

Um Estudo Sobre Identificação de Aglomerados de Galáxias

João Victor Ribeiro de Jesus
Laboratório de Inteligência Aplicada
Universidade do Vale do Itajaí – UNIVALI - SC
joaovictor@edu.univali.br

Rodrigo Lyra
Laboratório de Inteligência Aplicada
Universidade do Vale do Itajaí – UNIVALI – SC
rlyra@univali.br

ABSTRACT

Galaxy clusters are fundamental blocks of our universe, like stars and galaxies. The study of galaxy clusters has shown complexity due to a large number of stars superimposed on these clusters, which emit light and block the clear optical band visualization of the phenomenon. In this paper, we present the implementation of an algorithm capable of classifying and remove the stars of spatial images. The purpose is to improve not only the existing studies around galaxy clusters but also to assist the developing in astronomic researches with available treated spatial images.

KEYWORDS

Digital image processing, spatial image, galaxy cluster, classification algorithm, astronomical objects.

1 Introdução

Os aglomerados de galáxias constituem um importante campo de estudo na área da astronomia, astrofísica e da cosmologia, pois os mesmos auxiliam na compreensão de como a matéria escura se comporta e evolui, dando assim, pistas de como é constituída a estrutura do Universo como um todo [1].

Atualmente existem diversos métodos para o pré-processamento e processamento de imagens. Entre eles a limiarização, segmentação baseada em bordas e regiões, detecção de bordas, classificação e muitas outras [2][3]. Além dos métodos existentes, a área de Inteligência Artificial (IA) também tem um papel importante, tanto na segmentação, quanto na análise e reconhecimento das imagens com a utilização de métodos de aprendizado de máquina para, por exemplo, classificar aquilo que pode ser considerado estrela ou galáxia [4].

Tendo em vista a complexidade observacional dos aglomerados de galáxias e da relevância deles perante o entendimento das forças que regem o universo, é proposta uma solução que tem por objetivo principal facilitar o estudo desses aglomerados, por meio da implementação desses métodos de processamento de imagens, a fim de criar um algoritmo capaz de extrair estrelas em imagens espaciais para posteriores estudos dos aglomerados.

2 Solução Proposta

A proposta aqui apresentada visa desenvolver um algoritmo por meio da implementação de métodos de processamento de imagem, como o de segmentação, extração de características e classificação para o agrupamento de objetos com as mesmas

características, na tentativa de automatizar o processo de subtração de estrelas em imagens espaciais. A solução busca identificar galáxias e estrelas em imagens e remover os objetos classificados como estrelas da imagem final gerada.

3 Projeto

Para o desenvolvimento do algoritmo proposto será utilizada a linguagem de programação Python juntamente com as bibliotecas OpenCV (para o processamento de imagem), Pillow (como alternativa à biblioteca OpenCV), Scipy (para a programação de métodos científicos), Numpy (para a criação de matrizes e vetores, além da elaboração de cálculos matemáticos) e Matplotlib (para a plotagem de gráficos e representações visuais dos dados obtidos).

Anterior a implementação do algoritmo responsável pela remoção das estrelas, foi implementado o algoritmo responsável por converter as posições dos pixels da imagem em coordenadas RA e DEC para posterior validação dos resultados. As coordenadas RA e DEC dizem respeito ao sistema equatorial universal de coordenadas, muito usado na astronomia devido a este sistema ser capaz de indicar a posição de um astro de forma idêntica para todos os observadores na Terra, isto é, as coordenadas equatoriais (RA e DEC) de um dado astro celeste são idênticas para qualquer observador na superfície da terrestre [5].

A implementação do algoritmo responsável por remover as estrelas das imagens espaciais pode ser dividido em quatro etapas, sendo essas: Pré-processamento, Extração de características, Classificação e Limpeza.

Primeiramente, na etapa de pré-processamento, dá-se início à transformação da imagem original obtida em uma nova imagem em escala de cinza e, com base nesta, é aplicado o algoritmo de limiarização para a transformação desta, em uma nova imagem binária, onde os pixels brancos passam a representar parte dos objetos presentes na imagem e os pixels pretos passam a representar o plano de fundo, conforme ilustra a Figura 1.



Figura 1: Imagem original, imagem em escala de cinza e imagem limiarizada, respectivamente.

Após a etapa de pré-processamento foi calculado, para cada objeto segmentado, os valores de compactidade, excentricidade, retangularidade, os dois maiores eixos de cada objeto, assinatura, a variância das intensidades dos pixels que constituem cada objeto e, por fim, a coordenada da centroide de cada objeto.

A classificação foi realizada através de um algoritmo de agrupamento de dados não supervisionado denominado K-means com o objetivo de gerar dois grupos de objetos distintos a partir das características extraídas desses. Por fim, na etapa de limpeza foi atribuído à todos os pixels que constituem os objetos classificados como estrela, a cor preta apagando, assim, esses da imagem.

4 Resultados

Para a validação dos resultados foi realizada a aplicação do algoritmo em imagens com objetos previamente classificados como estrelas e galáxias. Primeiramente foi necessário a aplicação do algoritmo capaz de requisitar a imagem da base de dados Sloan Digital Sky Survey (SDSS-DR15) selecionando a coordenada central da imagem em RA e DEC, a altura e a largura da imagem. Em seguida, foi chamada a função responsável por converter as coordenadas dos pixels da imagem em coordenadas RA e DEC. Assim, para a centroide de cada objeto obteve-se, também, as coordenadas do mesmo em RA e DEC.

O algoritmo de validação buscou, a partir da base SDSS Sky Server (DR15), a classificação real de cada objeto passando por parâmetro as coordenadas RA e DEC de cada objeto e com base nelas retornou o valor STAR ou GALAXY de acordo com o valor na base de dados.

A aplicação do K-means foi realizada com diferentes conjuntos de características até ser encontrado o conjunto que gerou os melhores resultados e, com esse conjunto, foram definidas as características necessárias para a classificação.

A Tabela 1 apresenta os resultados em relação aos acertos da classificação obtida através da aplicação do K-means com as características de assinatura, compactidade, excentricidade, retangularidade e variância de cada objeto, para 3 imagens diferentes, ambas com tamanho de 800 x 800 pixels.

Tabela 1: Resultados obtidos para objetos classificados como estrelas e galáxias.

Imagem	Acertos	Erros	Acertos (%)
Imagem 1	109	20	84%
Imagem 2	109	17	86%
Imagem 3	88	11	88%

Para cada imagem foi gerada uma matriz de confusão, detalhando ainda mais os resultados obtidos pela classificação. A Tabela 2 apresenta os valores de Verdadeiro-Positivo, Verdadeiro-Negativo, Falso-Positivo e Falso-Negativo para os resultados aplicados em cada imagem.

Tabela 2: Matriz de Confusão para as três imagens utilizadas.

Imagem	VP	VN	FP	FN
Imagem 1	48	58	18	2
Imagem 2	59	50	16	1
Imagem 3	42	46	9	2

Na Tabela 2, VP é referente às galáxias classificadas como galáxias, VN referente às estrelas classificadas como estrelas, FP referente às estrelas classificadas como galáxias e FN referente às galáxias classificadas como estrelas.

O resultado do algoritmo proposto é apresentado na Figura 2, onde é apresentada a imagem original e a imagem resultante do processamento realizado. Além da remoção das estrelas na imagem o algoritmo também gera uma imagem com um retângulo envolvente em cada objeto classificados como galáxias, conforme mostra a Figura 3.

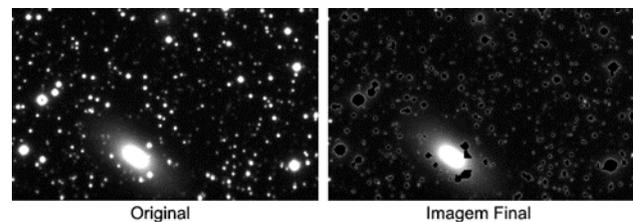


Figura 2: Imagem espacial com estrelas removidas.

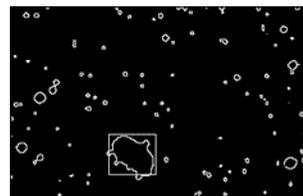


Figura 3: Imagem com retângulo envolvente em torno dos objetos classificados como galáxia.

5 Considerações Finais

Pode-se concluir, a partir dos resultados obtidos, que foi possível classificar de forma competente as galáxias e estrelas presentes nas imagens acreditando que a aplicação do algoritmo desenvolvido possa contribuir com a área da astronomia para a busca de novas galáxias e aglomerados de galáxias.

REFERÊNCIAS

- [1] E. Rasia, M. Meneghetti, R. Martino, S. Borgani, A. Bonafede, K. Dolag, S. Ettori, D. Fabjan, C. Giocoli, P. Mazzotta, J. Merten, M. Radovich, and L. ornatore. 2012. Lensing and x-ray mass estimates of clusters(simulations). *New Journal of Physics* 14, 5 (May 2012), 1 – 17. <https://doi.org/10.1088/1367-2630/14/5/055018>.
- [2] Hélio Pedrini and William Robson Schwartz (Eds.). 2000. *Análise de Imagens Digitais*. Cengage Learning, São Paulo.
- [3] Rafael C. Gonzales and Richard E. Woods (Eds.). 2000. *Processamento de imagens digitais*. Blucher, Brasil.
- [4] A. B. Krug, A. Parraga, F. Lorenzi, M. Nicolao, and A. S. Morales. 2008. Análise e Reconhecimento de Padrões Usando Processamento de Imagens e Inteligência Artificial. *Revista de Iniciação Científica da ULBRA* 7 (Jan 2008), 10.
- [5] Kepler de Souza Oliveira Filho and Maria de Fátima Oliveira Saraiva. 2019. *Sistemas de Coordenadas*. Retrieved October 25, 2019 from <http://astro.if.ufrgs.br/coord.htm>