

RESUMO I FÓRUM BRASILEIRO DE PECUÁRIA DE PRECISÃO, IA E
SUSTENTABILIDADE - 1.1) PECUÁRIA DE PRECISÃO E IA

**DETECÇÃO E CONTAGEM DE BOVINOS EM IMAGENS CAPTURADAS
POR VANT COM ESTRATÉGIAS DE INFERÊNCIA POR FATIAMENTO**

Roberto Pereira De Freitas Neto (neto2014roberto@gmail.com)

Max Hiroito Tieti (ra868222@ucdb.br)

André Gonçalves Da Silva (instagram.andre12@gmail.com)

Bianca Sabka De Queiroz (biancasabka@gmail.com)

Magnum Johanson De Abreu (magnumjabreuu@gmail.com)

Heron Leal Farias (leal.heron@hotmail.com)

Miguel Monteiro Cruz Matos (miguelmcmatos@gmail.com)

Maria Rita De Araújo (mariaritadearaujo3@gmail.com)

Allan Rios Bezerra (allanrb07@gmail.com)

O monitoramento automatizado de bovinos em áreas de pastagem é uma tarefa relevante para a pecuária de precisão, especialmente em regiões de grande produção como Mato Grosso do Sul, que concentra aproximadamente 7,9% do efetivo bovino nacional. O uso de Veículos Aéreos Não Tripulados associado a modelos de Aprendizado Profundo permite cobrir grandes áreas com baixo custo operacional. No entanto, a variação da altitude de voo e o redimensionamento das imagens reduzem a escala aparente dos animais, tornando a detecção de bovinos em imagens aéreas um problema desafiador

de detecção de objetos pequenos. Para lidar com esse problema, este trabalho avalia estratégias de inferência baseadas em fatiamento de imagens para a detecção e contagem de bovinos. São comparadas três configurações de inferência: imagem completa sem fatiamento, adotada como baseline; SAHI (Slicing Aided Hyper Inference), baseado em recortes sobrepostos de tamanho fixo; e ASahi (Adaptive Slicing-Aided Hyper Inference), uma abordagem adaptativa que ajusta dinamicamente o tamanho das fatias para controlar o número de partições conforme a resolução da imagem de entrada. Os experimentos foram conduzidos em imagens aéreas de pastagens de Mato Grosso do Sul, utilizando validação cruzada com cinco dobras. Foram avaliados três detectores representativos: YOLOv8, Faster R-CNN e DETR. Os resultados preliminares mostram que o ASahi melhora o desempenho dos detectores convolucionais em relação às demais estratégias. No YOLOv8, o $mAP@50$ aumentou de $63,29\% \pm 25,43\%$ no baseline para $74,76\% \pm 2,72\%$ com SAHI e $82,82\% \pm 4,87\%$ com ASahi. No Faster R-CNN, o ASahi também obteve o melhor resultado, alcançando $83,30\% \pm 1,84\%$ de $mAP@50$, contra $77,82\% \pm 5,11\%$ no baseline e $77,12\% \pm 1,42\%$ com SAHI. Além disso, o fatiamento elevou substancialmente o recall dos detectores convolucionais. No YOLOv8, a revocação passou de $64,65\% \pm 26,77\%$ no baseline para $91,44\% \pm 3,87\%$ com SAHI e $92,51\% \pm 3,20\%$ com ASahi. No Faster R-CNN, o recall aumentou de $82,40\% \pm 4,02\%$ para $96,26\% \pm 1,20\%$ com SAHI e $95,01\% \pm 1,10\%$ com ASahi. Contudo, o ganho em revocação veio acompanhado de aumento no Erro Médio Absoluto de contagem (MAE), indicando maior tendência à superestimação do rebanho. No Faster R-CNN, o MAE passou de $1,20 \pm 0,41$ no baseline para $7,99 \pm 0,31$ com SAHI, enquanto no YOLOv8 subiu de $2,52 \pm 1,83$ para $3,84 \pm 0,78$. Esse efeito sugere que a fragmentação das imagens pode ampliar falsos positivos associados a texturas do pasto ou artefatos de borda. Ainda assim, o ASahi mitigou parcialmente essa penalidade, reduzindo o MAE para $5,02 \pm 0,42$ no Faster R-CNN e $3,14 \pm 3,48$ no YOLOv8. Em contrapartida, o DETR apresentou melhor desempenho sem fatiamento, com $mAP@50$ de $57,42\% \pm 6,45\%$, enquanto SAHI e ASahi obtiveram $12,62\% \pm 5,05\%$ e $37,62\% \pm 8,86\%$, respectivamente. Esse comportamento sugere que detectores baseados em Transformers, por explorarem relações globais na imagem, podem ser prejudicados pela fragmentação espacial. Assim, os resultados indicam que o ASahi é promissor para detectores convolucionais, mas sua aplicação em arquiteturas baseadas em atenção deve considerar a preservação do contexto global da cena.

Palavras-chave: bovinos; vant ; detecção de objetos; sahi; asahi.