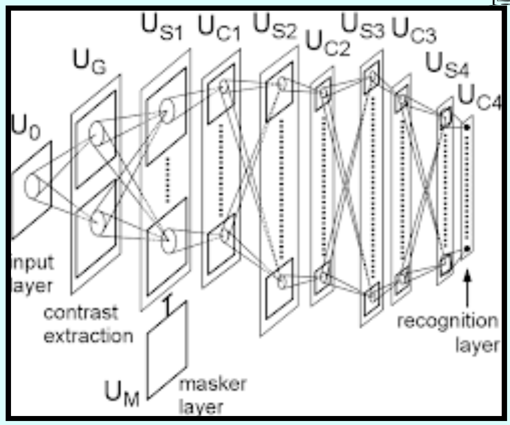
# *Deep learning for Galaxy Morphology*


*Diego Tuccillo*

*Marc Huertas- Company*

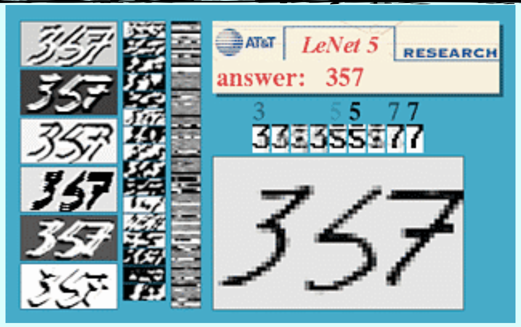

Sorrento, 20-24 October 2016

# Historical introduction in deep learning

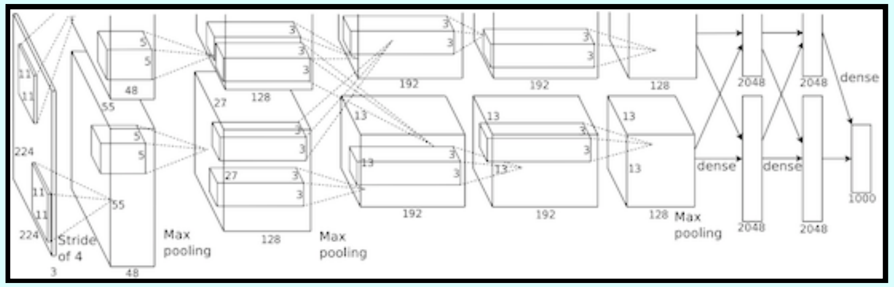



*Fukushima*
*Neocognitron*

*LeCun*
*LaNet 5*




1990's




Krizhevsky 2012
AlexNet

cited > 6,200 times!

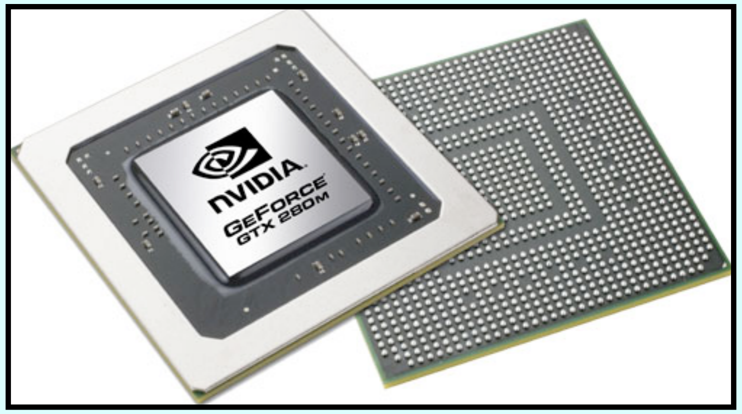
1980's

2000's

2010's

## Big come back of ANNs in the last years

*Tons of data*



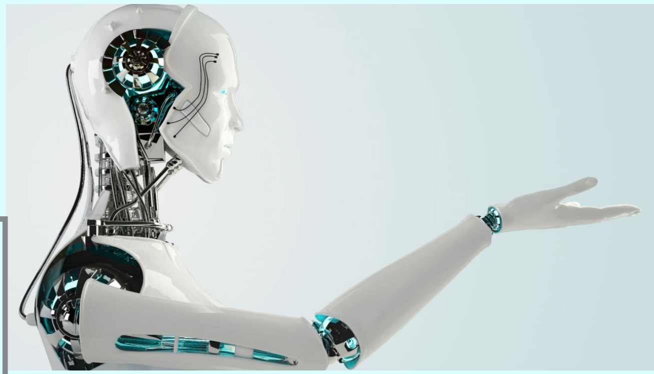

*cheap and fast GPUs*

deep learning shine
when there is lot of data
and complex problems

Computer vision
speech recognition
machine translation

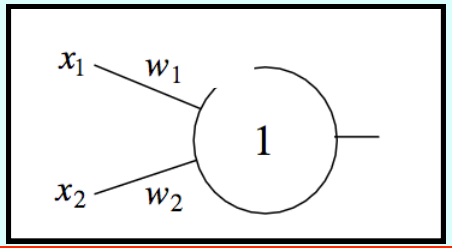
*) Machine learning born from the ambitious goal of ***Artificial intelligence***

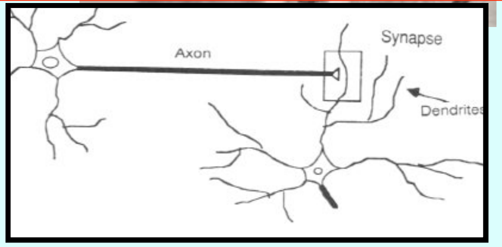
*) Artificial Neural Network founding project: The **Perceptron** (Frank Rosenblatt **1957**)



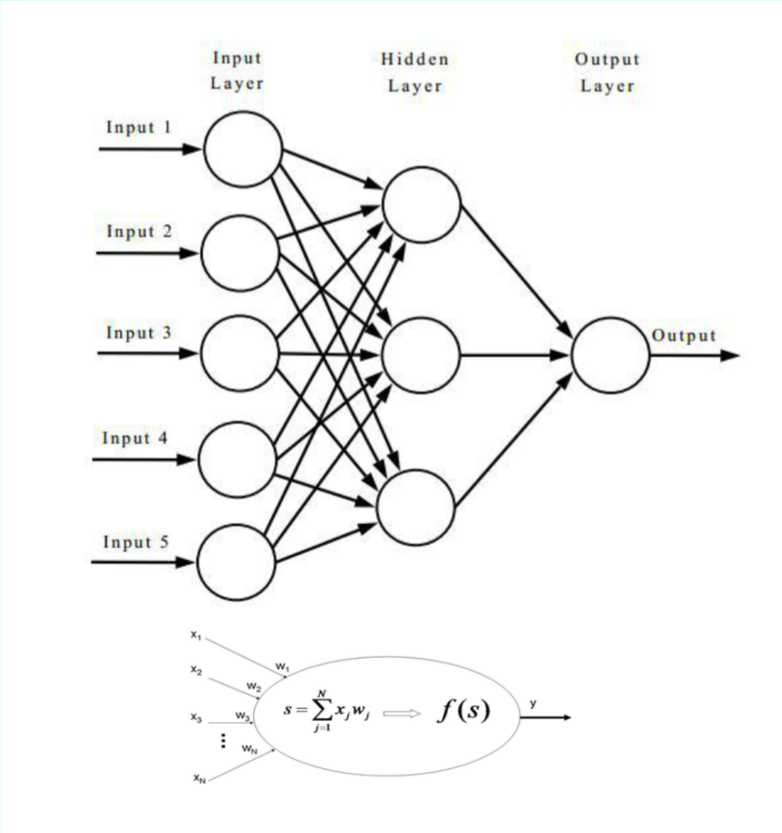
Basic unity **neuron**

***inspired by biological neural systems.***

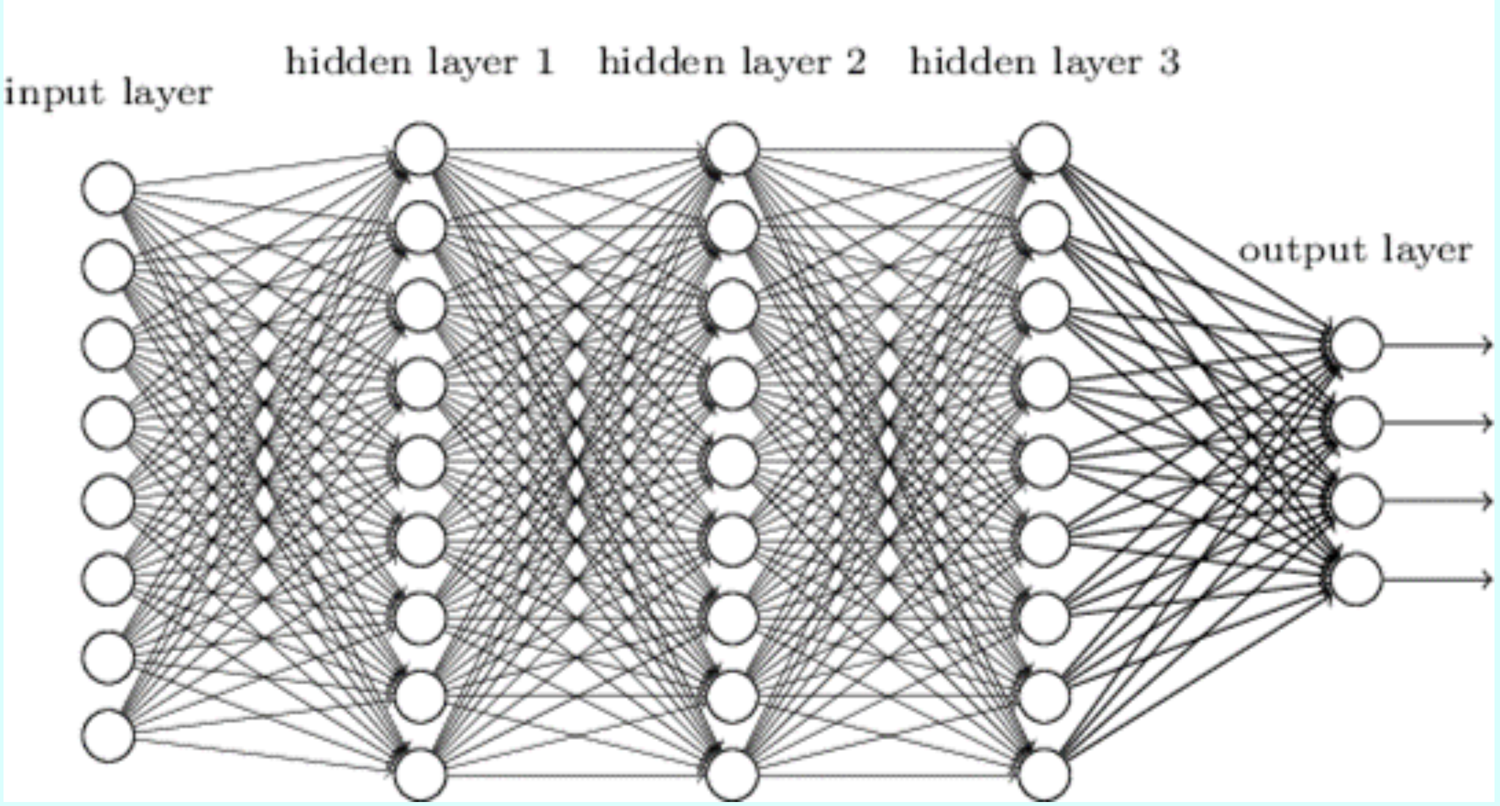



## *Classical ANN*

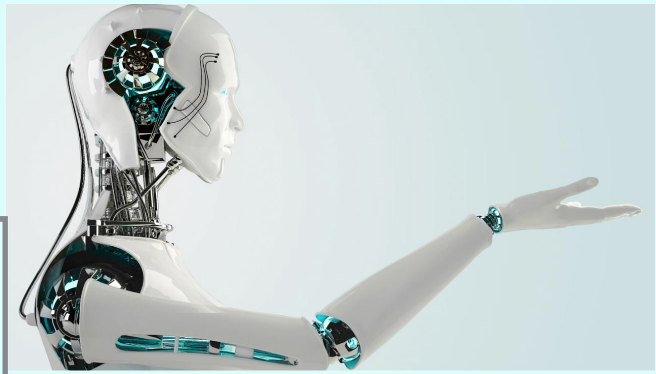



## *Deep ANN*



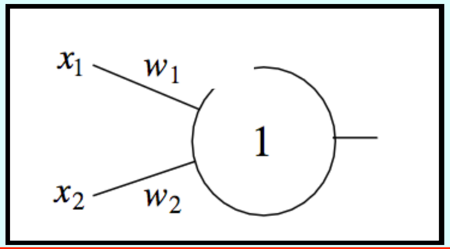
*) Machine learning born from the ambitious goal of ***Artificial intelligence***

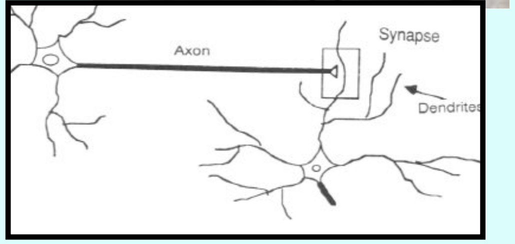
*) Artificial Neural Network founding project: The **Perceptron** (Frank Rosenblatt **1957**)



***inspired by the attempt of simulating biological neural systems.***

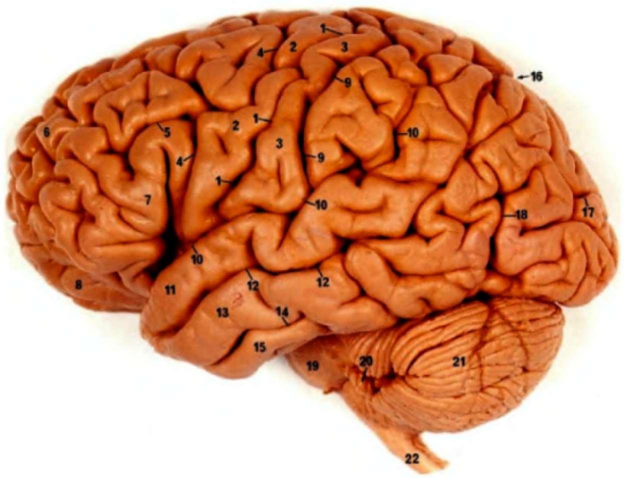
Basic unity **neuron**




An engineering perspective

- Compact
- Energy efficient (20 watts)
- $10^{12}$ Glial cells (power, cooling, support)
- $10^{11}$ Neurons (soma + wires)
- $10^{14}$ Connections (synapses)
- Volume = mostly wires.



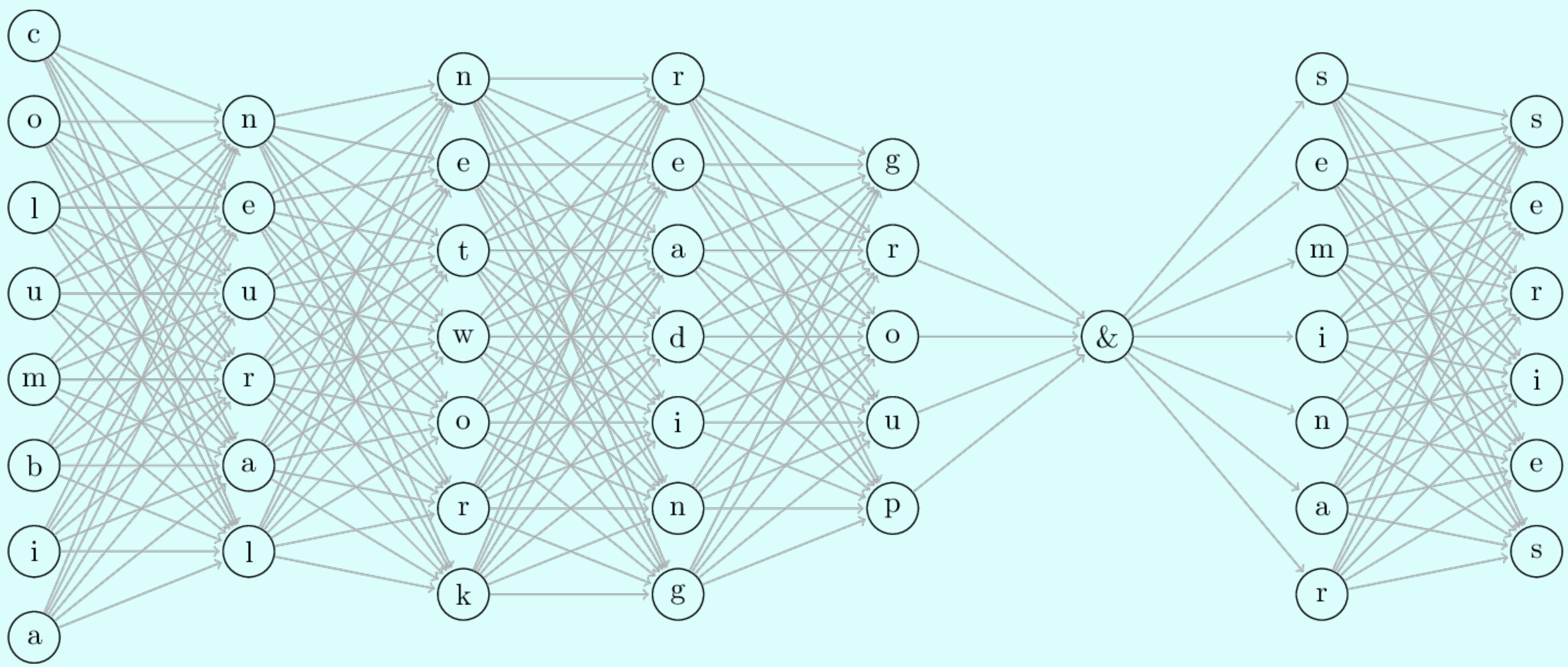
General computing machine?

- Slow for mathematical logic, arithmetic, etc.
- Very fast for vision, speech, language, social interactions, etc.
- Evolution: vision –> language –> logic.

*Considering both the architecture and better tasks is good at…*

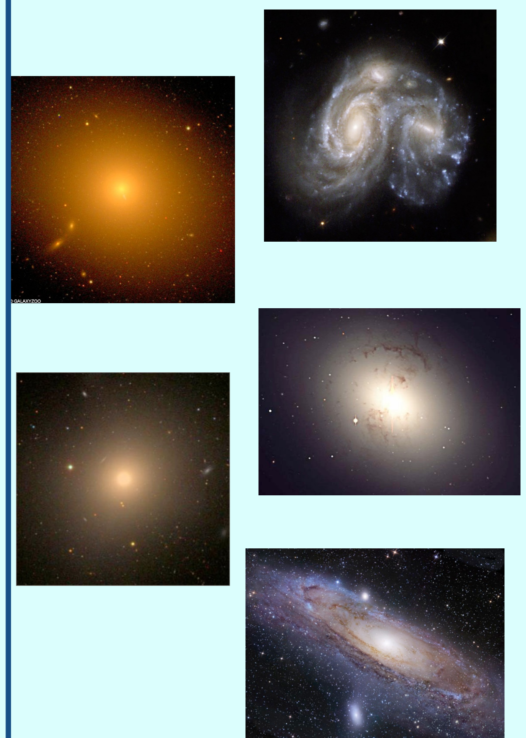
*…with DNN, we are getting closer to that original **dream**!*

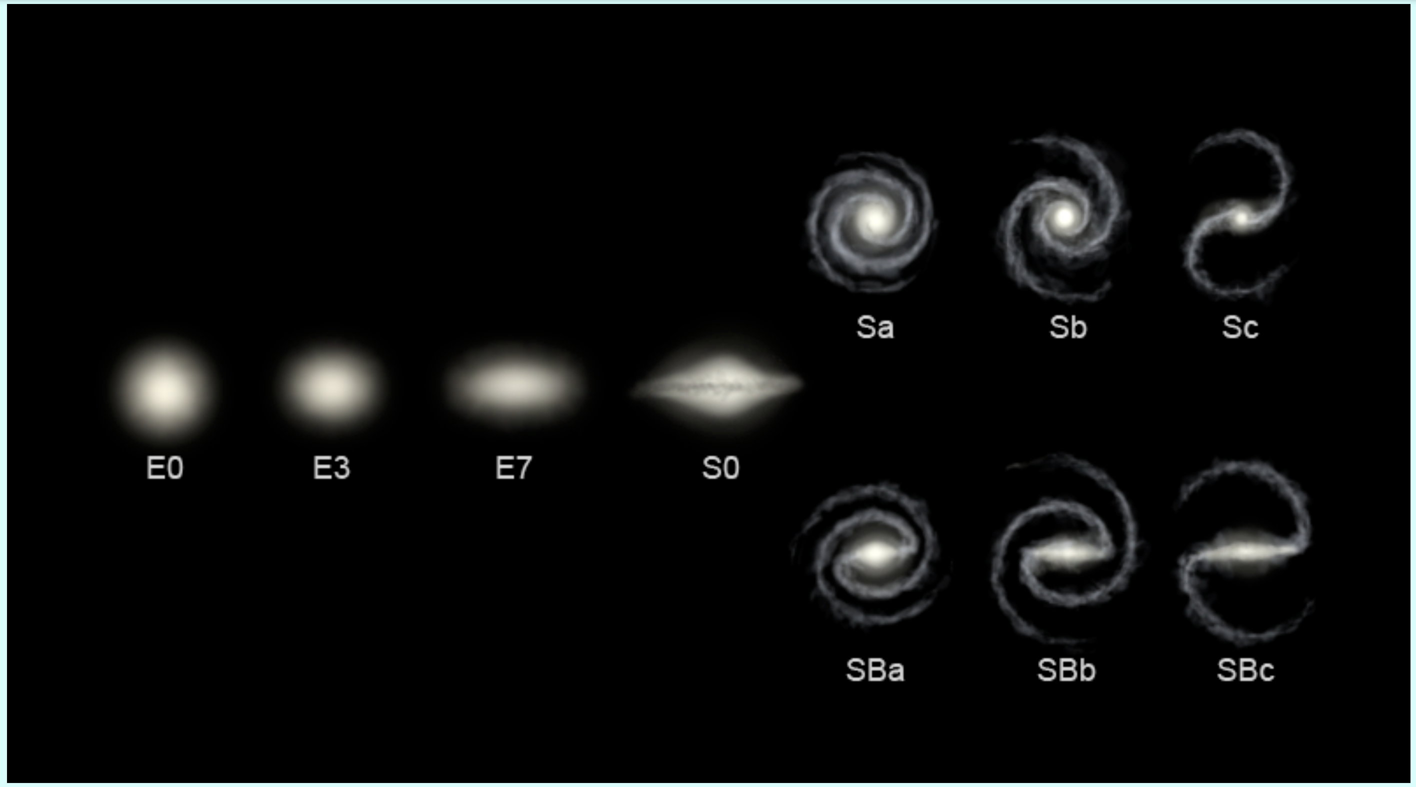
# DNN for Galaxy Morphology



# "Given a 2-D image of a galaxy, typical morphology studies:

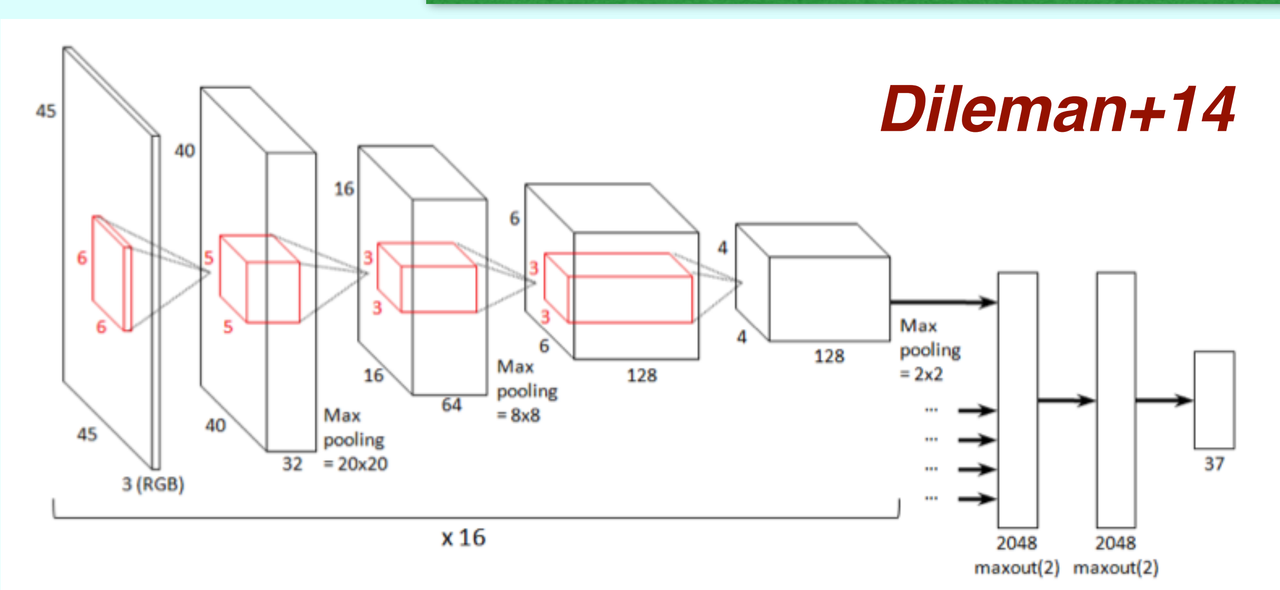
## Classification






"Objects in the same box experienced the same physics"

**Deep Neural Network proved to be very effective**

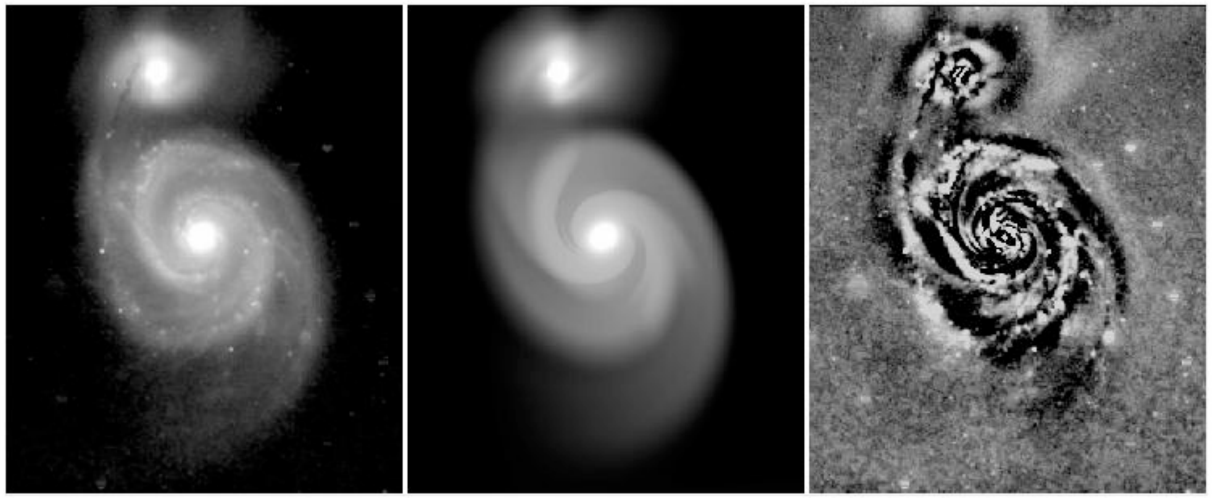



***Dileman+14***

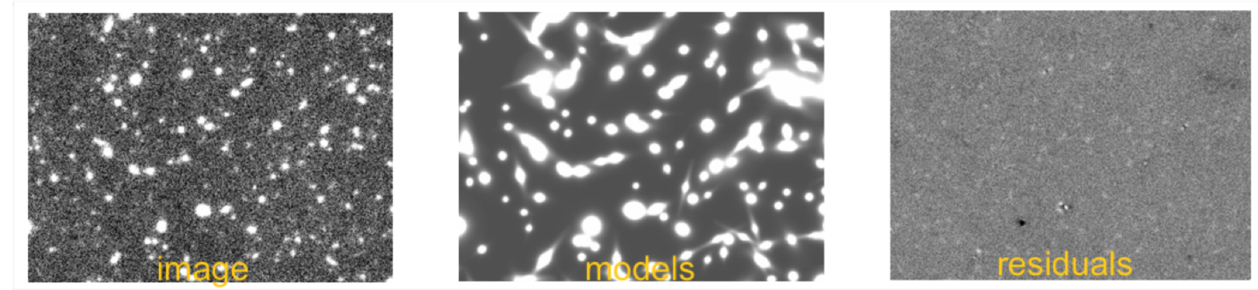
And at high-z with CANDELS FIELDS ***Huertas-Company +15***

accuracy > 90%

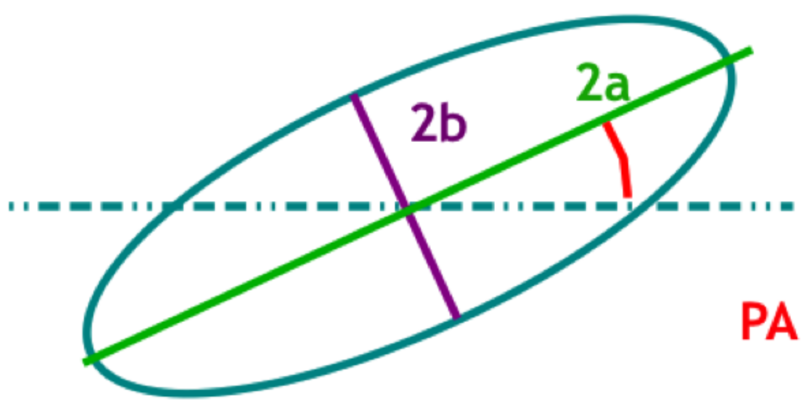
# Galaxy Fitting



model light perfiles using parametric functions




the features of interest are summarized into a small set of numbers




- magnitude
- effective radius
- PA = position angle
- ellipticity
- Sersic index

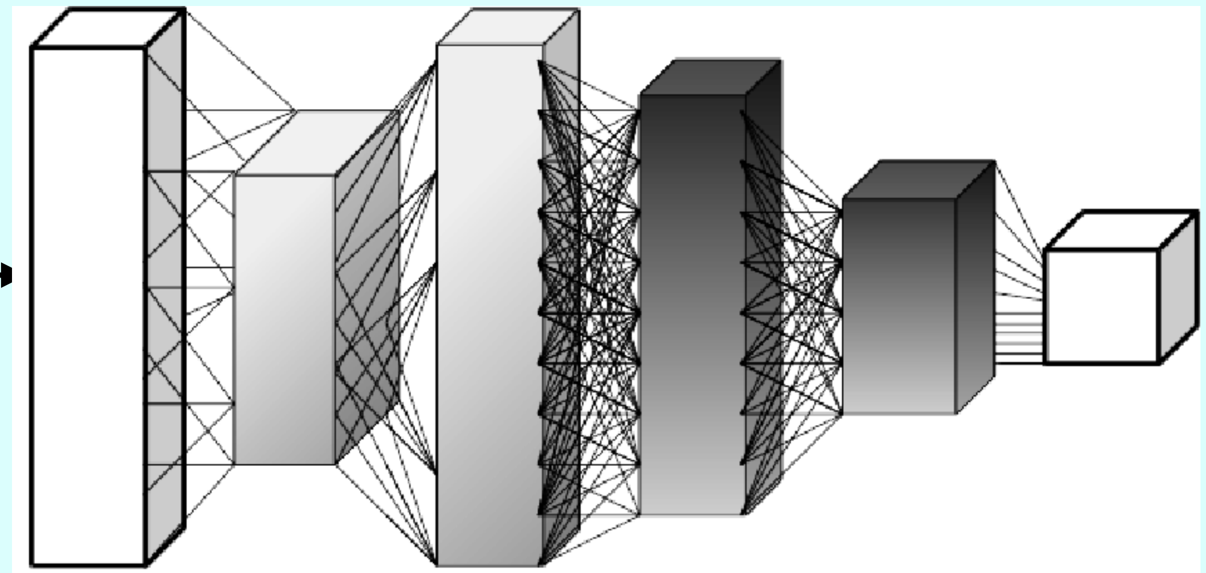
most popular codes:

- GALFIT (Peng et al., 2002, 2010)
- GIM2D (Simard et al, 2002)
- SExtractor (Bertin et al, TBD)

Our research: apply DNN for galaxy surface brightness profile fitting

# We developed a convolutional DNN algorithm to fit 1-component galaxies

input
128x128



Trained with 24,000

Validated with 6000

***Details on the true DNN architecture***

*(to be submitted by November)*

**Deep Neural Network for galaxy morphology**

D. Tuccillo, M. Huertas, E. Decenciere, S.

l'Observatoire de Paris

MINES ParisTech

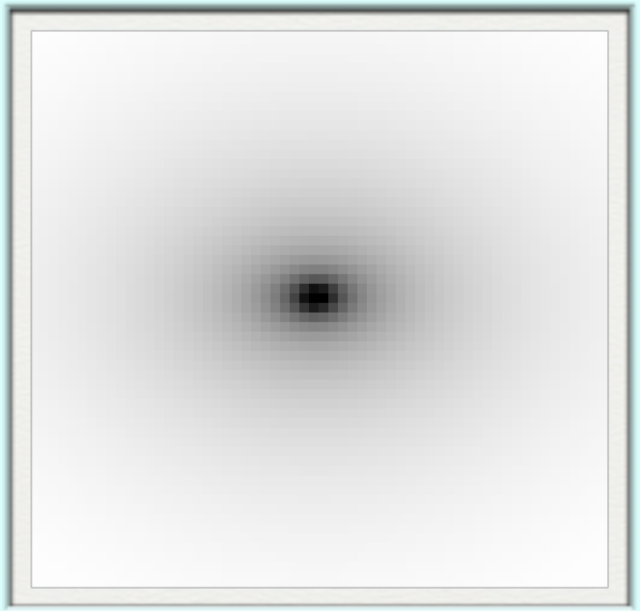
# We simulated 30,000 stamps of 1-component galaxies

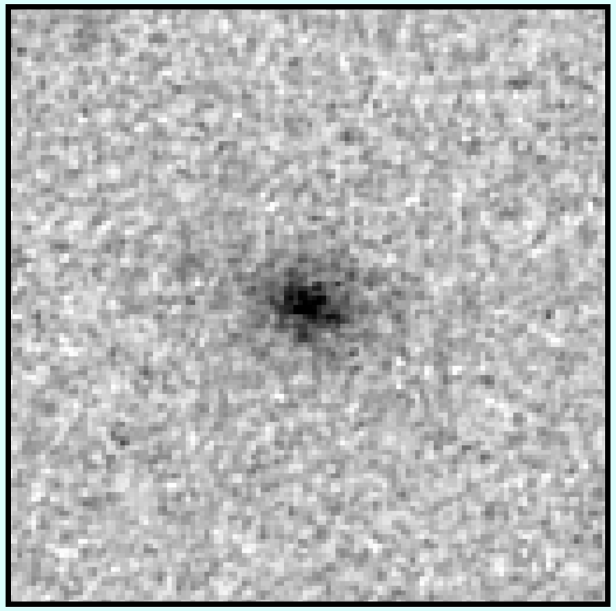
Single Sersic HST/CANDELS

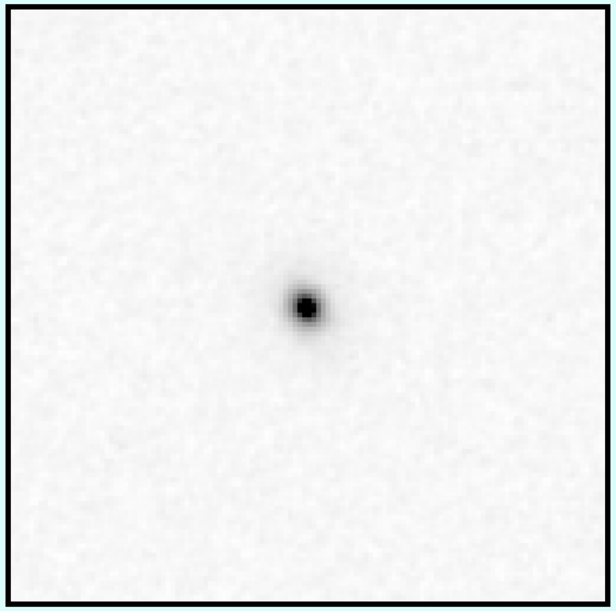
WITH REAL NOISE

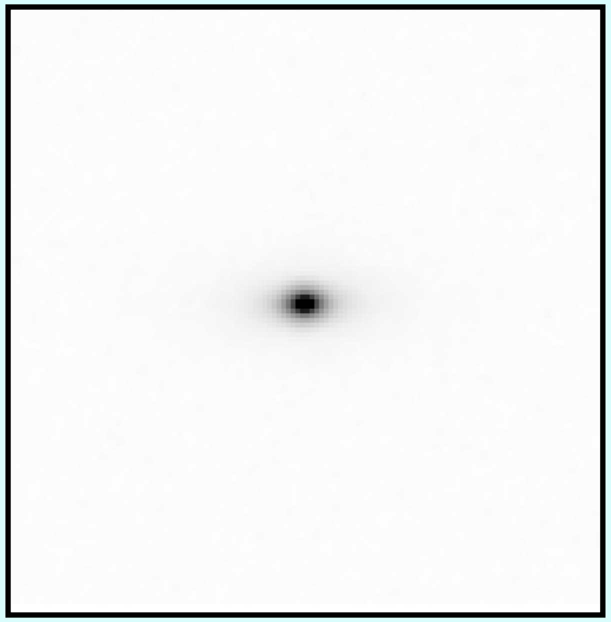
PIXEL scale 0.06''

REAL PSF CANDELS









0<RADIUS (arcsec)<1.9

18< MAG<23

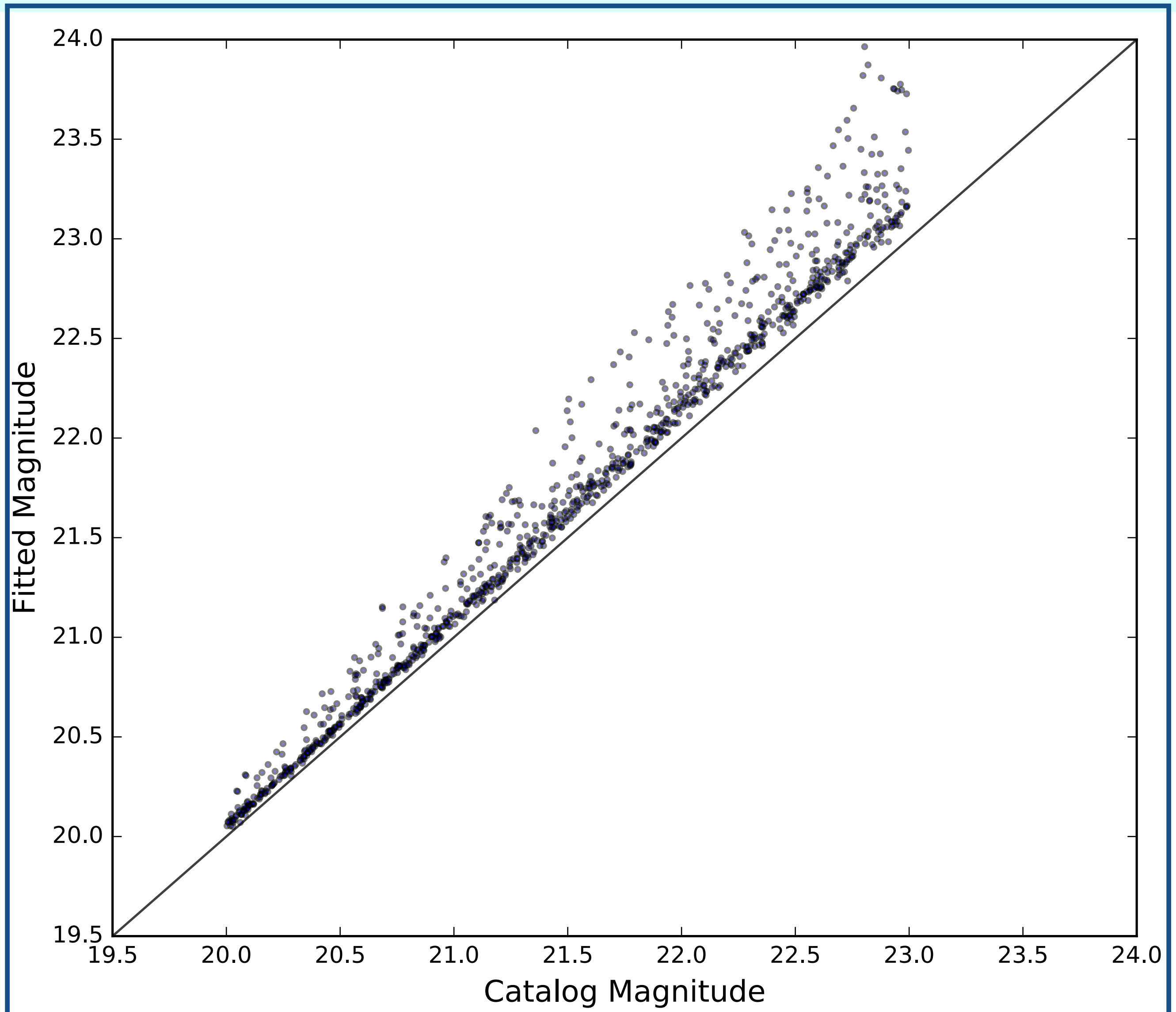
0.2<ELLIPTICITY<0.8

0.3<SERSIC INDEX<6.3

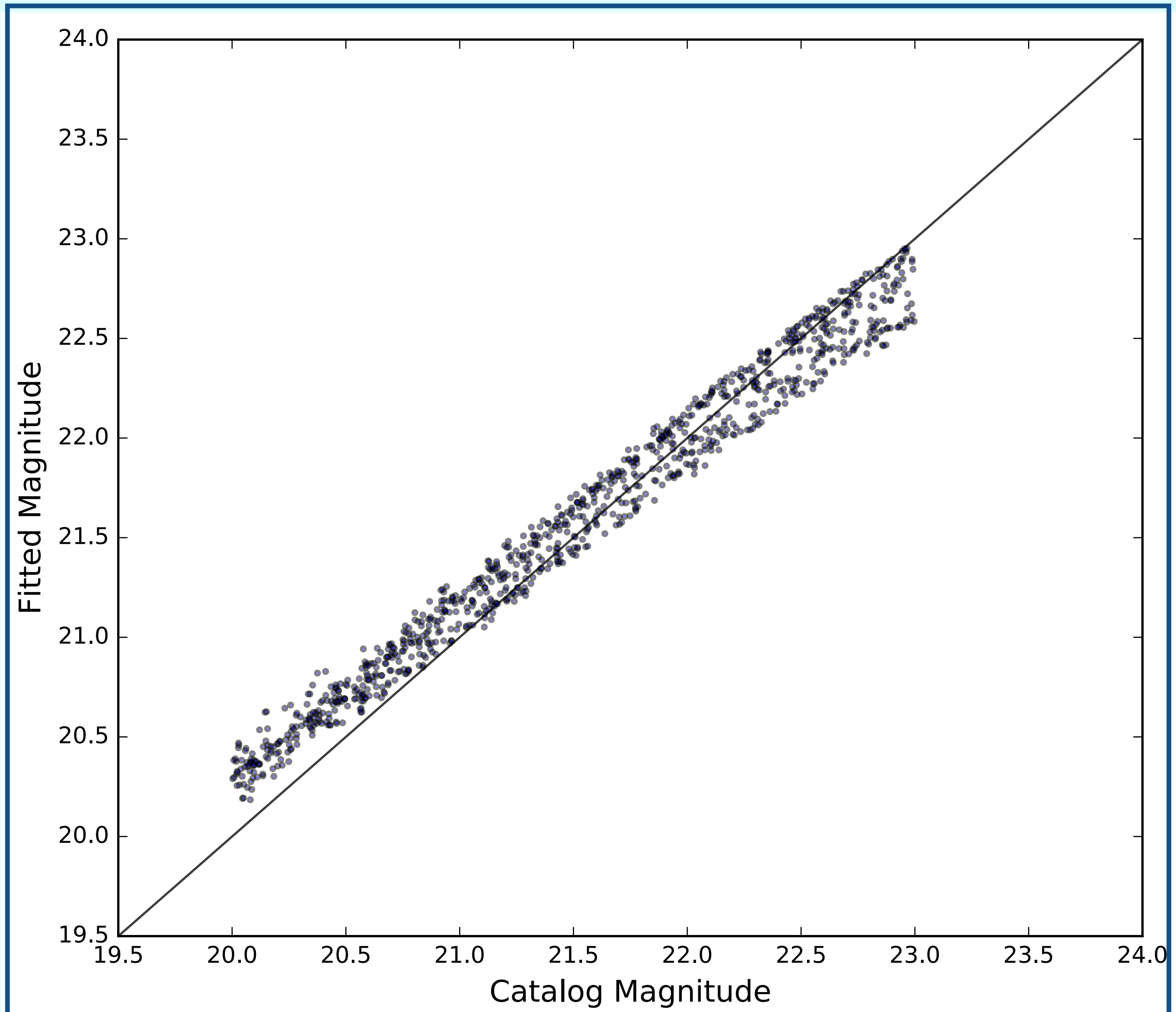
0<Position Angle (degree)<180

# Applied the result of learning on 1000 simulated galaxies

Magnitude



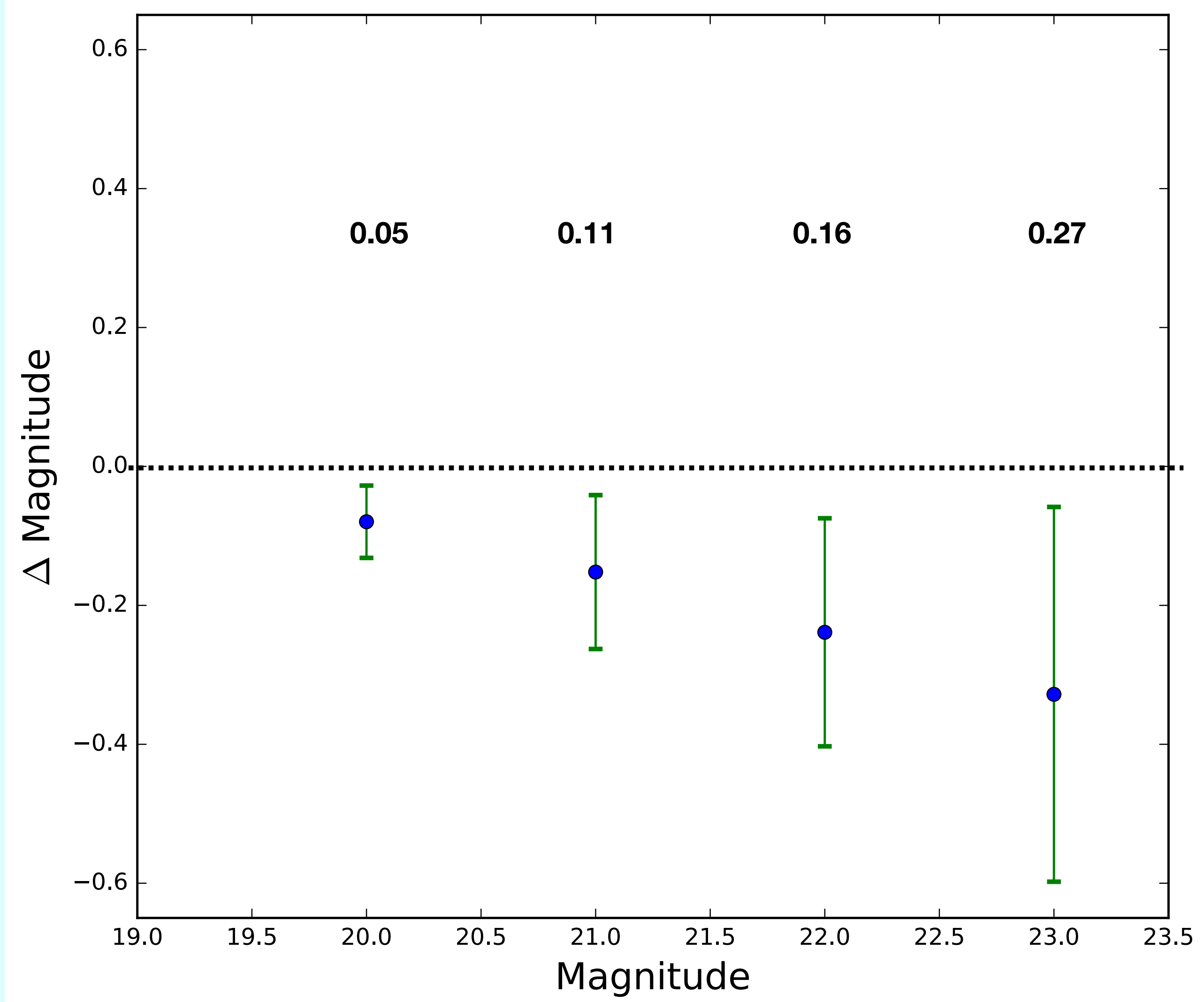

GALFIT-SExtractor



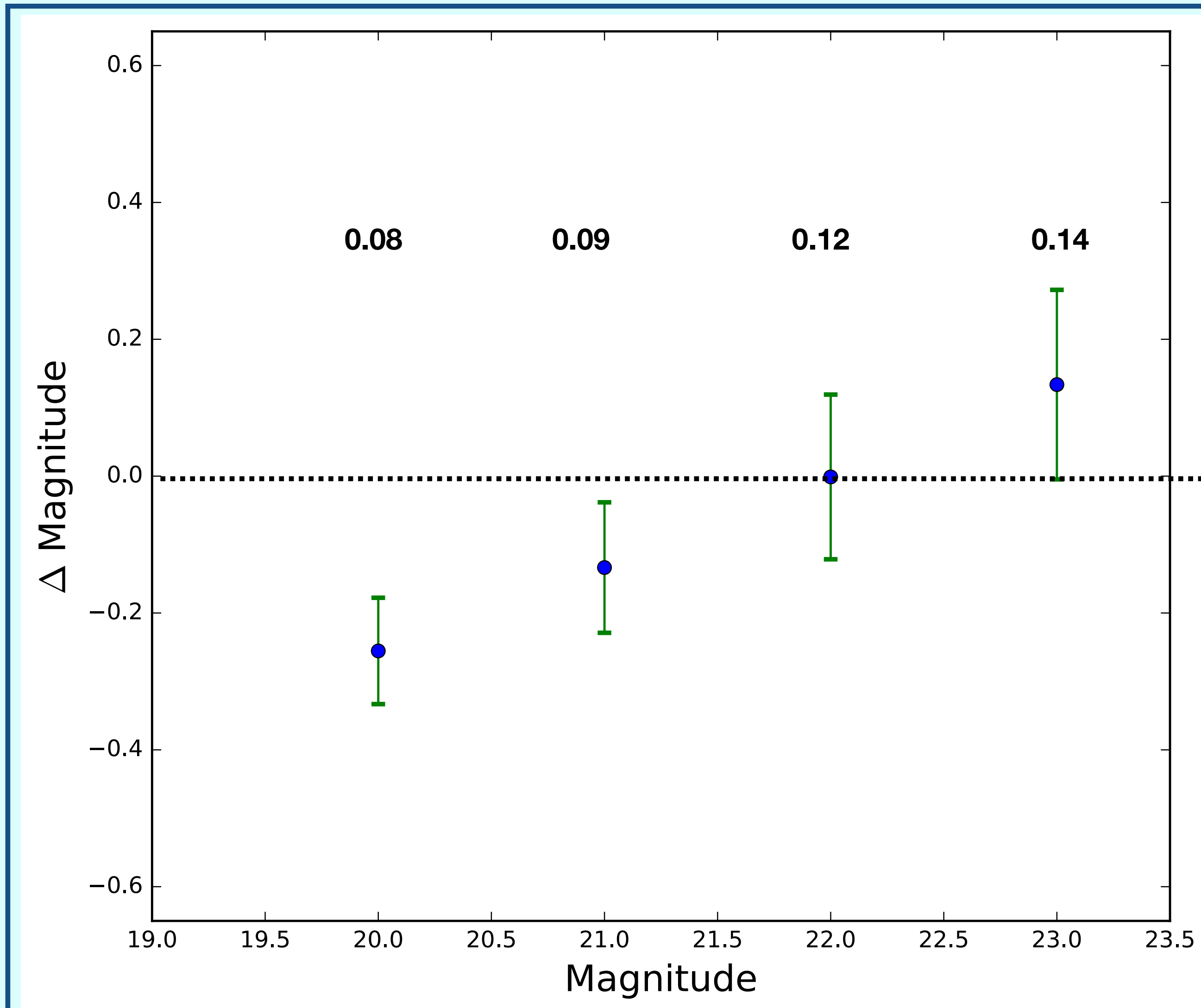

DNN

# Applied the result of learning on 1000 simulated galaxies

Magnitude



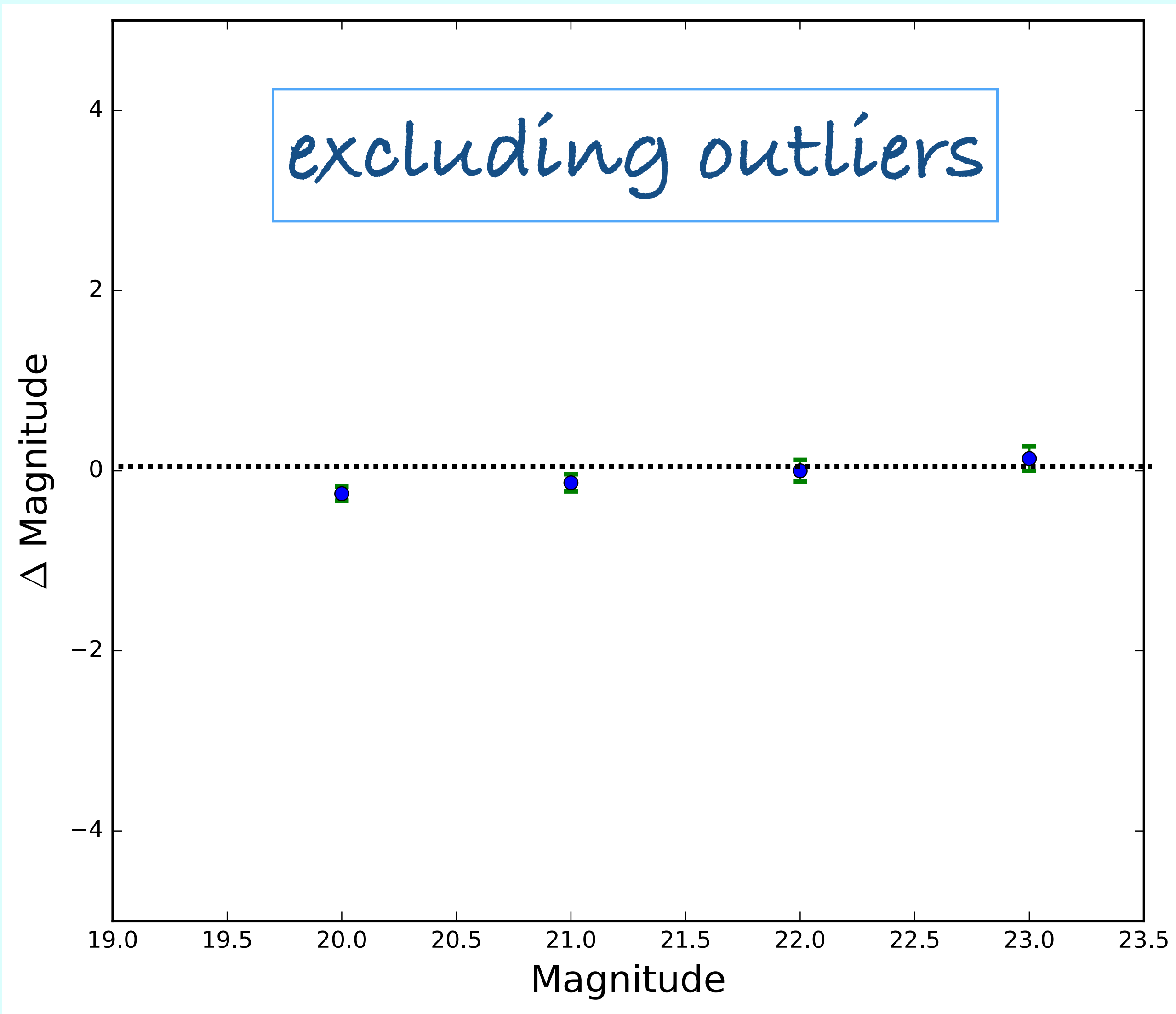

GALFIT-SExtractor



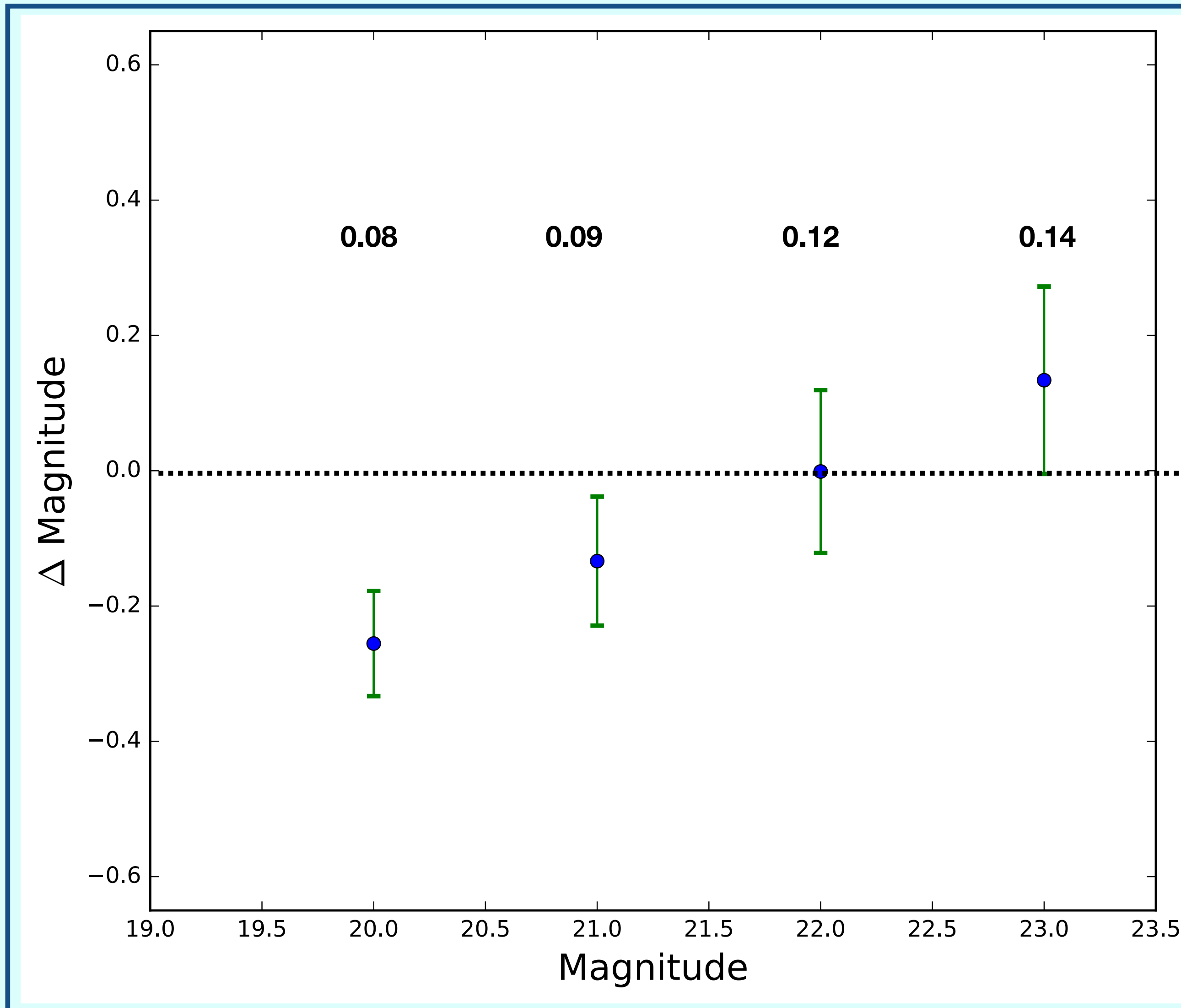

DNN

# Applied the result of learning on 1000 simulated galaxies

Magnitude



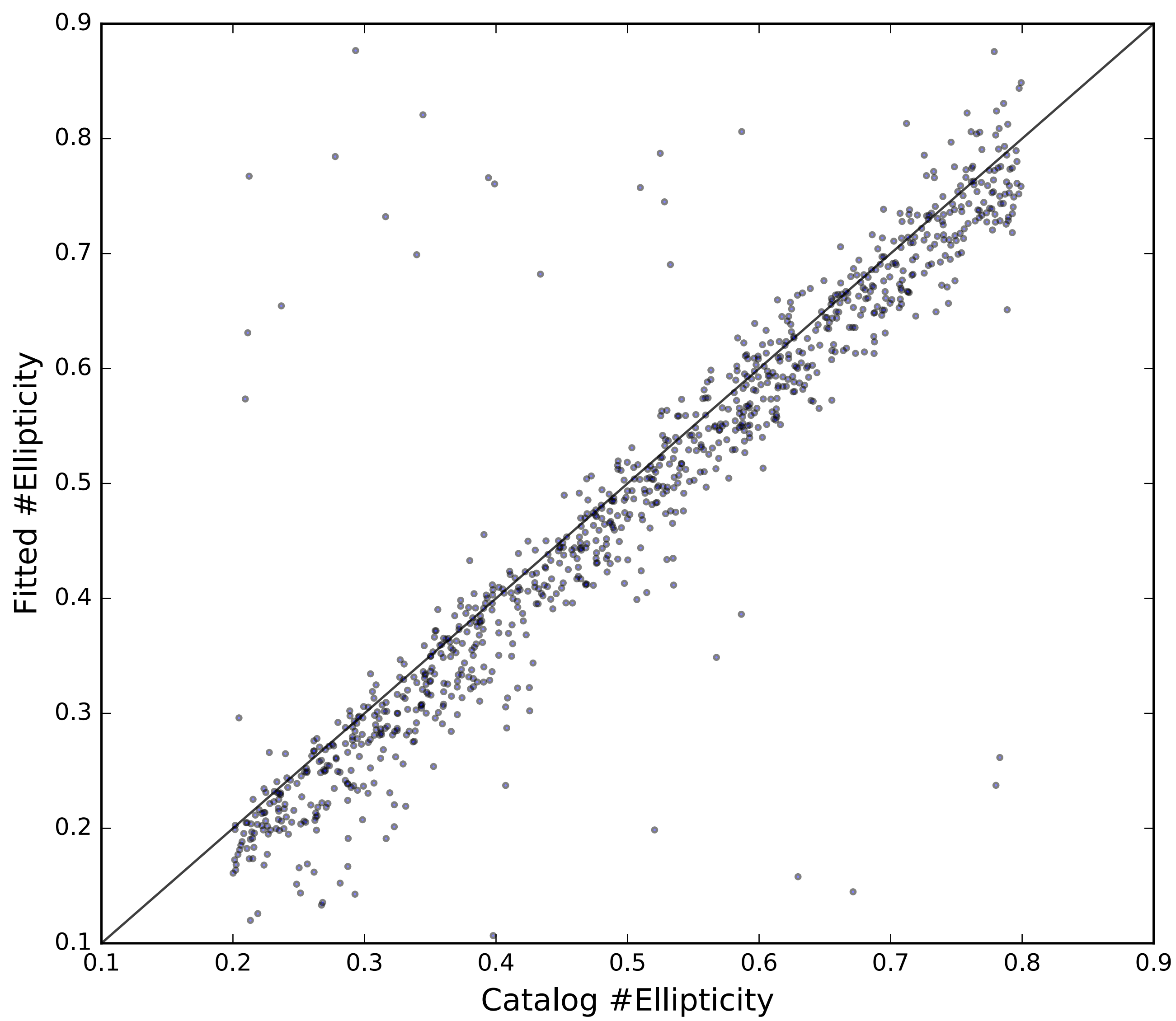

GALFIT-SExtractor



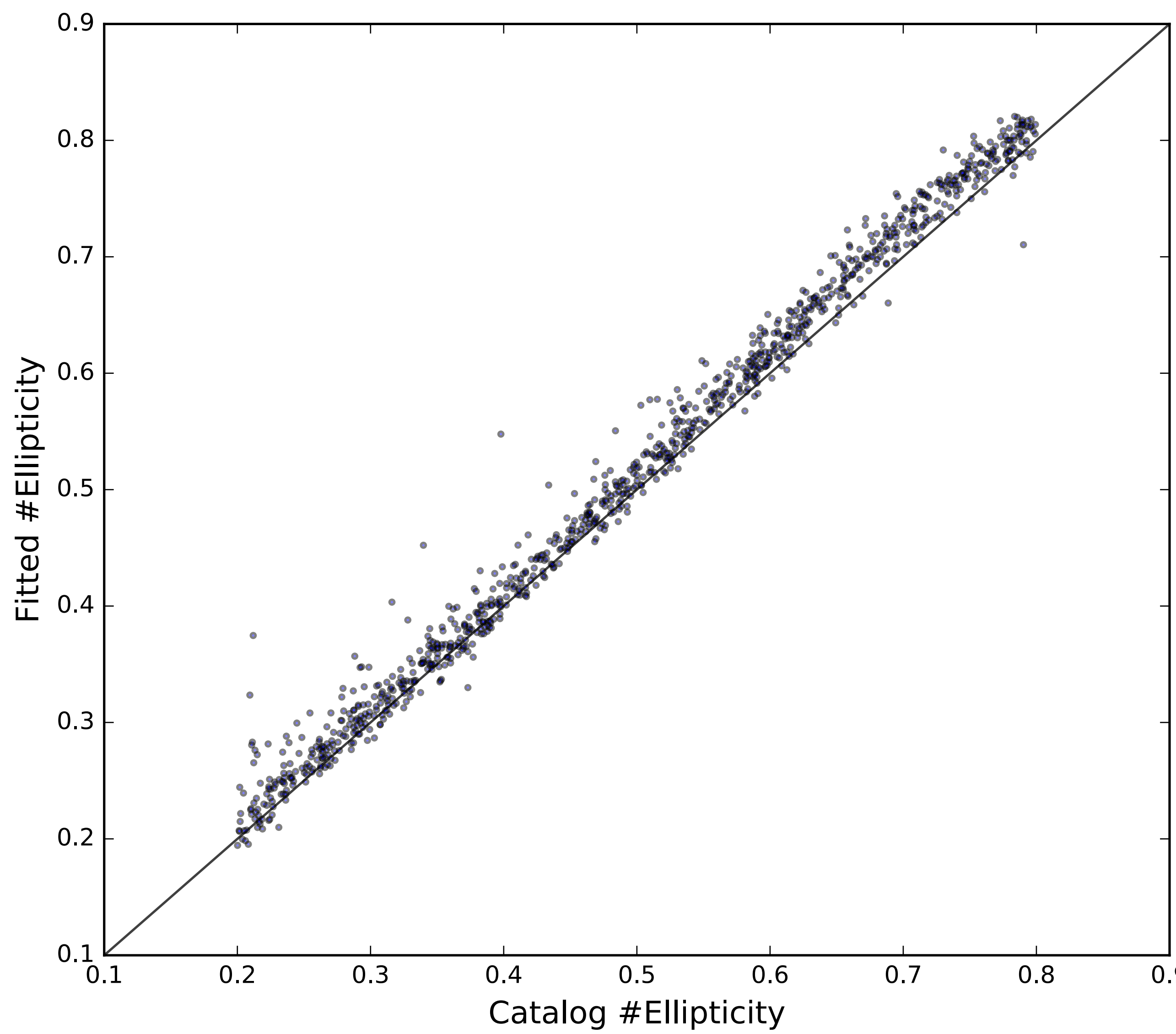

DNN

# Applied the result of learning on 1000 simulated galaxies

Ellipticity



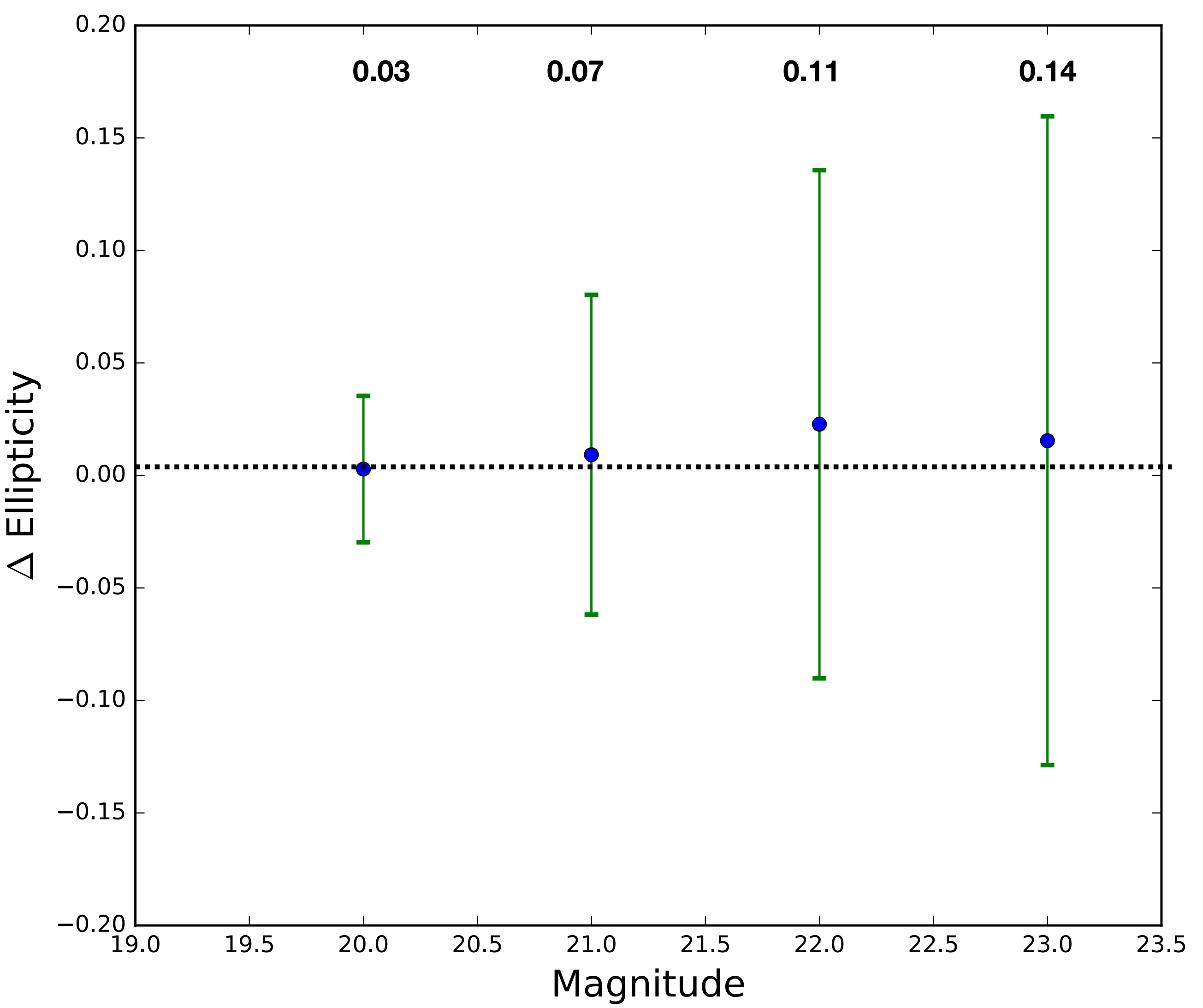

GALFIT-SExtractor



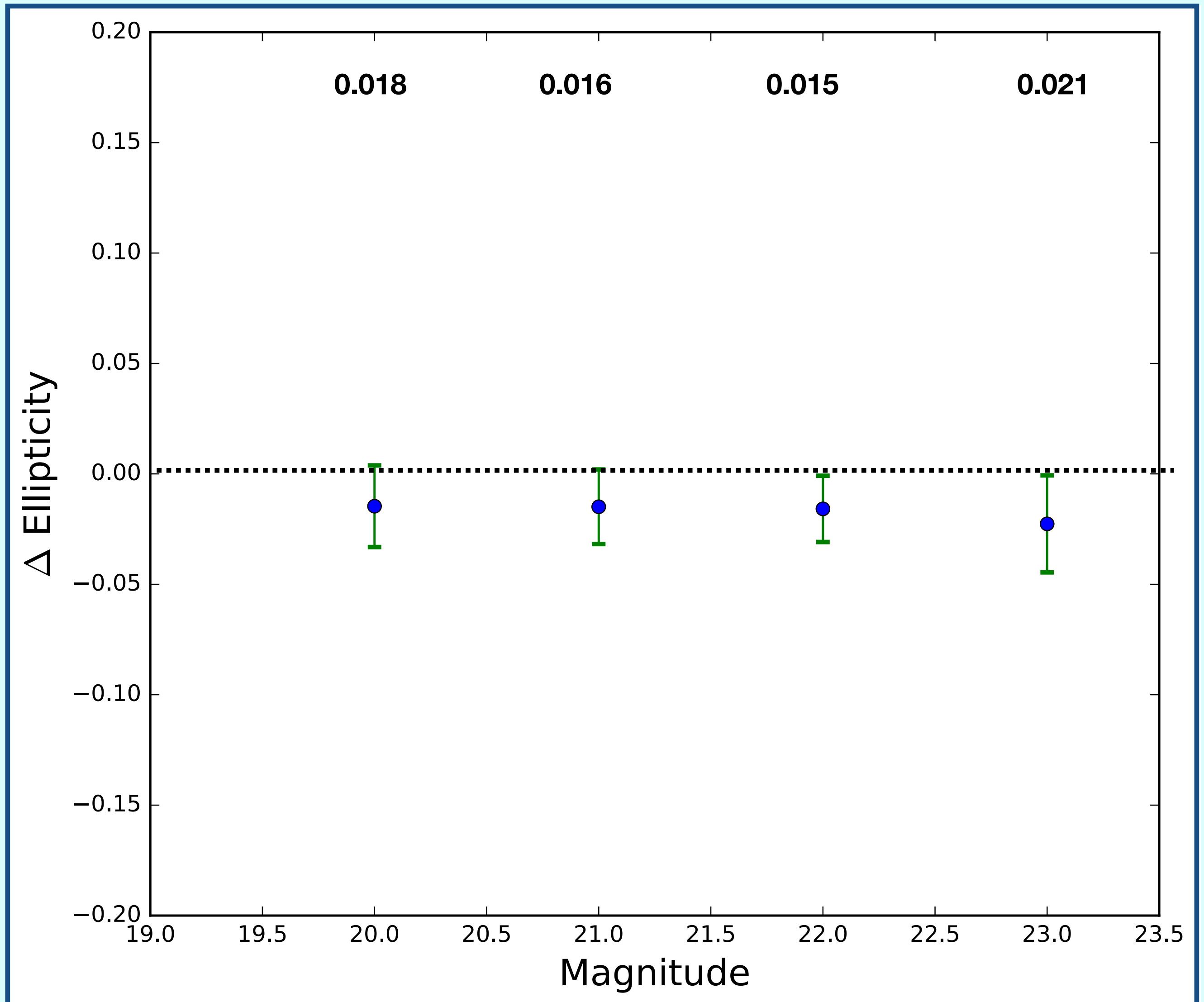

DNN

# Applied the result of learning on 1000 simulated galaxies

Ellipticity



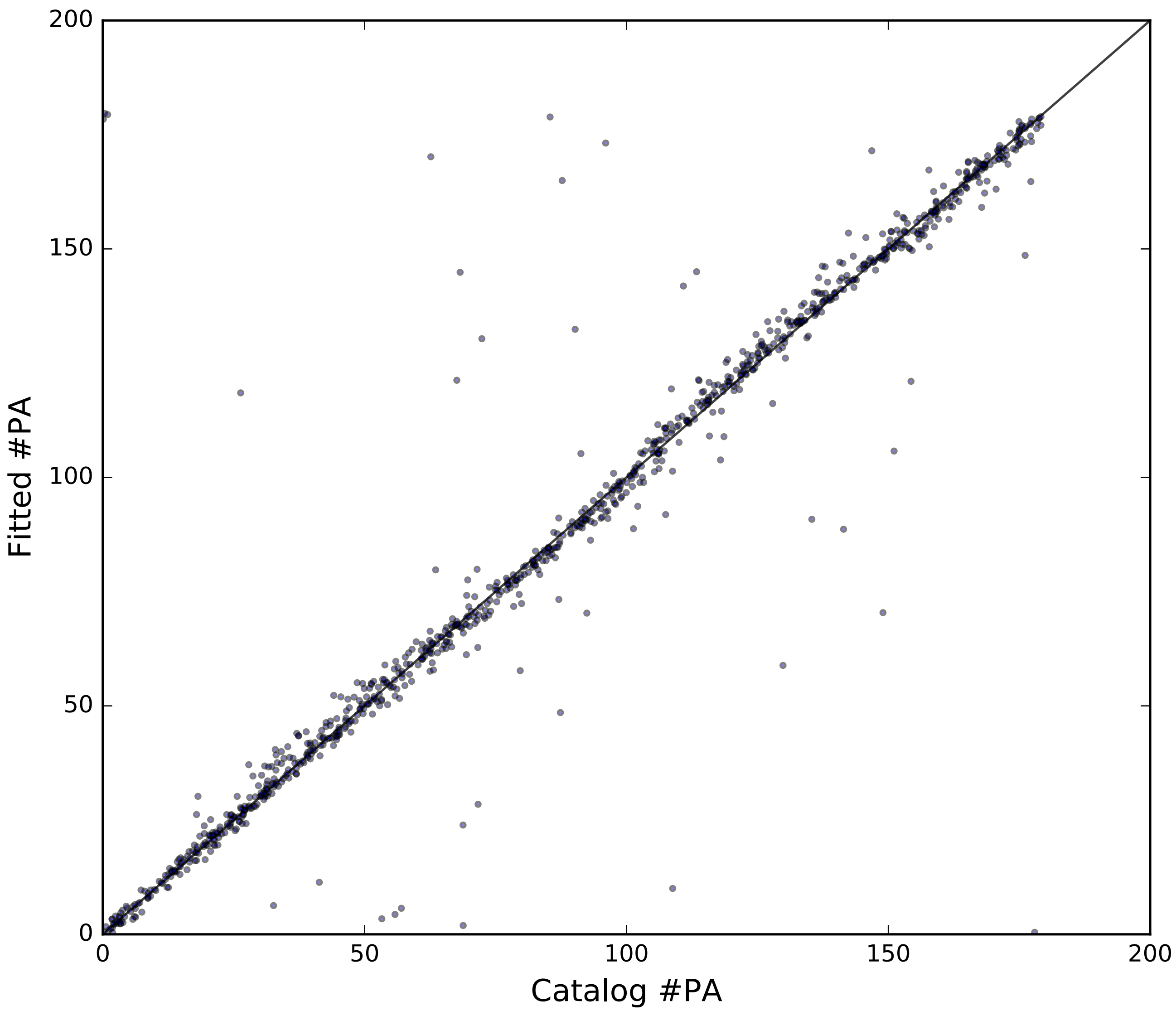

GALFIT-SExtractor



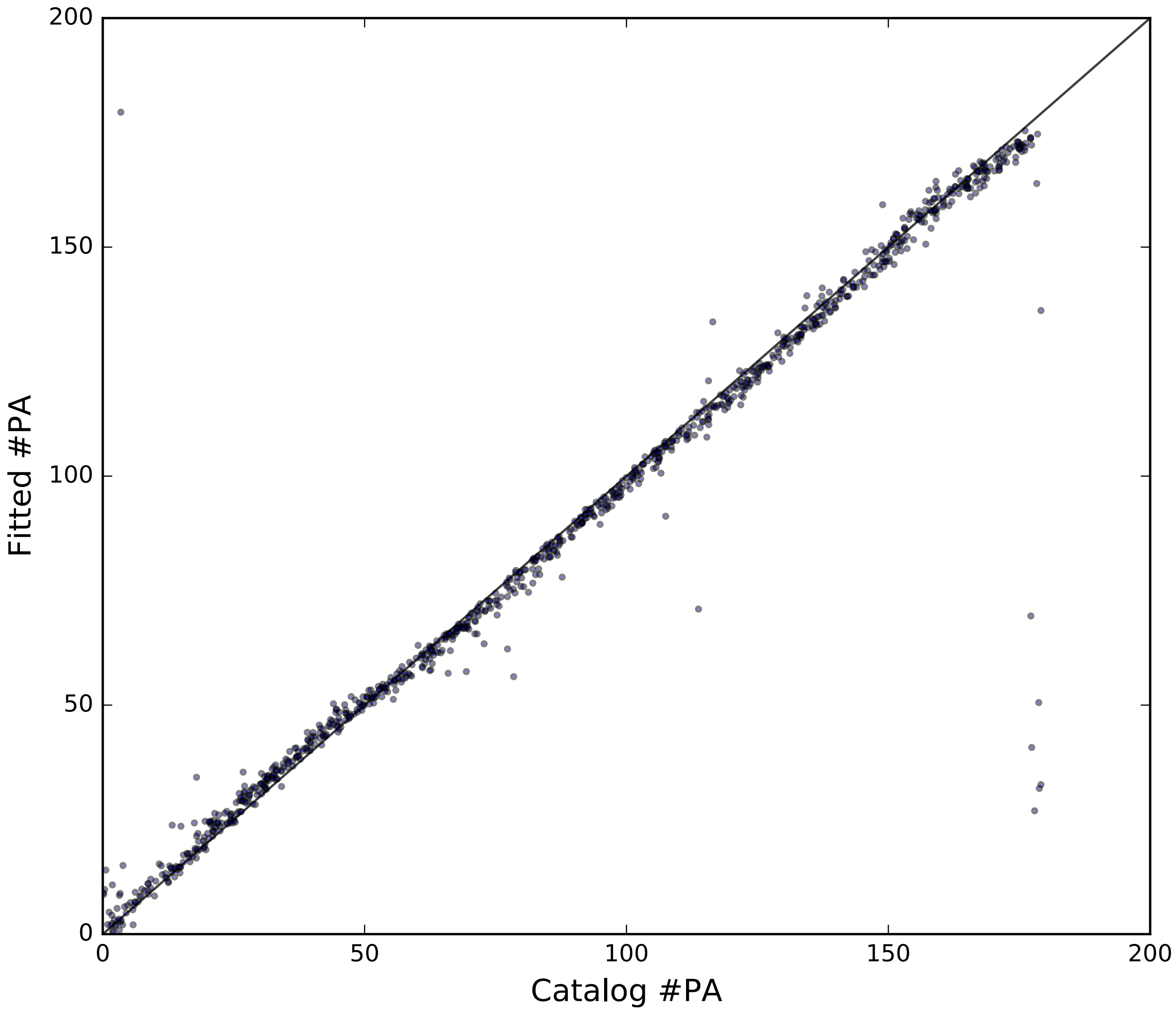

DNN

# Applied the result of learning on 1000 simulated galaxies

Position Angle



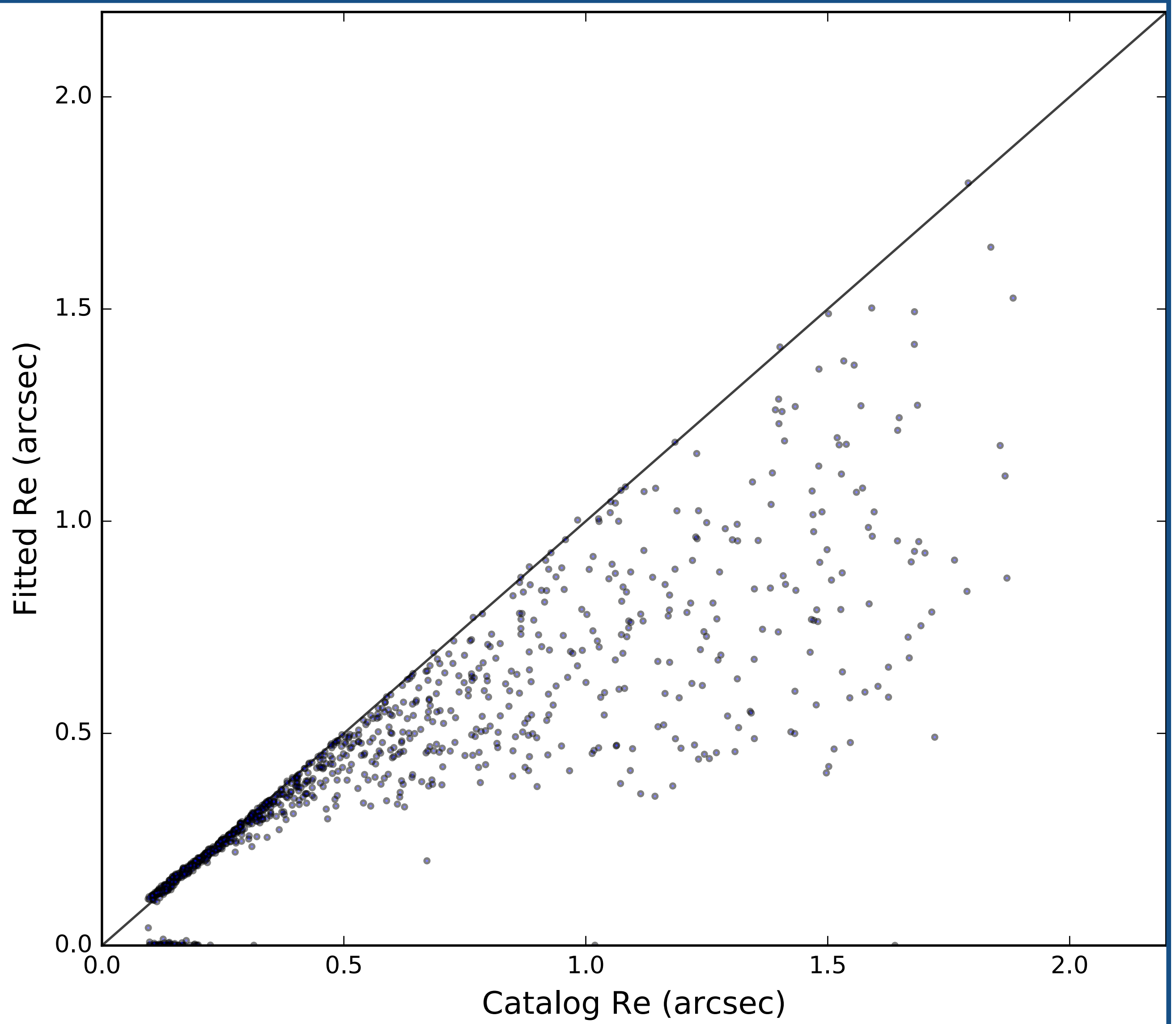

GALFIT-SExtractor



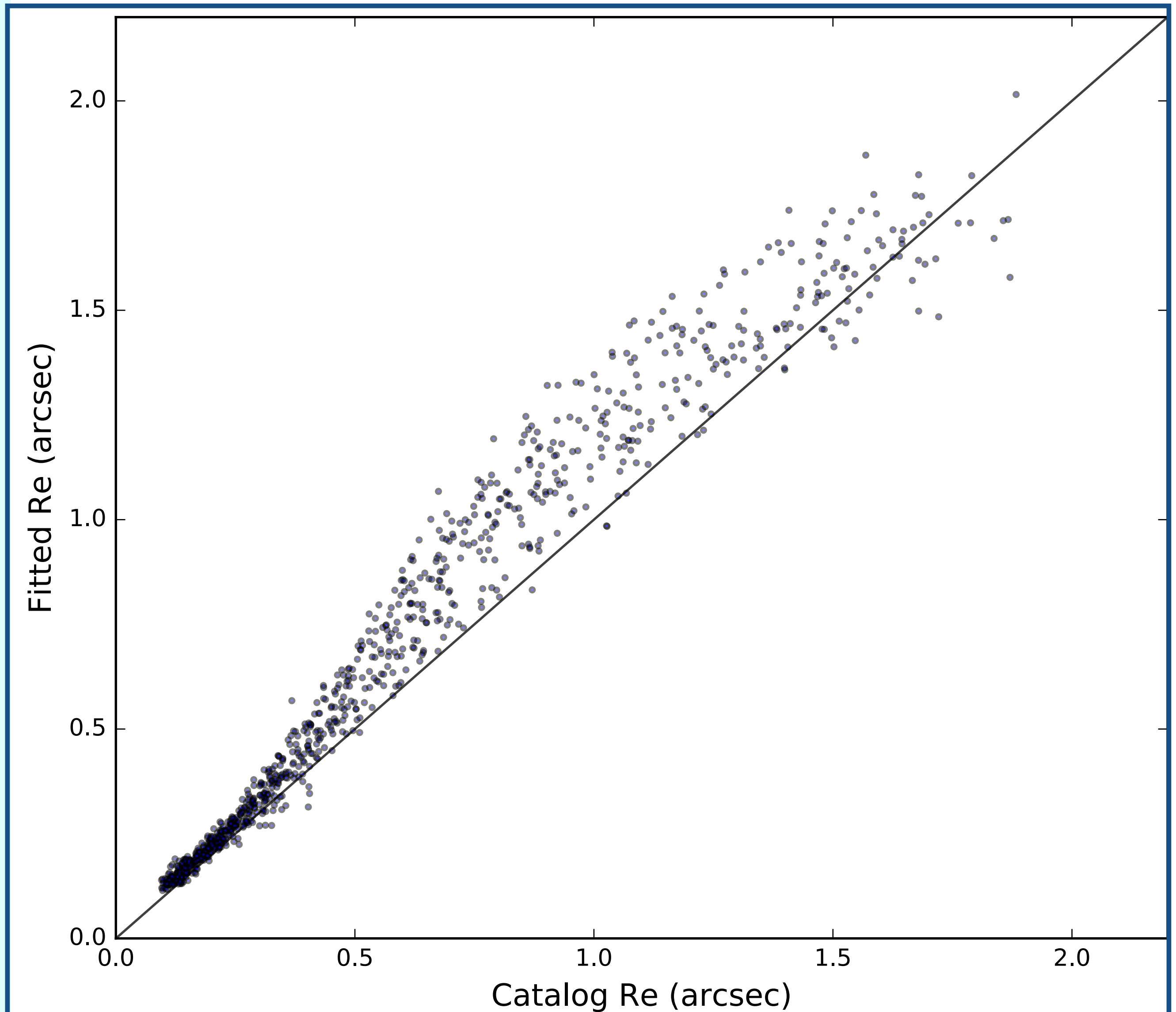

DNN

# Applied the result of learning on 1000 simulated galaxies

Radius



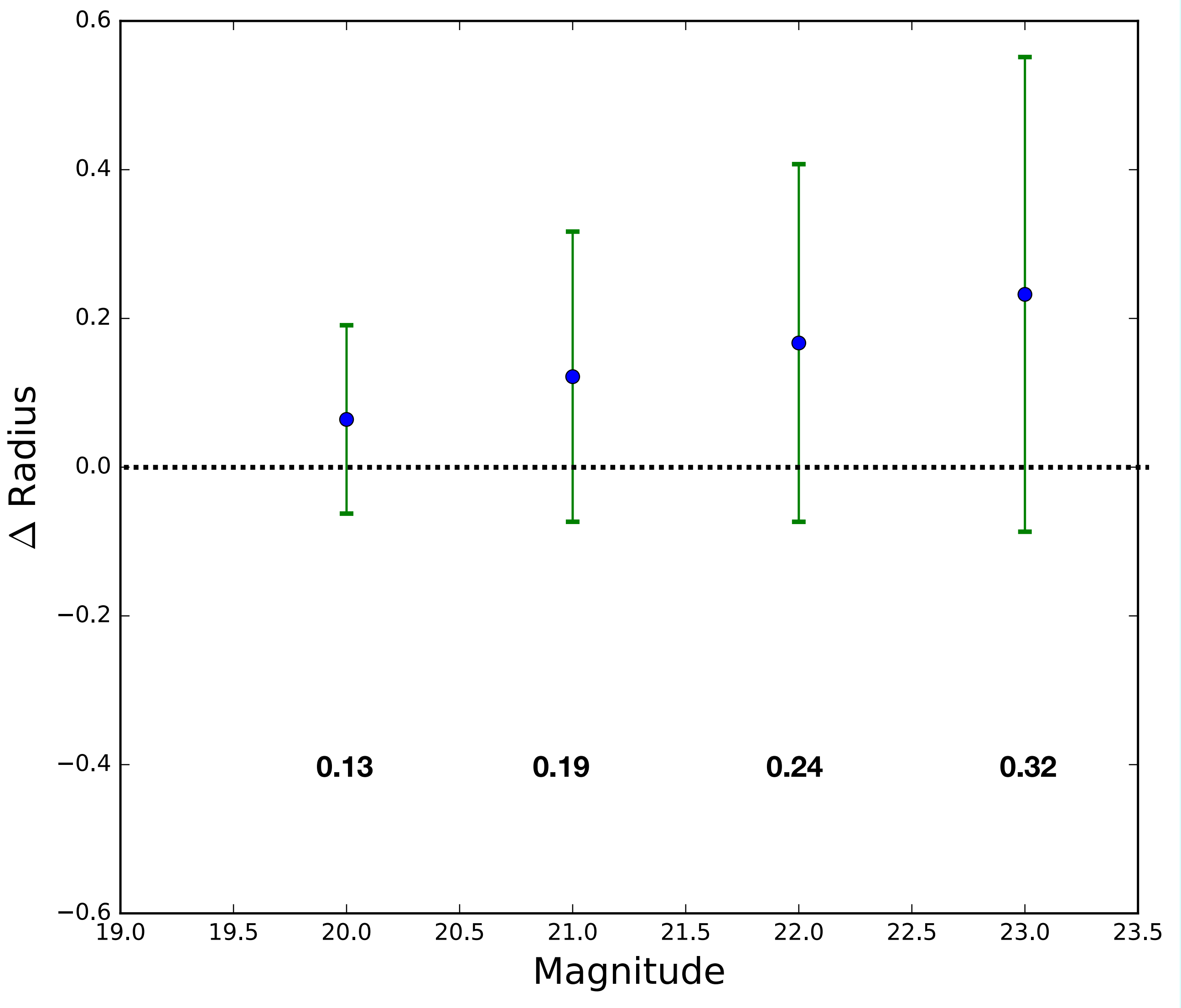

GALFIT-SExtractor



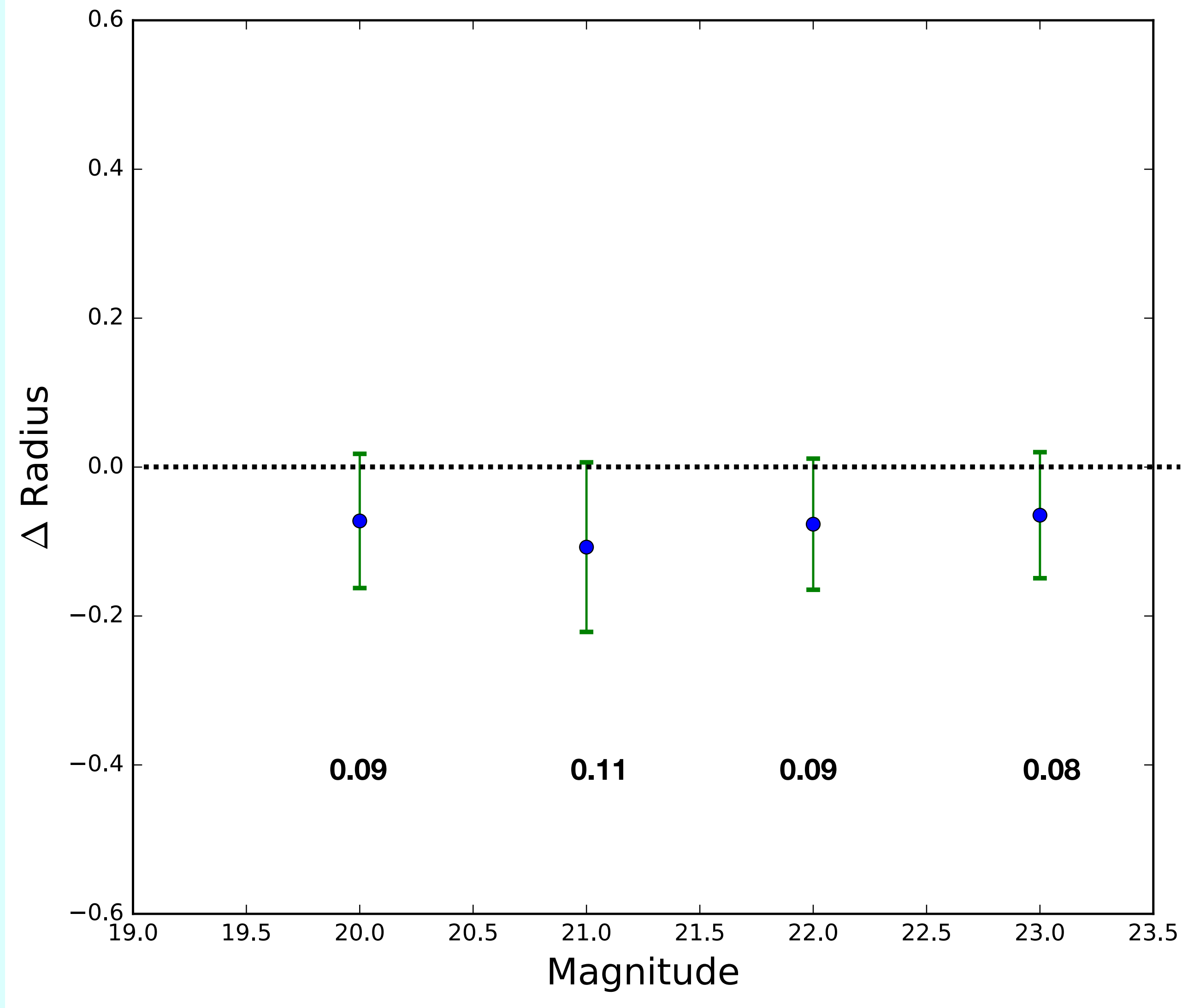

DNN

# Applied the result of learning on 1000 simulated galaxies

Radius



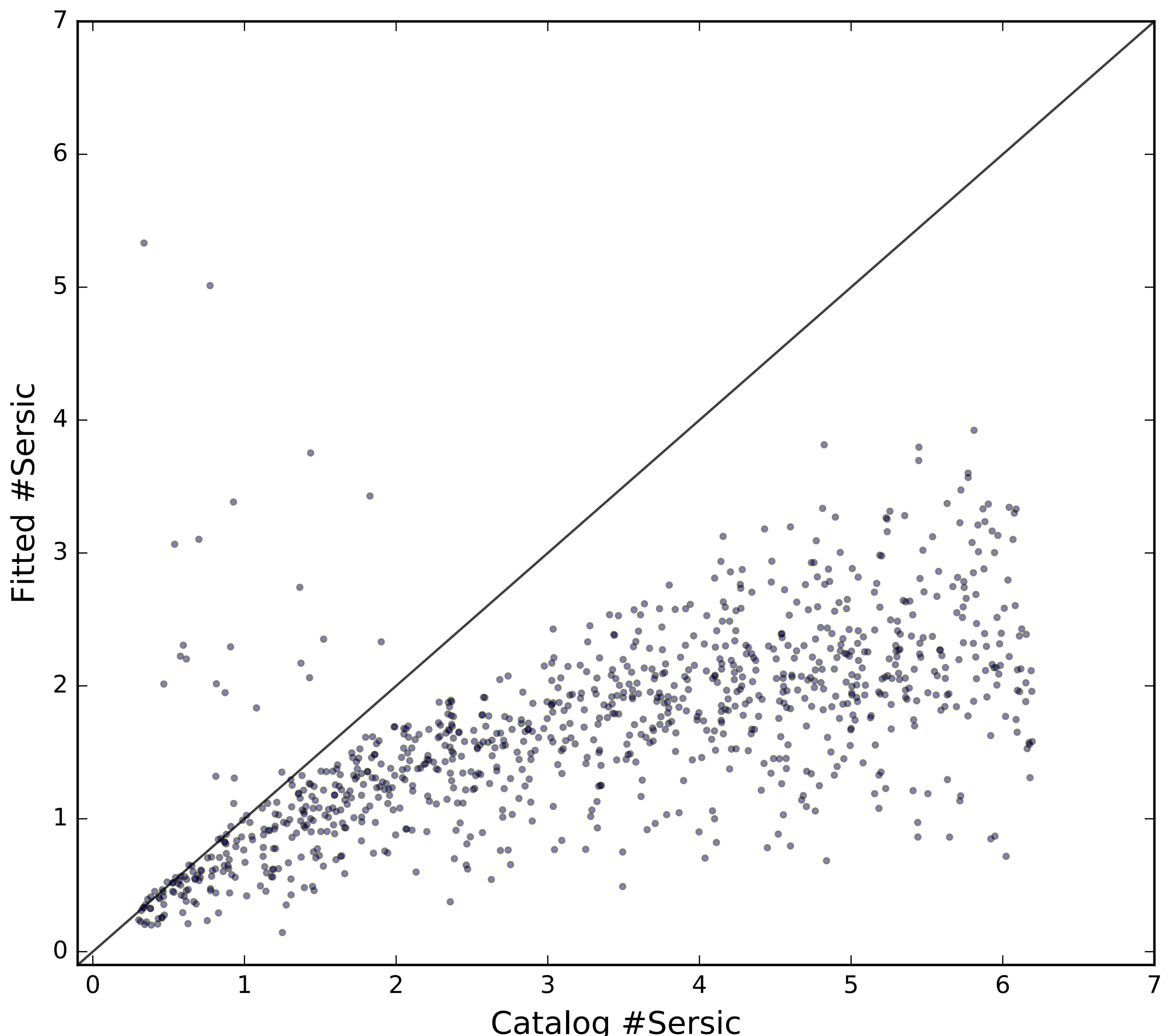

GALFIT-SExtractor



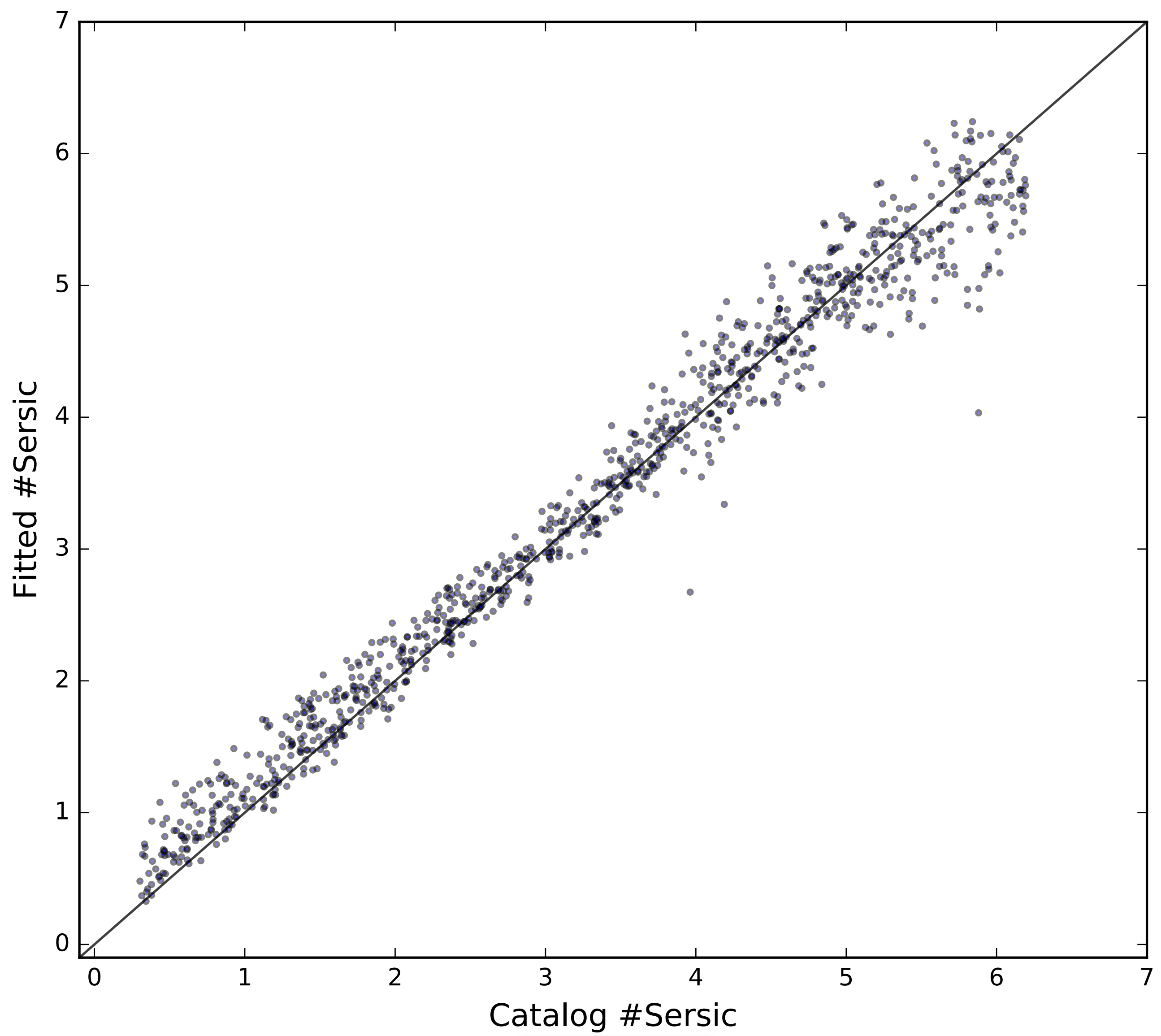

DNN

# Applied the result of learning on 1000 simulated galaxies

Sersic index



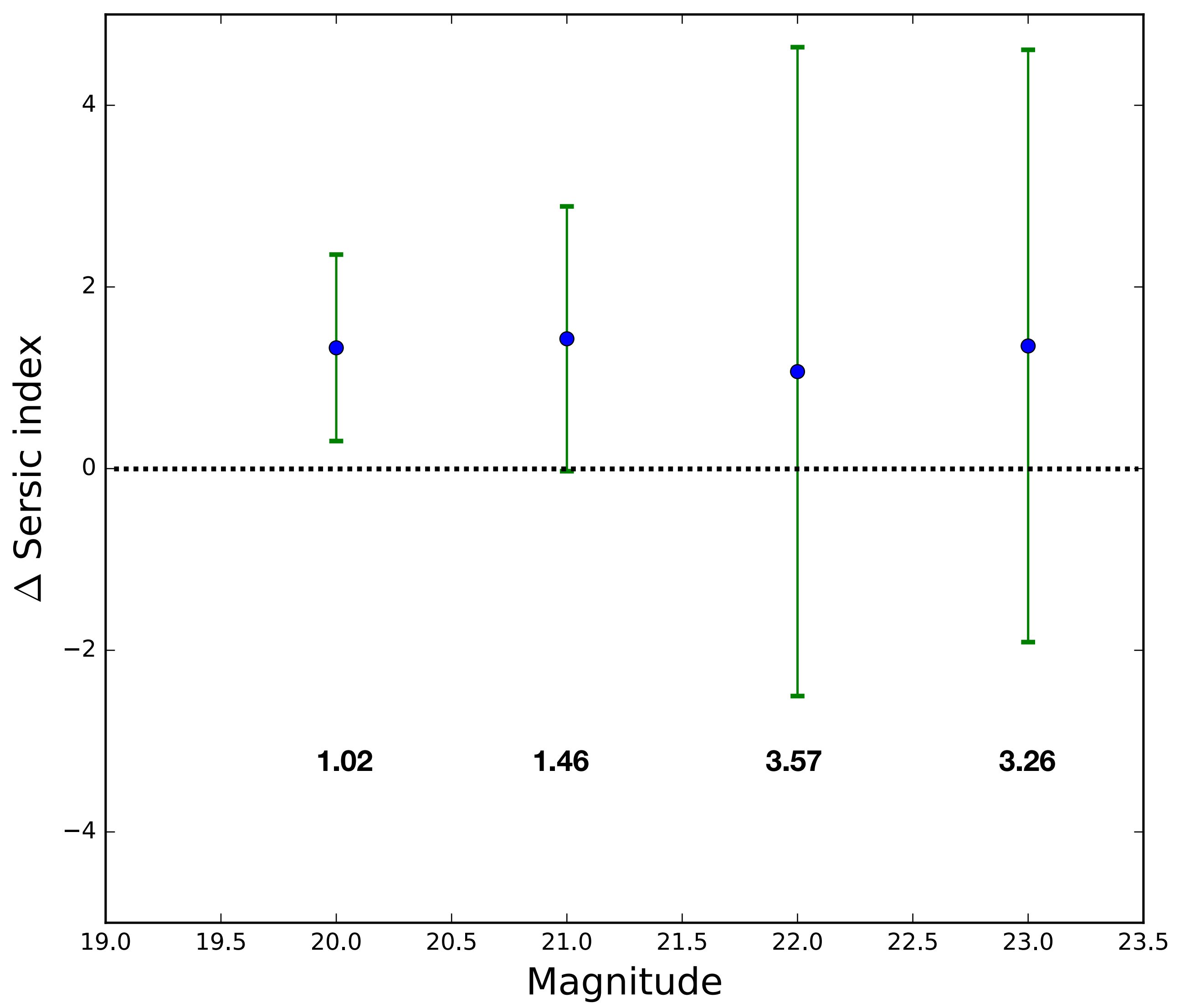

GALFIT-SExtractor



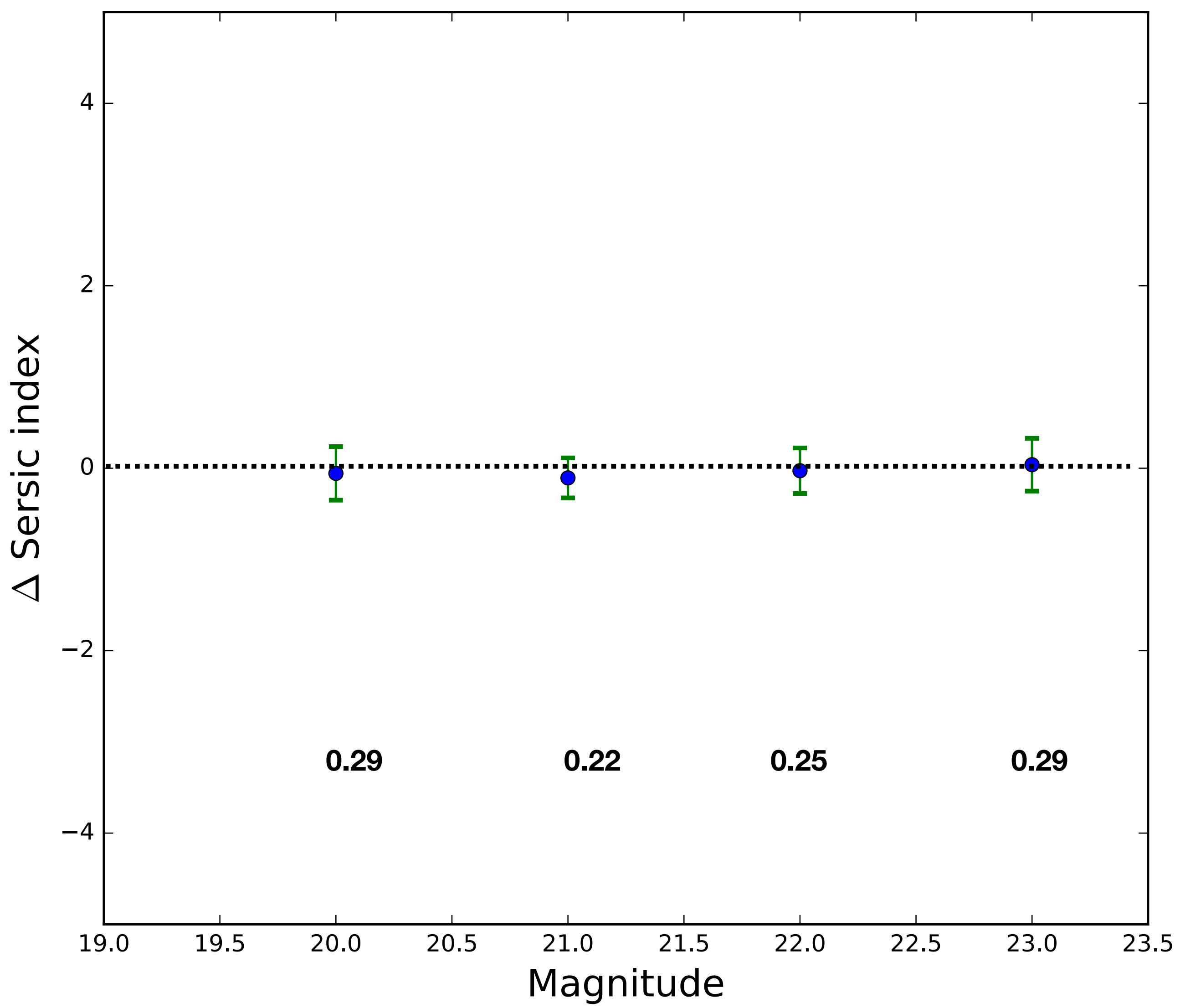

DNN

# Applied the result of learning on 1000 simulated galaxies

Sersic index



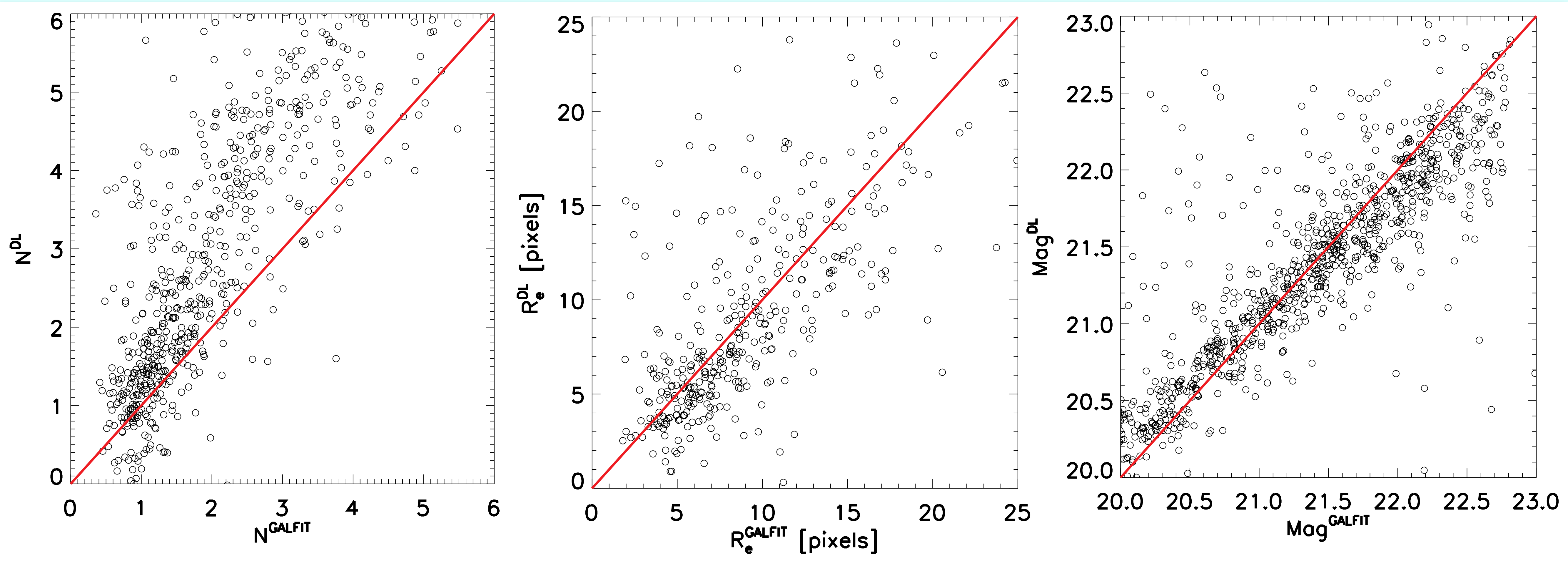

GALFIT-SExtractor




DNN

# Transfer learning

Applied what the DNN learned from simulation to 1000 **real HST galaxies z~1**




Sersic index

Radius

Mag

Parameters from "careful" GALFIT fitting

# *But, is it our motivation to improve results of old softwares?*

*….not really*

**Time**


~6000 sources fitted

MODEL FITTING WITH GALFIT/GALAPAGOS

Time: ~ 5 Hours

MODEL FITTING WITH SEXTRACTOR + PSFEX

Time: ~ 15 Minutes!! (do not fit neighbours)

# *But, is it our motivation to improve results of old softwares?*

OUR DNN

Takes ~2 minutes to train/validate 6000 sources
and, more important, once trained takes <5 seconds to fit 6000 sources

~6000
sources fitted

MODEL FITTING WITH GALFIT/GALAPAGOS

Time: ~ 5 Hours

MODEL FITTING WITH SEXTRACTOR + PSFEX

Time: ~ 15 Minutes!! (do not fit neighbours)

# But, is it our motivation to improve results of old softwares?

http://www.euclid-ec.org/?page_id=2581

| SURVEYS | | | | | |
|---|---|---|---|---|---|
| | Area (deg2) | Description | | | |
| Wide Survey | **15,000 deg$^2$** | Step and stare with 4 dither pointings per step. | | | |
| Deep Survey | **40 deg$^2$** | In at least 2 patches of > 10 deg$^2$ 2 magnitudes deeper than wide survey | | | |
| Wavelength range | 550– 900 nm | Y (920-1146nm), | J (1146-1372 nm) | H (1372-2000nm) | 1100-2000 nm |
| Sensitivity | 24.5 mag 10σ extended source | 24 mag 5σ point source | 24 mag 5σ point source | 24 mag 5σ point source | 3 $10^{-16}$ erg cm-2 s-1 3.5σ unresolved line flux |
| | Shapes + Photo-z of $n = 1.5\times10^9$ galaxies | | | $z$ of $n=5\times10^7$ galaxies | |

~6000 sources fitted

MODEL FITTING WITH GALFIT/GALAPAGOS

Time: ~ 5 Hours

MODEL FITTING WITH SEXTRACTOR + PSFEX

Time: ~ 15 Minutes!!

(do not fit neighbours)

# List of Tasks

**1) Morphological classification** ☑ DONE

***Dileman+14 Huertas-Company +15***

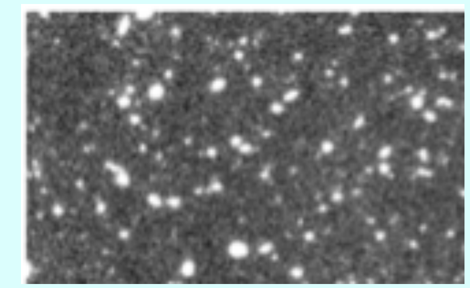
**2) DNN for 1 component Galaxies** ☑ DONE

**3) DNN for 2 component Galaxies**


**4) Improve transfer learning**

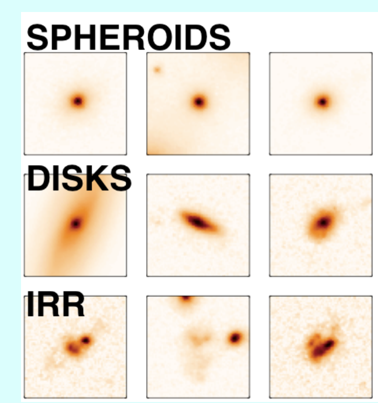
**5) Segmentation**

**6) A whole DNN machine:**



↓

Classification




↓ ↓ ↓

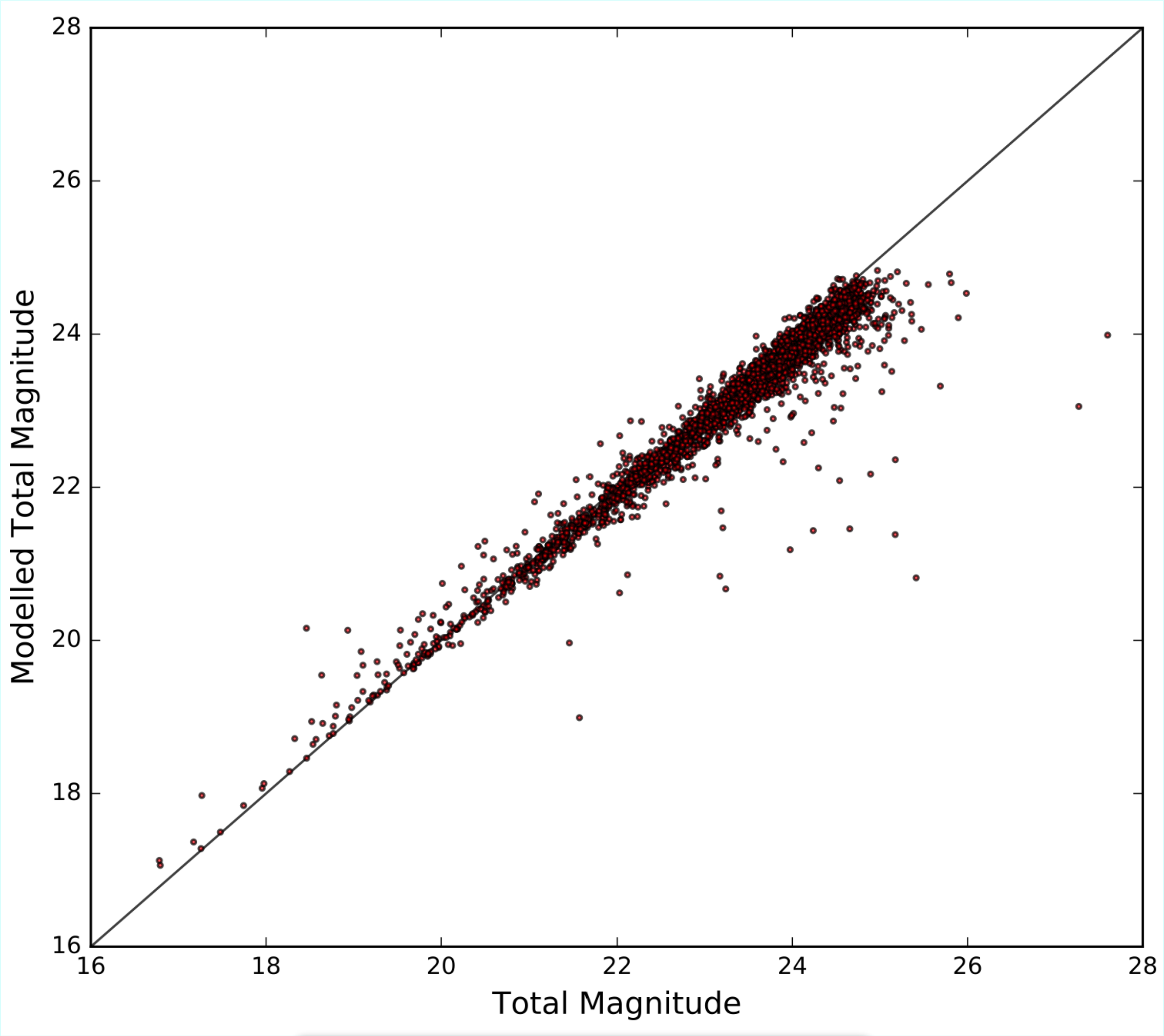
morphological parameters

# First results 2 components galaxies

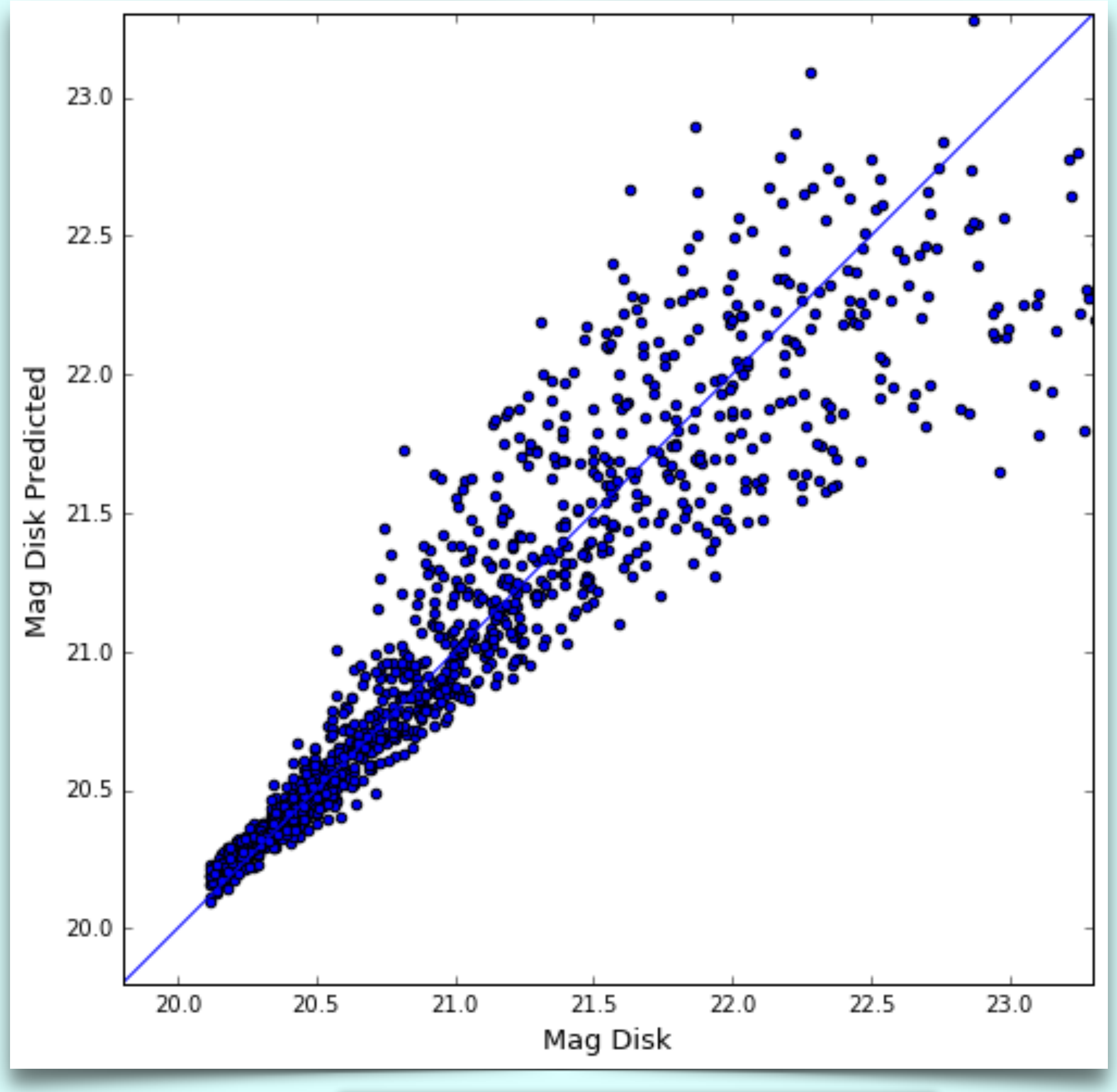
WITH REAL NOISE
REAL PSF CANDELS
PIXEL SCALE 0.06''

Double Sersic HST galaxies
VARIABLE BULGE/DISK RADIUS
VARIABLE B/T

18< MAG<24
VARIABLE SERCICS

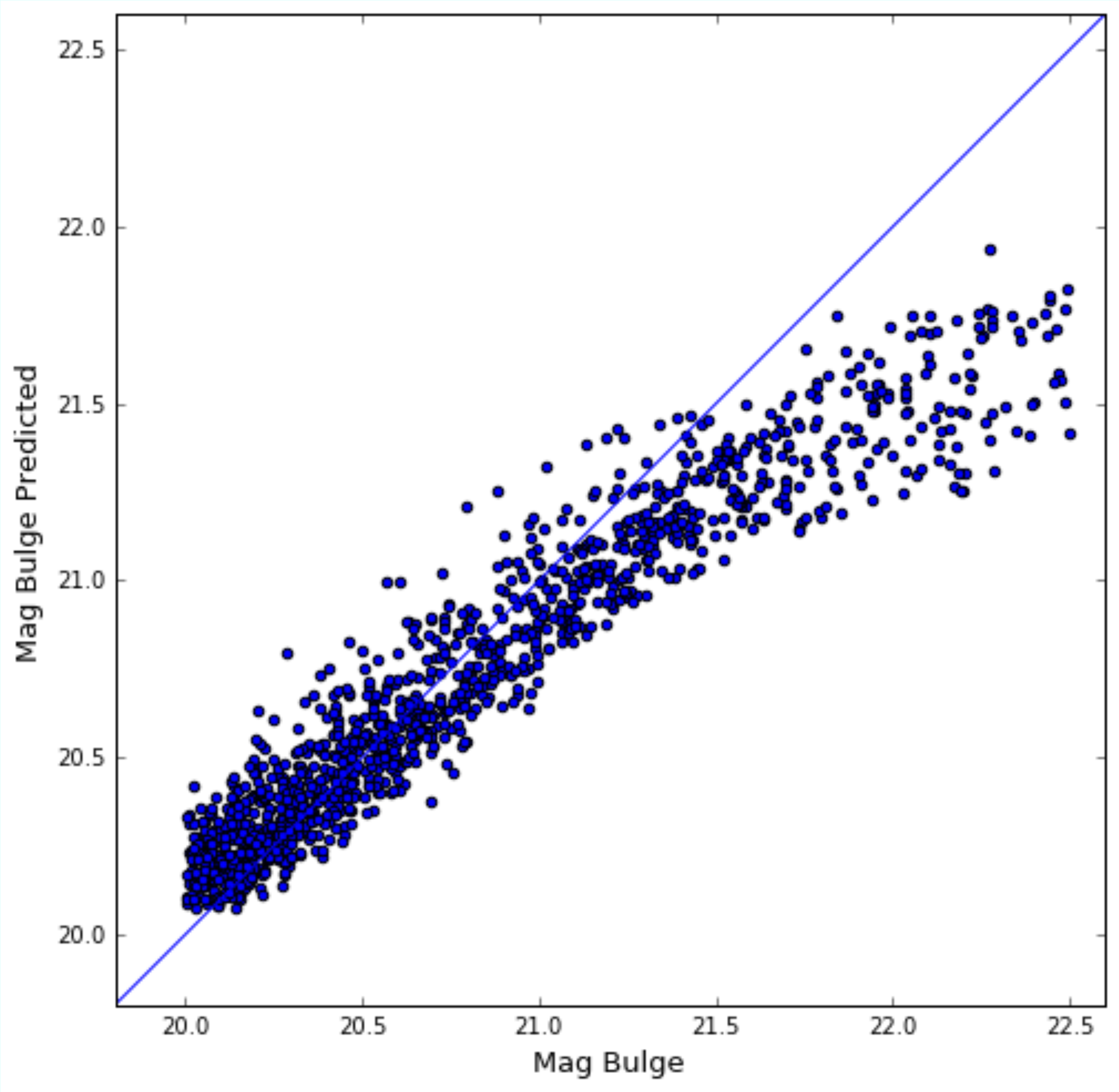


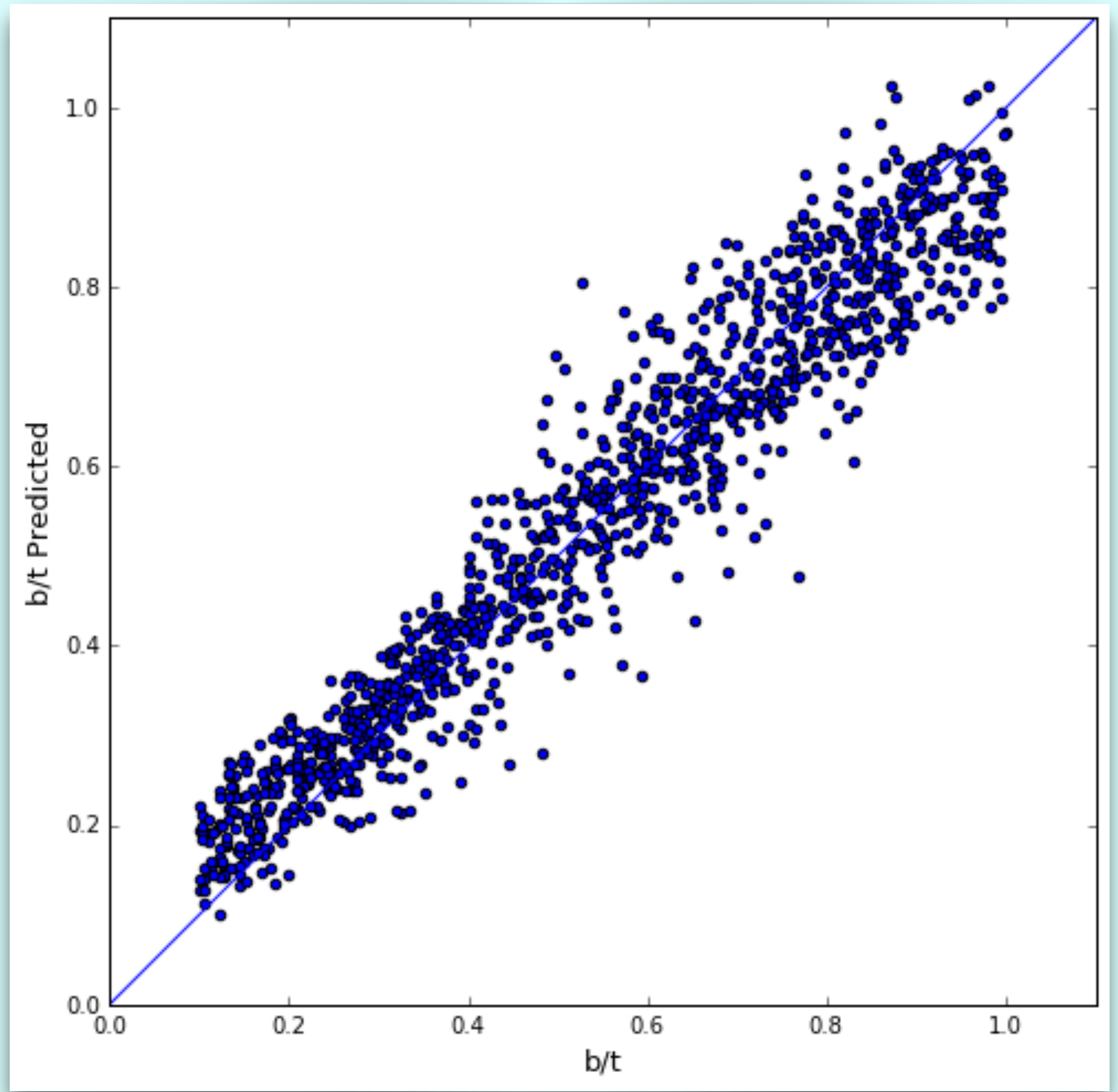

Total Magnitude




Disk Magnitude

# First results 2 components galaxies

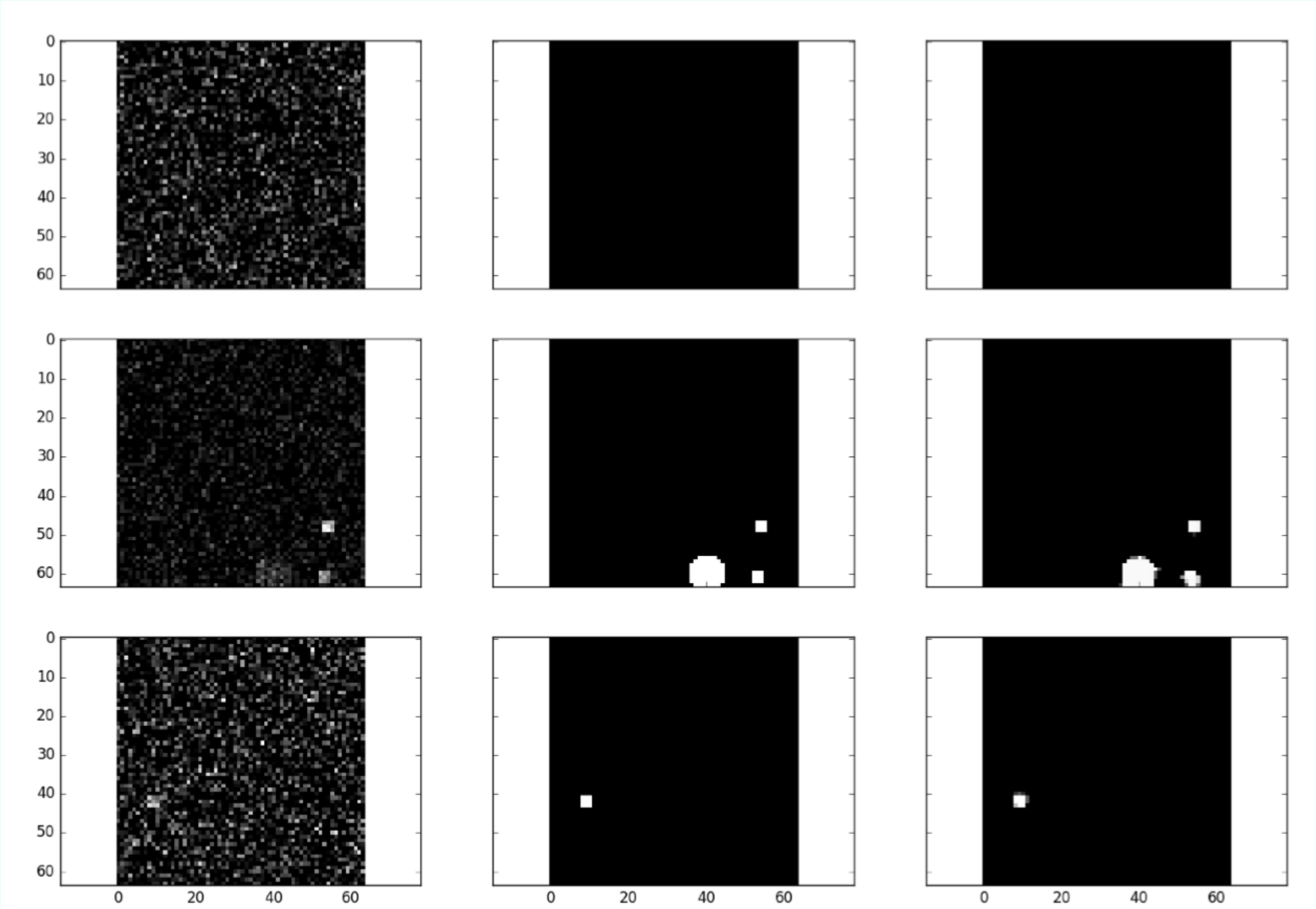



Bulge magnitude

B/T

# First results on segmentation




variable size

variable noise

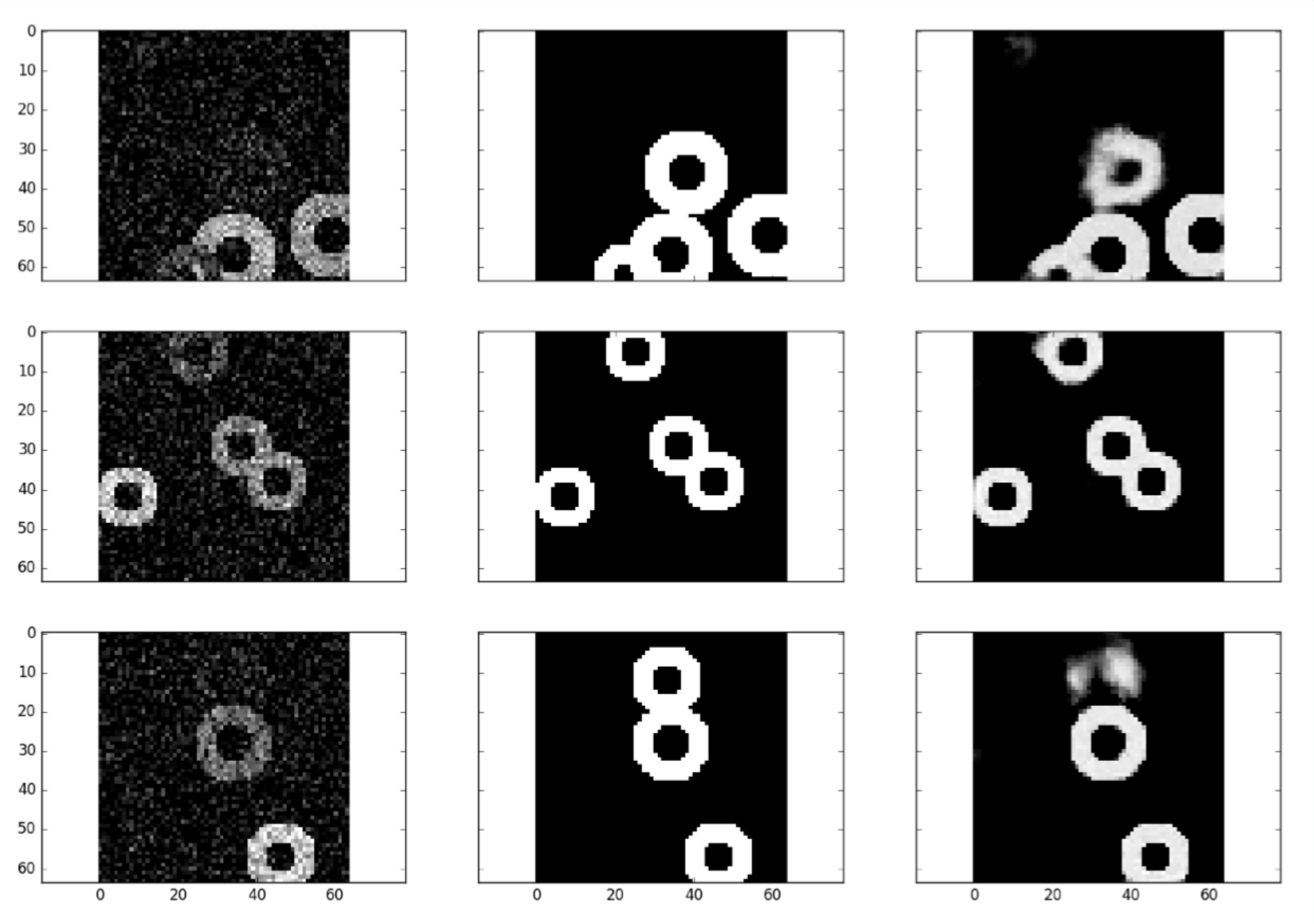
Results: disks

Fully convolutional networks for semantic segmentation [Long et al., 2015]

U-Net: convolutional networks for biomedical image segmentation (Ronneberger [Ronneberger et al., 2015]

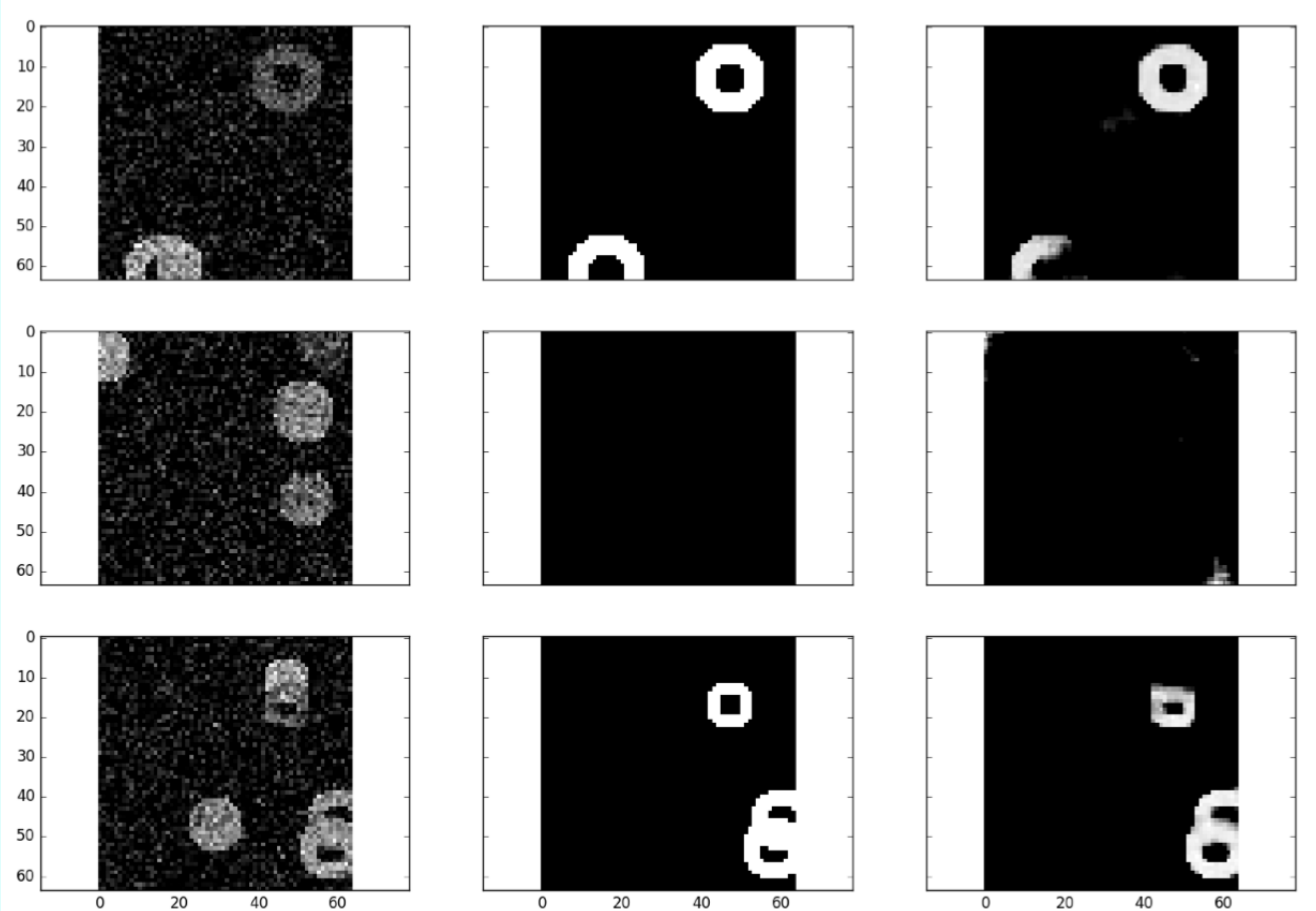
SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation [Badrinarayanan et al., 2015]

# *First results on segmentation*




Results: rings

# *First results on segmentation*




Results: rings and disks → task: find rings

# List of Tasks

1) *Morphological classification* DONE

*Dileman+14 Huertas-Company +15*

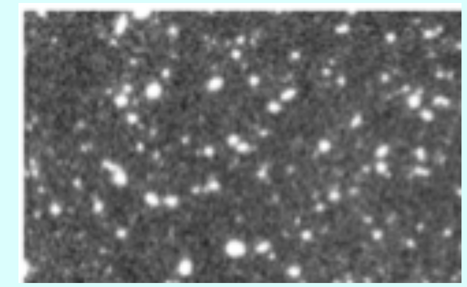
2) *DNN for 1 component Galaxies* DONE

3) *DNN for 2 component Galaxies*

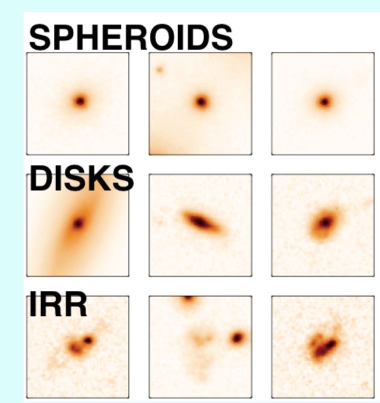

4) *Improve transfer learning*

5) *Segmentation*

6) *A whole DNN machine:*



Classification




morphological parameters

# List of Tasks

***1) Morphological classification*** DONE

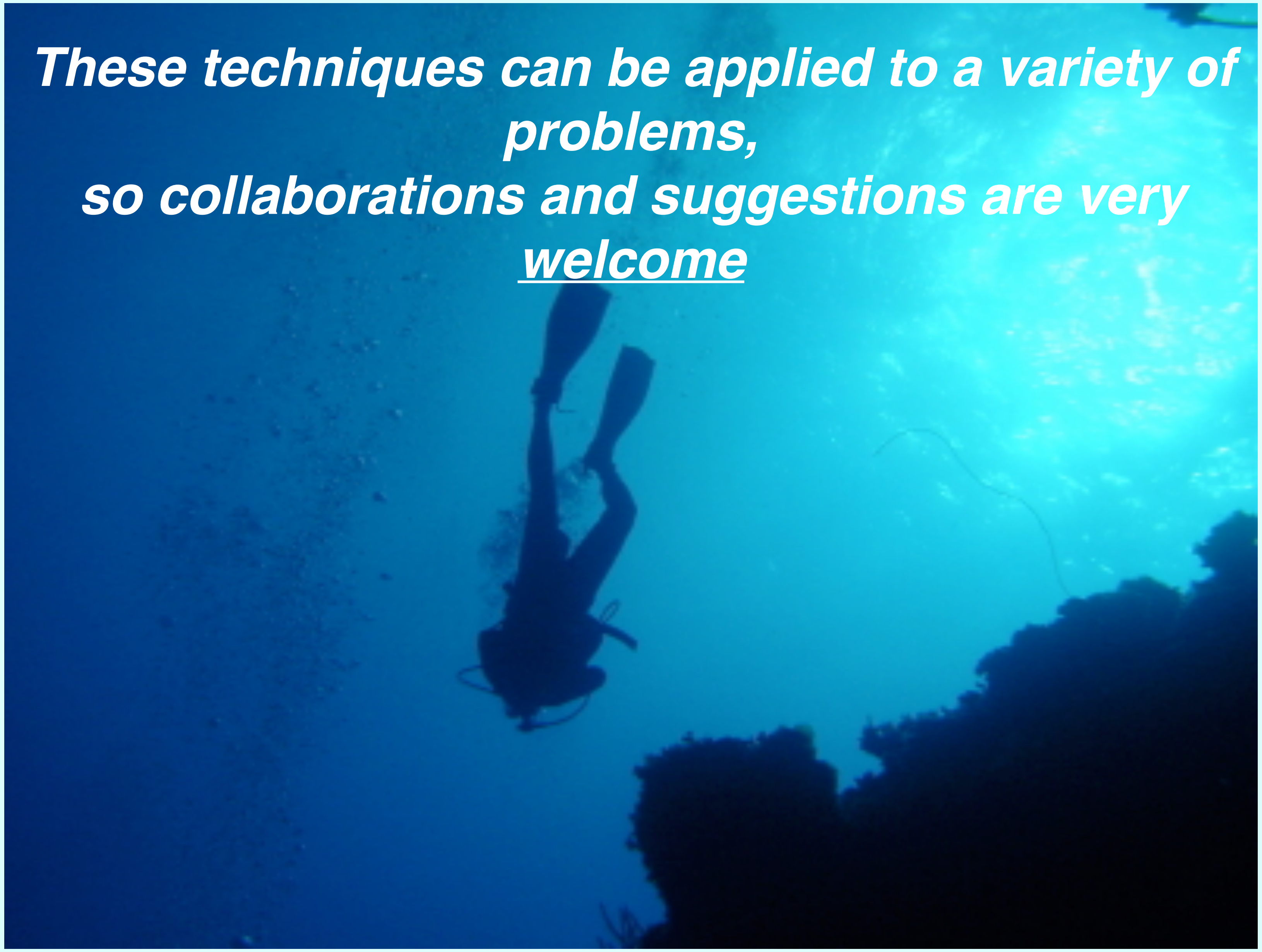
*Dileman+14 Huertas-Company +15*

***2) DNN for 1 component Galaxies*** DONE

***3) DNN for 2 component Galaxies*** WORK IN PROGRESS

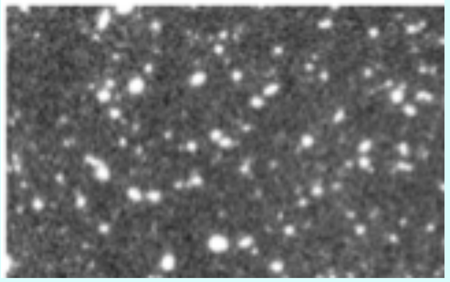
***4) Improve transfer learning*** WORK IN PROGRESS

***5) Segmentation*** WORK IN PROGRESS

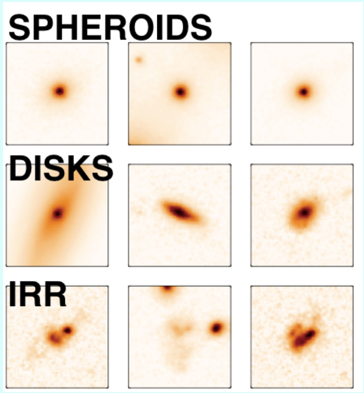


***These techniques can be applied to a variety of problems, so collaborations and suggestions are very welcome***

***6) A whole DNN machine:***



Classification



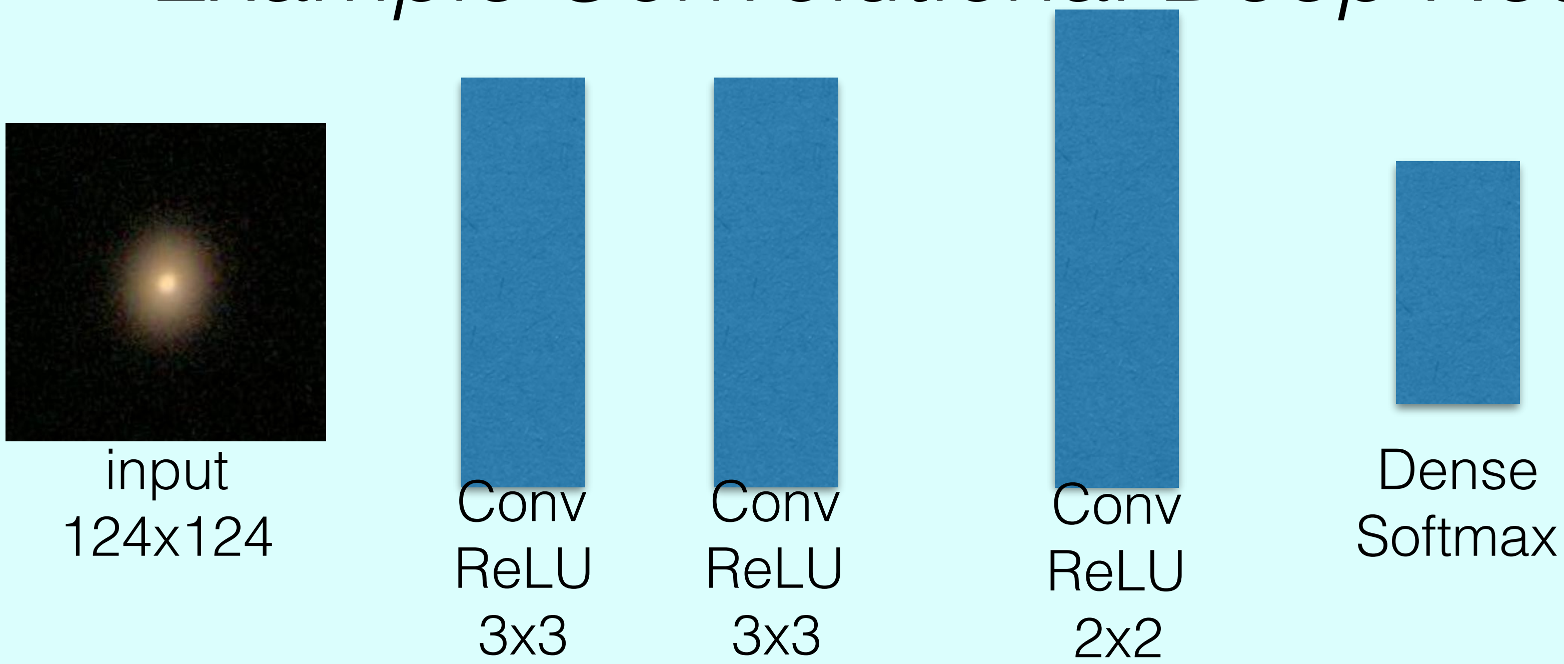

morphological parameters

# Thanks for listening!

# Thanks for listening!

# We developed a DNN algorithm.
# Used 30,000 stamps of 1-component galaxies

*Example Convolutional Deep Neural Network*




Trained with 24,000

Validated with 6000

***Details on the true DNN architecture***

*(to be submitted by November)*

**Deep Neural Network for galaxy morphology**

D. Tuccillo, M. Huertas, E. Decenciere, S.

l'Observatoire de Paris

MINES ParisTech