



Tema: Sistema Inteligente para Busca de Objetos Astronômicos por Similaridade Visual usando Aprendizagem Profunda

OBJETIVOS

Na astronomia moderna, onde o volume de dados gerados está sempre aumentando, a inspeção visual de objetos se torna uma tarefa muito demorada e trabalhosa. Portanto, propomos um Sistema Inteligente capaz de buscar similaridade visual, automatizando uma tarefa que, de outra forma, seria manual.

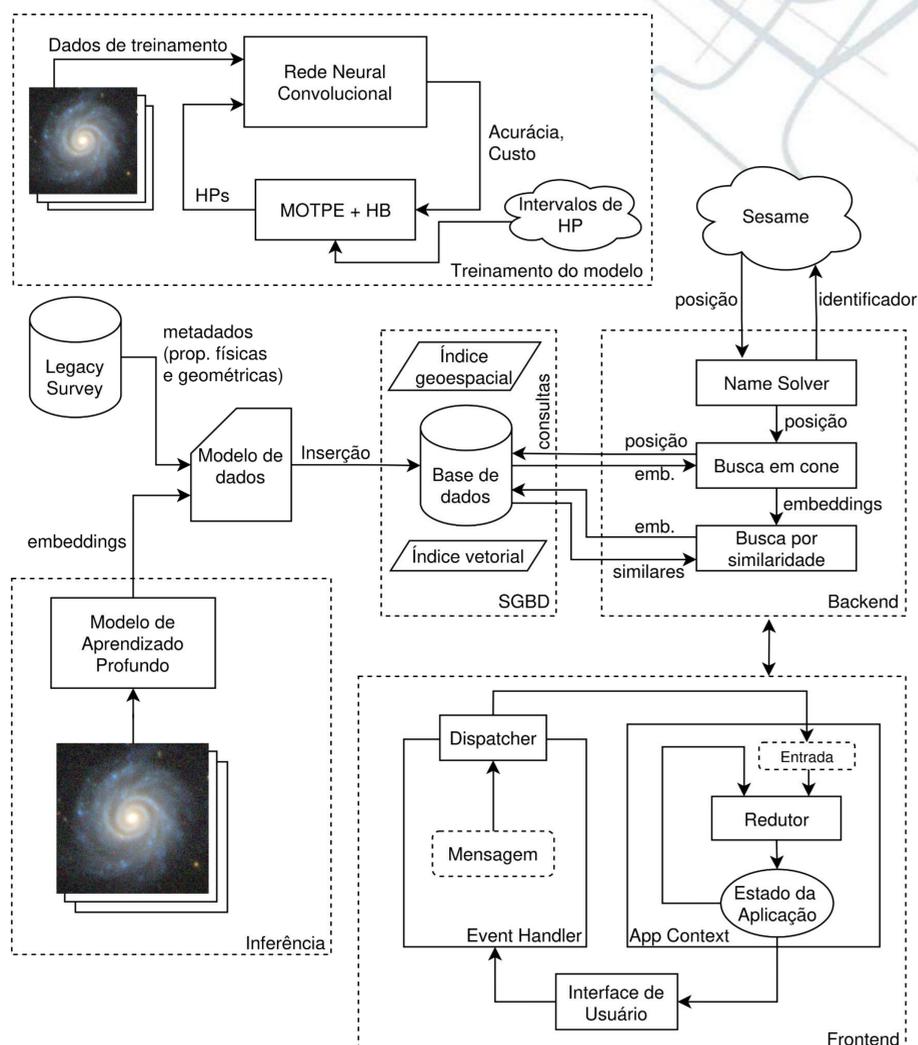
DADOS

O **conjunto de treinamento**, usado no treinamento supervisionado do modelo de aprendizado profundo, contém aproximadamente 400.000 galáxias, consistindo em imagens do Legacy Survey DR10 [1] com seus respectivos rótulos, provenientes do projeto de ciência cidadã GalaxyZoo [2].

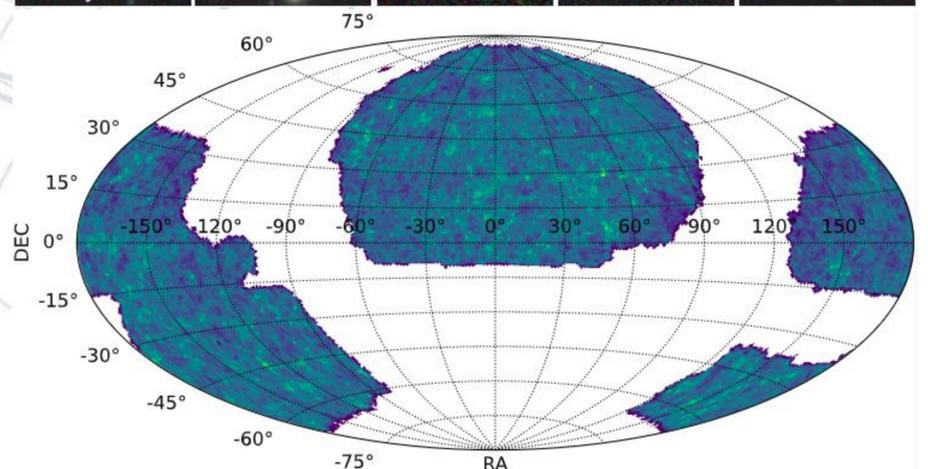
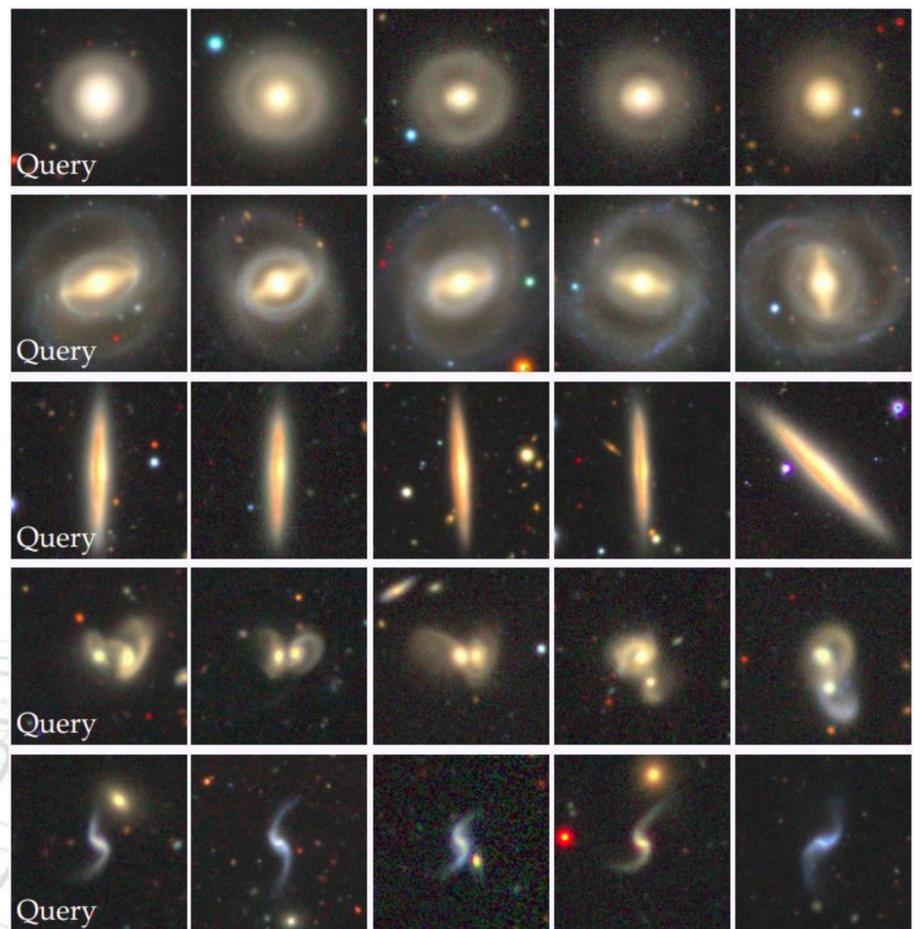
O **conjunto de inferência**, usado para gerar embeddings de imagens, tem aproximadamente 8 milhões de galáxias, consistindo apenas em imagens do Legacy Survey DR10.

MÉTODO

Um modelo de rede neural convolucional (CNN) é treinado usando AutoML e com a função de custo bayesiana Dirichlet-Multinomial. Os embeddings são gerados para o conjunto de inferência, que alimenta um banco de dados com indexação geoespacial e vetorial. O usuário utiliza a interface gráfica do frontend, construído como uma máquina de estados Mealy, para fazer a consulta. O frontend envia o termo de busca ao backend, que resolve o nome enviado, faz uma busca espacial para encontrar o embedding de referência e depois uma busca por similaridade para retornar as galáxias similares.



RESULTADOS



CONCLUSÃO

É possível usar uma CNN treinada a partir dos votos de cientistas cidadãos no projeto GalaxyZoo e com a função de custo bayesiana Dirichlet-Multinomial para gerar representações visuais de imagens, utilizadas para realizar buscas de similaridade em um sistema de recomendação que utiliza recuperação de imagem baseada em conteúdo. No futuro, será feita uma correlação com o levantamento S-PLUS [3]. Este sistema está disponível online em: <https://selene.natanael.net>.

Agradecimentos: recursos computacionais: IAG, CBPF; dados: Projeto GalaxyZoo, S-PLUS; financiamento: CAPES, CNPq, FAPESP (2019/26492-3);

REFERÊNCIAS

- [1] Arjun Dey et al. "Overview of the DESI legacy imaging surveys". In: The Astronomical Journal 157.5 (2019), p. 168.
- [2] Chris J. Lintott et al. "Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey". In: MNRAS (2008).
- [3] C. Mendes de Oliveira et al. "The Southern Photometric Local Universe Survey (S-PLUS): improved SEDs, morphologies, and redshifts with 12 optical filters". In: MNRAS (2019).

Integrantes: Natanael Magalhães Cardoso

Professor Orientador: Prof. Dr. Antonio Mauro Saraiva