

# Classificação morfológica de galáxias usando conjunto de redes neurais convolucionais

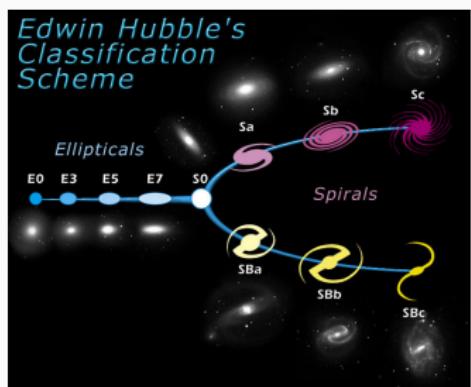
---

Natanael Magalhães Cardoso\*  
Profa. Cláudia Mendes de Oliveira†

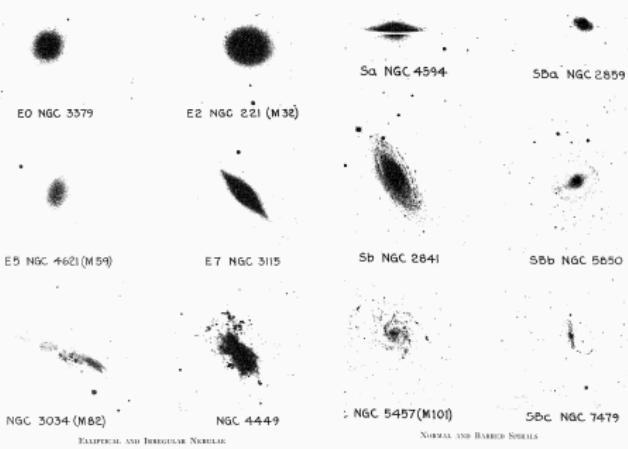
\* Departamento de Computação e Sistemas Digitais, Escola Politécnica, USP

† Departamento de Astronomia, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, USP

# BACKGROUND: ASTRONOMIA



**Figura 1:** Hubble's Tuning Fork



**Figura 2:** Extragalactic Nebulae; E. Hubble, 1926.

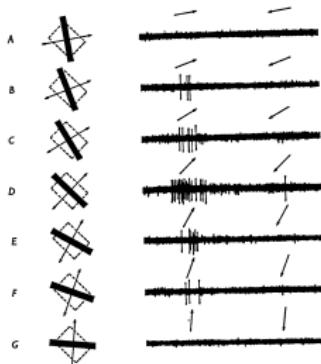
# BACKGROUND: NEUROCIÊNCIA

MONKEY STRIATE CORTEX

219

showed little or no directional preference. Even when responses were highly asymmetrical, the less effective direction of movement usually evoked some minimal response (see Text-fig. 2), but there were a few examples in which the maintained activity was actually suppressed.

Individual complex cells differed markedly in their relative responsiveness to slits, edges, or dark bars. The majority responded very much better to one than to the other two, but some reacted briskly to two of them, and a few to all three. For a cell that was sensitive to slits, but not to edges, the



Text-fig. 2. Responses of a complex cell in right striate cortex (layer IV A) to various orientations of a moving black bar. Receptive field in the left eye indicated by the interrupted rectangles; it was approximately  $\frac{1}{2} \times 1^\circ$  in size, and was situated  $4^\circ$  below and to the left of the point of fixation. Ocular-dominance group 4. Duration of each record, 2 sec. Background intensity  $1.3 \log_{10}$  cd/m<sup>2</sup>; dark bars  $0.0 \log$  cd/m<sup>2</sup>.

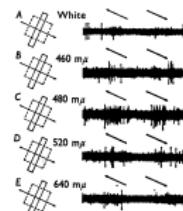
responses increased as slit width was increased up to some optimal value,

224

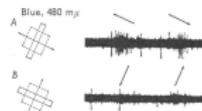
D. H. HUBEL AND T. N. WIESEL

mentioned, favouring monochromatic slits over white, but also had hyper-complex characteristics in that a long slit was distinctly less effective than one of limited length.

*Cells with concentric fields and dual-opponent systems.* Under this heading we group a very few cells with centre-surround receptive-field organization, but with more complex behaviour than anything we have seen in



Text-fig. 5. Complex cell with colour coded properties recorded in layer II of striate cortex. Responses to movement of centrally oriented slits of white light and monochromatic light at various wavelengths. Monochromatic light made by interposing interference filters in a beam of white light; stimulus energies are greatest for A, and progressively less for E, D, C and B. None of the responses was improved by lowering the intensity. Size of receptive field  $\frac{1}{2} \times 1^\circ$ . Ocular dominance group 1. Background and white stimulus intensities as in Text-fig. 4. Time for each record 5 sec.



Text-fig. 6. Same cell as Text-fig. 5. Responses to two orthogonal stimulus orientations at 450 nm.

**Figura 3: Monkey Striate Cortex; D. H. Hubel & T. N. Wisel, 1968.**

# BACKGROUND: VISÃO COMPUTACIONAL

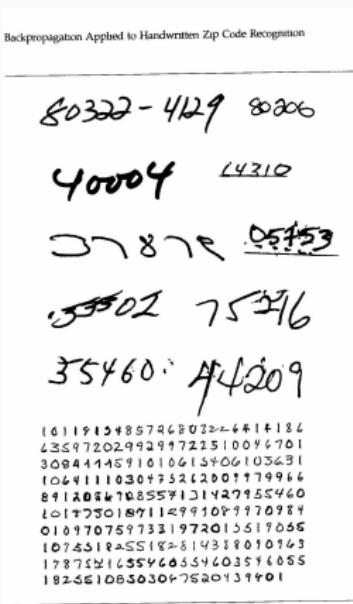


Figure 1 Examples of original zip codes (top) and normalized digits from testing set (bottom).

548 LeCun, Boser, Denker, Henderson, Howard, Hubbard, and Jackel

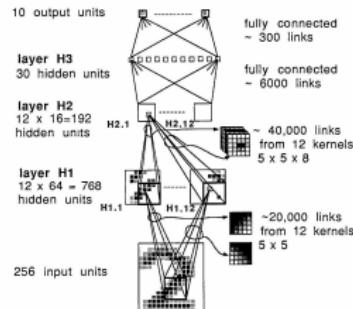


Figure 3 Log mean squared error (MSE) (top) and raw error rate (bottom) versus number of training passes

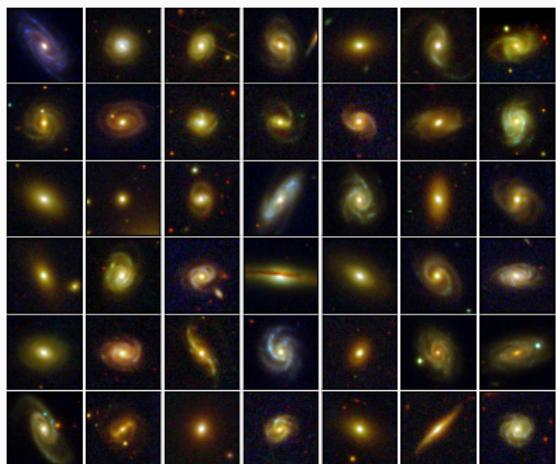
training set, 8.1% misclassifications on the test set, and 19.4% rejections for 1% error rate on the remaining test patterns. A full comparative study will be described in another paper

**5.1 Comparison with Other Work.** The first several stages of processing in our previous system (described in Denker et al., 1989) involved convolutions in which the kernels were both laterally and depthwise hand designed. In the present system, the first two layers of the network are convolutional and depthwise separable, but the system automatically learns the coefficients that make up the kernels. This "constrained backpropagation" is the key to success of the present system: it not only builds in shift-invariance, but vastly reduces the entropy, the Vapnik-Chervonenkis dimensionality, and the number of free parameters, thereby proportionately reducing the amount of training data required to achieve a given level

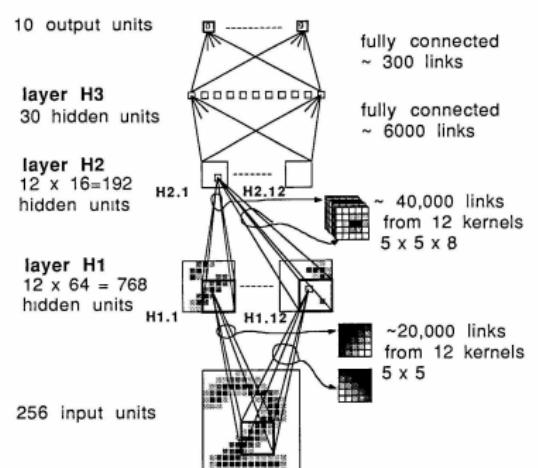
**Figura 4:** Backpropagation applied to Handwritten Zip Code Recognition; Y. LeCun et al., 1989.

# OBJETIVO

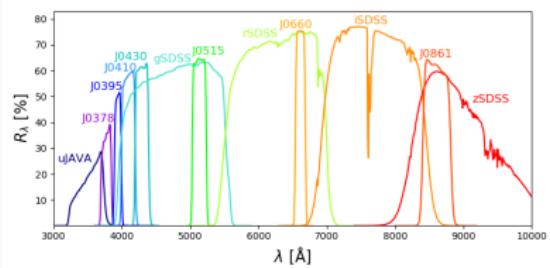
## Problema



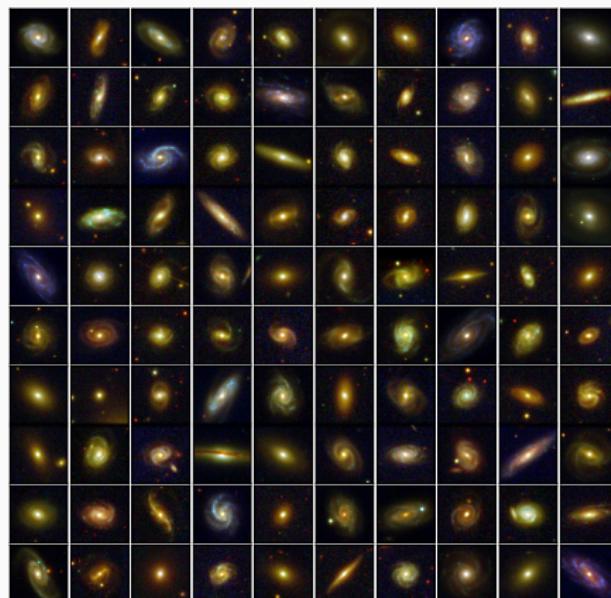
## Método



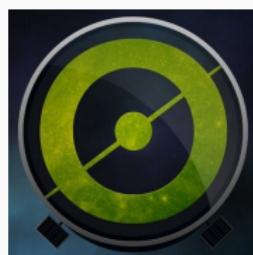
# S-PLUS: SOUTHERN PHOTOMETRIC LOCAL UNIVERSE SURVEY



# CONJUNTO DE DADOS

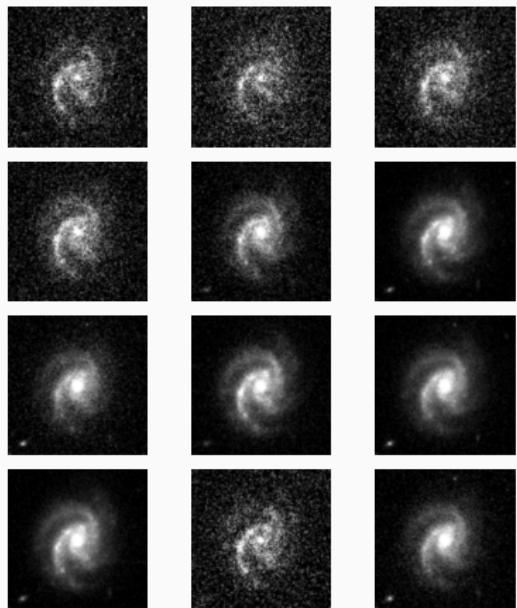


Classe	Treino	Validação	Teste
Elíptica	714	79	88
Espiral	1517	169	188
Total	2231	248	276



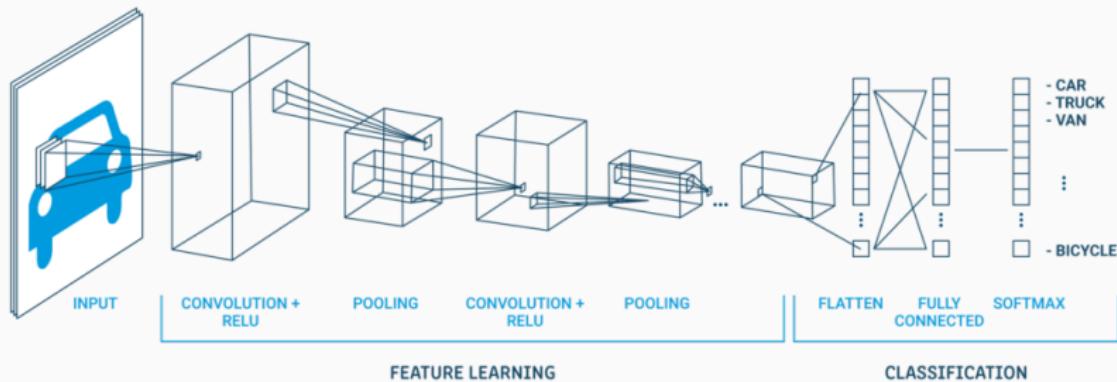
Imagens do S-PLUS e Classificações do GalaxyZoo

## PRÉ-PROCESSAMENTO: REDUÇÃO



**Figura 5:** Imagem em cada banda do S-PLUS e respectiva redução para RGB usando Trilogy.

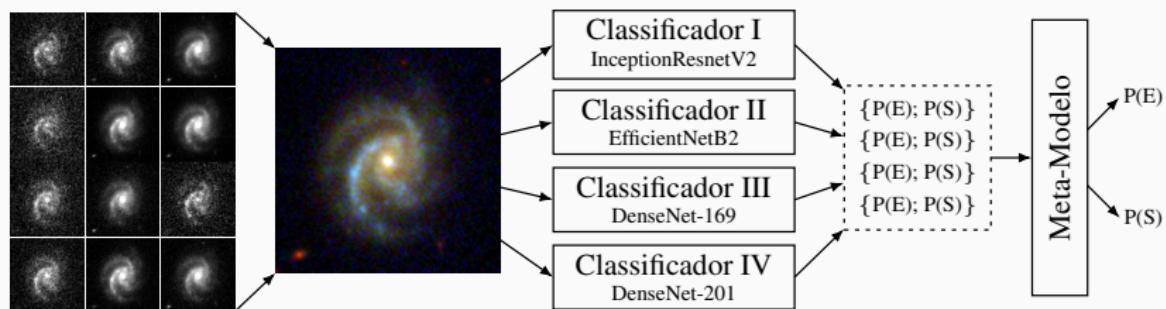
# REDES NEURAIS CONVOLUCIONAIS



**Figura 6:** Estrutura simplificada de uma rede neural convolucional

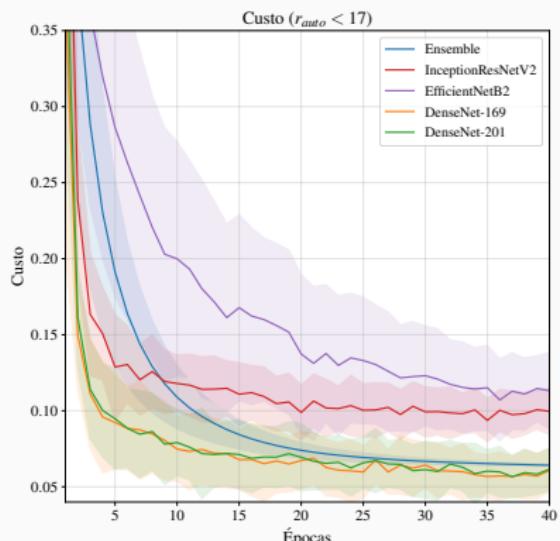
# CONJUNTO DE REDES NEURAIS CONVOLUCIONAIS

Stacking Ensemble (Stacked generalization; D. H. Wolpert, 1992)

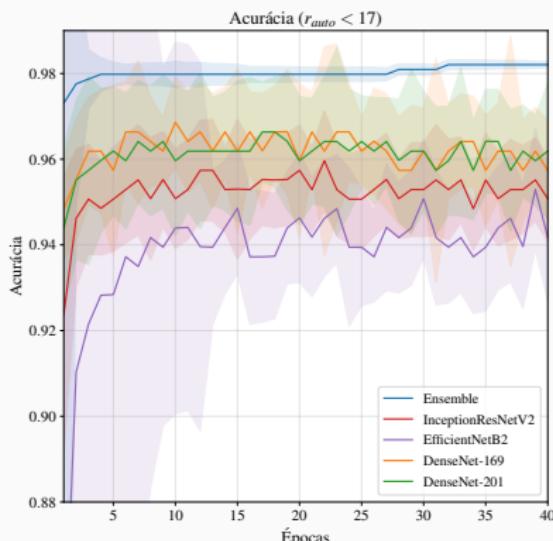


**Figura 7:** Estrutura do modelo proposto.

# PERFORMANCE NO TREINAMENTO

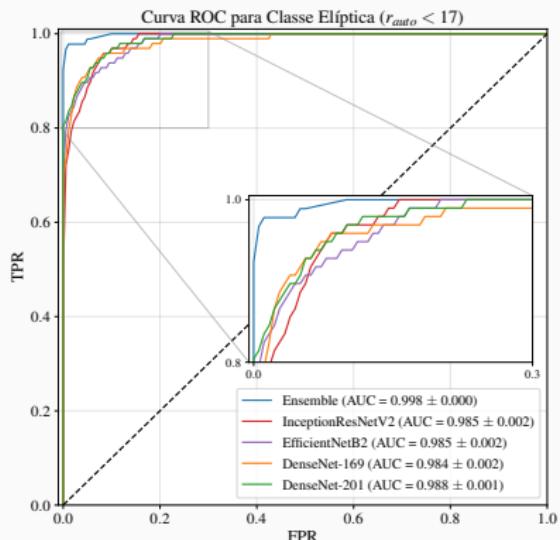


**Figura 8:** Treinamento

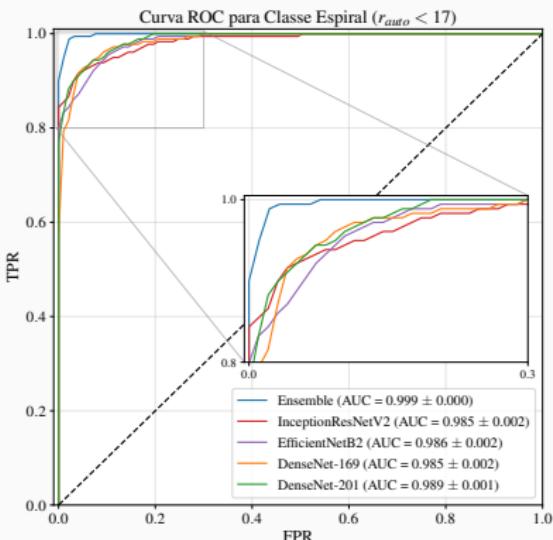


**Figura 9:** Validação

# PERFORMANCE NO CONJUNTO DE TESTE: CURVA ROC

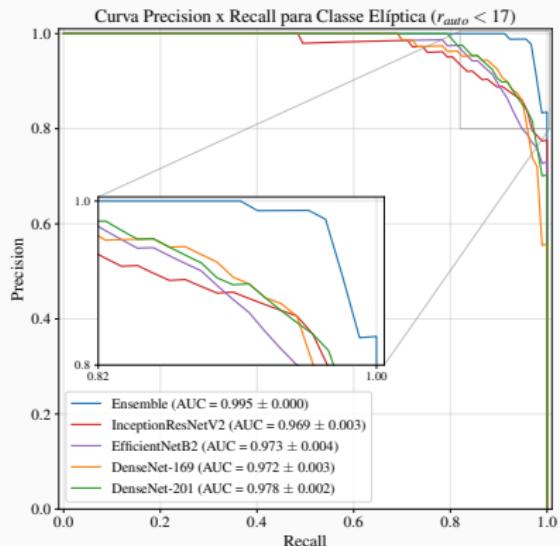


**Figura 10:** Classe Elíptica

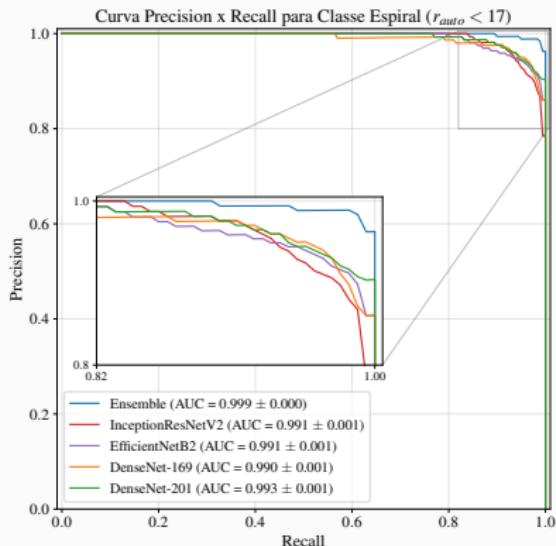


**Figura 11:** Classe Espiral

# PERFORMANCE NO CONJUNTO DE TESTE: CURVA PR

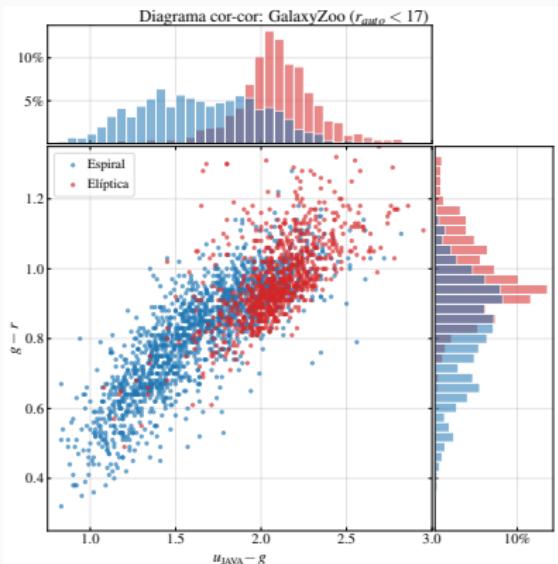


**Figura 12:** Classe Elíptica

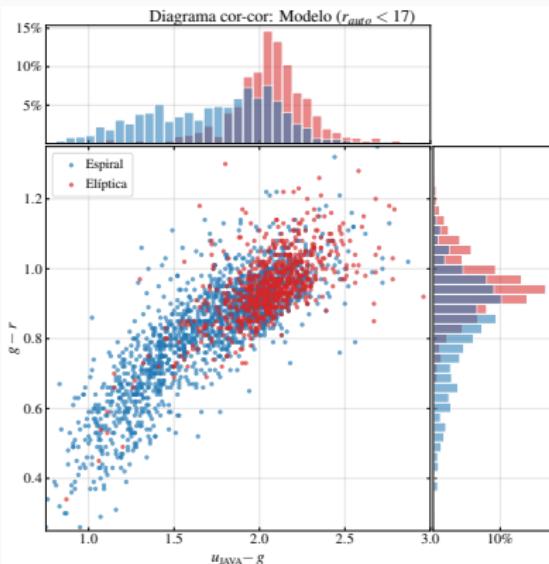


**Figura 13:** Classe Espiral

# PERFORMANCE NO CONJUNTO BLIND

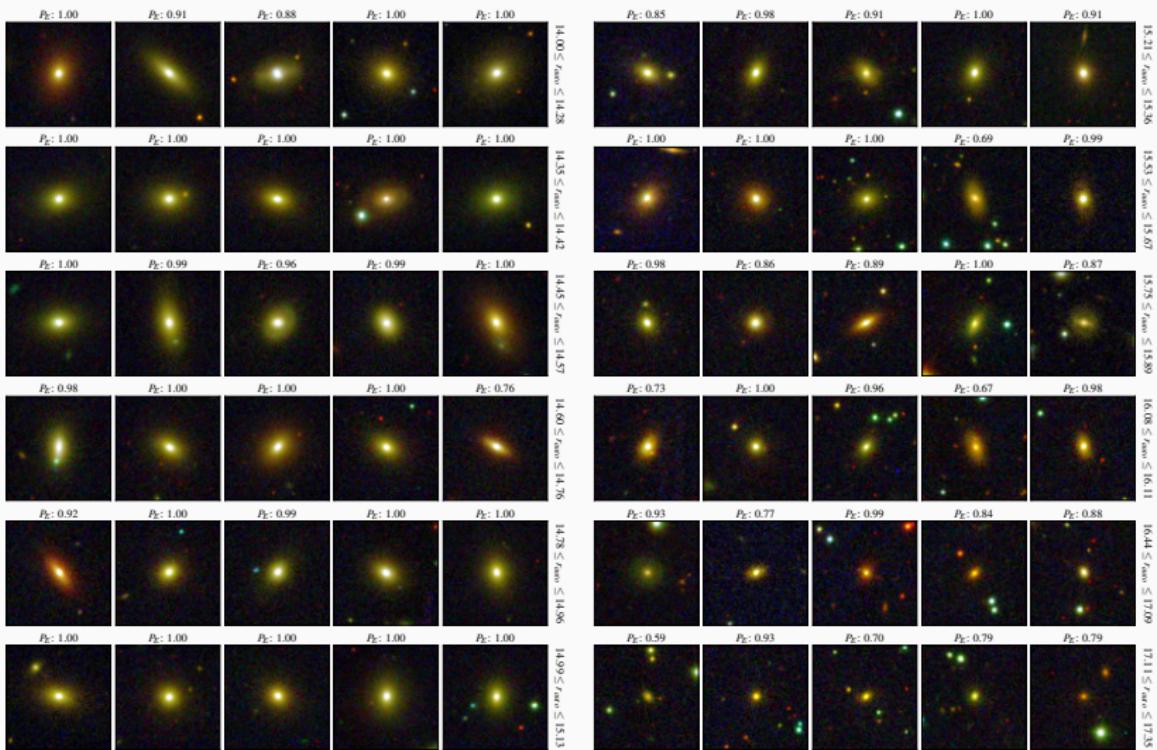


**Figura 14:** Classificação humana

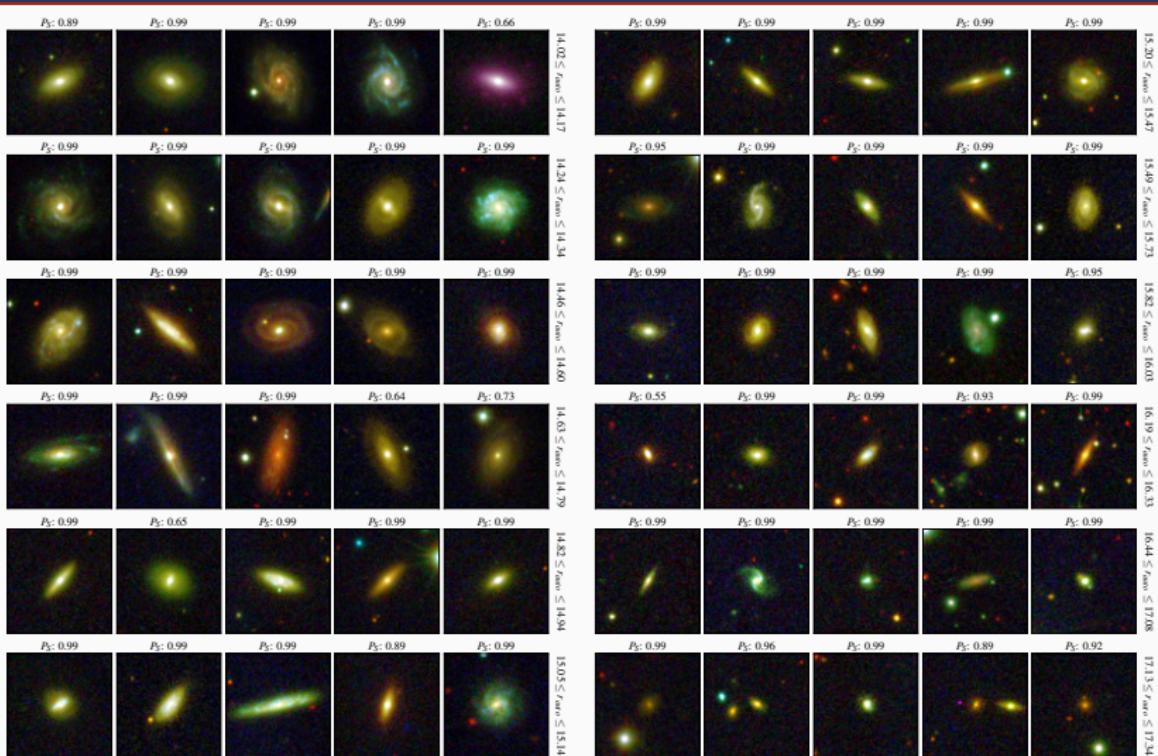


**Figura 15:** Classificação automática

# GALÁXIAS ELÍPTICAS DO CONJUNTO BLIND



# GALÁXIAS ESPIRAIS DO CONJUNTO BLIND



# CONSLUSÃO

- Modelo de *Deep Learning* para classificação E/S.
- Catálogo com classificação de 4000 galáxias ainda não classificadas.
- Artigo em estágio final de escrita.

CRPP-NT-000000

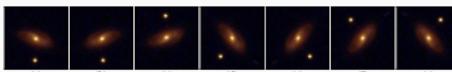


Figura 5: Exemplo do aumento artificial de dados em uma imagem original, mostrada no painel (a). Os painéis (b), (c), (d), (e), (f) e (g) contêm os resultados da equação (3) substituindo  $M$  por diferentes combinações das transformações da equação (2). Em (b)  $M = V$ , em (c)  $M = H$ , em (d)  $M = R(30^\circ)$ , em (e)  $M = VR(30^\circ)$ , em (f)  $M = HR(30^\circ)$  e em (g)  $M = HRV(30^\circ)$ .

### 3.1.1. Agrupamento das bandas para configuração das imagens RGB

Como as imagens do S-PLUS foram obtidas em 12 bandas fotométricas (listadas em [9]), para representá-las no sistema de cor RGB é necessário agrupar as bandas. O SPLUS divide as imagens em 12 bandas, mas o dataset G as gSDSS 00515 e 00660 e cinco bandas mais anais u\_JAVA, J0373, J0395, J0410 e J0430 (as características destes filtros são dadas na Tabela 1 de [9]). Na configuração das bandas em cada pixel, foi feita uma média simples dos valores dos pixels. Depois de realizadas a tais bandas, as imagens são usadas como entrada do programa Trilegy [11].

### 3.1.2. ImageNet

Como já mencionado em trabalhos anteriores, a inicialização dos pesos provenientes de uma rede pré-treinada usando a base de dados ImageNet<sup>8</sup> traz uma grande melhoria na precisão dos resultados da classificação. Essa base de dados possui uma grande variedade de objetos cotidianos, já foi utilizada especificamente para a classificação de objetos astronômicos (veja por exemplo [12]), com excelentes resultados.

O uso de dados pré-treinados para previsões específicas de processamento utilizados principalmente pelos autores de cada rede, este procedimento este foi crucial para garantir um fit competitivo, i.e., não banchmarking original deles redes para o seu uso. Para a rede Trilegy (Seção 3.1), a ordem das bandas foi invertida de RGB para ImageNet, e as bandas foi centralizada zero em relação à ImageNet, sem escalonamentos, ou seja, os pixels de cada banda tiveram o valor da média da banda subtraída e o resultado foi dividido pelo desvio padrão da banda. Para a rede EfficientNetV2 (Seção 3.4), os pixels de entrada foram escalonados entre -1 e 1 em relação a amostra de treino. Para a rede EfficientNet (Seção 3.5), os pixels de entrada foram escalonados entre -1 e 1 em relação a amostra de treino. E, para a rede DenseNet (Seção 3.6), os pixels de entrada foram escalonados entre 0 e 1 e cada banda foi padronizada em relação à ImageNet, isto é, os pixels de cada banda tiveram o

valor da média subtraída e o resultado foi dividido pelo desvio padrão da distribuição da respectiva banda da ImageNet.

### 3.2. Aumento artificial de dados

Aumento artificial de dados [13] é a aplicação de transformações artificiais no conjunto de dados de treinamento, por exemplo rotação, reflexão, translação e mudança de escala. As matrizes da equação (2) definem as transformações usadas.

$$\begin{aligned} R(\theta) &= \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ H &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad V = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (2)$$

onde  $R(\theta)$  é a transformação rotação por um ângulo  $\theta$ ,  $H$  é a transformação reflexão horizontal e  $V$  é a transformação vertical.

Seja  $M$  a matriz das transformações combinadas,  $(x,y)$  a coordenada do pixel da imagem original e  $(x',y')$  a coordenada transformada do pixel, as transformações nas imagens são realizadas através da equação (3), onde  $(x_0, y_0)$  é a coordenada do centro da imagem e as matrizes ao lado de  $M$  são as matrizes translação. Isto é feito para que a transformação  $M$  tenha o centro da imagem como ponto de simetria.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} M \begin{bmatrix} 1 & 0 & -t_x \\ 0 & 1 & -t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3)$$

Além disso, é aplicada uma interpolação bilinear como anti-aliasing [14, 15]. Durante o treinamento da rede, novas imagens de entrada são geradas a cada época a partir da transformação das imagens originais. A Figura 5 mostra a transformação das imagens originais (no painel (a)), e diversas transformações, nos demais painéis, aplicadas substituindo  $M$  da equação (3) por combinações (multiplicação matrixial) das transformações da equação (2). Tais transformações são

<sup>8</sup><https://www.tensorflow.org/tutorials/images/classification>  
<sup>9</sup><https://www.image-net.org/>

Introdução  
○○○○

Conjunto de Dados  
○○○

Métodos  
○○

Resultados  
○○○○○○

Conclusão  
○●

**FIM**

---

**natanael.net**