Classificação e Contagem de Bovinos em Imagens Aéreas utilizando Visão Computacional e Aprendizagem de Máquina

Juliana Midlej do Espírito Santo[1], Jorge Lima de Oliveira Filho[1]

[1] Departamento de Ciências Exatas e Tecnológicas
Universidade Estadual de Santa Cruz (UESC) – Ilhéus, BA – Brazil

jmesanto91@gmail.com, jlofilho@uesc.br

Abstract. *This article exposes a method that uses computer vision and machine learning to allow detection, classification and counting of white cattle in aerial images obtained using a drone. The Python programming language was used with the OpenCV (Open Source Computer Vision) library for image pre-processing and segmentation and the scikit-learn library for image classification. The method was tested on images divided between the dry and rainy periods to assist in the pre-processing and segmentation stages and obtained efficiency of 95.2% in the detection and accuracy of 90,1% in the classification of white cattle.*

Resumo. *Este artigo expõe um método que faz uso da visão computacional e da aprendizagem de máquina para possibilitar a detecção, classificação e contagem de bovinos brancos em imagens aéreas obtidas através de um drone. Foram utilizadas a linguagem de programação Python e as bibliotecas OpenCV (Open Source Computer Vision) para o pré-processamento e segmentação das imagens e scikit-learn para a classificação com regressão logística. O método foi testado em imagens divididas entre os períodos seco e chuvoso para auxiliar nas etapas de pré-processamento e segmentação e obteve eficiência de 95,2% na detecção e acurácia de 90,1% na classificação de bovinos brancos.*

1. Introdução

A criação de bovinos no Brasil está presente desde o início de sua colonização até os dias atuais. Segundo Batistella et al. (2011), em meados dos anos 1970, ela representava uma das mais importantes atividades econômicas da agropecuária brasileira. A bovinocultura ocupa cerca de um quinto do território nacional e o país possui um dos maiores rebanhos mundiais do produto, sendo também um dos líderes em exportação e consumo.

O desenvolvimento de um método capaz de classificar e contar bovinos brancos através de imagens aéreas obtidas por meio de drones tem por objetivo tornar mais fácil o controle dos bovinos nas propriedades. Além disso, segundo Oliveira et al. (2002), raças indianas, como a Nelore, possuem destaque na pecuária brasileira e tiveram uma melhor adaptação ao ambiente tropical do que raças europeias também introduzidas no território. Um grande número dos bovinos no Brasil pertence a raça Nelore, a qual possui pelagem curta e branca e está amplamente presente nas imagens estudadas.

*Os autores gostariam de agradecer o auxílio financeiro da FAPESB (pedido nº 4470/2019)

O método desenvolvido atua sobre dois conjuntos de imagens aéreas retiradas entre 20 e 46 metros de altura divididas de acordo com a cor do solo entre imagens do período seco ou do período chuvoso. No método, foram utilizados filtros provenientes da biblioteca de visão computacional OpenCV [Bradski 2000] para o pré-processamento das imagens. Os filtros realizam os processos de redução dos ruídos, limiarização e detecção de contornos que auxiliam na segmentação da imagem. Após a segmentação, foi realizada uma classificação das imagens com o uso de regressão logística [Venticinque 2007]. A regressão logística estuda a relação entre uma variável resposta e uma ou mais variáveis independentes. Na regressão logística as variáveis dependentes estão dispostas em categorias e a resposta é expressa por meio de uma probabilidade de ocorrência.

Este artigo está organizado em cinco seções contando com esta introdução. A segunda seção traz trabalhos correlatos realizados na área da detecção e classificação de objetos a partir de imagens aéreas. A terceira seção apresenta a metodologia para realização do método em cada uma de suas etapas, desde a aquisição das imagens até sua classificação. A quarta seção apresenta os resultados obtidos e as dificuldades enfrentadas. E a quinta e última seção discute as conclusões deste artigo e os trabalhos futuros a serem desenvolvidos.

2. Trabalhos Correlatos

Há diversos trabalhos na linha de pesquisa da detecção e classificação de objetos. Dentre eles, Ribeiro et al. (2019) realizou um trabalho sobre a utilização da visão computacional para detecção de bovinos através de imagens aéreas. Na pesquisa realizada por eles foram utilizados filtros de diminuição de ruídos, limiarização e detecção de bordas da biblioteca OpenCV para realizar o pré-processamento e a segmentação das imagens e detectores de objetos binários para realizar a contagem das áreas segmentadas. Os resultados alcançados pelos autores foram obtidos para oito diferentes imagens, dentre elas, imagens de campo aberto, com bovinos próximos ou distantes e bovinos em confinamento. Os melhores resultados foram alcançados para a detecção de bovinos brancos em campo aberto. A pesquisa desenvolvida neste trabalho foi realizada com 108 imagens de bovinos em campo aberto divididas nos períodos seco e chuvoso para auxiliar em seu pré-processamento e detecta apenas bovinos brancos. Além disso, após a segmentação das imagens, foi realizada uma classificação utilizando aprendizagem de máquina.

A automatização da detecção e contagem de animais por fotografias aéreas realizada por Sirmacek et al. (2012) também propõe o uso de técnicas de processamento de imagens e visão computacional. Em sua pesquisa, foi utilizado o método FAST (*features from accelerated segment test*), especialmente desenvolvido para detecção de bordas por Rosten et al. (2010), e utilizado para extrair locais de interesse em imagens. Os locais detectados são utilizados como pontos de observação para gerar uma função de densidade de probabilidade. Após o uso dessa função em uma imagem, é realizada a segmentação e aplicação de uma teoria de grafos baseada em detecção de animais para obter o número real de animais presentes.

Groom et al. (2007) mapearam aves de diferentes espécies e gêneros por meio de fotografias retiradas a partir de uma câmera digital, desenvolveram e compararam a precisão dos métodos de mapeamento de imagem com base em pixel e objeto. Para o mapeamento com base em pixel, devido sua simplicidade, foi utilizado apenas o processo de *level-slicing*. Para as imagens mapeadas com base em objeto foi utilizado o processo de segmentação para derivar os locais de interesse da imagem, além da definição de características do ambiente, como a forma, cor e textura, utilizadas para rotular os objetos encontrados em classes pertencentes.

Por fim, Maussang et al. (2015) apresentaram em seu artigo diferentes abordagens para observar e classificar o ambiente marinho, composto, por exemplo, por aves, tartarugas e pedras, a partir de imagens aéreas e utilizando algoritmos de aprendizagem. Dentre essas abordagens, destacaram-se duas formas para melhor resolução do problema proposto, a primeira é baseada na extração de características que torne possível seu reconhecimento usando um modelo de classificação e a segunda consiste na aprendizagem profunda. O primeiro método apresentado é semelhante ao algoritmo desenvolvido neste trabalho, o qual visa extrair da imagem, após a fase de pré-processamento, locais de interesse para serem classificados através de regressão logística.

3. Metodologia

Para desenvolver o método apresentado neste trabalho foi levado em consideração uma série de passos que serão apresentados nesta seção. Para facilitar o entendimento, os passos serão discutidos em subseções que são, respectivamente, a aquisição e seleção das imagens, o pré-processamento, a segmentação e a classificação. O método desenvolvido utilizou a linguagem de programação Python, a biblioteca de visão computacional OpenCV e a biblioteca de aprendizagem de máquina scikit-learn [Pedregosa et al. 2011].

3.1 Aquisição e Seleção das Imagens

As imagens aéreas foram obtidas através de um drone e foram selecionadas considerando sua qualidade, ângulo e altura relativa. Foram priorizadas as imagens com ângulo mais próximo a 90 graus em relação ao solo e com altura entre 20 e 46 metros. A altura relativa da imagem foi obtida através do acesso aos metadados das imagens aéreas utilizando a ferramenta Exiftool [Harvey 2013]. A seleção dos períodos aos quais a imagem pertence foi realizada de forma manual e foram observadas para a seleção a cor do solo e vegetação presentes. A Figura 1 mostra duas imagens retiradas pelo drone e selecionadas, respectivamente, para os períodos seco e chuvoso.

Figura 1. Imagens selecionadas para os períodos seco e chuvoso

3.2 Pré-processamento

Segundo Marengoni e Stringhini (2009), para o melhor funcionamento dos processos de visão computacional, tais como a segmentação, é necessária uma etapa de pré-processamento que envolve a utilização de filtros para remoção de ruídos e conversão da imagem para o formato necessário. As etapas do pré-processamento do algoritmo desenvolvido são explicadas a seguir:

1) Suavização da imagem com utilização de um filtro gaussiano. Os filtros gaussianos são filtros de média e são utilizados para suavizar a imagem de forma ponderada e simétrica [Marengoni e Stringhini 2009];
2) Conversão do sistema de cores RGB (vermelho, verde, azul) para o sistema de cores HSV (matiz, saturação, valor). O sistema de cores HSV é um modelo para representar o espaço de cores semelhante ao modelo de cores RGB. Como o canal matiz modela o tipo de cor, é muito útil em tarefas de processamento de imagem que precisam segmentar objetos com base em sua cor;
3) Limiarização da imagem, a qual destaca o objeto de interesse ao escolher valores mínimos e máximos nos canais do sistema HSV os quais apenas o objeto desejado seja destacado em relação ao restante da imagem;
4) Erosão e dilatação da imagem. Os processos de erosão e dilatação causam a remoção de ruídos restantes após a limiarização [Rosebrock 2015];
5) Detecção de contornos. Os contornos podem ser explicados simplesmente como uma curva que une todos os pontos contínuos (ao longo do limite), com a mesma cor ou intensidade. Os contornos são uma ferramenta útil para análise de formas e detecção e reconhecimento de objetos.

Para obter o fluxograma de pré-processamento que alcançasse bons resultados na detecção e segmentação de bovinos brancos em todo o conjunto de imagens, foram realizados diversos testes com diferentes parâmetros e filtros da biblioteca OpenCV. Para determinar os parâmetros de limiarização, ou seja, os valores máximos e mínimos dos canais no sistema HSV para destacar locais de interesse, os testes foram realizados em um conjunto de 20 imagens escolhidas manualmente de forma que melhor representassem todo o universo de imagens estudado. Os parâmetros que obtiveram os melhores resultados para essas imagens foram escolhidos para o método desenvolvido.

No método proposto, os parâmetros de limiarização foram utilizados de forma mais abrangente, destacando regiões com bovinos brancos e regiões com características próximas, o que foi essencial para realizar o treinamento do algoritmo de classificação pelo aumento do número de imagens segmentadas e, portanto, pela melhoria da proporção entre bovinos e não bovinos. Esse uso da limiarização e a diferença entre o solo dos períodos seco e chuvoso justifica a existência dessa divisão. Para abranger um maior número de regiões no período chuvoso, assim como no seco, e manter uma boa demarcação da área do bovino, as variações nos parâmetros da limiarização devem ser diferentes.

A Figura 2 demonstra todo o pré-processamento de uma imagem, sendo a imagem 0 a entrada do nosso algoritmo, uma imagem sem modificações. A partir da

entrada, segue-se: a imagem 1 mostra a suavização; a imagem 2 mostra a conversão para o sistema de cores HSV; a imagem 3 mostra a limiarização utilizando parâmetros escolhidos através de alguns testes apenas para demonstração da função; a imagem 4 mostra o resultado dos processos de erosão e dilatação, após isso nota-se uma diminuição no ruído presente na imagem anterior; a imagem 5 destaca os contornos obtidos a partir da imagem 4 sobre a entrada do nosso algoritmo.

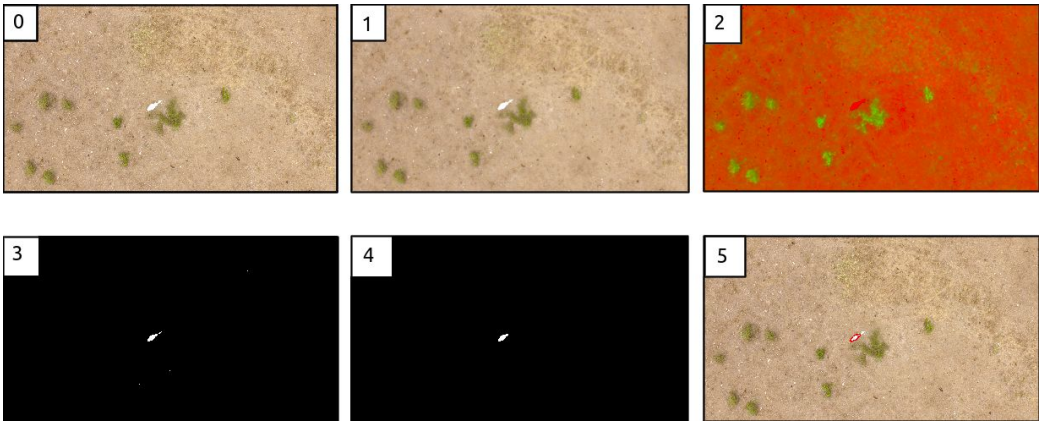

Figura 2. Etapas do pré-processamento da imagem

Na Figura 3 podemos observar o fluxograma do método proposto apresentando os processos pelos quais a imagem passa a partir do pré-processamento até a sua classificação. A divergência no fluxograma ocorre no processo de limiarização e refere-se aos limites escolhidos nos canais do sistema HSV de acordo com o período.

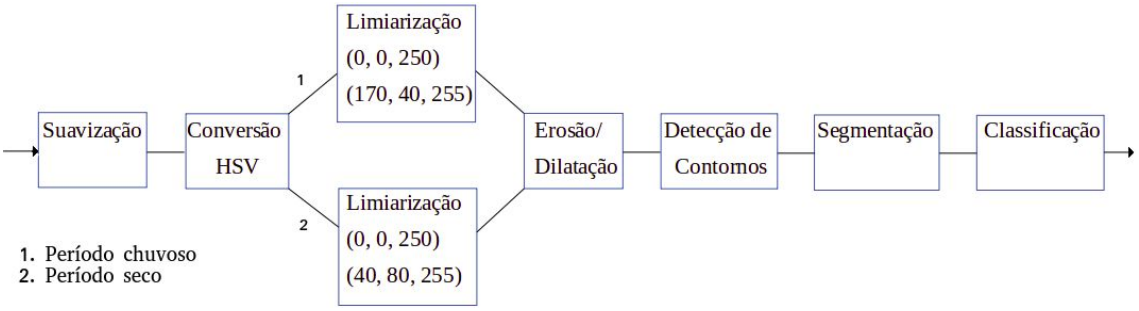

Figura 3. Fluxograma de detecção e classificação de bovinos brancos

3.3 Segmentação

A segmentação refere-se ao processo de dividir a imagem em regiões que, esperançosamente, correspondem a unidades estruturais na cena ou distinguem objetos. Com a extração de locais de interesse e divisão da imagem, melhora-se o desempenho do algoritmo de classificação utilizado e diminui-se o esforço computacional necessário. No método proposto neste trabalho, para cada contorno encontrado após o pré-processamento é verificado se o tamanho em pixels do raio do círculo circunscrito a ele corresponde a um objeto de interesse, se corresponder, a região é segmentada e passará para etapa de classificação.

3.4 Classificação

A classificação dos objetos foi realizada através de regressão logística após a segmentação das imagens. As imagens resultantes do processo de segmentação representam regiões de interesse menores, viabilizando assim o uso e melhor eficácia da regressão. Antes de iniciar a classificação, todas as imagens resultantes são rotuladas manualmente com o auxílio do editor labelImg [Tzutalin 2015] e redimensionadas de forma a se obter um padrão. Neste trabalho, as variáveis independentes da regressão correspondem aos pixels das imagens segmentadas e a resposta obtida para essas variáveis é a probabilidade dessa imagem ser um bovino.

Foram utilizadas 3701 imagens como entrada no algoritmo de regressão logística divididas entre imagens de treino e teste nas quais o número de imagens de teste corresponde a um quinto do número de imagens total. Dentre as 3701 imagens, 72.3% são imagens de bovinos brancos e o restante corresponde principalmente ao solo ou a partes menores do animal separadas durante o pré-processamento, as quais não são classificadas como bovinos para que um único animal não seja classificado mais de uma vez. As imagens passam pelo processo de estandardização e a avaliação da regressão é feita através do método de *cross-validation*.

4. Resultados Obtidos

Para validar o algoritmo proposto, foram utilizadas imagens aéreas de 20 a 46 metros de altura, todas em campo aberto, totalizando 67 imagens no período de chuva e 41 imagens no período de seca. A Tabela 1 exibe os resultados obtidos na detecção dos bovinos com este conjunto de imagens. O número de bovinos real é representado por Bovinos Existentes, o número de bovinos que não foram encontrados na etapa de pré-processamento é representado por Não Encontrados e os bovinos que foram unidos a outros nesta mesma etapa são os Bovinos Próximos, de forma que, se há n bovinos unidos em uma imagem, o valor somado a Bovinos Próximos é $n-1$. Tal cálculo visa facilitar o estudo dos bovinos que podem ser corretamente detectados e classificados, o que não acontece no caso dos Bovinos Próximos devido a classificação ser binária, mesmo que hajam 3 bovinos detectados juntos em uma imagem, apenas 1 pode ser classificado.

Tabela 1. Detecção de bovinos em imagens aéreas com altura de 20 a 46 metros

Período	Bovinos Existentes	Bovinos Próximos	Não Encontrados
Chuvoso	730	14	1
Seco	1909	60	51

Apesar de possuir um menor número de images, as imagens do período seco possuem, em média, mais bovinos do que as imagens do período chuvoso, o que causou o aumento desproporcional do número de Bovinos Próximos nesse período. Além disso, as imagens do período seco possuem também um maior número de bezerros que, devido ao seu tamanho, são mais difíceis de detectar, causando também um aumento significativo e desproporcional dentre os bovinos Não Encontrados em relação ao número de Bovinos Existentes quando comparado com o período chuvoso.

O percentual de reconhecimento de bovinos no pré-processamento e segmentação da imagem, ou seja, o número total de bovinos que conseguimos detectar sobre o total de bovinos real corresponde a 98%. Entretanto, como a classificação é binária e os bovinos detectados juntos só podem ser classificados como zero ou um, é importante notar que o percentual de bovinos detectados que podem ser classificados é, na verdade, 95,2%. Neste último cálculo, consideramos bovinos que foram unidos a outros, Bovinos Próximos, como bovinos Não Encontrados. A regressão logística treinada e testada sobre os dois períodos juntos resulta em uma acurácia de 90.1% e precisão de 92,7%.

O trabalho realizado obtém os melhores resultados para detecção de bovinos brancos em imagens nas quais o bovino está mais disperso e o ângulo da fotografia é próximo de 90 graus em relação ao solo. As maiores dificuldades e principais causas de bovinos não encontrados estão em casos de detecção de bezerros ou quando obstáculos, como árvores, impedem a visão completa do animal. Além disso, a proximidade dos bovinos pode fazer com que eles sejam demarcados juntos e como a classificação é binária ela não lida com esses casos de maneira eficiente. Outro problema comum é a separação da cabeça do bovino de seu corpo, o que torna aprendizagem mais complexa, pois para manter a contagem de bovinos correta e impedir que o mesmo bovino seja contado duas vezes fez-se necessário que apenas a cabeça não representasse um bovino.

5. Conclusão e Trabalhos Futuros

O método proposto neste artigo consegue abranger uma boa margem de altura e é muito eficiente na detecção e classificação de bovinos brancos nos casos em que as imagens apresentam bovinos dispersos e o ângulo da fotografia é próximo a 90 graus. A execução desse método em imagens dos diferentes períodos permitiu ver também as principais dificuldades enfrentadas na construção e aprimoramento de um algoritmo para detecção e classificação de bovinos e sua eficiência, o que é de extrema relevância para continuidade desta pesquisa. Os resultados obtidos com a utilização desse método podem ser melhorados e as principais dificuldades podem ser estudadas e, possivelmente, resolvidas com o uso da aprendizagem profunda.

Como trabalho futuro é prevista a utilização de redes neurais convolucionais para, após o pré-processamento e segmentação das imagens, assegurar a veracidade da detecção com métodos de classificação binária mais eficientes. É possível também, havendo tempo e dados suficientes, fazer uso de uma rede neural convolucional que realize a detecção de objetos através do deslizamento de janelas sobre a imagem, como a rede neural YOLO (You Only Look Once) [Redmon et al. 2016]. Dessa maneira, a proximidade dos bovinos e a classificação dos objetos que podem ser confundidos com eles poderão ser melhor tratadas e a eficiência do método tende a melhorar.

Referências

Batistella, M. Andrade, R. G., Bolfe, E. L., Victoria, D. D. C. e Da Silva, G. B. S. (2011). Geotecnologias e gestão territorial da bovinocultura no Brasil. Em Embrapa Territorial-Artigo em periódico indexado (ALICE), páginas 251-260.

Bradski, G. (2000). The OpenCV Library. Em Dr. Dobbs Journal.

Groom, G., Petersen, I. K. e Fox A. D. (2007). Sea bird distribution data with object based mapping of high spatial resolution image data. Em Challenges for earth observation-scientific, technical and commercial. Proceedings of the Remote Sensing and Photogrammetry Society Annual Conference.

Harvey, P. (2013). “ExifTool: Read, write and edit meta information”, https://exiftool.org, Fevereiro.

Marengoni, M. e Stringhini, S. (2009). Tutorial: Introdução à visão computacional usando opencv. Em Revista de Informática Teórica e Aplicada, páginas 125-160.

Maussang, F., Guelton, L., Garello, R., e Chevallier, A. (2015). Marine life observation using classification algorithms on ocean surface photographs. Em OCEANS 2015-Genova. IEEE.

Oliveira, J. H. F., Magnabosco, C. D. U. e Borges, A. D. S. (2002), Nelore: Base Genética e Evolução Seletiva no Brasil, Planaltina, 1a edição.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... e Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Em Journal of machine Learning research.

Redmon, J., Divvala, S., Girshick, R. e Farhadi, A. (2016). You only look once: Unified, real-time object detection. Em Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, páginas 779-788.

Ribeiro, N. G. Vi., Guedes, G. B. e Barbieri, T. T. (2019). Aplicação de algoritmos de visão computacional na contagem de gado por meio de processamento de imagens aéreas. Em Revista Eletrônica Argentina-Brasil de Tecnologias da Informação e da Comunicação.

Rosebrock, A. (2015) “OpenCV Track Object Movement”, https://www.pyimagesearch.com/2015/09/21/opencv-track-object-movement, Março.

Rosten E., Porter R. e Drummond T., "Faster and Better: A Machine Learning Approach to Corner Detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105-119, Jan. 2010, doi: 10.1109/TPAMI.2008.275.

Sirmacek, B., Wegmann, M., Cross, A. D. P., Hopcraft, J. G. C., Reinartz, P. e Dech, S. (2012). Automatic population counts for improved wildlife management using aerial photography. Em 6th International Congress on Environmental Modelling and Software.

Tzutalin, D. (2015). “LabelImg”, https://github.com/tzutalin/labelImg, Abril.

Venticinque, E. M., Carneiro, J. S., Moreira, M. P., & Ferreira, L. (2007). O uso de regressão logística para espacialização de probabilidades. Em Megadiversidade.