



# Centro Brasileiro de Pesquisas Físicas



## Deep Learning aplicado a Astronomia e Astrofísica

Bernardo Fraga - [bernardo@cbpf.br](mailto:bernardo@cbpf.br)

Clecio R. Bom

*Oficina de Instrumentação e Inovação*

*O2I - CBPF, 16/11/2022*

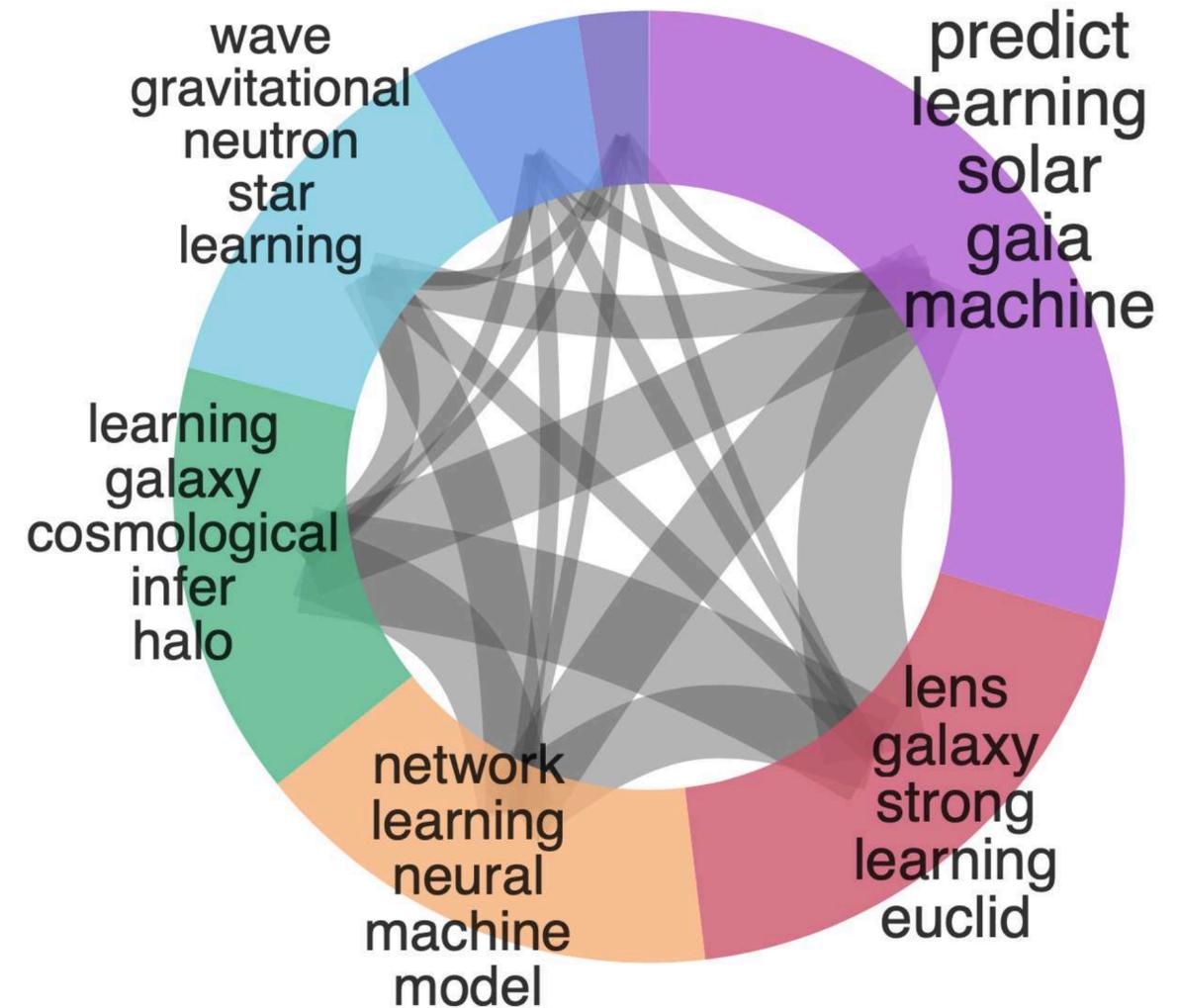
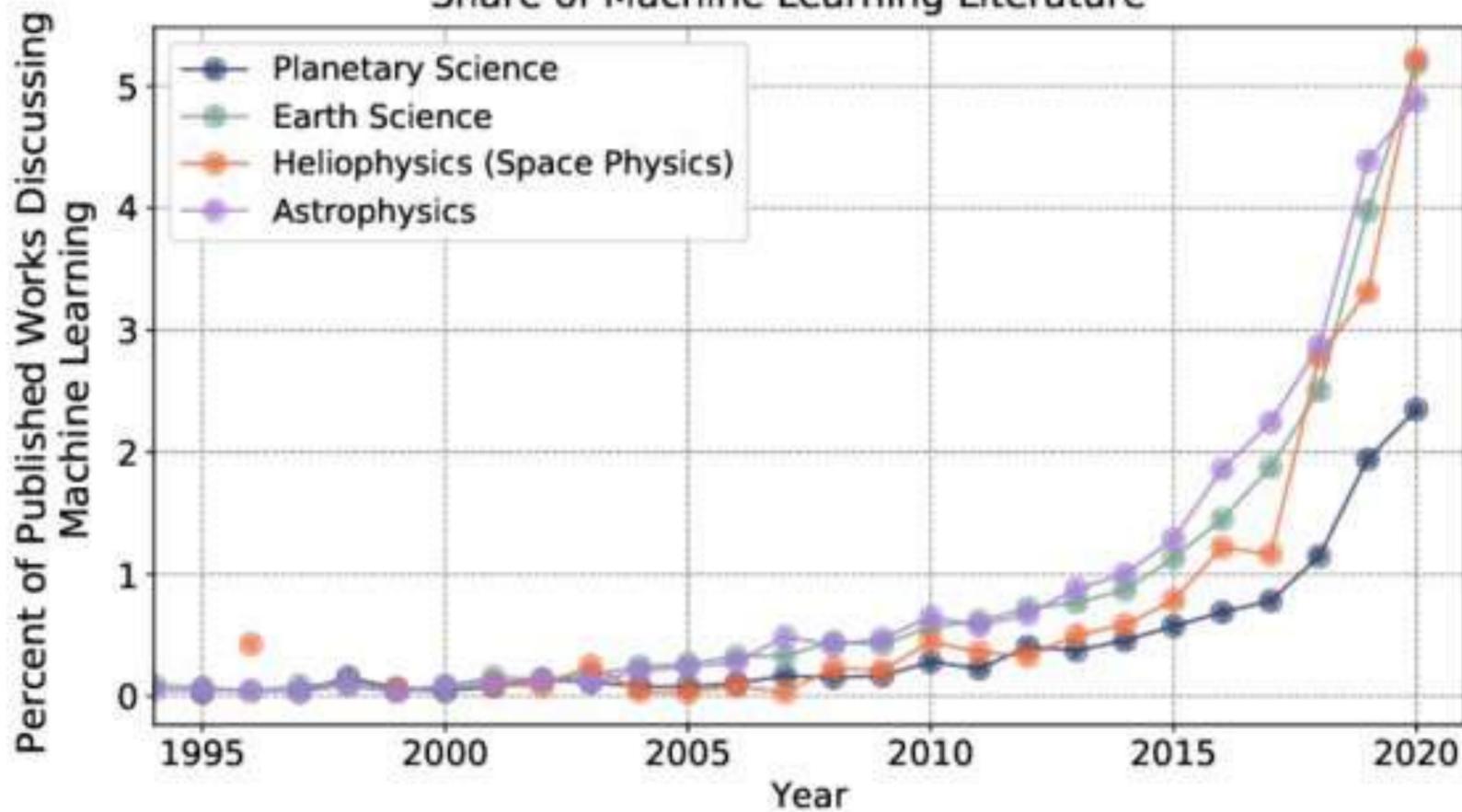


# ML e DL em A&A

( abs:"neural network" OR abs:"deep learning" OR abs:"machine learning") AND year:2022

Your search returned **1,901** results

NASA Science Directorate Topics  
Share of Machine Learning Literature



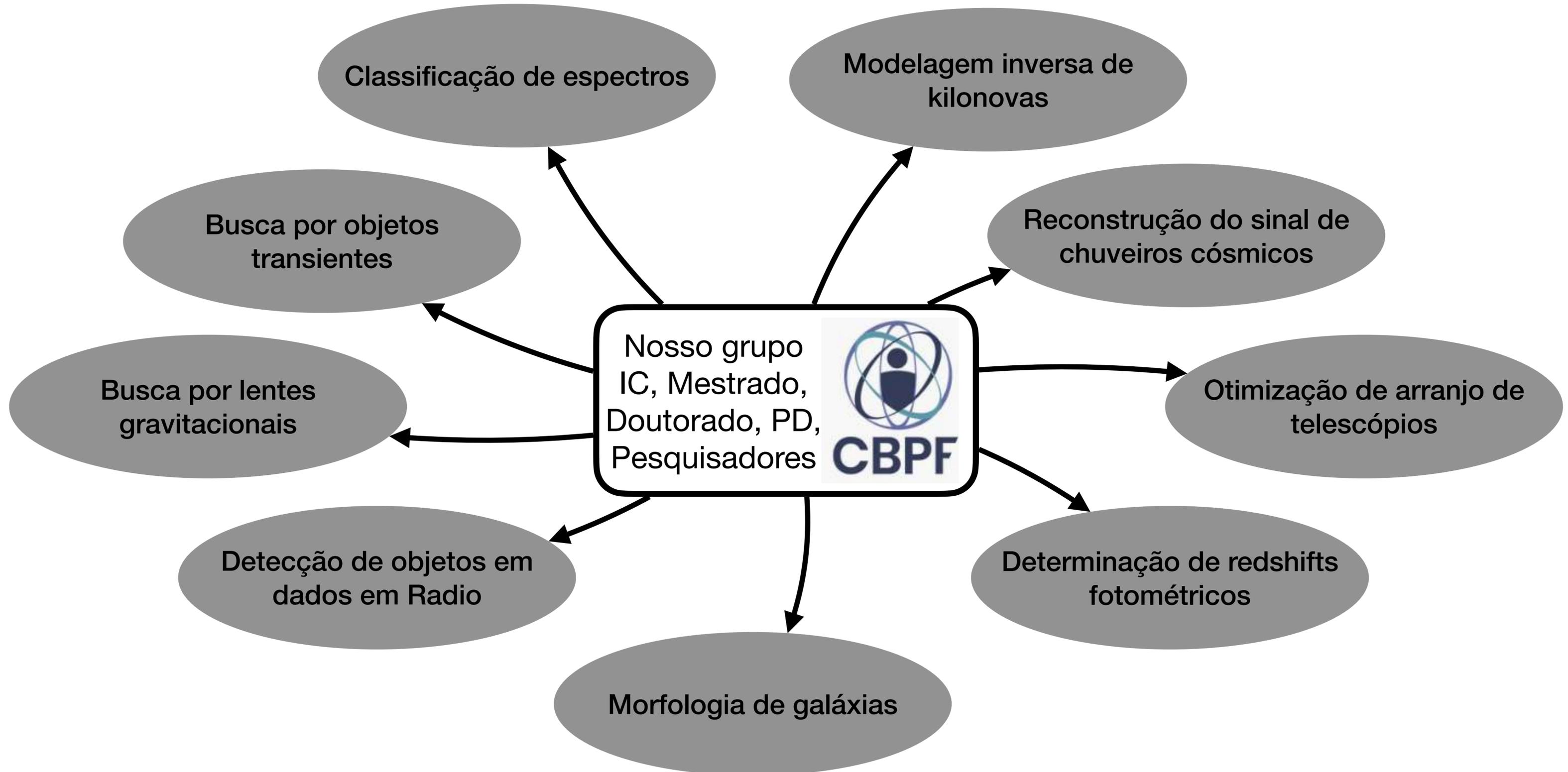
# Astrofísica e DL no CBPF

# Astrofísica e DL no CBPF

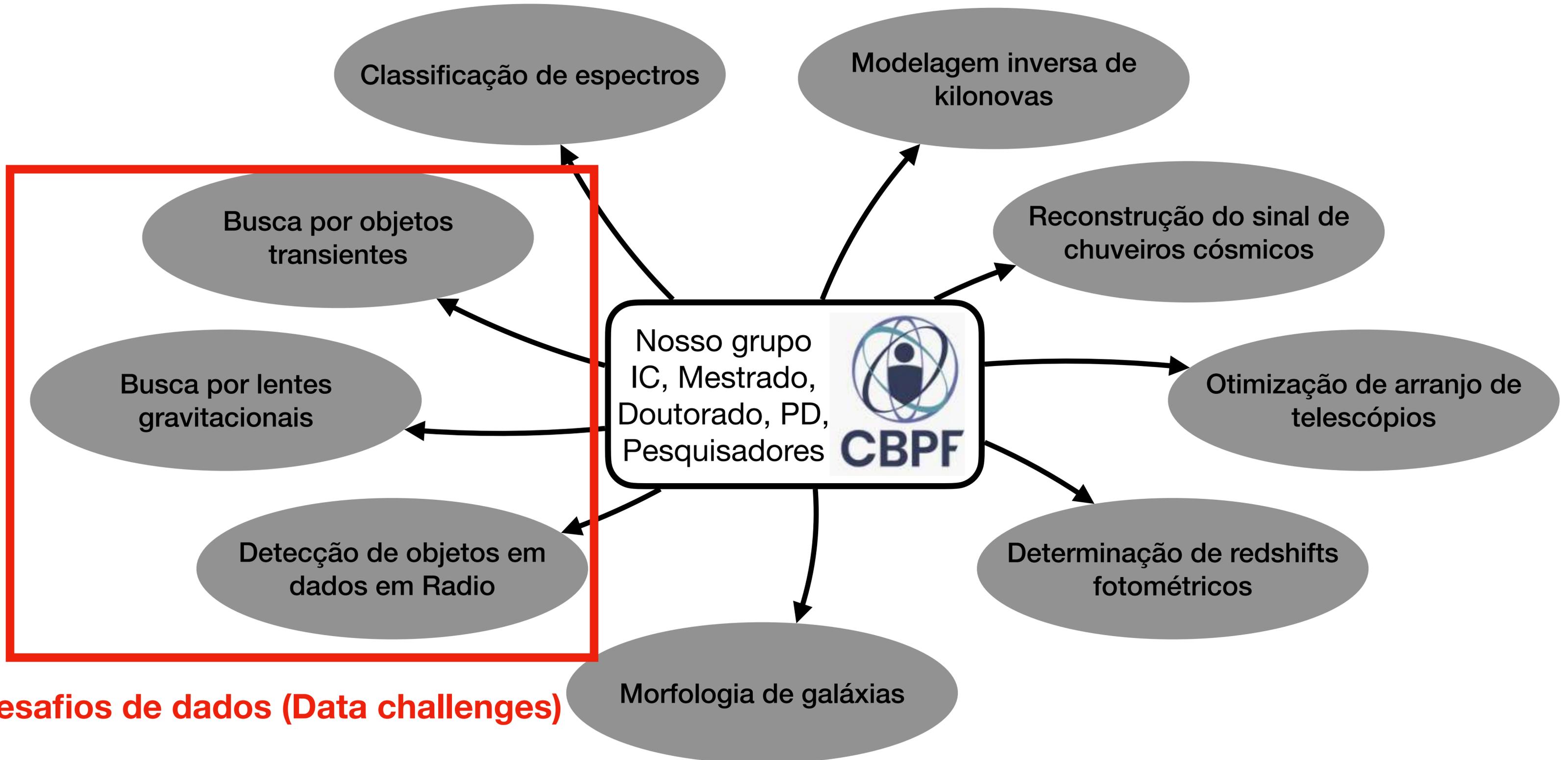
Nosso grupo  
IC, Mestrado,  
Doutorado, PD,  
Pesquisadores



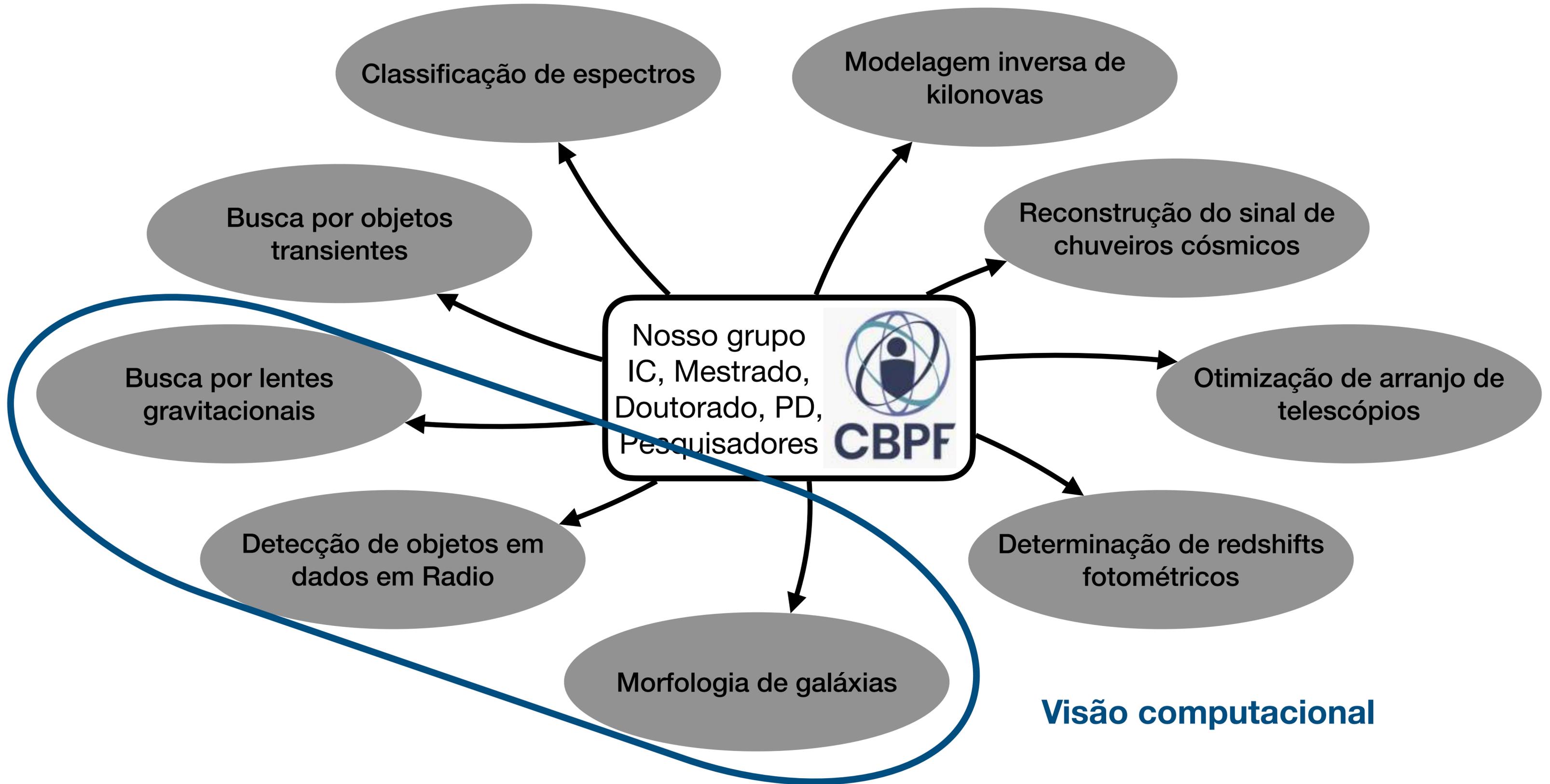
# Astrofísica e DL no CBPF



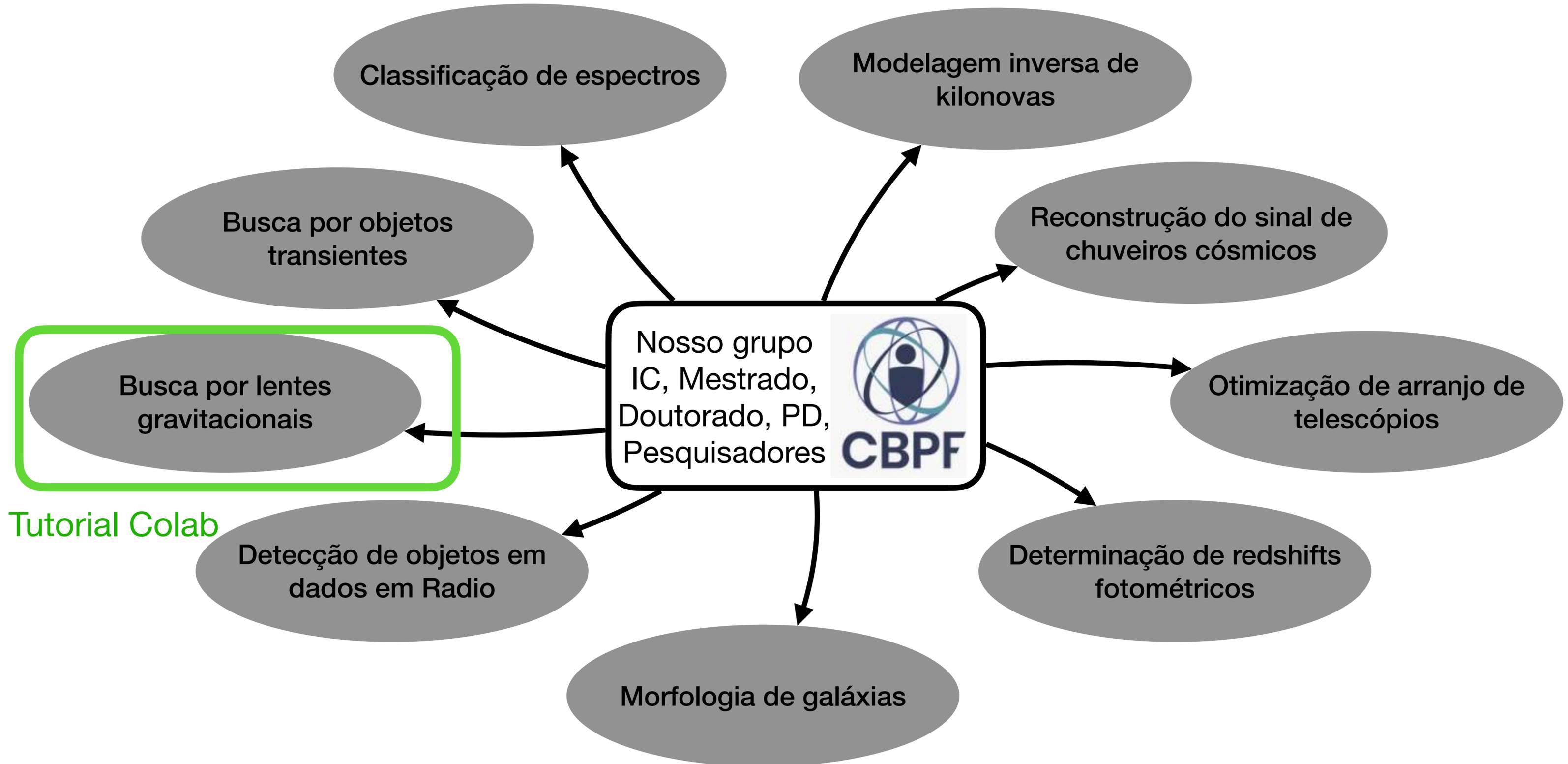
# Astrofísica e DL no CBPF



# Astrofísica e DL no CBPF



# Astrofísica e DL no CBPF



# Astrofísica e DL no CBPF

## Trabalhos recentes

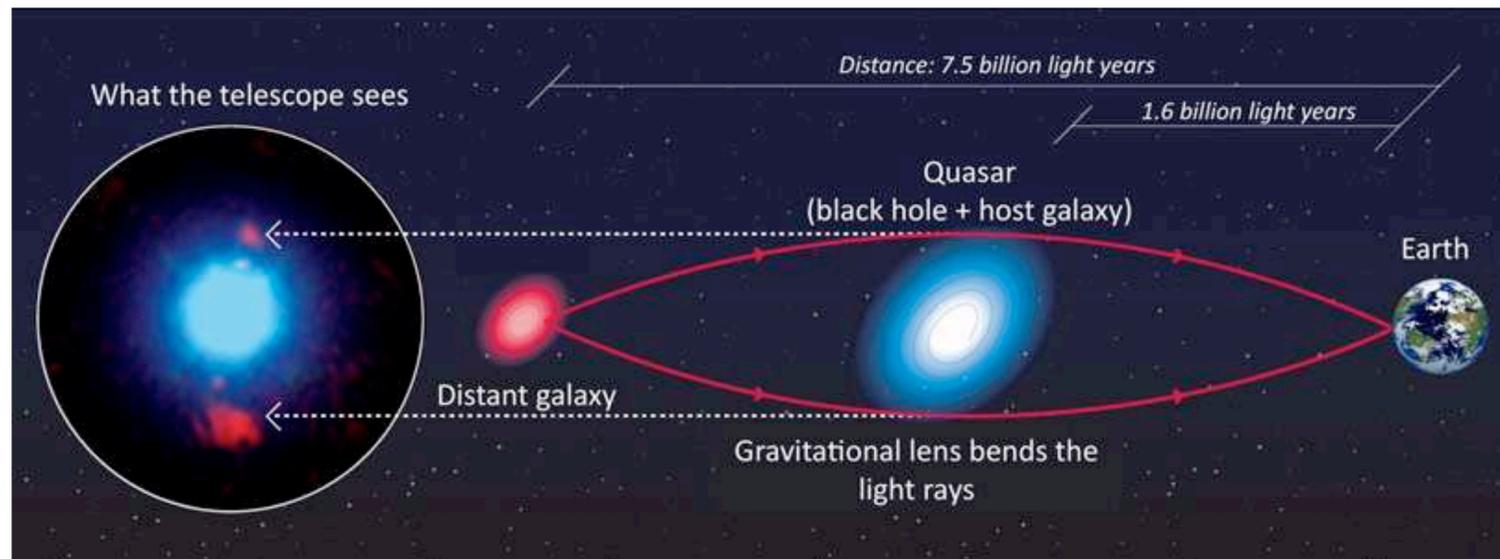
1. Uso de redes bayesianas para reconstruir parâmetros de chuveis atmosféricos no SWGO (regressão/tabular);
2. Encontrar o melhor arranjo de telescópios usando Aprendizado por Reforço (Não supervisionado);
3. Classificação de blazares/não blazares usando redes neurais recorrentes (classificação/tabular);
4. Classificação da morfologia de galáxias com dados do S-PLUS (classificação/imagens);
5. Detecção de objetos em dados de telescópios em Radio (segmentação 3D e regressão/imagens).



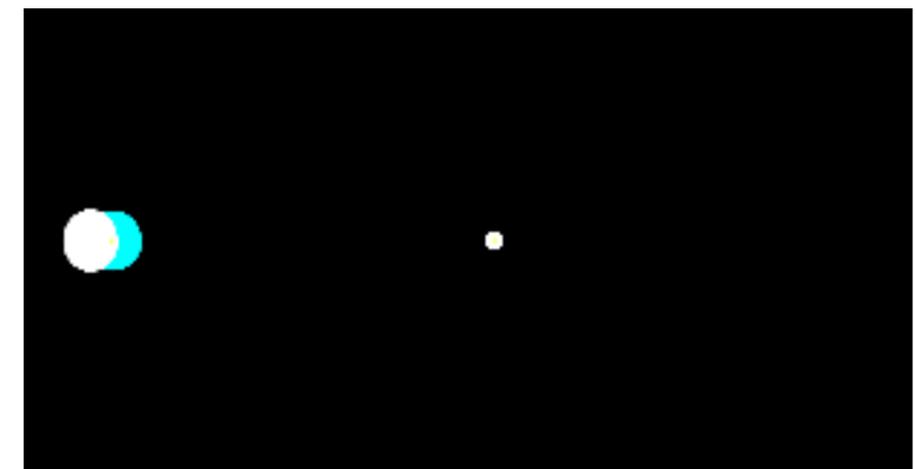
# II Strong Gravitational Lens Challenge

## O que são lentes gravitacionais?

- Lentes gravitacionais ocorrem quando a luz de um objeto distante é defletida por um objeto mais perto;
- Lenteamento forte ocorre quando a lente é compacta e massiva -> imagens múltiplas, arcos, anéis, magnificação do brilho;
- Efeitos normalmente podem ser vistos a olho nu;
- Importantes: informação sobre a distribuição de matéria (inclusive matéria escura), permite ver objetos distantes.



<https://sites.astro.caltech.edu/~george/qsolens/>

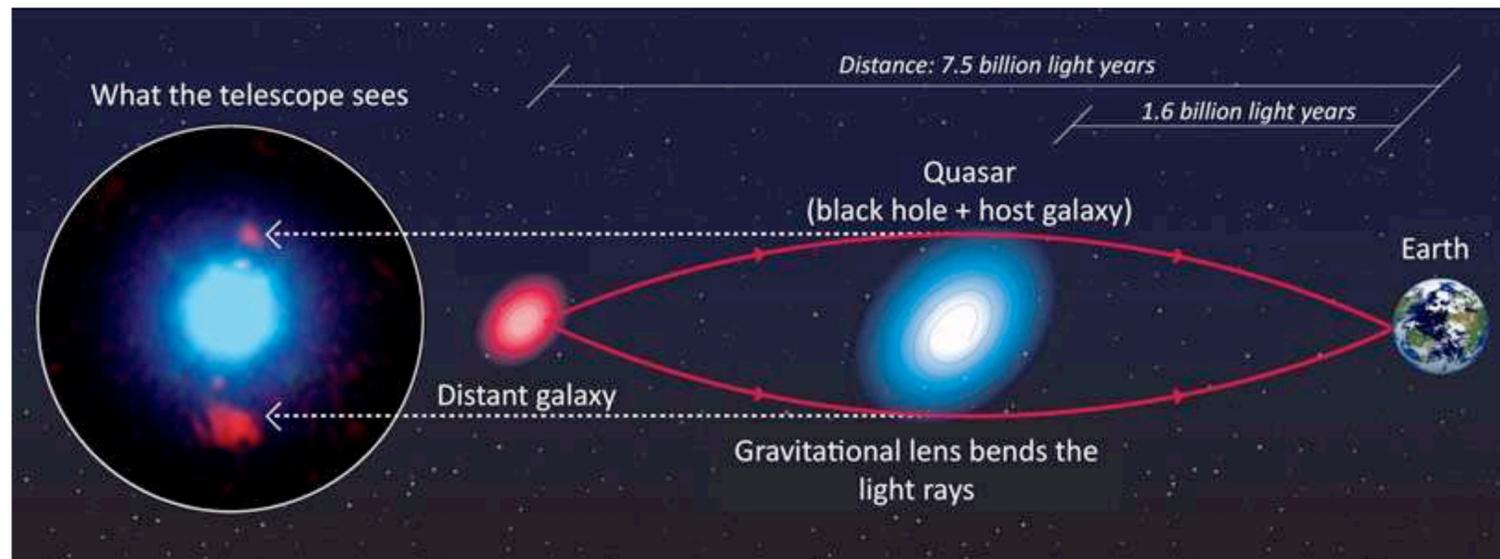


By G. Mikaberidze - Own work, CC BY-SA 4.0,  
<https://commons.wikimedia.org/w/index.php?curid=50502914>

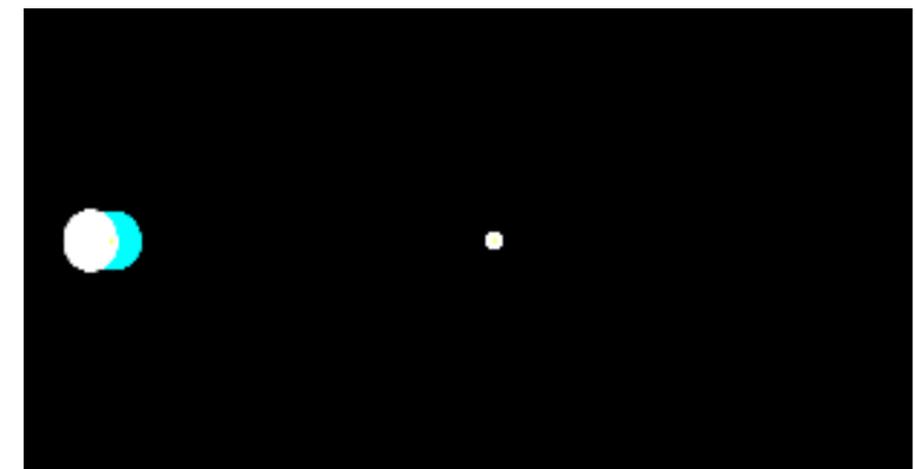
# II Strong Gravitational Lens Challenge

## O que são lentes gravitacionais?

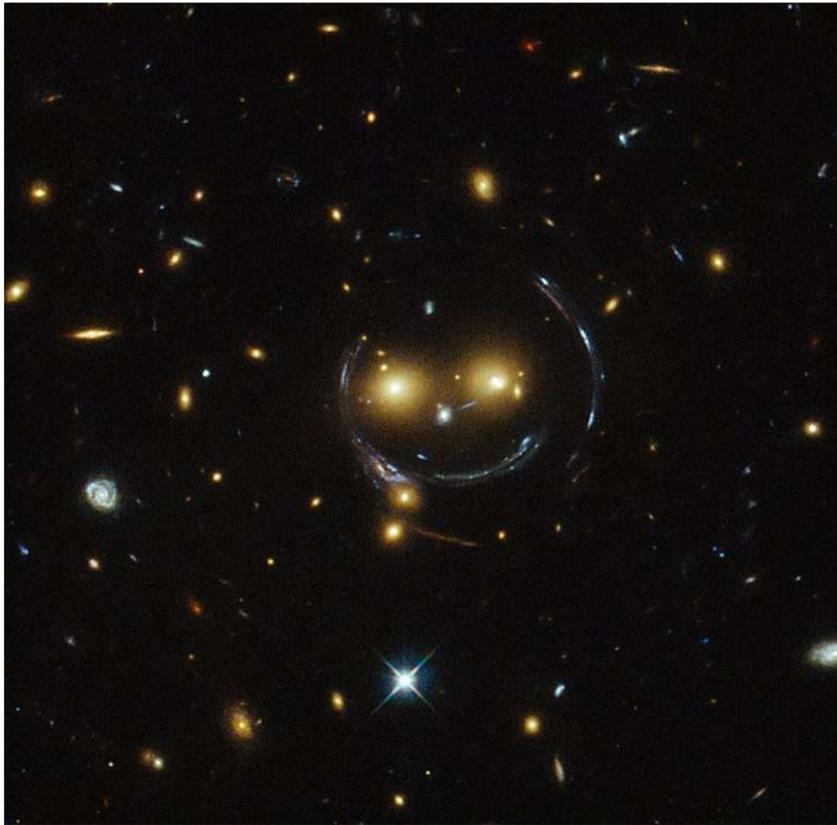
- Lentes gravitacionais ocorrem quando a luz de um objeto distante é defletida por um objeto mais perto;
- Lenteamento forte ocorre quando a lente é compacta e massiva -> imagens múltiplas, arcos, anéis, magnificação do brilho;
- Efeitos normalmente podem ser vistos a olho nu;
- Importantes: informação sobre a distribuição de matéria (inclusive matéria escura), permite ver objetos distantes.



<https://sites.astro.caltech.edu/~george/qsolens/>



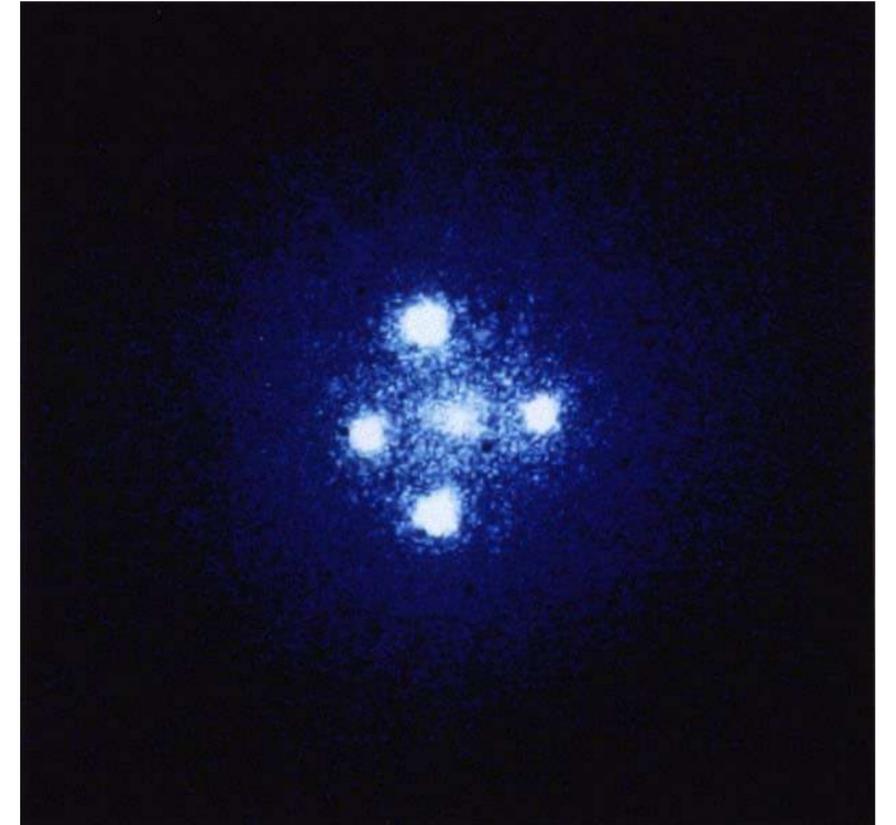
By G. Mikaberidze - Own work, CC BY-SA 4.0,  
<https://commons.wikimedia.org/w/index.php?curid=50502914>



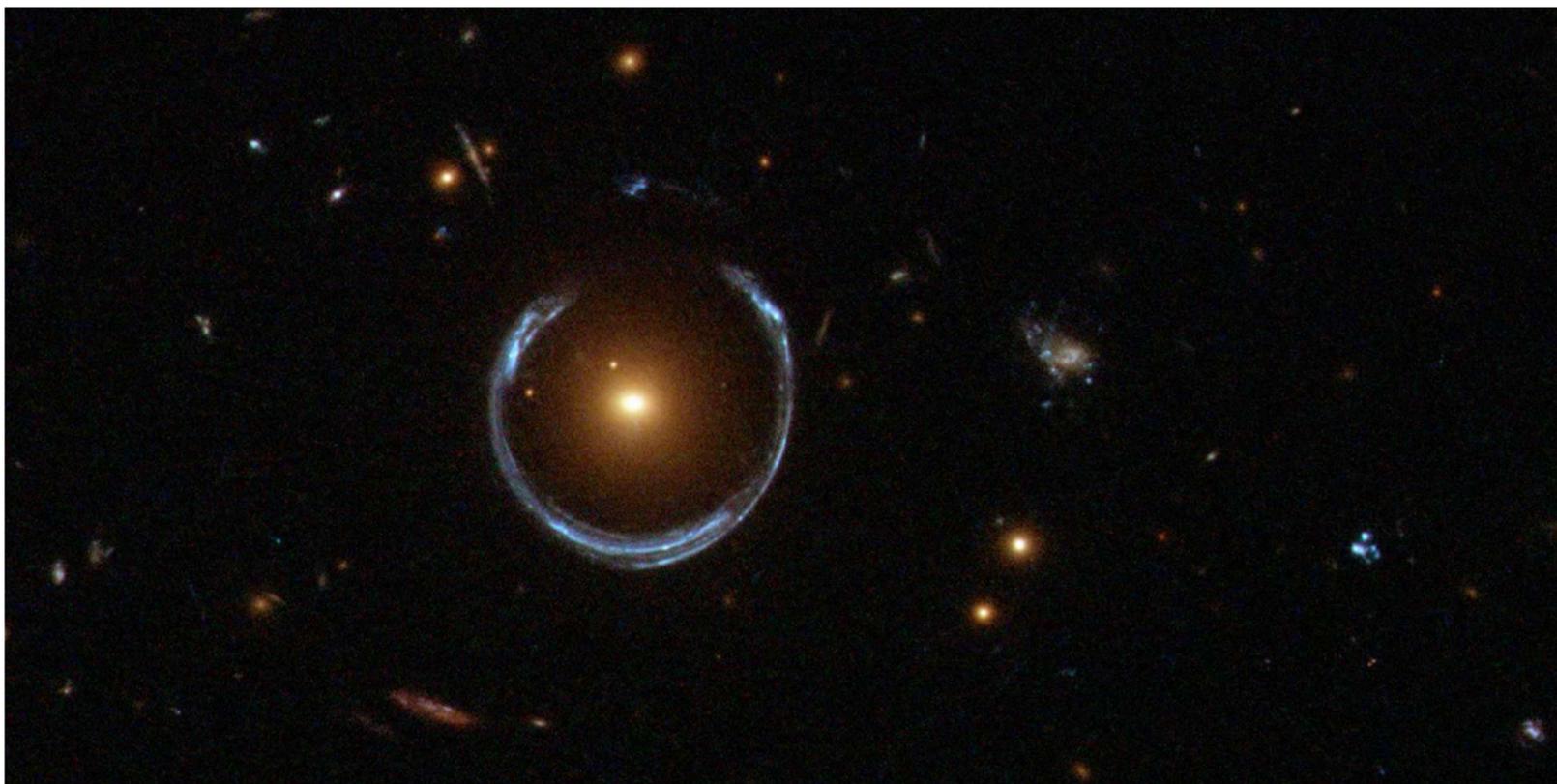
Hubble: [http://www.nasa.gov/sites/default/files/thumbnails/image/15861603283\\_3579db3fc6\\_o.jpg](http://www.nasa.gov/sites/default/files/thumbnails/image/15861603283_3579db3fc6_o.jpg)



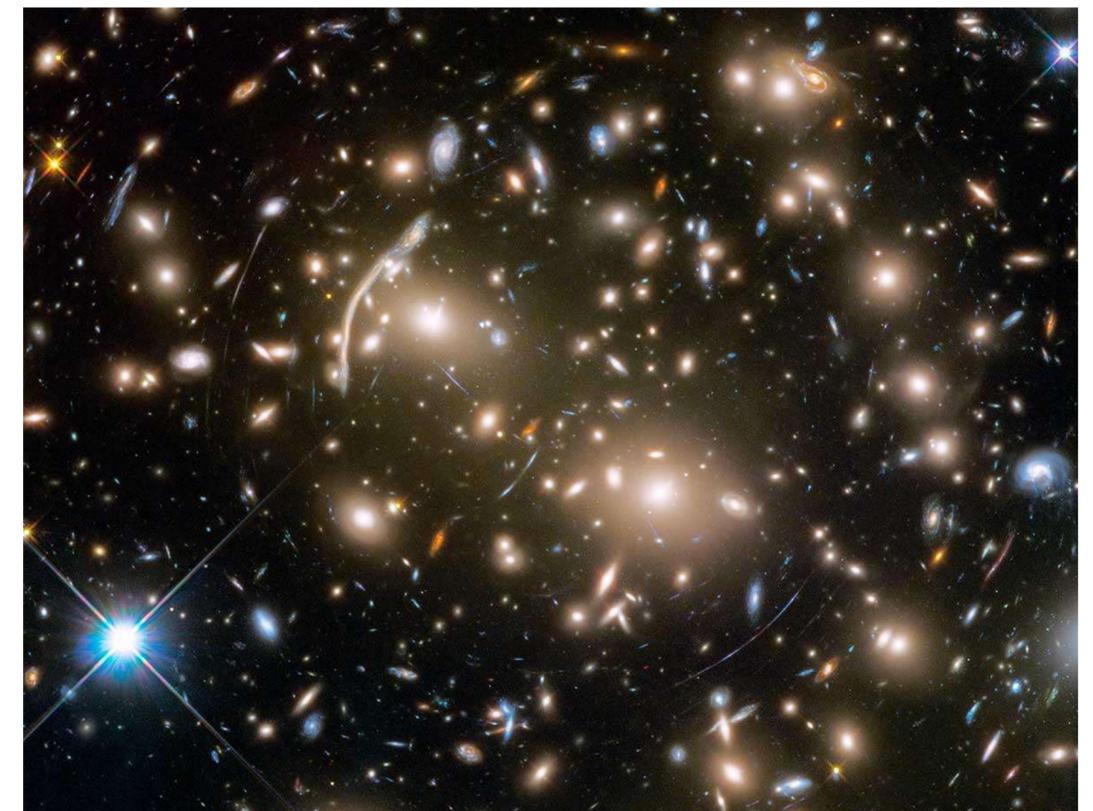
ESA/Hubble, NASA, Suyu et al.



NASA, STScI



<https://hubblesite.org/contents/articles/gravitational-lensing>



<https://hubblesite.org/contents/articles/gravitational-lensing>

# II Strong Gravitational Lens Challenge

## O Desafio

- Classificar imagens - com lente/sem lente (classificação binária);
- Dataset de treino - 100.000 imagens em 4 filtros (H, J, Y, VIS), disponível aos participantes;
- Dataset de teste - 100.000 imagens diferentes, 4 filtros. Conjunto cego;
- 1 filtro tem resolução diferente! 200x200 (VIS) vs 66x66 (H,J,Y);
- Métricas de avaliação (no dataset de teste):
  - Precisão (pureza): lentes preditas corretamente / total de lentes preditas;
  - Completeza: lentes preditas corretamente / total de lentes;
  - F\_score: combinação entre precisão e completeza.

# II SGLC

## Inspeção dos dados e pré processamento

# II SGLC

## Inspeção dos dados e pré processamento

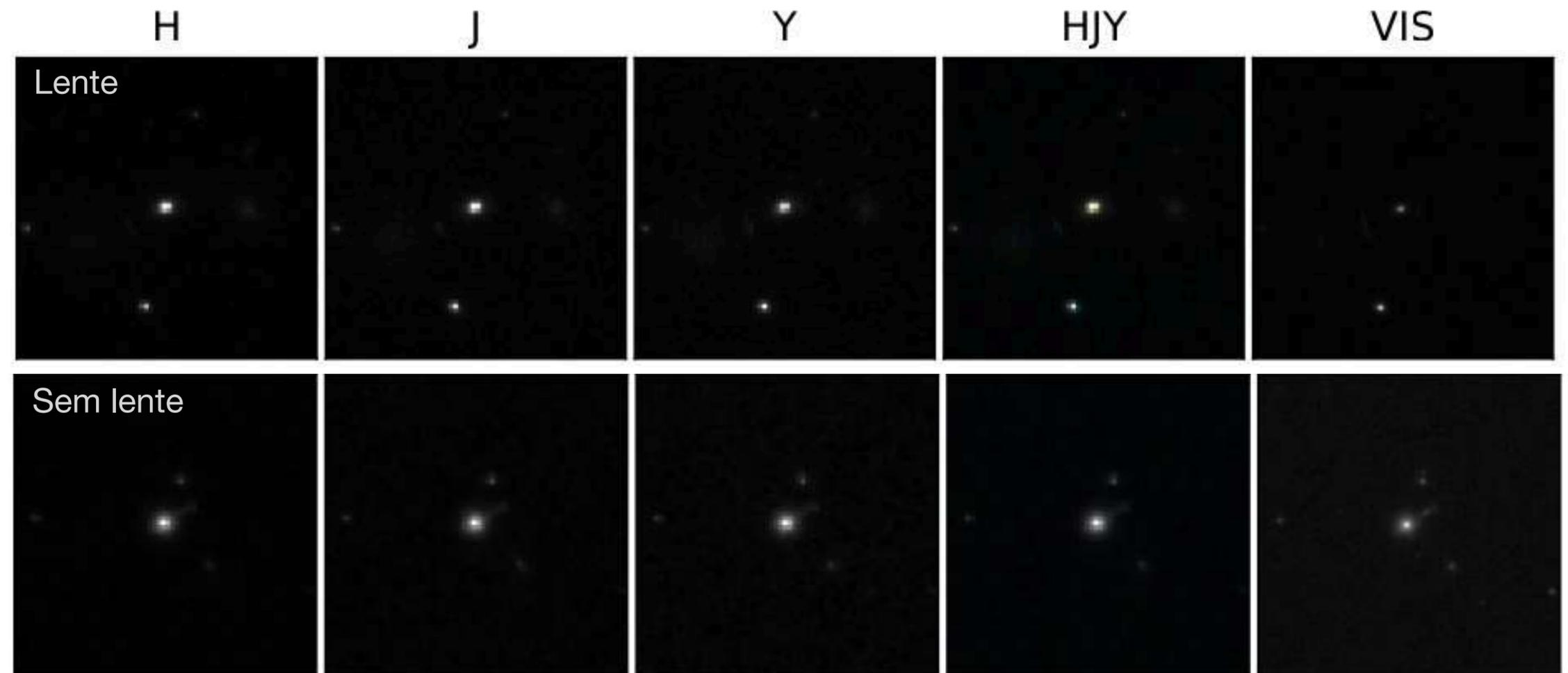
- 1º passo: inspeção visual de uma parte dos dados

# II SGLC

## Inspeção dos dados e pré processamento

- 1º passo: inspeção visual de uma parte dos dados

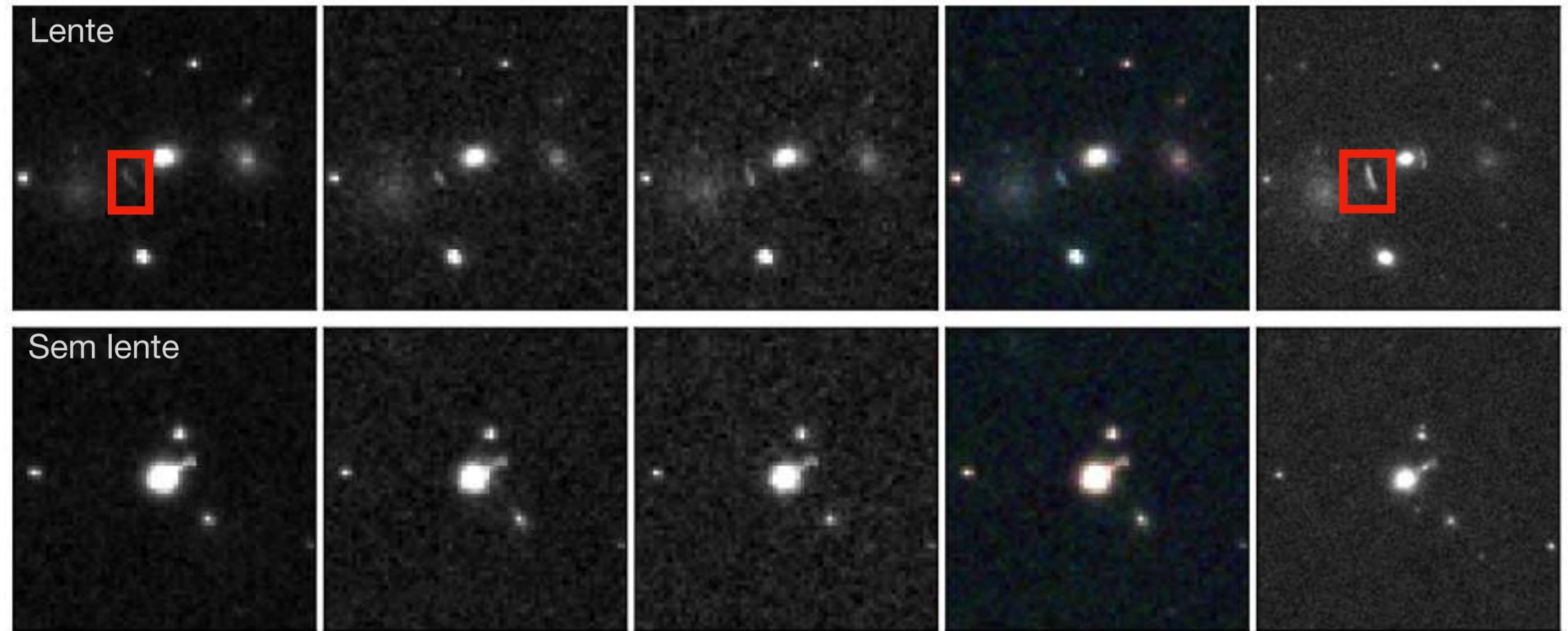
**Características dos objetos totalmente escondidas!**



# II SGLC

## Inspeção dos dados e pré processamento

- 1º passo: inspeção visual de uma parte dos dados
- Reprocessar as imagens para melhor visualização
  - Clipar as imagens: limitar valores de pixel muito altos ou muito baixos



# II SGLC

## Inspeção dos dados e pré processamento

- 1º passo: inspeção visual de uma parte dos dados
- Reprocessar as imagens para melhor visualização
  - Clipar as imagens: limitar valores de pixel muito altos ou muito baixos
- Pré processamento é muito importante!
  - Ainda assim muitas lentes difíceis de discernir

# II SGLC

## Inspeção dos dados e pré processamento

- 1º passo: inspeção visual de uma parte dos dados
- Reprocessar as imagens para melhor visualização
  - Clipar as imagens: limitar valores de pixel muito altos ou muito baixos
- Pré processamento é muito importante!
  - Ainda assim muitas lentes difíceis de discernir
- Como combinar os diferentes filtros?
  - Reduzir o tamanho da VIS? Perda de dados
  - Aumentar o tamanho de H, J, Y? Criar artefatos
  - H+J+Y+VIS? HJY+VIS? ...

# II SGLC

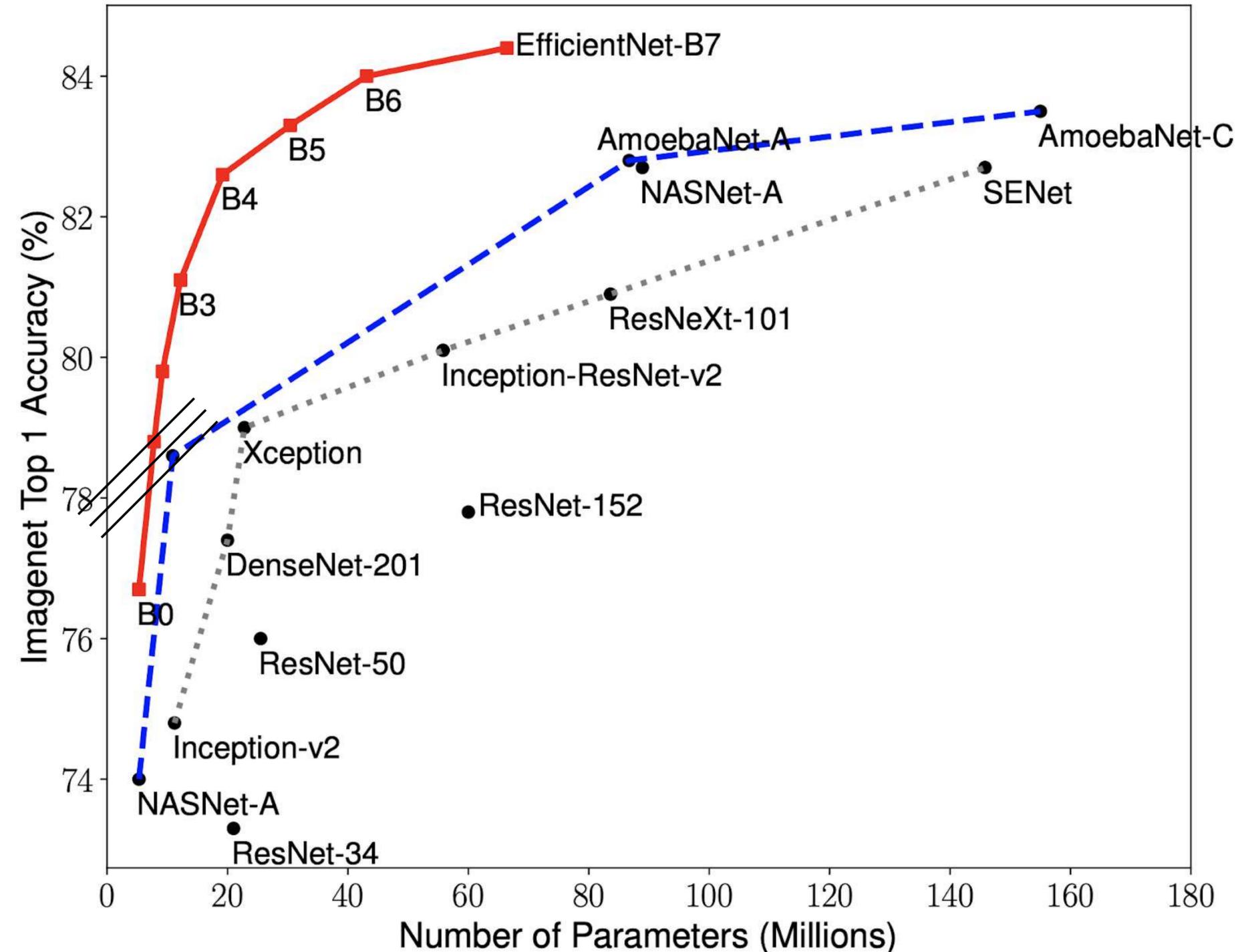
## Inspeção dos dados e pré processamento

- 1º passo: inspeção visual de uma parte dos dados
- Reprocessar as imagens para melhor visualização
  - Clipar as imagens: limitar valores de pixel muito altos ou muito baixos
- Pré processamento é muito importante!
  - Ainda assim muitas lentes difíceis de discernir
- Como combinar os diferentes filtros?
  - Reduzir o tamanho da VIS? Perda de dados
  - Aumentar o tamanho de H, J, Y? Criar artefatos
  - H+J+Y+VIS? HJY+VIS? ...
- Testar!

# II SGLC

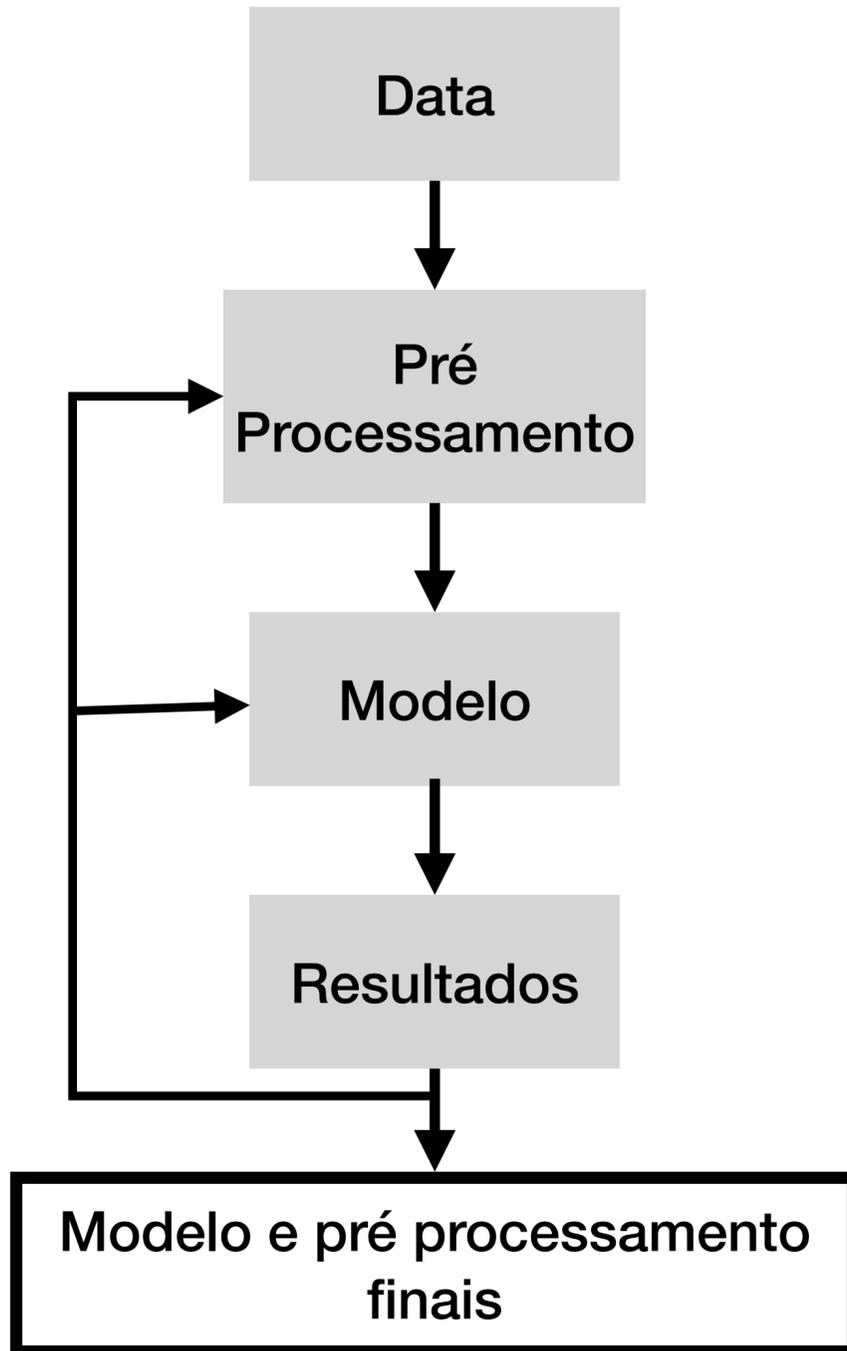
## Escolha da rede

- Criar um modelo do zero;
- Usar modelos já testados em outros datasets;
- ImageNet: >1 milhão de imagens, 1000 classes (gato, cachorro, cavalo...). Um dos Datasets mais usados para benchmark de modelos de visão computacional;
- Transfer learning -> usar o aprendizado do modelo em outro dataset como base;
- Ajuste entre performance e custo computacional e de tempo;
- Testar!



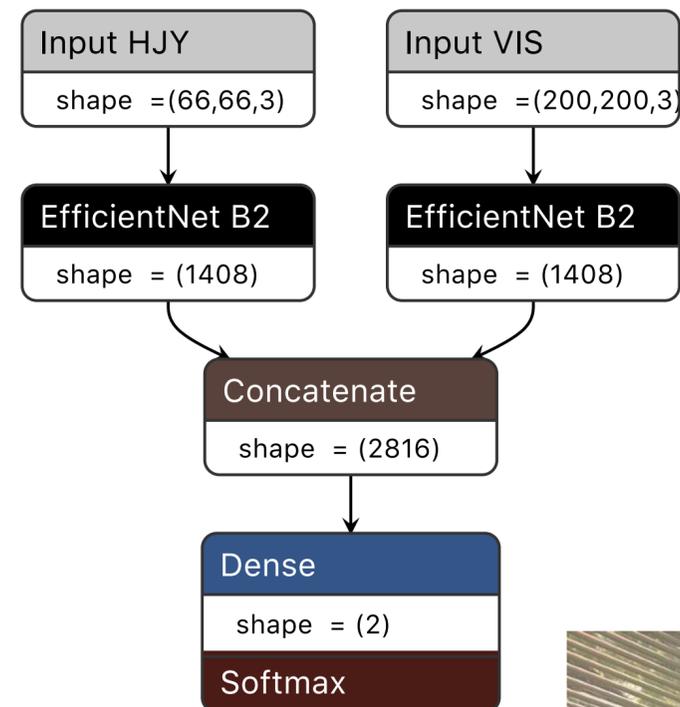
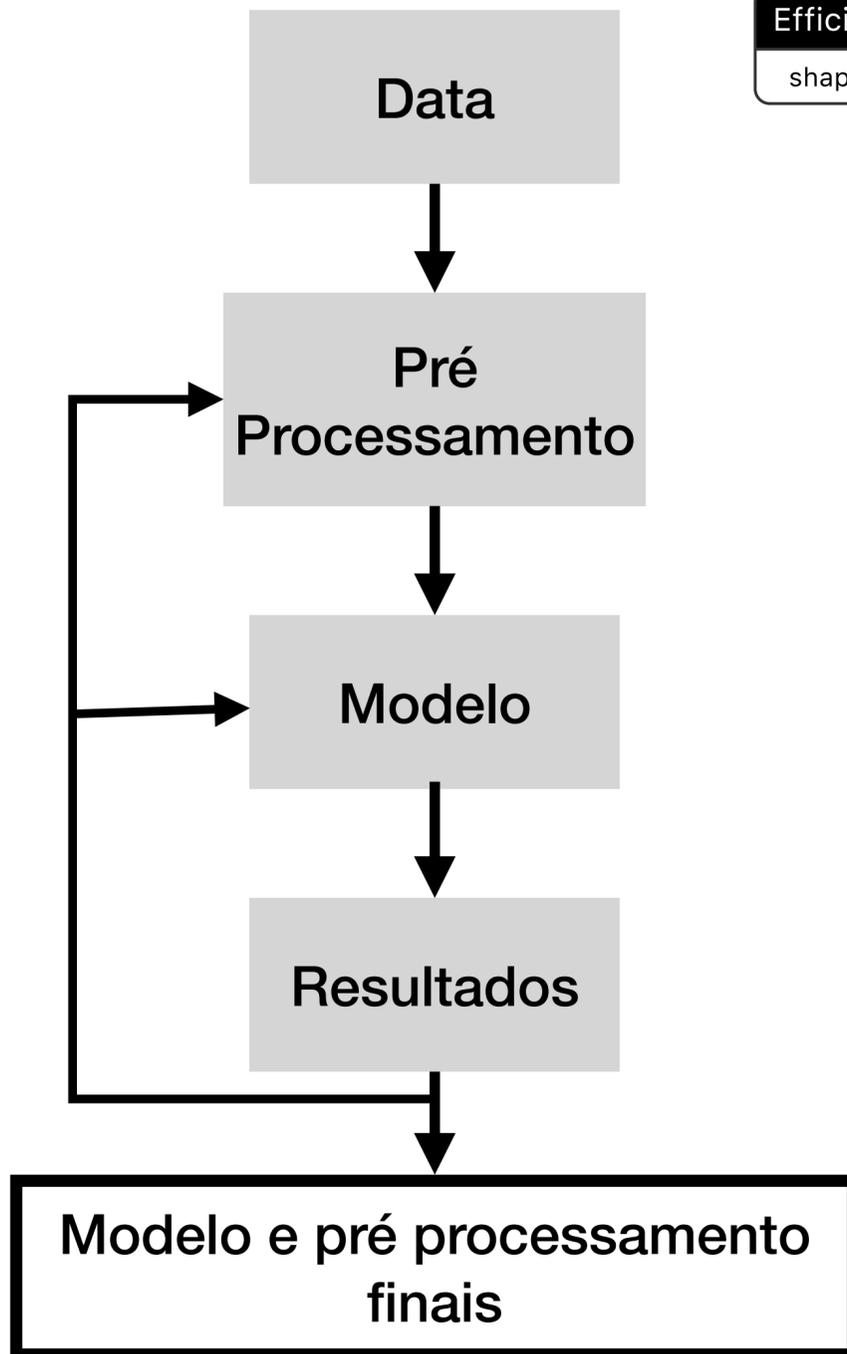
# II SGLC

## Fase de testes



# II SGLC

## Fase de testes

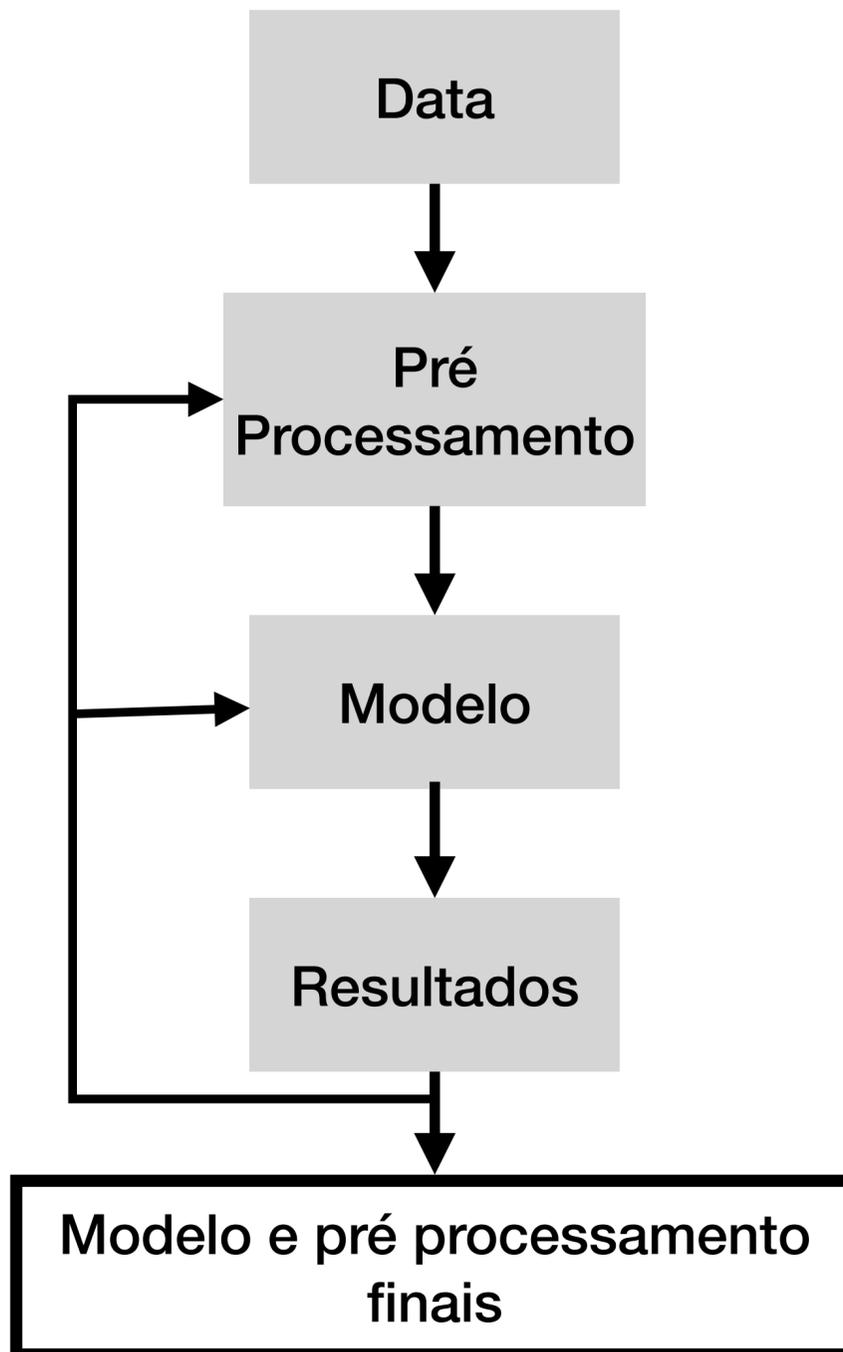


Aumentação de dados:  
Rotações, reflexões, zoom



# II SGLC

## Fase de testes

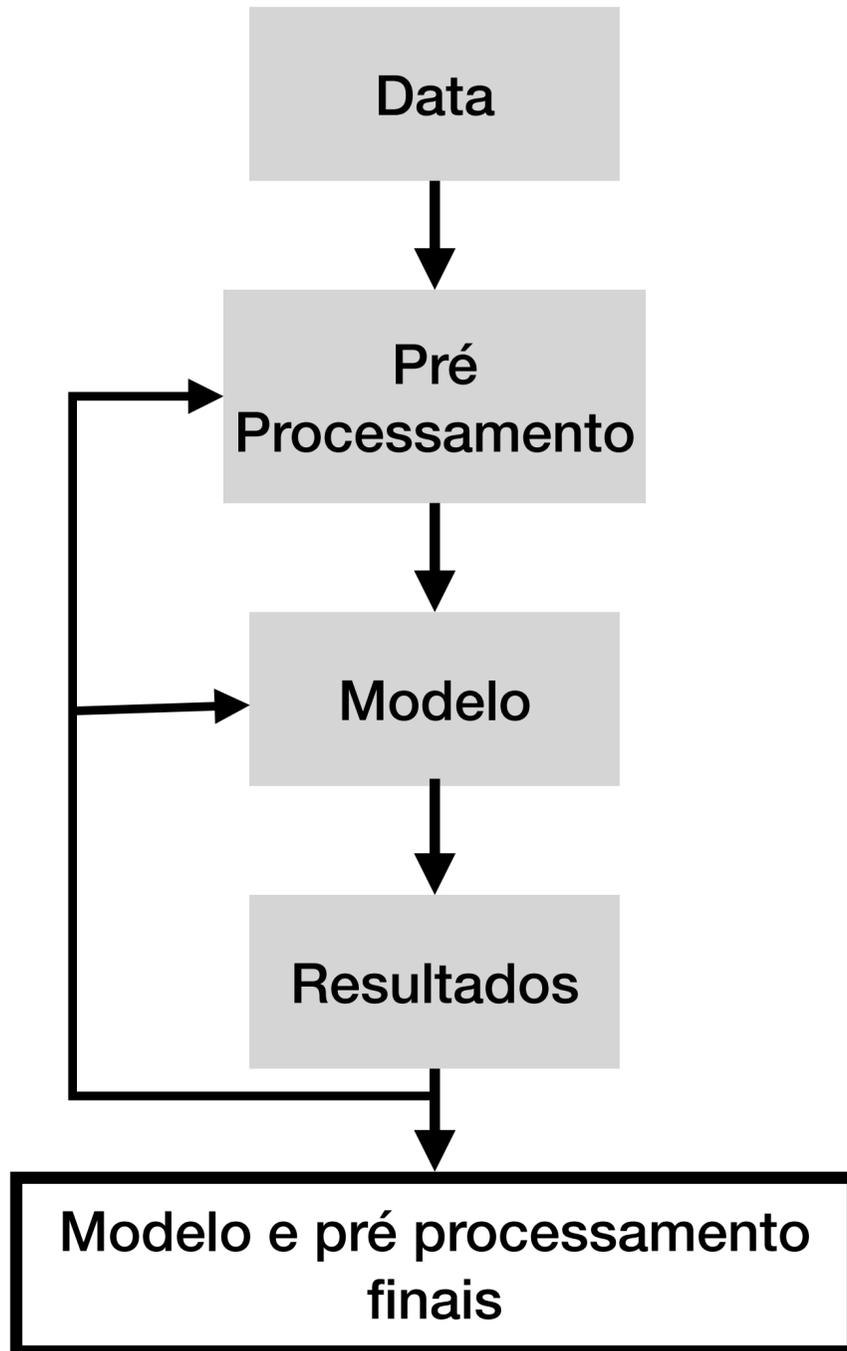


	F	Precision	Recall	names
<b>1</b>	<b>0.993468</b>	<b>0.997855</b>	<b>0.168824</b>	<b>Cast_efficient_auc Clecio Bom. et al.</b>
2	0.993456	0.996666	0.216973	CNN-napoli-groningen Gentile F., et al.
3	0.993415	0.997041	0.197052	Cast_custom_res_auc Clecio Bom. et al.
4	0.992514	0.997135	0.161404	Cast_efficient_f Clecio Bom, et al.
5	0.992502	0.997300	0.156404	cnnLF Rui Li et al.
6	0.992460	0.996276	0.188806	Cast_custom_res_f Clecio Bom, et al.
7	0.992210	0.997031	0.155718	GAMOCLASS2 Diego Tuccillo et al.
8	0.990452	0.996228	0.133105	manchester1 Neal Jackson
9	0.989814	0.995895	0.127168	OU-VIS-JYH-200 Joshua Wilde et al.
10	0.987513	0.992573	0.148197	GAMOCLASS Diego Tuccillo, at al.
11	0.985876	0.990229	0.167554	;-L-Finding-Challenge-2 Joshua Wilde et al.
12	0.985229	0.994180	0.089544	BarSanCNN-N1 Iberto Manjón García, et al.
13	0.984689	0.991219	0.118356	OU-JYH-VIS-66 Joshua Wilde et al.
14	0.982051	0.991908	0.081559	BarSanCNN-N2 Iberto Manjón García, et al.
15	0.981997	0.989571	0.103315	I-ASTRO-FINDER Flodie Savary et al.



# II SGLC

## Fase de testes



	F	Precision	Recall	names
1	0.993468	0.997855	0.168824	Cast_efficient_auc Clecio Bom. et al.
2	0.993456	0.996666	0.216973	CNN-napoli-groningen Gentile F., et al.
3	0.993415	0.997041	0.197052	Cast_custom_res_auc Clecio Bom. et al.
4	0.992514	0.997135	0.161404	Cast_efficient_f Clecio Bom, et al.
5	0.992502	0.997300	0.156404	cnnLF Rui Li et al.

Monthly Notices  
of the  
ROYAL ASTRONOMICAL SOCIETY

MNRAS **515**, 5121–5134 (2022)  
Advance Access publication 2022 July 22

<https://doi.org/10.1093/mnras/stac2047>

## Developing a victorious strategy to the second strong gravitational lensing data challenge

C. R. Bom <sup>ID</sup>, <sup>1,2</sup>★ B. M. O. Fraga <sup>ID</sup>, <sup>1</sup> L. O. Dias, <sup>1</sup> P. Schubert, <sup>1</sup>† M. Blanco Valentin, <sup>1</sup>‡ C. Furlanetto <sup>ID</sup>, <sup>3</sup>  
M. Makler <sup>ID</sup>, <sup>1,4</sup> K. Teles, <sup>1</sup> M. Portes de Albuquerque <sup>1</sup> and R. Benton Metcalf <sup>5,6</sup>

		No Preprocessing	II SGLC Preprocessing	Alternative Preprocessing
HJY	ROC	0.4994 ± 0.00135	0.5389 ± 0.01891	0.5350 ± 0.02414
	PR	0.4950 ± 0.0155	0.5535 ± 0.0210	0.5342 ± 0.0303
	$F_\beta$	0.620002 ± 0.102507	0.918411 ± 0.069823	0.800559 ± 0.098771
VIS (repeated)	ROC	0.5040 ± 0.0108	0.6589 ± 0.0423	0.6266 ± 0.0445
	PR	0.5129 ± 0.0779	0.6936 ± 0.0366	0.6634 ± 0.0434
	$F_\beta$	0.526797 ± 0.066389	0.982306 ± 0.007477	0.979158 ± 0.011016
VIS (zeros)	ROC	0.5001 ± 0.0101	0.7002 ± 0.0306	0.6534 ± 0.0414
	PR	0.4884 ± 0.0384	0.7326 ± 0.0256	0.6864 ± 0.0387
	$F_\beta$	0.548153 ± 0.066831	0.986763 ± 0.005173	0.979588 ± 0.010142
HJY+VIS (zeros)	ROC	0.5017 ± 0.0069	0.8018 ± 0.0173	0.8147 ± 0.0161
	PR	0.4916 ± 0.0074	0.8239 ± 0.0128	0.8359 ± 0.0132
	$F_\beta$	0.613814 ± 0.094273	0.990140 ± 0.003437	0.991610 ± 0.003873
HJY+VIS (repeated)	ROC	0.4932 ± 0.0102	0.8016 ± 0.0186	0.8295 ± 0.0088
	PR	0.4921 ± 0.0175	0.8230 ± 0.0149	0.8469 ± 0.0081
	$F_\beta$	0.619433 ± 0.106441	0.990402 ± 0.004087	0.992070 ± 0.003427
Y+VIS (zeros)	ROC	0.5011 ± 0.0174	0.8015 ± 0.0186	0.8149 ± 0.0144
	PR	0.4939 ± 0.0124	0.8243 ± 0.0135	0.8369 ± 0.0120
	$F_\beta$	0.613662 ± 0.125263	0.989978 ± 0.004941	0.990235 ± 0.004237





```
*****-Inspiron-N4010:~$ python3
Python 3.8.5 (default, May 27 2021, 13:30:53)
[GCC 9.3.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import this
The Zen of Python, by Tim Peters

Beautiful is better than ugly.
Explicit is better than implicit.
Simple is better than complex.
Complex is better than complicated.
Flat is better than nested.
Sparse is better than dense.
Readability counts.
Special cases aren't special enough to break the rules.
Although practicality beats purity.
Errors should never pass silently.
Unless explicitly silenced.
In the face of ambiguity, refuse the temptation to guess.
There should be one-- and preferably only one --obvious way to do it.
Although that way may not be obvious at first unless you're Dutch.
Now is better than never.
Although never is often better than *right* now.
If the implementation is hard to explain, it's a bad idea.
If the implementation is easy to explain, it may be a good idea.
Namespaces are one honking great idea -- let's do more of those!
>>> █
```

# Agora, mãos à obra!