



REDE NEURAL CONVOLUCIONAL PARA CLASSIFICAÇÃO MORFOLÓGICA DE GALÁXIAS

Josué Gabriel dos Santos, Rodrigo Lyra.

Engenharias e Ciências Agrárias, Exatas e da Terra
Ciência da Computação - Metodologia e Técnicas da Computação

Atualmente estimasse que existem mais de 170 bilhões de galáxias no universo observável, que são o foco de estudo na astrofísica e cosmologia, áreas quem tem como objetivo o estudo da origem e evolução do universo. A comunidade astronômica já captou imagens de um quarto do céu por meio de poderosos telescópios e a classificação desses objetos tem sido realizada por especialistas em astronomia, e projetos como Galaxy Zoo, que usa crowdsourcing, uma técnica que consiste em utilizar do conhecimento coletivo para resolução de um problema, o Galaxy Zoo realiza isso através de seu site onde eles disponibilizaram questões de múltiplas escolhas referente a morfologia da galáxia apresentada, assim possibilitando que os voluntários possam contribuir nas classificações. O problema é que os métodos anteriormente citados de classificação, já não conseguem atender a grande demanda de dados gerados hoje, que tende a aumentar pelos próximos anos com a construção de novos telescópios como o Large Synoptic Survey Telescope no Chile, que pretende detectar mais de 20 milhões de galáxias, ou Integral Field Units que já gera cerca de 60 GB de dados por noite, que evidencia a necessidade da automatização desse processo. Por causa disso, a solução proposta neste trabalho é a modelagem e treinamento de uma Rede Neural Convolutiva profunda tendo como base a arquitetura da VGG16, pois, é um modelo que tem apresentado bons resultados, visto que ele atribui pesos em regiões da imagem referente a aspectos relevantes. O trabalho também buscou trazer técnicas de aprendizado de máquina e pré-processamento de imagens que visam o destacar a morfologia e consequentemente diminuindo a quantidade de pixels na entrada da rede, para trazer mais eficiência no treinamento e desempenho na classificação. A base de dados usada no projeto é a do Galaxy Zoo disponibilizada em conjunto com Sloan Digital Sky Survey durante um desafio denominado Galaxy Zoo: The Galaxy Challenge, o conjunto contém 61 mil imagens para treinamento com 37 classes de galáxias ao todo. Neste projeto foi utilizado apenas seis classes, por esse motivo foi feito um filtro na base para selecionar as galáxias pertencentes a essas classes, totalizando 53 mil imagens, sabendo que uma galáxia pode pertencer até a 8 classes diferentes segundo o diagrama do Galaxy Zoo e 20% da base é destinada à validação do modelo. No pré-processamento as imagens são redimensionadas, tendo uma diminuição de cerca 70% do seu tamanho inicial, realizado de modo a preservar o objeto de estudo no centro da imagem. Após isso é aplicado diversos filtros como interpolação e labeling, responsáveis pela remoção do fundo, ruídos e outros objetos presentes nas imagens. E por fim é utilizado uma técnica de aumento de dados, por meio da rotação das imagens pelos seguintes ângulos: 0°, 90°, 180° e 270°, tendo como objetivo auxiliar o aprendizado do modelo impedindo que ele fique dependente do ângulo em que as imagens são



apresentadas a rede. Para avaliar a técnicas de pré-processamento utilizadas, é criado uma cópia da base, denominado conjunto A e conta com o redimensionamento e aumento de dados sem os demais filtros, diferentemente do conjunto B que possui todas as técnicas anteriormente citadas. O treinamento do conjunto A levou 18 horas e o B apenas 6 horas, provando que de fato os filtros contribuíram no desempenho do treinamento, os modelos apresentaram resultados satisfatórios com uma precisão de 80%, contudo, as técnicas se provaram não tão robusta, pois, apesar das métricas de precisão e acurácia apresentarem resultados similares a sensibilidade apresentou melhor resultado no conjunto que não usava os filtros.

Palavras-chave: Processamento de imagem; Classificação morfológica de galáxias; Rede Neural Convolutacional.

Programa UNIEDU - Bolsa de Pesquisa Art. 170 e Art. 171 / Governo de Santa Catarina / UNIVALI