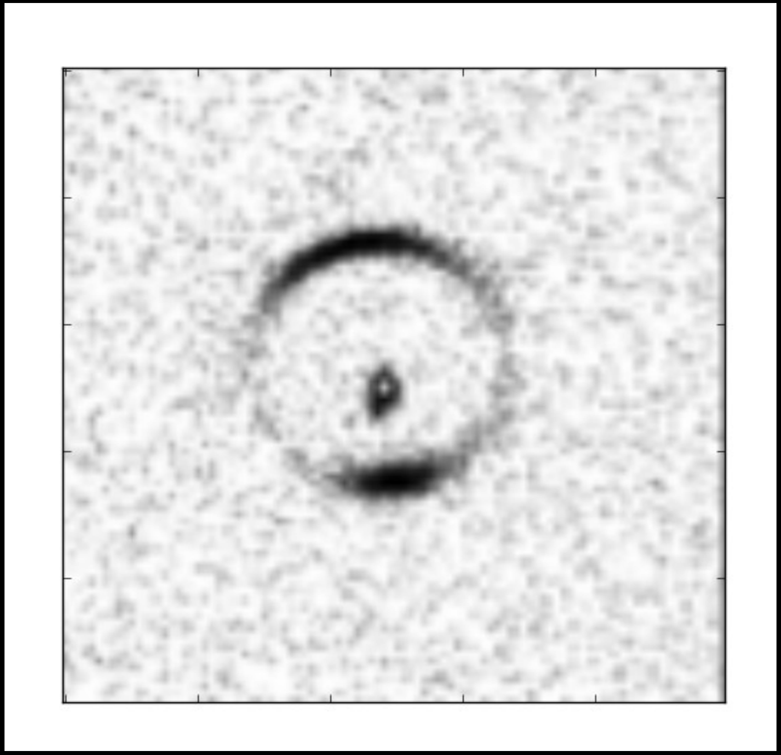


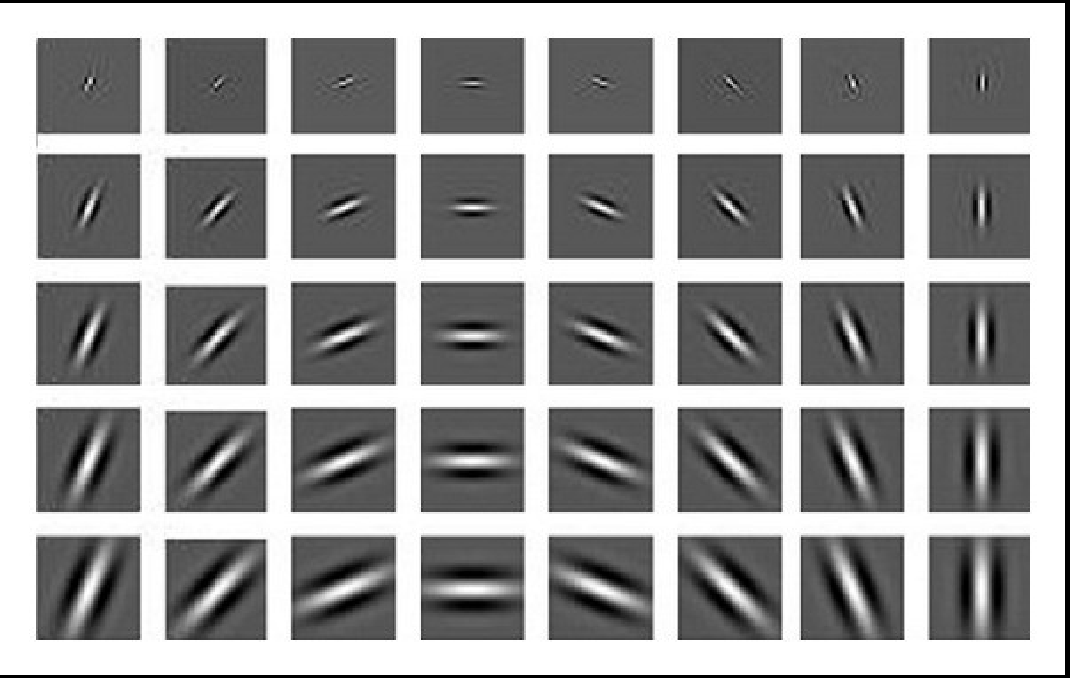
100 * 100 pixels = **10 000** features



Feature extraction

Apply **Gabor filters**: model the simple cells of the mammillian visual cortex (Marcelja 1980)

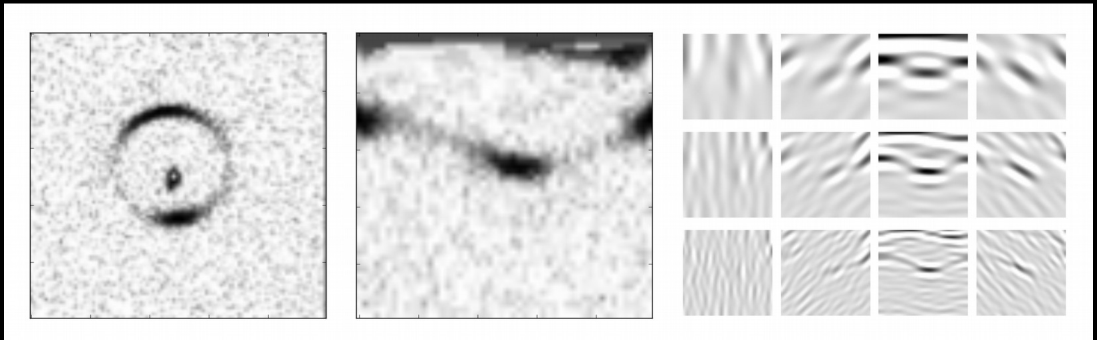




$$G_c[i,j] = Be^{-\frac{(i^2+j^2)}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda}(i\cos\theta + j\sin\theta)\right)$$

Feature extraction

Training sample → polar transform → Gabor filterbank → calculate moments

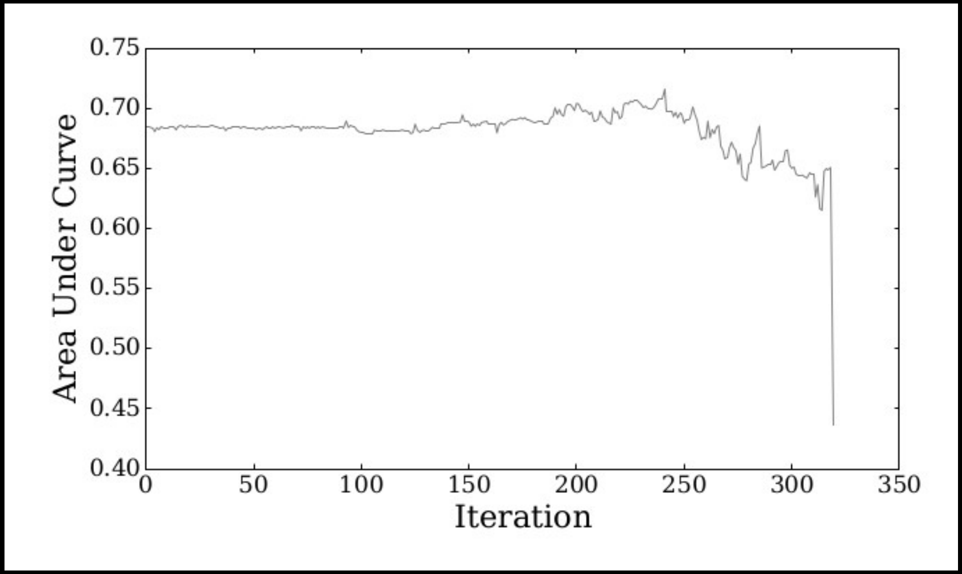


Mean	$\mu_1(x_1, \ldots, x_N) = \frac{1}{N}\sum_{j=1}^{N} x_j$
Variance	$\mu_2(x_1, \ldots, x_N) = \frac{1}{N-1}\sum_{j=1}^{N}(x_j - \mu_1)^2$
Skew	$\mu_3(x_1, \ldots, x_N) = \frac{1}{N}\sum_{j=1}^{N}\left[\frac{x_j-\mu_1}{\mu_2}\right]^3$
Kurtosis	$\mu_4(x_1, \ldots, x_N) = \left\{\frac{1}{N}\sum_{j=1}^{N}\left[\frac{x_j-\mu_1}{\mu_2}\right]^4\right\}$
Local energy	$E_s(x_1, \ldots, x_N) = \sum_{j=1}^{N} x_j^2$

Hartley et al. 2017 MNRAS

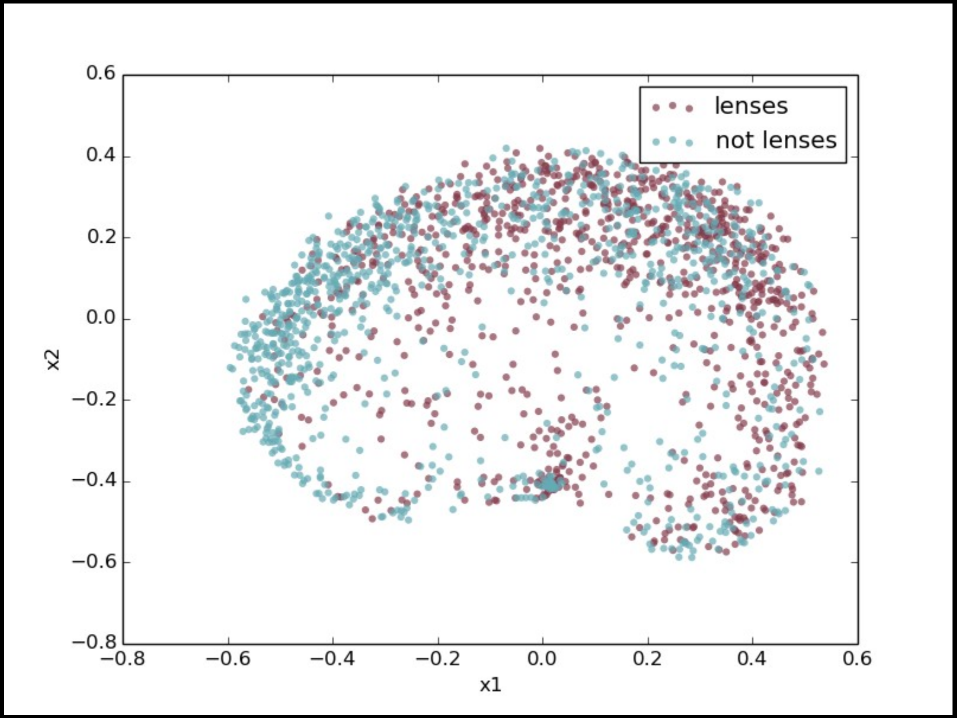
4 bands * 9 kernel frequencies * 7 kernel rotations * 5 moments = **1260** features

Feature selection



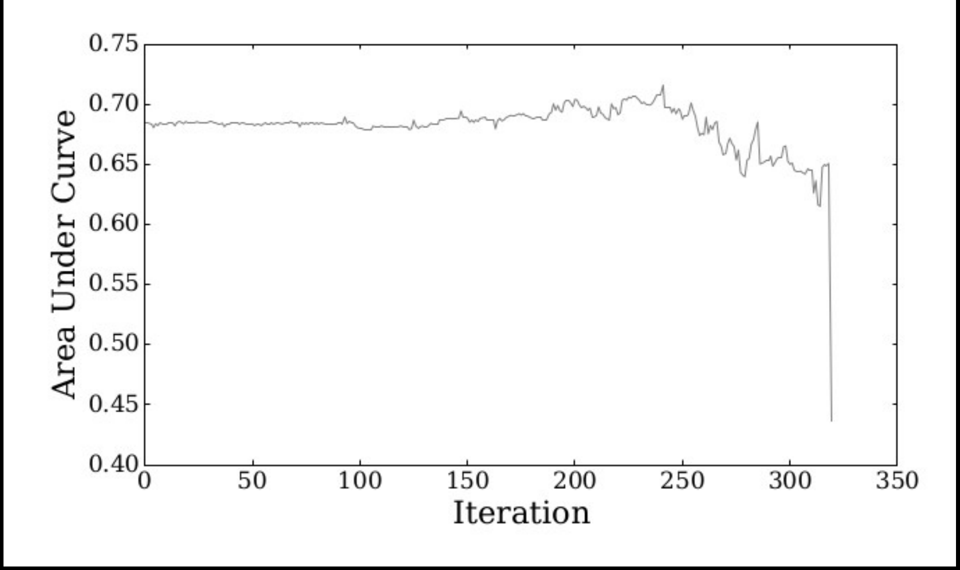
Recursive feature elimination

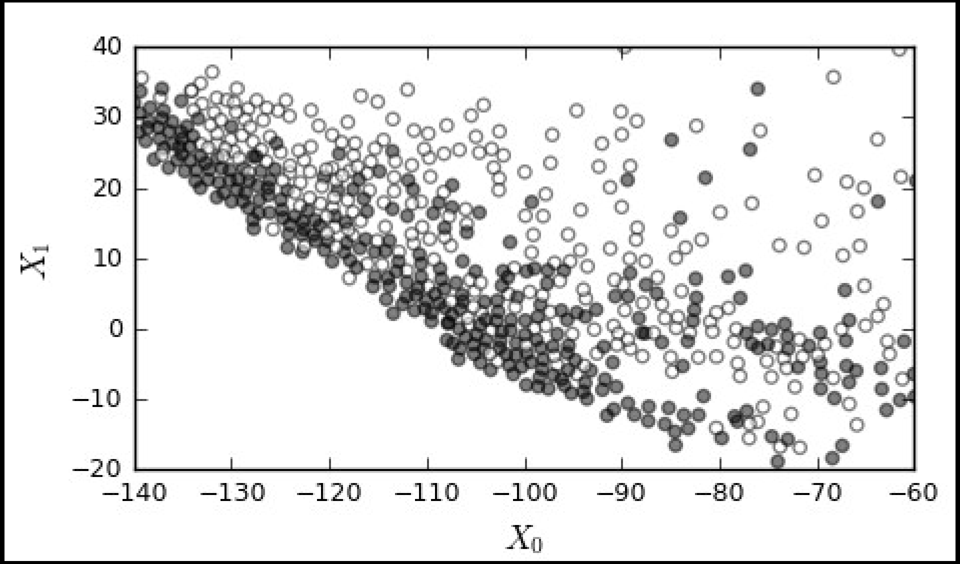
Simple, but can be unstable



Principle component analysis

Feature selection





Recursive feature elimination

Simple, but can be unstable

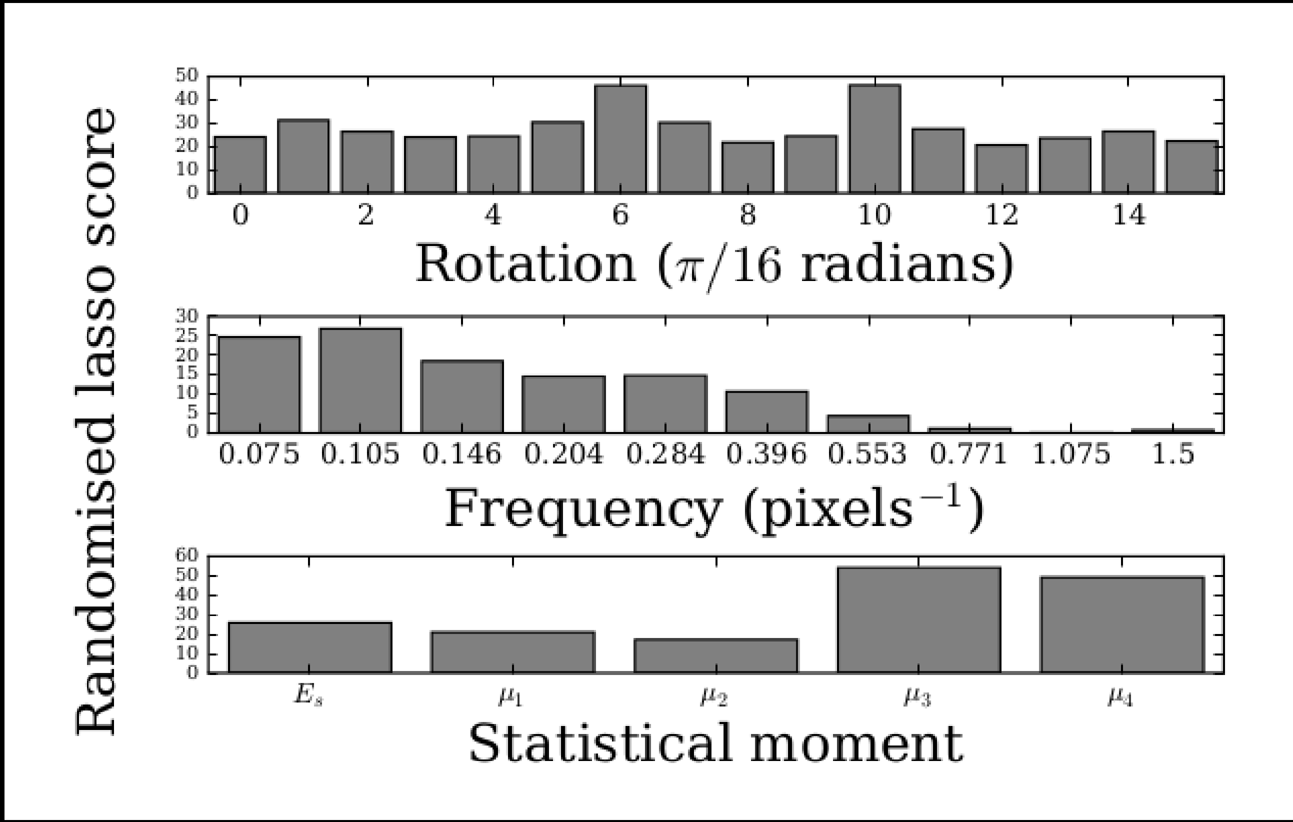
t-distributed stochastic neighbour embedding (t-SNE)

Like principle component analysis but able to respresent non-linear relationships

Feature selection

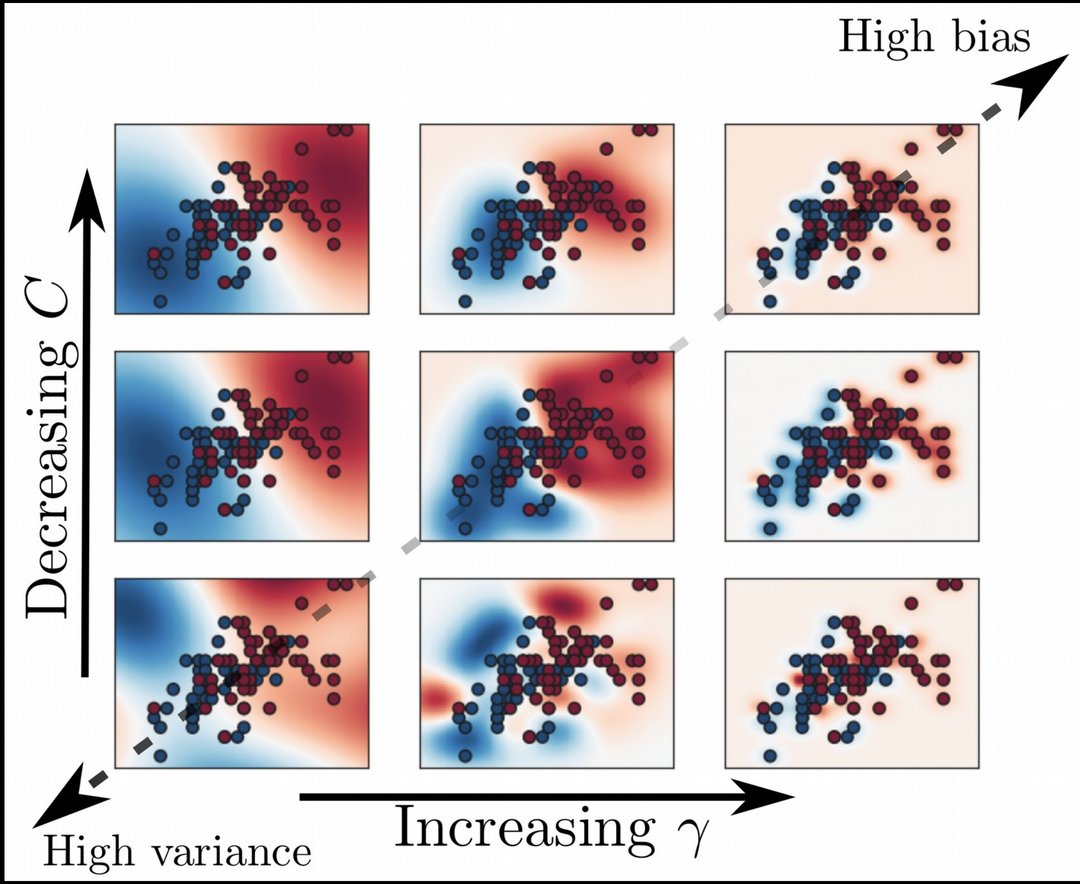
Stability selection:

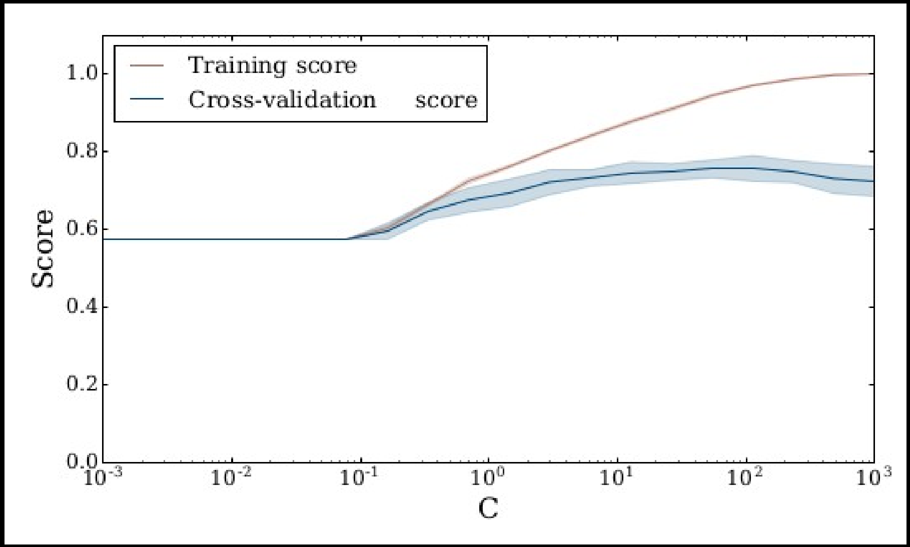
Features are subsampled and performance evaluated

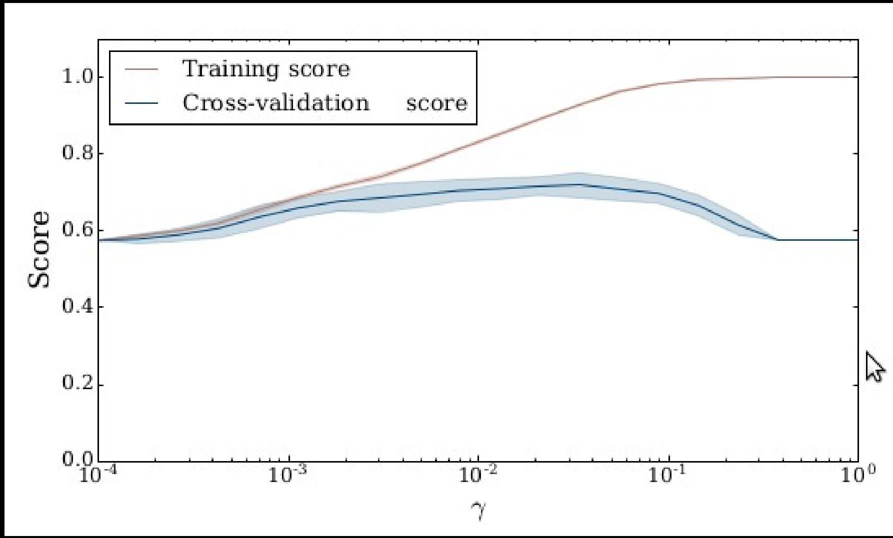


Model selection

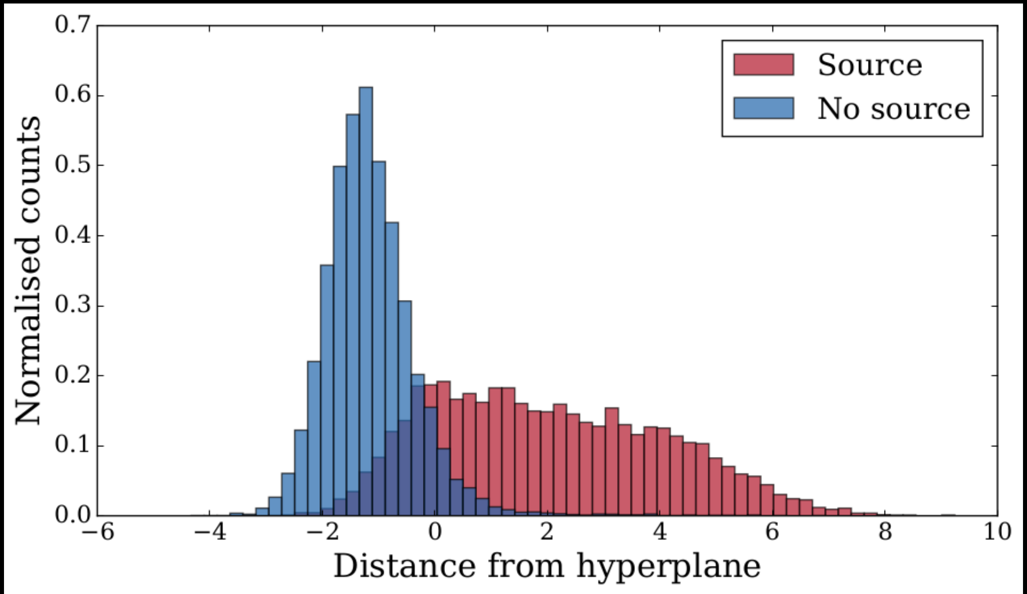
Regularisation parameters



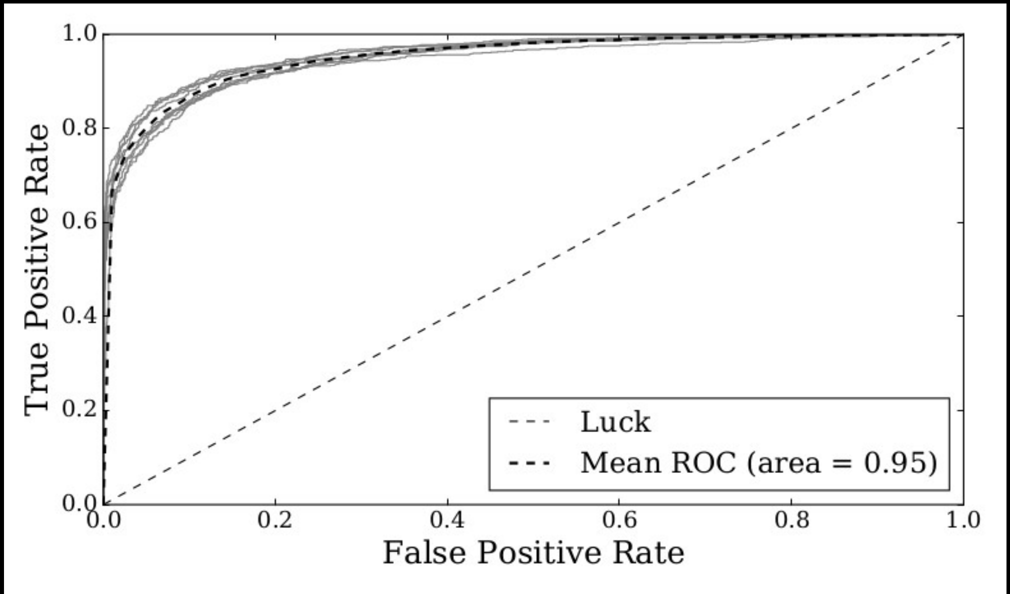




Results

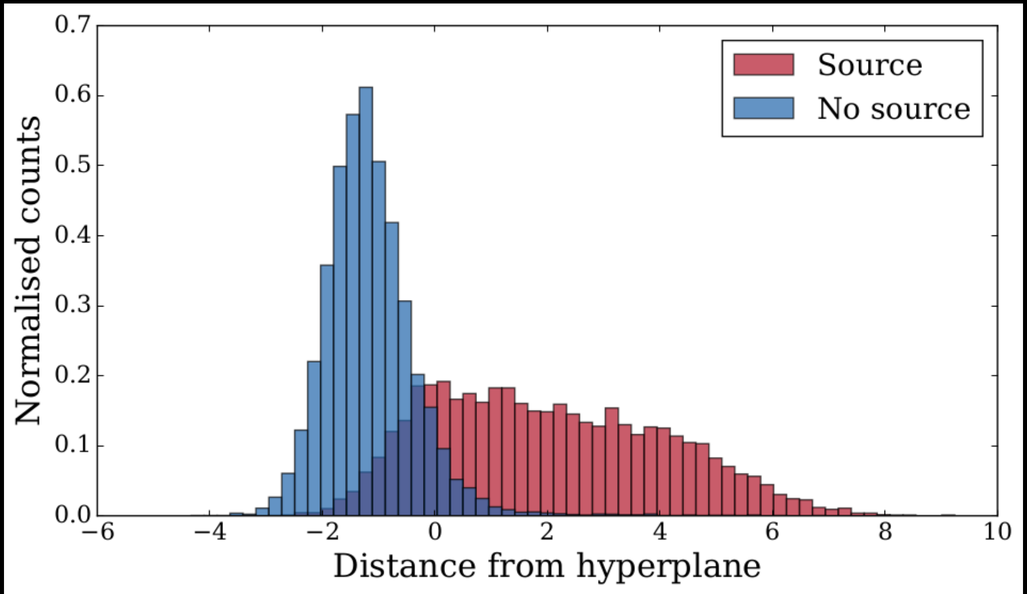


Raw score from SVM

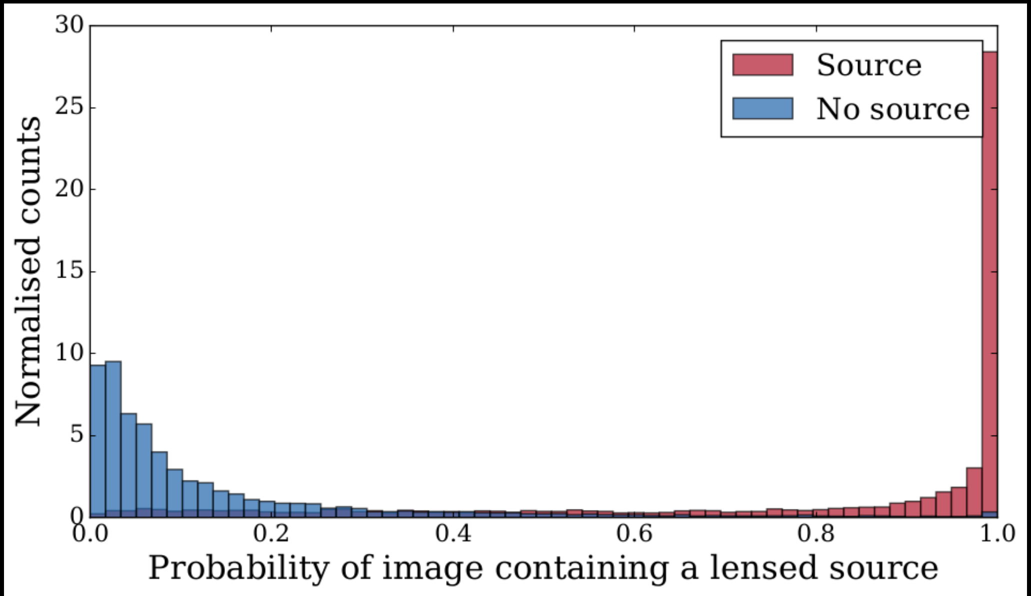


Receiver operating characteristic (ROC) curve

Results: Platt scaling



Raw score from SVM

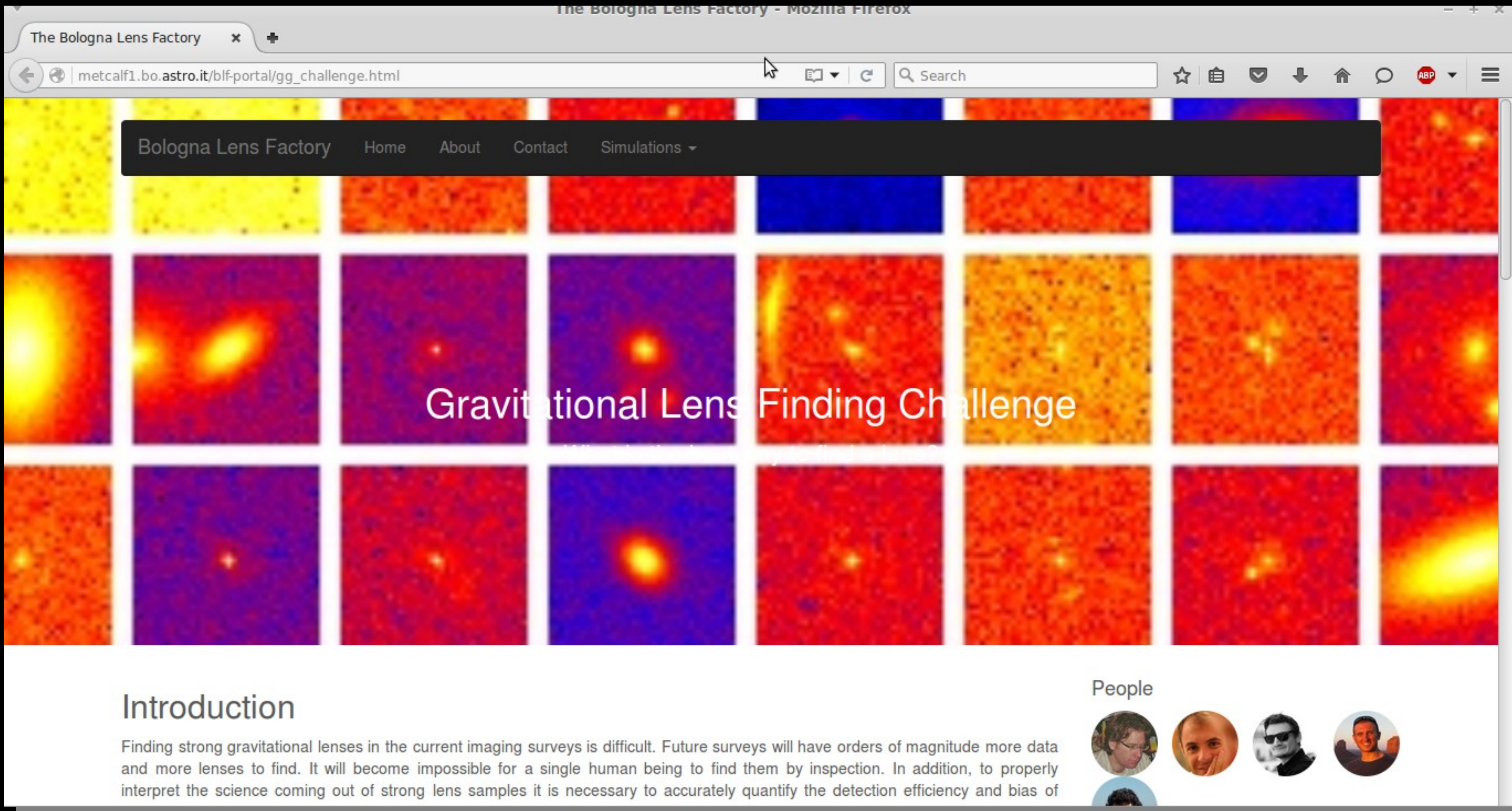


Convert output into calibrated probabilities

Overview

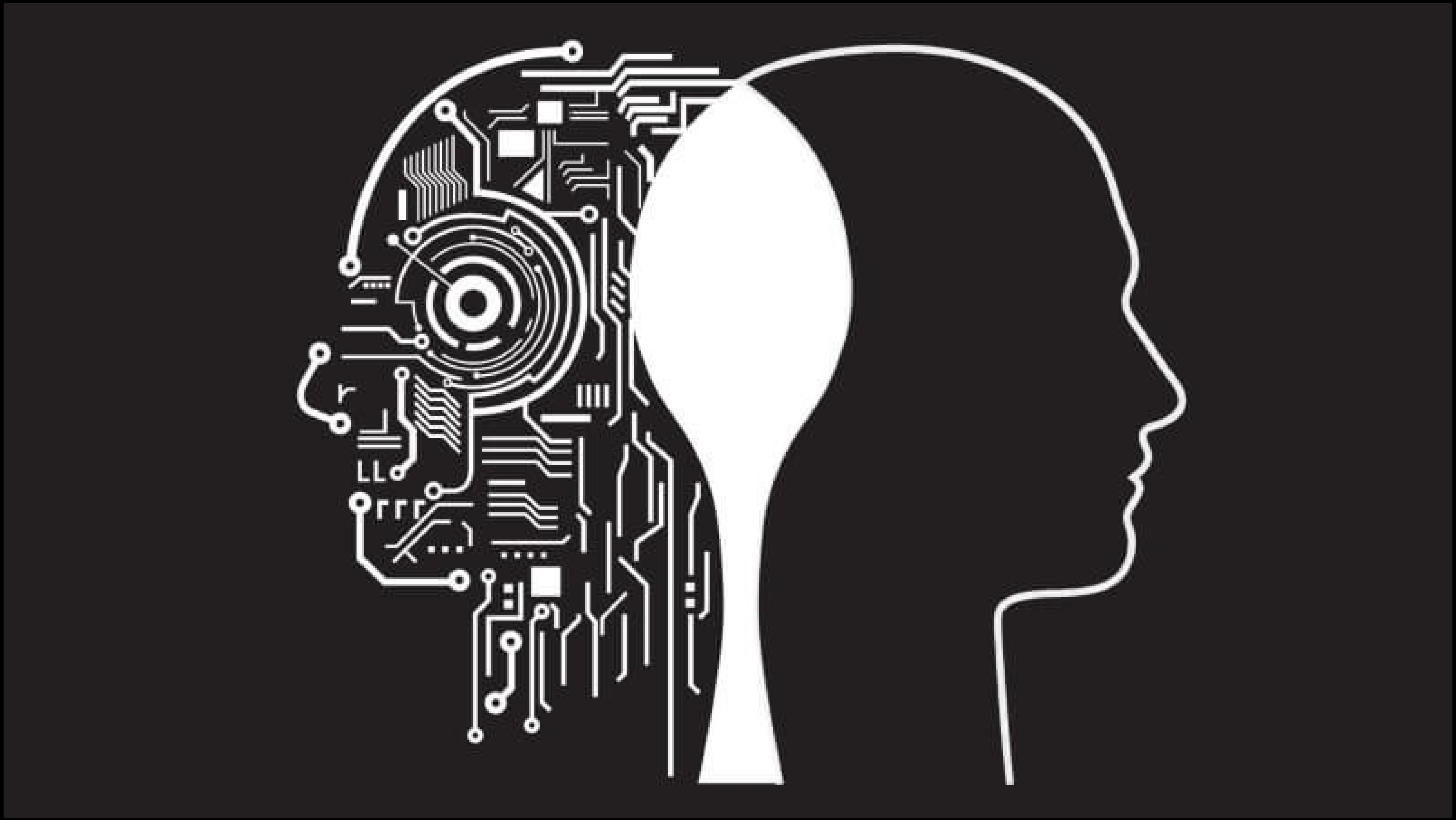
- **Lensing overview**
  - Lensing history
  - Types of gravitational lensing
- **Science from strong lenses**
  - Current applications of lenses
  - Euclid sky survey: more lenses
- **An automatic lens finder**
  - Support vector machine approach
- **Application of the lens finder**
  - Lens finding challenge
  - Kilo Degree Survey
- **The future….**

Lens Finding Challenge



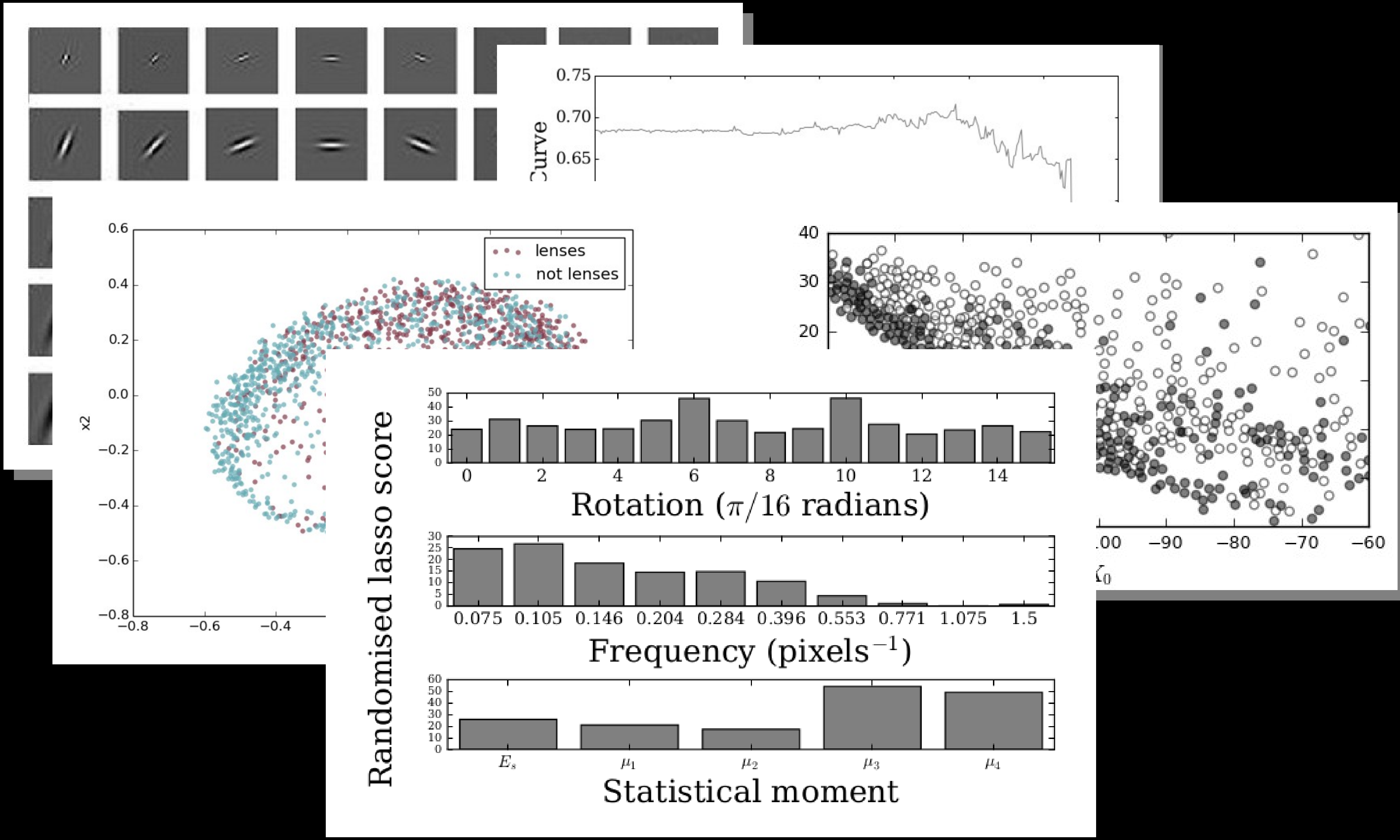
Metcalf et al., 2019

Machine vs human

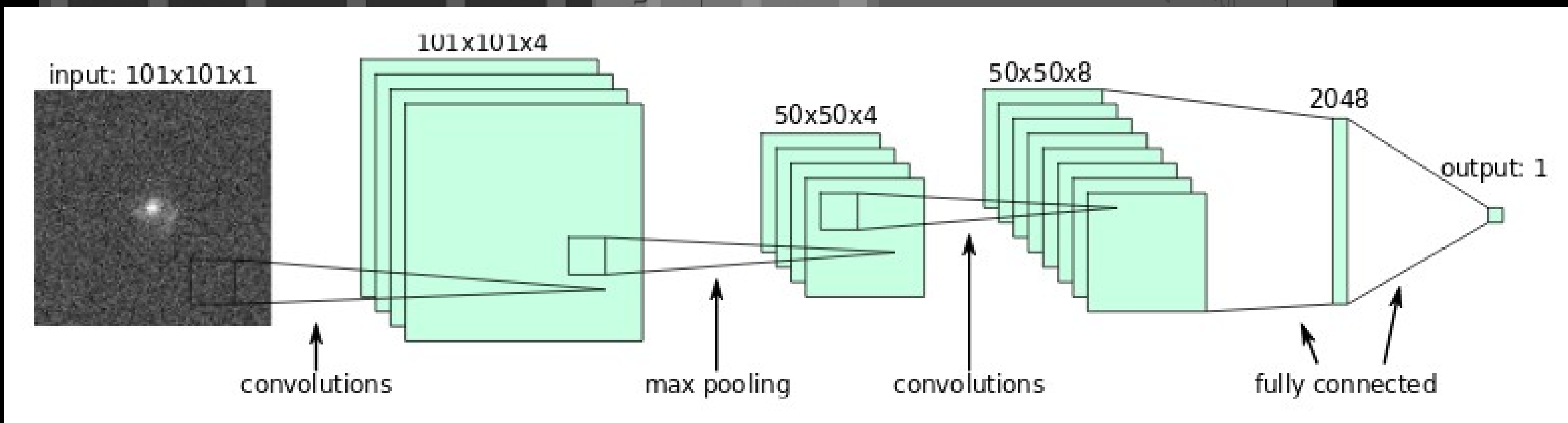


100 000 simulated images, 48 hours

Lens Finding Challenge



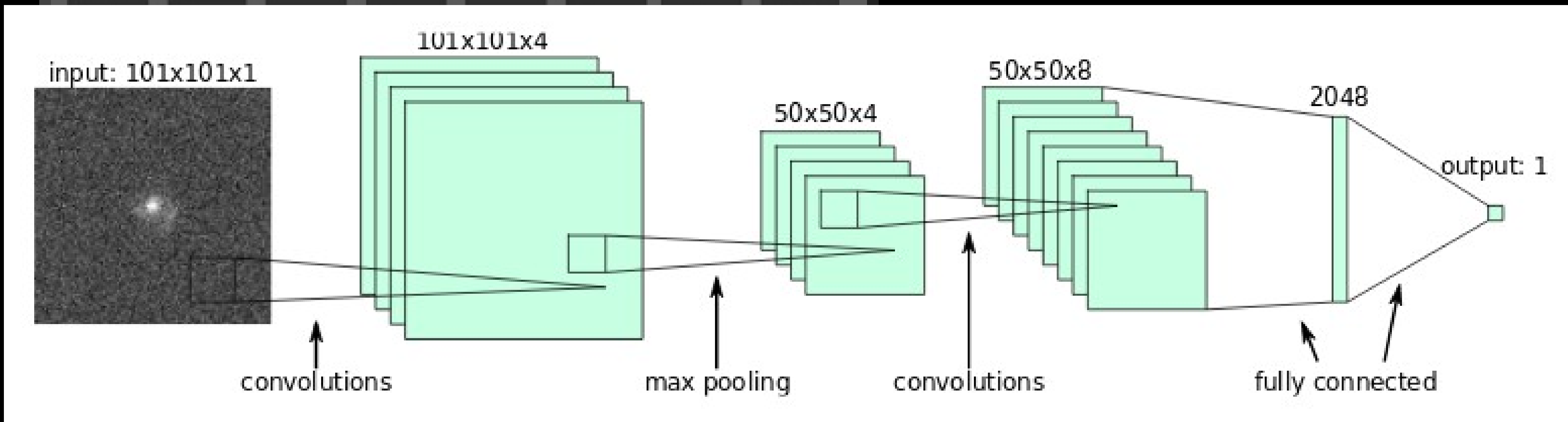
Lens Finding Challenge

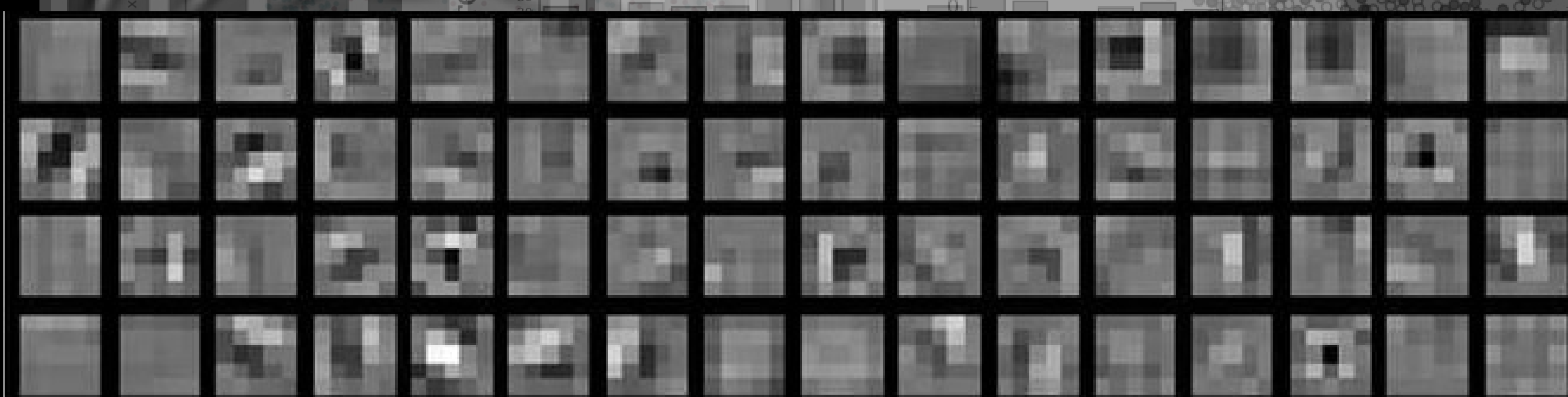


Convolutional neural network
Credit: Schaefer et al. 2017

Lens Finding Challenge

Credit: Schaefer et al. 2017

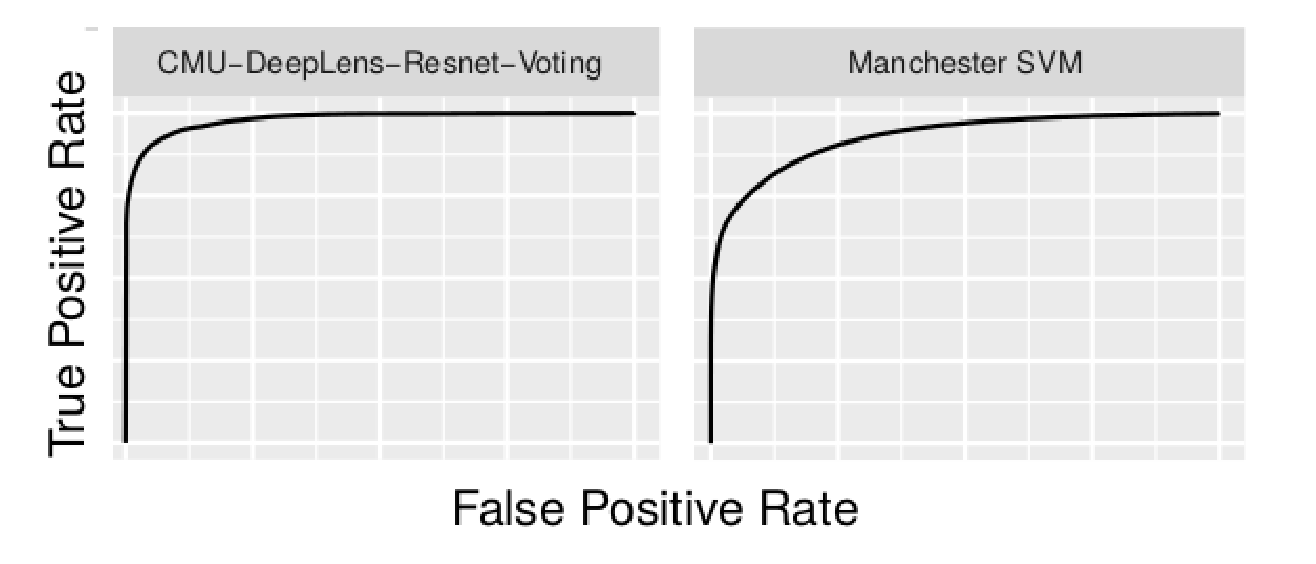




Philippa Hartley

Lens Finding Challenge: results

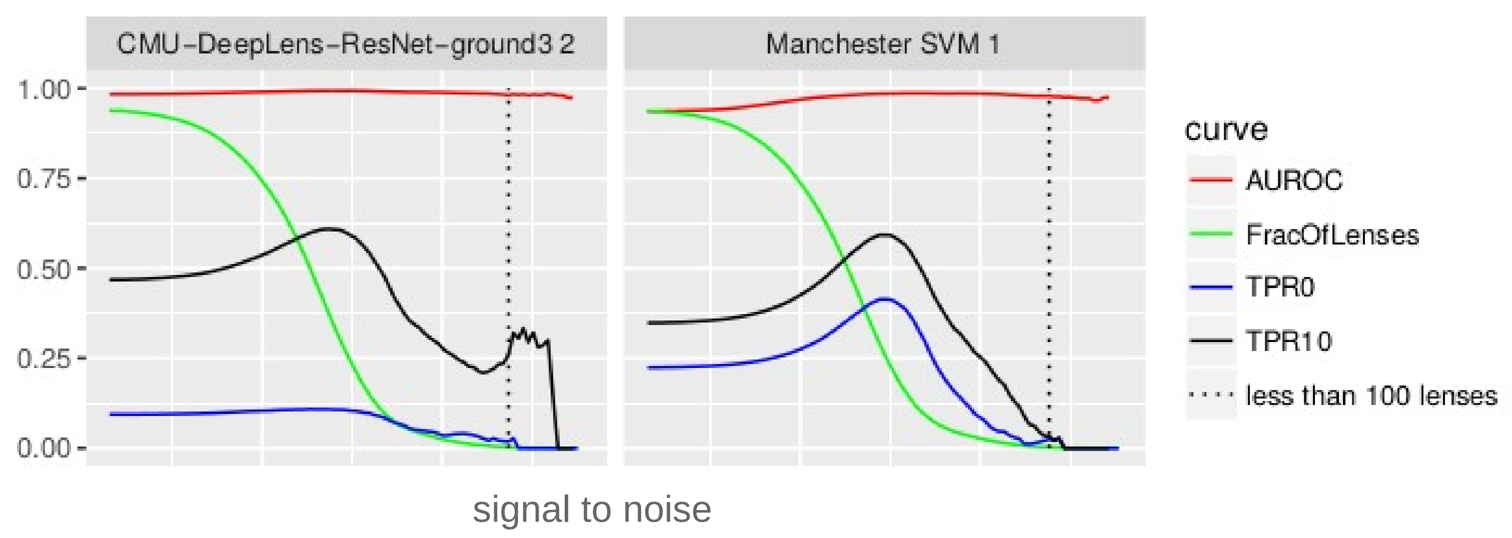
Name	type	AUROC	TPR_0	TPR_{10}	short description
CMU-DeepLens-ResNet-ground3	Ground-Based	0.98	0.09	0.45	CNN
CMU-DeepLens-Resnet-Voting	Ground-Based	0.98	0.02	0.10	CNN
LASTRO EPFL	Ground-Based	0.97	0.07	0.11	CNN
CAS Swinburne Melb	Ground-Based	0.96	0.02	0.08	CNN
AstrOmatic	Ground-Based	0.96	0.00	0.01	CNN
Manchester SVM	Ground-Based	0.93	0.22	0.35	SVM / Gabor
Manchester-NA2	Ground-Based	0.89	0.00	0.01	Human Inspection
ALL-star	Ground-Based	0.84	0.01	0.02	edges/gradiants and Logistic Reg.
CAST	Ground-Based	0.83	0.00	0.00	CNN / SVM
YattaLensLite	Ground-Based	0.82	0.00	0.00	SExtractor



Metcalf et al., in preparation 2017

Lens Finding Challenge: results

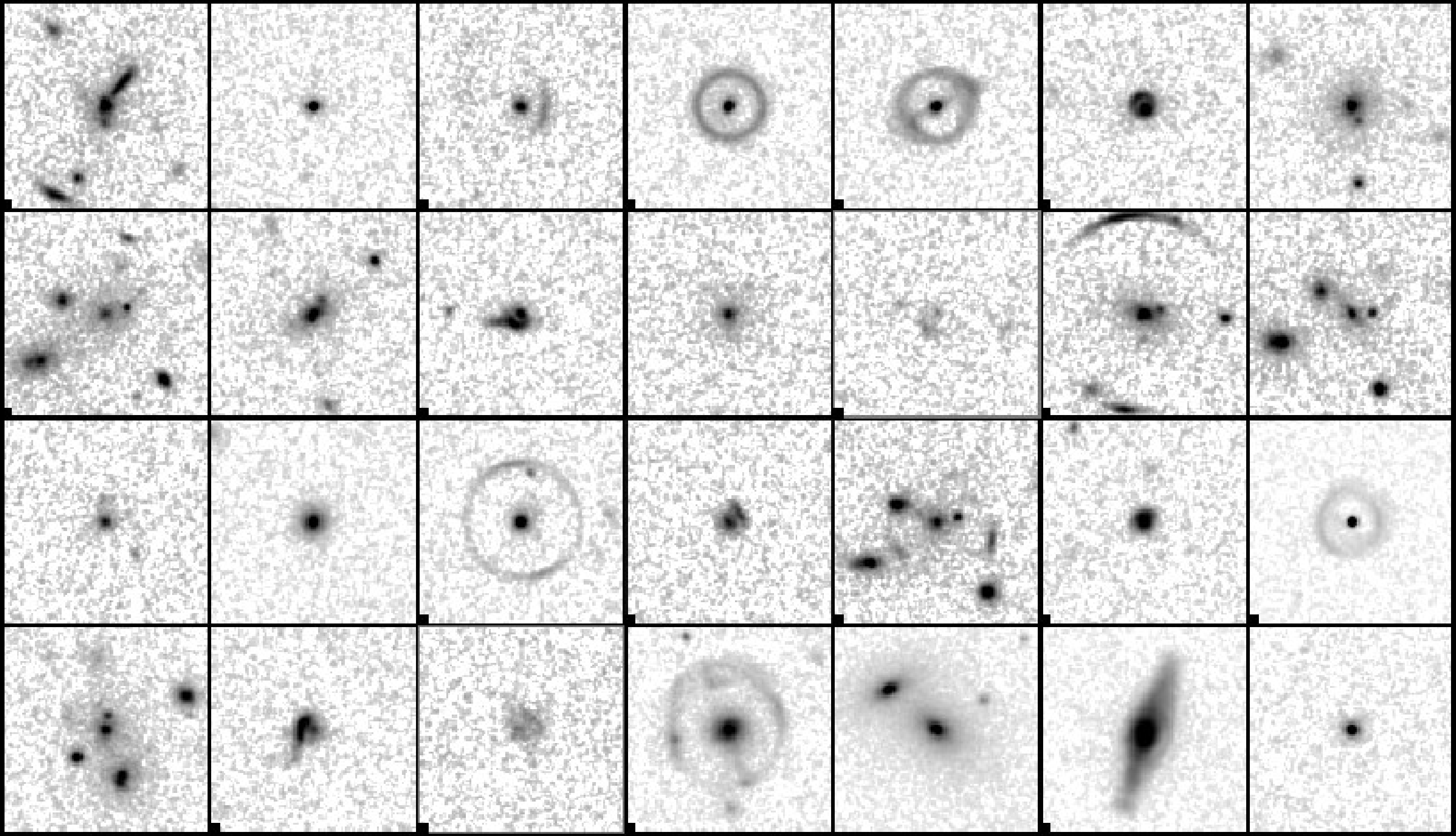
Name	type	AUROC	TPR_0	TPR_{10}	short description
Manchester SVM	Ground-Based	0.93	0.22	0.35	SVM / Gabor
CMU-DeepLens-ResNet-ground3	Ground-Based	0.98	0.09	0.45	CNN
LASTRO EPFL	Ground-Based	0.97	0.07	0.11	CNN
CMU-DeepLens-Resnet-Voting	Ground-Based	0.98	0.02	0.10	CNN
CAS Swinburne Melb	Ground-Based	0.96	0.02	0.08	CNN
ALL-star	Ground-Based	0.84	0.01	0.02	edges/gradiants and Logistic Reg.
Manchester-NA2	Ground-Based	0.89	0.00	0.01	Human Inspection
YattaLensLite	Ground-Based	0.82	0.00	0.00	SExtractor
CAST	Ground-Based	0.83	0.00	0.00	CNN / SVM
AstrOmatic	Ground-Based	0.96	0.00	0.01	CNN



Metcalf et al., in preparation 2017

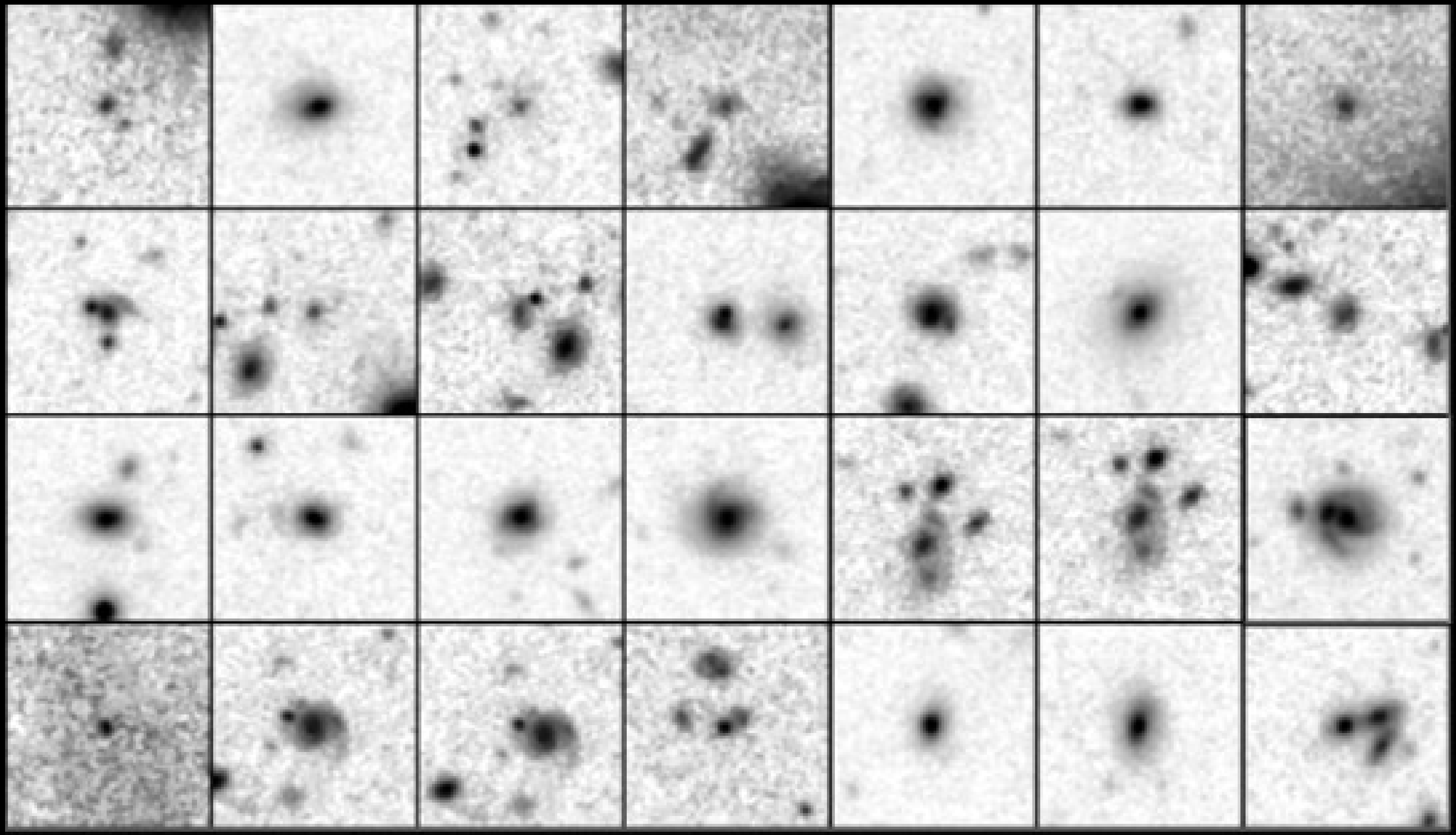
Real life data: Kilo Degree Survey

Domain adaptation



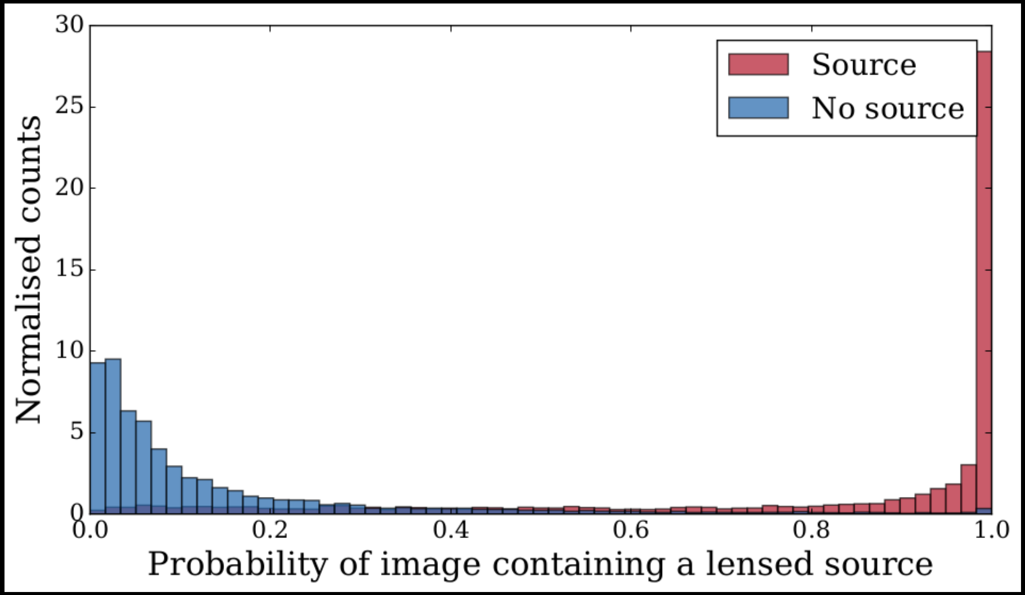
Real life data: Kilo Degree Survey

Domain adaptation

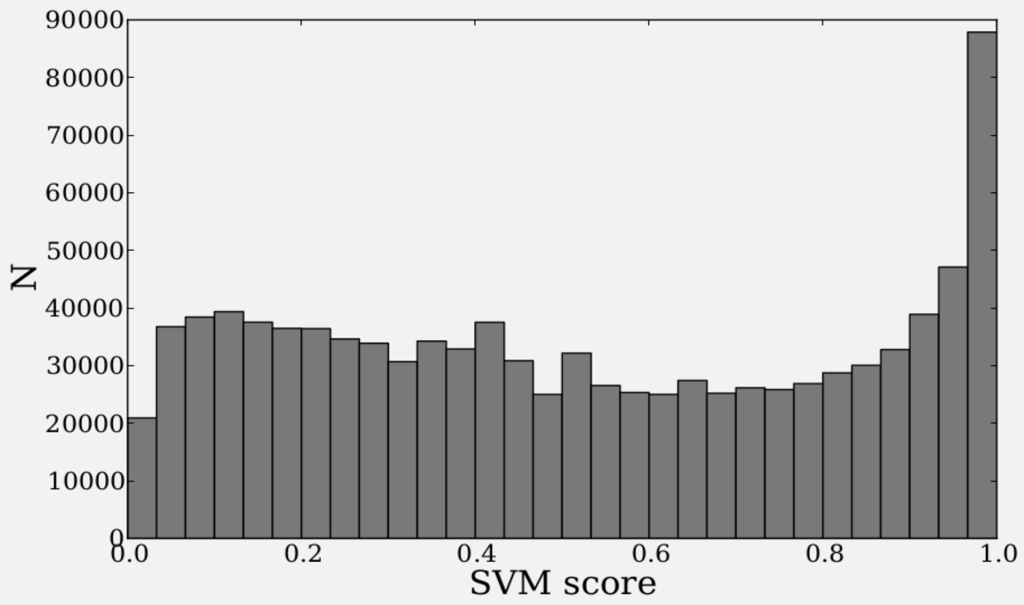


Real data: Kilo Degree Survey

1 000 000 real images after pre-selection



Classification of test mock data

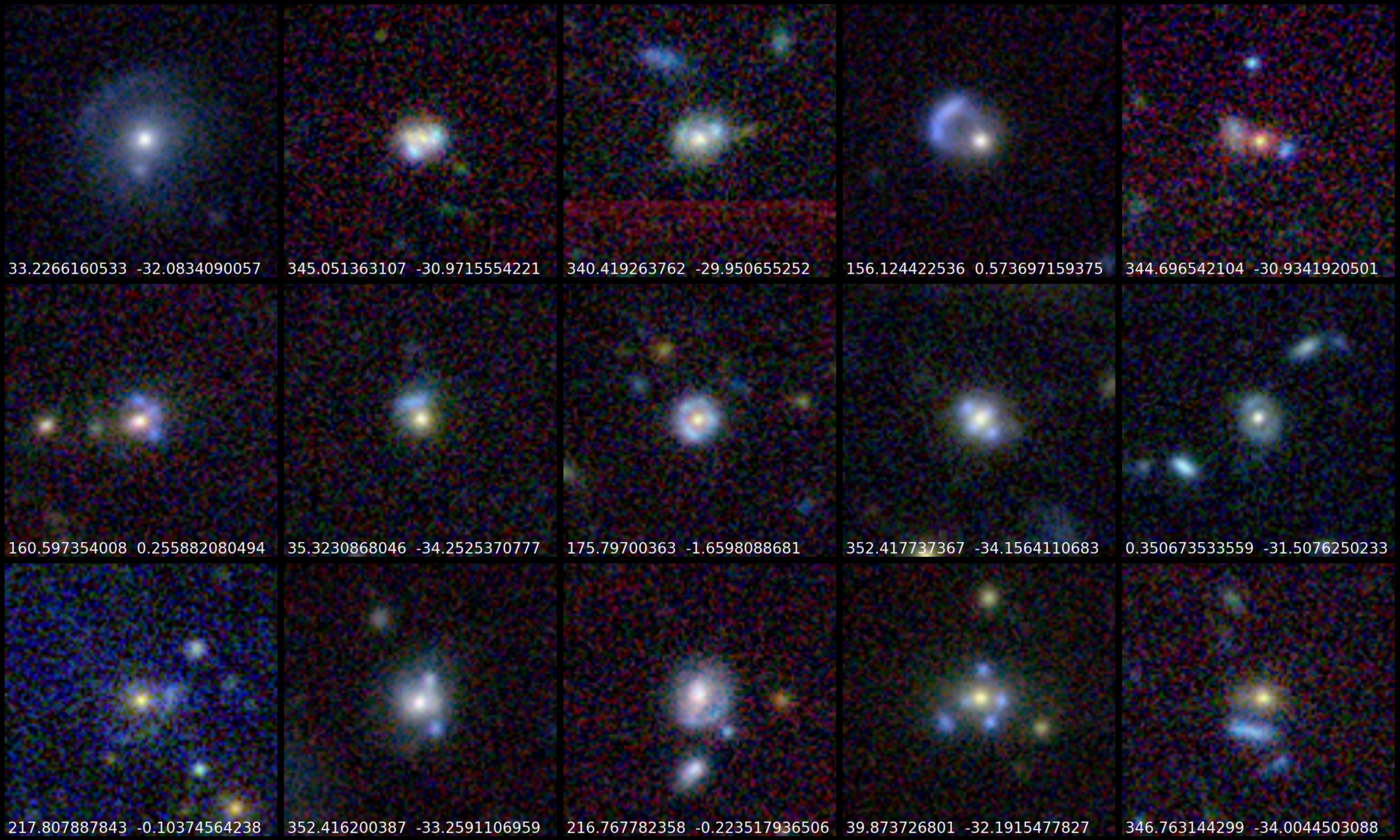


Classification of KiDS real data

Real data: Kilo Degree Survey

Candidate lenses

Hartley et al. 2017 MNRAS



Real data: Kilo Degree Survey
False positive

A ‘smoke ring’ galaxy and compact companion

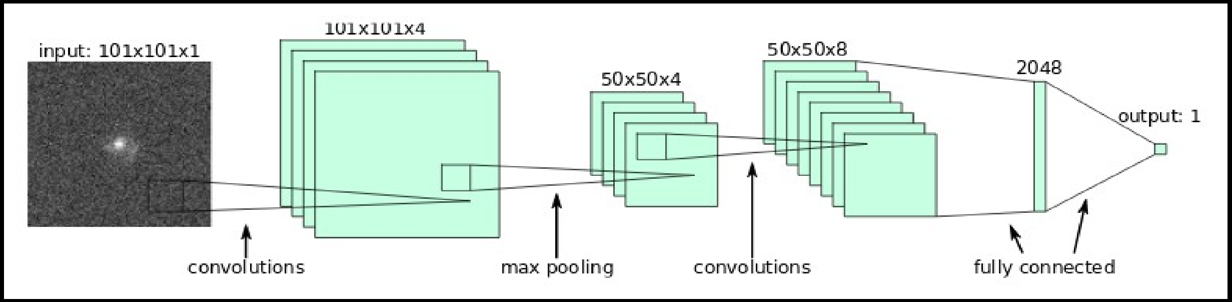
Both located at redshift ~0.4

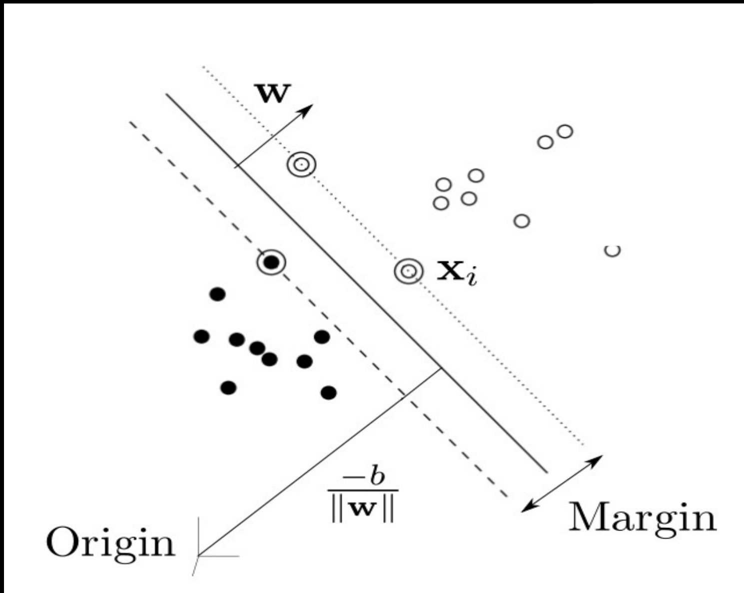
Possibly result of collision

Hartley et al. 2017 MNRAS

Conclusions

- More lenses needed in order to exploit full scientific potential
- Machines now surpass humans in finding lenses
- Surprising strength of SVMs when false postives are a problem
- Domain adaptation: limited by quality of training data
- The best architecture might be: CNN + SVM





P. Hartley, R. Flamary, N. Jackson, A. S. Tagore, R. B. Metcalfe, MNRAS 471 (3): 3378-3397, 2017


Credit: NASA/ESA
Thank you!