

# PLANO DE TRABALHO

**Título do Projeto:** Visão Computacional e Aprendizagem Automática para Aplicações em Agropecuária e Ciências Forenses.

**Título do Plano de Trabalho:** Desenvolvimento de um protótipo de software para a contagem de bovinos por meio do método Faster R-CNN.

**Orientador:** Hemerson Pistori ([pistori@ucdb.br](mailto:pistori@ucdb.br))

**Coorientador:** Gilberto Astolfi ([gilberto.astolfi@ifms.edu.br](mailto:gilberto.astolfi@ifms.edu.br))

**Orientado:** João Vitor de Andrade Porto **RA:** 170291

**Curso:** Engenharia de Computação **Semestre:** 5º semestre

## *Resumo*

A proposta do plano de trabalho contempla o desenvolvimento de tecnologias aplicadas à agropecuária, buscando aumentar a qualidade e rapidez na contagem de bovinos, que atualmente é feita de maneira manual ou com auxílio de sensores. Com base nisso, propomos o desenvolvimento de um software baseado em visão computacional e redes neurais de aprendizagem profunda para contagem automática de bovinos. Pretendemos realizar a contagem por meio de imagens capturadas por VANTs.

## **1. Antecedentes e Justificativa do problema a ser abordado**

O mercado de carne bovina brasileiro é um dos setores que mais gera renda para o país, hoje, este mercado ocupa o segundo lugar mundial em relação à exportação de carne. Analisando o rebanho bovino brasileiro, nas últimas quatro décadas, houve um crescimento considerável do número animais, contrapondo-se às áreas de pastagem que não acompanharam a demanda para oferta de carne e espaço de criação. Este desequilíbrio, mesmo que favorecido pelo avanço tecnológico da produção, afeta diretamente na qualidade do produto oferecido como também aumenta o preço para o consumidor final (EMBRAPA, 2018).

No Brasil a contagem dos bovinos se dá de forma manual por meio de afunilamento do gado para dentro de um curral necessitando de muita mão de obra para minimizar as falhas,

que se fazem cada vez mais presentes com o aumento do número de cabeças de gado em espaços cada vez menores (BERNARDES, 1995).

O emprego da tecnologia de visão computacional pode ajudar a minimizar esse problema, uma vez que um número grande de bovinos podem ser contados de maneira instantânea e automática usando Veículo Aéreo Não Tripulado (VANT) e, dependendo da precisão do método adotado, falhas na contagem podem ser evitadas. Além disso, os custos com mão de obra podem ser reduzidos, já que não haveria mais a necessidade de recolher os bovinos em um curral. Essa proposição se baseia no fato de que a visão computacional já vem sendo usada com sucesso na biologia para classificação e contagem de indivíduos (WEINSTEIN, 2018).

Levando em consideração a importância comercial da carne bovina para o cenário nacional e a maneira como a contagem é realizada, foi levantada a hipótese da utilização de drones para facilitar a obtenção de imagens dos bovinos nos latifúndios por meio de capturas aéreas de imagens, agilizando o processo e apresentando menor gasto junto de maior versatilidade que os métodos tradicionais (CASSEMIRO e PINTO, 2014).

Analisando os estudos já existentes relacionados a contagem de bovinos por meio de redes neurais convolucionais (CNN), foi levantada a hipótese do uso das Faster R-CNNs (redes neurais convolucionais mais rápidas focadas em proposta de região) para o mesmo procedimento de contagem, pois, tais redes neurais, permitem a detecção e classificação de múltiplos objetos em uma imagem mesma imagem através de diversas camadas de processamento em paralelo, que oferecem componentes de modelagem com informações geométricas essenciais reduzindo a influência do ser humano e provocando aumento na velocidade e precisão do processo de contagem de bovinos (MENDES, 2018, GIRSHICK, 2015, RIVAS et al, 2018).

## **2. Objetivos Gerais e Específicos**

### **2.1 Objetivo Geral**

O objetivo deste trabalho é desenvolver um protótipo de software que realiza contagem de bovinos usando técnicas de visão computacional a partir de imagens capturadas por um VANT.

## **2.2 Objetivo Específico**

1. Criar um banco de imagens aéreas de bovinos capturas por um VANT.
2. Propor um método de treinamento para uma rede neural Faster R-CNN .
3. Desenvolver um protótipo de software, baseado em Faster R-CNN, para contar bovinos a partir de imagens capturadas por um VANT.

## **3. Revisão de Literatura**

### **3.1 Produção de bovinos**

A exportação de carne bovina apresenta grande importância para o cenário econômico nacional, em 2015 representou 3% das exportações brasileiras que proporcionou um faturamento de 6 bilhões de reais, representando 30% do PIB do Agronegócio e 6% do PIB nacional (EMBRAPA, 2018).

Levando em consideração a dimensão do ganho através da exportação deste produto é notável a necessidade de melhoria constante na produção, no cenário atual esta melhoria está diretamente ligada às tecnologias que nela são aplicadas priorizando sempre o ganho de produtividade para retornos econômicos com consistência e sustentabilidade no tempo (EMBRAPA, 2017).

### **3.2 O VANT**

A princípio havia o problema sobre como capturar imagens do gado no pasto de maneira rápida e precisa, a partir desse problema foi levantada a hipótese do uso de VANTs (Veículos Aéreos Não Tripulados) definidos pela ANAC como uma aeronaves remotamente pilotadas para fins experimentais, comerciais ou institucionais. Tais aparatos possuem grande flexibilidade, podendo ser facilmente manuseados para a operação desejada, além de apresentarem baixo custo de aquisição e manutenção, fato esse que se faz bem notável ao comparar a captura de imagem utilizando VANTs com a captura de imagens por satélite. Tais veículos têm obtido sucesso na área da agricultura, mais especificamente nas lavouras de soja, tornando possível a rápida classificação das doenças foliares e assim permitem que peritos e agricultores tomem melhores decisões de gerenciamento (ANAC, 2017, DA LUZ et al, 2017 , SOUZA, 2018, TETILA et al, 2017).

### 3.3 Visão Computacional e Redes Neurais

A visão computacional pode ser definida como a geração de novas representações ou decisões utilizando ferramentas computacionais a partir de uma imagem estática ou em movimento que, por meio de valores extraídos de seus pixels, infere o conteúdo da imagem e cria identidade para certos grupos de pixels, podendo assim automatizar tal processo (WEINSTEIN, 2018, KAEHLER e BRADSKI, 2016).

Esta área da computação busca descrever o mundo a nossa volta através de imagens para reconstruir seus detalhes digitalmente, sendo altamente usada onde há a necessidade de transformar informações visuais em dados que o computador possa entender e processar, um exemplo se dá na detecção de objetos em imagens ou vídeos, onde se faz necessária a obtenção dos valores dos pixels nas regiões denominadas como de interesse para que os mesmo sejam agrupados permitindo assim identificar objetos em várias imagens que apresentam agrupamentos semelhantes a estes pixels (ROWLEY, 1999).

A junção da visão computacional com o reconhecimento de padrões é comumente usada nas redes neurais, definidas como unidades computacionais paralelas com interconexão parcial ou total através de unidades funcionais nomeadas de “neurônios artificiais”, que realizam operações determinadas e transmitem seus resultados entre si pelo processo de *feed forward*. Por meio de um treinamento, a rede consegue reconhecer e identificar padrões em várias imagens (PERELMUTER, 1995).

### 3.4 Redes Neurais Convolucionais

Antes da criação das redes neurais convolucionais (*CNN*), um dos métodos clássicos para se classificar imagens de qualquer tipo se dava pela extração de características visuais da imagem, geração de um pacote de palavras representando tais características e classificação das mesma através do uso de uma Máquina de Vetores de Suporte (*Support Vector Machine*) associada ao pacote previamente gerado. Com a criação das *CNNs*, os três processos foram unificados gerando alta capacidade de treinamento. Geralmente, estas redes têm sido usadas para a compreensão de imagens, que podem ser integradas e convertidas em algoritmos capazes de gerar informações para uma identificação seletiva (LENC e VEDALDI, 2015, KARPATY et al, 2014).

As *CNNs* são focadas na classificação das imagens como um todo, ignorando os possíveis objetos contidos nela. Através deste conceito foi desenvolvida uma rede neural

convolucional focada em regiões, a R-CNN, que trabalha com regiões dentro da imagem classificando seus objetos separadamente. Apesar da melhora na classificação de objetos a R-CNN ainda apresenta complexidade notável devido às várias localizações que podem ter um objeto, localizações estas chamadas de propostas presentes durante toda a execução do método, gerando a necessidade de um maior processamento para remover regiões desnecessárias, sacrificando a velocidade ou acurácia. Por esta razão foi desenvolvida a *Fast R-CNN* que busca melhorar o algoritmo de aprendizagem de região já presente na R-CNN a fim de simplificar e agilizar o treinamento e classificação (GIRSHICK, 2014, GIRSHICK, 2015).

Mantendo a sequência de melhorias já feitas na *Fast R-CNN*, o método das redes neurais convolucionais mais rápidas baseadas em propostas de região (*Faster R-CNN*) usam duas redes neurais convolucionais multicamadas executadas em paralelo. A primeira sendo essencialmente uma *Fast R-CNN*, responsável pela extração de características das regiões anotadas pela segunda rede, sendo esta a Rede Neural para Proposta de Região (RPN) que é baseada na *Fast R-CNN* porém apresenta camadas adicionais para auxiliar na proposição de regiões retangulares com pontuações associadas na imagem, onde cada uma destas determina a probabilidade que a região apresenta de conter um dado objeto (GIRSHICK, 2015, HE et al, 2017, REN, 2015).

#### **4. Metodologia**

Nesta seção será discutida a metodologia a ser seguida para que os objetivos gerais e específicos definidos por este trabalho sejam alcançados.

A primeira etapa deste trabalho se dá pela criação do banco de imagens dos bovinos através do uso de VANTs que percorrerão o campo capturando imagens em intervalos regulares de tempo. As imagens serão gerenciadas por uma aplicação de software para uma maior facilidade na manipulação das mesmas. Por meio do banco de imagens serão formados de forma aleatória três grupos contendo parte dos elementos: o primeiro grupo servirá para treinamento da rede neural contendo 60% do total de indivíduos, o segundo grupo servirá de teste com os indivíduos restantes 20% do total de indivíduos e por último um o grupo de validação contendo os indivíduos restantes (20%).

Com os bancos construídos, o foco nessa etapa será na anotação do mesmo por meio da ferramenta *LabelImg*<sup>1</sup>, gerando arquivos .XML únicos para cada imagem contendo regiões de interesse retangulares marcadas. Tais arquivos .XML serão usados para o treinamento da rede neural Faster R-CNN.

O processo de treinamento será realizado quatro vezes juntamente com o teste por meio dos conjuntos de imagem de treinamento e teste respectivamente, cada vez que a rede for treinada serão usadas uma quantidade X de épocas variando de 25 a 100 com incrementos de 25 a cada processo de treino, ajustando os hiperparâmetros da rede ao fim de cada teste com intuito de se obter a melhor classificação possível de imagens levando em consideração as características aprendidas, porém sempre tentando evitar o sobreajuste da rede neural.

Para cada uma das quatro vezes que a rede será treinada haverá a execução do processo de validação, deixando a rede identificar e contar os indivíduos presentes nas imagens do conjunto de dados de validação. Para avaliar o desempenho de cada etapa de treinamento serão usadas as métricas de taxa de classificação correta (TCC), medida-F e área sob a curva ROC. As métricas geradas para os grupos de teste e validação serão submetidas à análise de variância (ANOVA) e avaliadas pelo teste t de student no nível de significância de 5%.

Com os procedimentos de aprendizagem, teste e validação concluídos e a análise estatística finalizada, será desenvolvido um protótipo de ferramenta na linguagem python que realiza a contagem de bovinos em imagens através da técnica apresentada no escopo deste plano de trabalho.

## 5. Atividades e Cronograma de Execução

Atividade	12 meses					
	2	4	6	8	10	12
Revisão de literatura	x	x				
Coleta de imagens			x			

<sup>1</sup> Disponível em: <https://github.com/tzutalin/labelimg>

Treinamento, teste e validação da rede neural			x	x	x	
Análise estatística dos dados coletado					x	
Produção dos relatórios parcial e final, juntamente com o protótipo			x			x

## 6. Referência:

LENC, Karel; VEDALDI, Andrea. R-cnn minus r. **arXiv preprint arXiv:1506.06981**, 2015. Disponível em: <<https://arxiv.org/abs/1506.06981>> Acesso em 17 mai. 2019.

EMBRAPA. Posicionamentos da Embrapa. BRASIL. 2018. Disponível em: <<https://www.embrapa.br/qualidade-da-carne/posicionamentos-da-embrapa>>. Acessado em: 17 mai. 2019.

WEINSTEIN, Ben G. A computer vision for animal ecology. **Journal of Animal Ecology**, v. 87, n. 3, p. 533-545, 2018. Disponível em: <<https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/1365-2656.12780>> Acessado em: 22 mai. 2019.

CASSEMIRO, Guilherme Henrique Medeiros; PINTO, Hugo Borges. Composição e processamento de imagens aéreas de alta-resolução obtidas com Drone. **Trabalho de Conclusão de Curso, Universidade de Brasília, Brasília**, 2014. Disponível em: <[https://fga.unb.br/articles/0000/7686/TCC2\\_GuilhermeCassemiro\\_090115465\\_e\\_HugoBorges\\_090116461\\_v2.pdf](https://fga.unb.br/articles/0000/7686/TCC2_GuilhermeCassemiro_090115465_e_HugoBorges_090116461_v2.pdf)> Acessado em: 13 mai. 2019.

MENDES, Marina de Souza. Aprendizado em profundidade na descrição semântica de imagens. 2018. Disponível em: <<http://www.monografias.ufop.br/handle/35400000/1590>> Acessado em: 15 mai. 2019.

REN, Shaoqing et al. Faster r-cnn: Towards real-time object detection with region proposal networks. In: **Advances in neural information processing systems**. 2015. p. 91-99. Disponível em: <<http://papers.nips.cc/paper/5638-faster-r-cnn-towards-real-time-object-detection-with-region-proposal-networks.pdf%20obs:%20ImageNet>> Acessado em: 08 mai. 2019.

RIVAS, Alberto et al. Detection of Cattle Using Drones and Convolutional Neural Networks. **Sensors**, v. 18, n. 7, p. 2048, 2018. Disponível em: <<https://www.mdpi.com/1424-8220/18/7/2048/htm>> Acessado em: 10 mai. 2019.

EMBRAPA. A sustentabilidade da pecuária brasileira. BRASIL. 2017. Disponível em: <[https://www.embrapa.br/documents/10180/21470602/GBMJ\\_ArtigoPecu%C3%A1ria-Embrapa\\_29mar17\\_v.rev.pdf/49d67c99-9162-7d79-0baa-6dcac25a520d](https://www.embrapa.br/documents/10180/21470602/GBMJ_ArtigoPecu%C3%A1ria-Embrapa_29mar17_v.rev.pdf/49d67c99-9162-7d79-0baa-6dcac25a520d)> Acessado em: 11 mai. 2019.

DA LUZ, Cristhyano Cavali; ANTUNES, Alzir Felipe Buffara; RATTON, Philipe. Aplicabilidade da tecnologia VANT na atualização de bases de dados cartográficos. **Boletim Paranaense de Geociências**, v. 73, n. 1, 2017. Disponível em: <<https://revistas.ufpr.br/geociencias/article/view/41428/32814>> Acessado em: 11 mai. 2019.

SOUZA, João Pedro Carvalho de et al. Pouso autônomo de VANTs baseado em rede neural artificial supervisionada por lógica fuzzy. 2018. Disponível em: <<https://repositorio.ufjf.br/jspui/handle/ufjf/6512>> Acessado em: 11 mai. 2019.



KAEHLER, Adrian; BRADSKI, Gary. **Learning OpenCV 3: computer vision in C++ with the OpenCV library.** " O'Reilly Media, Inc.", 2016. Disponível em: [https://books.google.com.br/books?hl=pt-BR&lr=&id=LPm3DQAAQBAJ&oi=fnd&pg=PP1&dq=openvc3&ots=2vMqVf9hAe&sig=GLRkRh4l\\_yyNVPn63in-MrYLeE#v=onepage&q=openvc3&f=false](https://books.google.com.br/books?hl=pt-BR&lr=&id=LPm3DQAAQBAJ&oi=fnd&pg=PP1&dq=openvc3&ots=2vMqVf9hAe&sig=GLRkRh4l_yyNVPn63in-MrYLeE#v=onepage&q=openvc3&f=false)> Acessado em: 12 mai. 2019.

ROWLEY, Henry A. **Neural network-based face detection.** CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER SCIENCE, 1999. Disponível em: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a366182.pdf>> Acessado em: 13 mai. 2019.

PERELMUTER, Guy et al. Reconhecimento de imagens bidimensionais utilizando Redes Neurais Artificiais. **Anais do VIII Sibgrapi**, p. 197-203, 1995. Disponível em: <http://sibgrapi.sid.inpe.br/col/sid.inpe.br/sibgrapi/2013/02.18.15.52/doc/25%20Reconhecim ento%20de%20imagens%20bidimensionais.pdf>>. Acessado em: 13 mai. 2019.

KARPATHY, Andrej et al. Large-scale video classification with convolutional neural networks. In: **Proceedings of the IEEE conference on Computer Vision and Pattern Recognition.** 2014. p. 1725-1732. Disponível em: [https://www.cv-foundation.org/openaccess/content\\_cvpr\\_2014/papers/Karpathy\\_Large-scale\\_Video\\_Classification\\_2014\\_CVPR\\_paper.pdf](https://www.cv-foundation.org/openaccess/content_cvpr_2014/papers/Karpathy_Large-scale_Video_Classification_2014_CVPR_paper.pdf)>. Acessado em: 14 mai. 2019.

GIRSHICK, Ross et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition.** 2014. p. 580-587. Disponível em: [http://openaccess.thecvf.com/content\\_cvpr\\_2014/papers/Girshick\\_Rich\\_Feature\\_Hierarchies\\_2014\\_CVPR\\_paper.pdf](http://openaccess.thecvf.com/content_cvpr_2014/papers/Girshick_Rich_Feature_Hierarchies_2014_CVPR_paper.pdf)>. Acessado em: 14 mai. 2019.

GIRSHICK, R. Fast R-CNN. In: **Processing of IEEE International Conference on Computer Vision. Santiago**, p. 13-16, 2015. Disponível em: [http://openaccess.thecvf.com/content\\_iccv\\_2015/papers/Girshick\\_Fast\\_R-CNN\\_ICCV\\_2015\\_paper.pdf](http://openaccess.thecvf.com/content_iccv_2015/papers/Girshick_Fast_R-CNN_ICCV_2015_paper.pdf)> Acessado em: 18 mai. 2019.

HE, Kaiming et al. Mask r-cnn. In: **Proceedings of the IEEE International Conference on Computer Vision.** 2017. p. 2961-2969. Disponível em:

<[http://openaccess.thecvf.com/content\\_ICCV\\_2017/papers/He\\_Mask\\_R-CNN\\_ICCV\\_2017\\_paper.pdf](http://openaccess.thecvf.com/content_ICCV_2017/papers/He_Mask_R-CNN_ICCV_2017_paper.pdf)> Acessado em 18 mai. 2019.

TETILA, Everton Castelão et al. Identification of soybean foliar diseases using unmanned aerial vehicle images. **IEEE Geoscience and Remote Sensing Letters**, v. 14, n. 12, p. 2190-2194, 2017. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/8091106>> Acessado em 30 mai. 2019.

BERNARDES, Carmo. O gado e as larguezas dos Gerais. **Estudos avançados**, v. 9, n. 23, p. 33-58, 1995. Disponível em: <<http://www.scielo.br/pdf/ea/v9n23/v9n23a04.pdf>> Acessado em 03 jun. 2019.

ANAC. Drones. BRASIL. 2017. Disponível em: <<https://www.anac.gov.br/assuntos/paginas-tematicas/drones>> Acessado em 03 jun. 2019.