

João dos Santos Neto

Orientador: Romuere Rodrigues Veloso e Silva

Co-orientador: Júlio Vitor Monteiro Marques

**Análise Comparativa de Modelos de Detecção
para Identificação de Pequenos Ruminantes em
Ambientes Naturais**

Picos - PI

426 de junho de 2025

João dos Santos Neto
Orientador: Romuere Rodrigues Veloso e Silva
Co-orientador: Júlio Vitor Monteiro Marques

Análise Comparativa de Modelos de Detecção para Identificação de Pequenos Ruminantes em Ambientes Naturais

Modelo de Trabalho de Conclusão de Curso em Bacharelado em Sistemas de Informação na Universidade Federal do Piauí. Este modelo está em conformidade com as normas ABNT.

Universidade Federal do Piauí
Campus Senador Heuvídio Nunes de Barros
Bacharelado em Sistemas de Informação

Picos - PI
426 de junho de 2025

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

S237a

Santos Neto, João dos.

Análise comparativa de modelos de detecção para identificação de pequenos ruminantes em ambientes naturais / João dos Santos Neto – 2025.
32 f.

1 Arquivo em PDF

Indexado no catálogo *online* da biblioteca José Albano de Macêdo-CSHNB
Aberto a pesquisadores, com restrições da Biblioteca

Trabalho de Conclusão de Curso (Graduação) – Universidade Federal do Piauí,
Curso de Bacharelado em Sistemas de Informação, Picos, 2025.

“Orientador: Romuere Rodrigues Veloso e Silva”.

“Co-orientador: Júlio Vitor Monteiro Marques”.

1. Sistemas informacionais. 2. Monitoramento automatizado – rebanhos. 3. Pecuária moderna. I. Santos Neto, João dos. II. Veloso e Silva, Romuere Rodrigues. III. Marques, Júlio Vitor Monteiro. IV. Título.

CDD 005.7

Elaborada por Maria Letícia Cristina Alcântara Gomes
Bibliotecária CRB nº 03/1835

Análise Comparativa de Modelos de Detecção para Identificação de Pequenos Ruminantes em Ambientes Naturais

João dos Santos Neto

Monografia aprovada como exigência parcial para obtenção do grau de Bacharel em Sistemas de Informação.

Data de Aprovação

Picos – PI, 23 de Junho de 2025

Documento assinado digitalmente
 **ROMUERE RODRIGUES VELOSO E SILVA**
Data: 08/07/2025 15:09:07-0300
Verifique em <https://validar.iti.gov.br>

Prof. Romuere Rodrigues Veloso e Silva

Documento assinado digitalmente
 **JOAO VITOR MOREIRA PASSOS**
Data: 08/07/2025 12:36:23-0300
Verifique em <https://validar.iti.gov.br>

Prof. João Vitor Moreira Passos

Documento assinado digitalmente
 **ANTHONY IRLAN MARQUES LUZ**
Data: 08/07/2025 12:26:05-0300
Verifique em <https://validar.iti.gov.br>

Prof. Anthony Irlan Marques Luz

Agradecimentos

Primeiramente, gostaria de agradecer a Deus por sempre estar comigo nos momentos de tristeza e de alegria, e pela oportunidade de ter conseguido chegar até aqui. Agradeço à minha mãe, Eristene Santos Gomes, e ao meu pai, José Emeraldó Silva Gomes, que sempre acreditaram em mim, me ajudaram e apoiaram minhas decisões.

Gostaria de agradecer aos meus amigos "Os Miudos" que estiveram comigo desde o início da minha jornada. E em especial a Lara Gabriela de Moura Rocha, Clara Emilly Sousa Sátiro, Anna Beatriz Barbosa de Godoy, Bryan Martins Silva e Nara Raquel Dias Andrade.

Agradeço à professora Juliana, uma mulher que inspira muitos alunos; ao meu professor e orientador, Romuere, que talvez não tenha noção do quanto contribuiu para o meu crescimento; e à professora Alcilene, por ensinar a arte de ser uma pessoa autêntica.

E agradeço à minha segunda família: Luiz e Nelson dos Santos Lima, Arthur Sabino Santos e Isla Nathanaelly Silva Pereira Sousa. Não há palavras que possam expressar o quanto eu amo vocês.

E se você olhar, durante muito tempo, para um abismo, o abismo também olhará para dentro de você.

Friedrich Nietzsche

Resumo

O monitoramento automatizado de rebanhos é uma estratégia essencial para a pecuária moderna, pois permite otimizar recursos, reduzir falhas humanas e melhorar a tomada de decisão no campo. Este trabalho tem como objetivo comparar diferentes arquiteturas de redes neurais convolucionais aplicadas à detecção de animais em ambientes naturais, visando identificar a solução mais eficiente. Foram avaliados quatro modelos amplamente utilizados na literatura: YOLOv8, SSD, RetinaNet e Faster R-CNN. O desempenho foi medido por meio de métricas padronizadas, como precisão, recall, mAP e tempo de inferência. Os resultados demonstraram que o YOLOv8 obteve o melhor desempenho geral, com mAP de 0,98, precisão de 0,75 e recall de 0,91, destacando-se pela alta acurácia e velocidade de processamento. Conclui-se que o YOLOv8 é a abordagem mais adequada para aplicações em pecuária de precisão com foco em detecção visual em tempo real.

Palavras-chaves: Detecção. CNN. YOLO. Ambientes Naturais.

Abstract

Automated herd monitoring is a key strategy in modern livestock farming, enabling resource optimization, reduced human error, and improved decision-making in the field. This study aims to compare different convolutional neural network architectures applied to animal detection in natural environments, with the goal of identifying the most efficient solution. Four widely adopted models were evaluated: YOLOv8, SSD, RetinaNet, and Faster R-CNN. Performance was assessed using standard metrics such as precision, recall, mean Average Precision (mAP), and inference time. The results showed that YOLOv8 achieved the best overall performance, with a mAP of 0.98, precision of 0.75, and recall of 0.91, standing out for its high accuracy and fast processing speed. It is concluded that YOLOv8 is the most suitable approach for precision livestock farming applications focused on real-time visual detection.

Lista de ilustrações

Figura 1 – Representação de subcampos entre IA, ML e DL. Fonte: (NOBRE, 2021)	13
Figura 2 – Representa ilustrativa de modelos <i>One-Stage</i> e <i>Two-Stage</i> Fonte:(YUN, 2021)	14
Figura 3 – Detecção de Espécies de Pássaros em Câmeras Selvagens.Fonte: (FERGUS et al., 2024)	15
Figura 4 – Detecção de Bovinos em Pastos. Fonte: (RURAL, 2023)	15
Figura 5 – Fluxograma da metodologia proposta, composta pelas etapas de: (i) aquisição da base de dados, (ii) realização dos experimentos para análise comparativa de desempenho, (iii) detecção e seleção das redes neurais convolucionais (CNNs) e (iv) avaliação dos modelos com base em métricas padronizadas.	18
Figura 6 – Estrutura de coleta de dados: (a) representação esquemática da passagem equipada com sensor Kinect V2, posicionada acima do animal para captura de imagens; (b) implementação da estrutura em ambiente real de manejo.	19
Figura 7 – Imagem de exemplo da base de dados.	19
Figura 8 – Exemplo de Detecção - YOLOv8	20
Figura 9 – Arquitetura RetinaNet	21
Figura 10 –Arquitetura YOLOv8.	21
Figura 11 –Exemplo de Saídas: (a) representa uma saída com acerto; (b) representa uma saída com falso positivo	24

Lista de tabelas

Tabela 1 – Quadro comparativo de trabalhos relacionados	16
Tabela 2 – Desempenho comparativo dos modelos de detecção.	23
Tabela 3 – Parâmetros de configuração e tempo de inferência dos modelos.	23
Tabela 4 – Comparação do desempenho do YOLOv8 com trabalhos da literatura.	24

Lista de abreviaturas e siglas

DL	Deep Learning
CNNs	Redes Neurais Convolucionais
YOLO	<i>You Only Look Once</i>
SSD	<i>Single Shot MultiBox</i>
VC	Visão Computacional
IA	Inteligência Artificial
ML	<i>Machine Learning</i>
TA	Taxa de Aprendizado
RPA	<i>Region Proposal Network</i>
BB	<i>Bounding Boxes</i>
FL	<i>Focal Loss</i>

Lista de símbolos

Σ	Somatório
i	Índice ou variável genérica

Sumário

1	Introdução	11
1.1	Objetivos	11
1.2	Organização do Trabalho	12
1.3	Publicações	12
2	Referencial Teórico	13
2.1	Visão Computacional e Machine Learning	13
2.2	Modelos de Detecção de Objetos	14
2.3	Aplicações de Detecção de Animais em Ambientes Naturais	15
3	Trabalhos Relacionados	16
4	Metodologia	18
4.1	Aquisição dos dados	18
4.2	Experimentos	19
4.3	Detecção	20
4.4	Avaliação	21
5	Resultados	23
5.1	Limitações	25
6	Conclusão	26
	Referências	27
	Apêndices	31
	APÊNDICE A Apêndice	32

1 Introdução

A identificação e o monitoramento de animais em ambientes naturais têm se tornado cada vez mais relevantes, tanto para pesquisas científicas quanto para aplicações práticas na agricultura e conservação ambiental (SILVA et al., 2021). Nesse sentido, a detecção precisa desses animais em cenários naturais pode contribuir para a gestão eficiente de rebanhos, o monitoramento da saúde animal, a prevenção de perdas e a otimização de práticas de pastoreio (NGUYEN et al., 2017).

Nos últimos anos, avanços em técnicas de visão computacional e aprendizado de máquina, especialmente com o uso de *Deep Learning* (DL) e modelos de Redes Neurais Convolucionais (CNNs), têm revolucionado a capacidade de detectar e identificar objetos em imagens e vídeos (LECUN; BENGIO; HINTON, 2015). Esses modelos vêm sendo aplicados com sucesso em diversas áreas, incluindo a detecção de animais em ambientes naturais (BJERGE et al., 2023).

No entanto, a complexidade dos ambientes naturais, marcada por variações de iluminação, vegetação, topografia e interações entre os animais, representa um desafio significativo para o desenvolvimento de sistemas de detecção robustos e confiáveis (NEUPANE; SATO; GAUTAM, 2022). Além disso, a seleção do modelo mais adequado para essa tarefa exige uma análise criteriosa de fatores como precisão, velocidade de processamento, resistência a variações ambientais e custo computacional (PARK; SACCHI, 2020). Diante da diversidade de abordagens disponíveis, uma avaliação comparativa torna-se fundamental para identificar as soluções com melhor desempenho nesse contexto.

Neste cenário, este trabalho propõe uma análise comparativa de diferentes abordagens de detecção aplicadas à identificação de ovelhas em ambientes naturais. Serão avaliados quatro modelos de detecção de objetos: *You Only Look Once* (YOLO), *Single Shot MultiBox*(SSD), RetinaNet e Faster R-CNN. O comparativo considera o desempenho dos modelos em tarefas de detecção, com base na avaliação de métricas como precisão, capacidade de generalização e eficiência computacional.

Vale destacar que a detecção automática de animais representa um passo fundamental dentro de sistemas mais amplos de monitoramento inteligente, podendo integrar soluções para rastreamento, análise comportamental, controle sanitário e suporte à tomada de decisão na pecuária de precisão e em projetos de conservação ambiental.

1.1 Objetivos

O objetivo geral deste trabalho é realizar uma análise comparativa de modelos de detecção com a finalidade de identificar ovelhas em ambientes naturais.

Os objetivos específicos dessa pesquisa são:

- Revisar a literatura sobre métodos de detecção de objetos aplicados à identificação de animais em ambientes naturais, com ênfase nos modelos YOLO, SSD, RetinaNet e Faster R-CNN.
- Analisar a robustez dos modelos diante de variações naturais como iluminação, oclusão e movimentação dos animais, a partir de testes com diferentes condições de imagem.
- Comparar os resultados obtidos de forma quantitativa e qualitativa, identificando os pontos fortes e limitações de cada modelo para a detecção de ovelhas em ambientes naturais.
- Sugerir o modelo mais adequado para aplicação prática em contextos de monitoramento animal, com base nos resultados obtidos e nas necessidades do cenário agrícola ou ambiental.

1.2 Organização do Trabalho

O restante do trabalho está estruturado da seguinte forma:

- O Capítulo 2 se refere ao referencial teórico e contempla conceitos fundamentais ao campo da área de visão computacional.
- O Capítulo 3 apresenta os trabalhos relacionados, nos quais são discutidas as principais contribuições e avanços na área de detecção de animais em ambientes naturais.
- o Capítulo 4 aborda a metodologia, destacando as etapas de aquisição da base de dados, detecção, experimentos e avaliação.
- O Capítulo 5 apresenta os resultados obtidos a partir da metodologia proposta, além de uma comparação com outros trabalhos presentes na literatura.
- O Capítulo 6 sintetiza os principais achados do estudo e propõe possíveis direções para pesquisas futuras.

1.3 Publicações

- **Análise Comparativa de Modelos de Detecção para Identificação de Ovelhas em Ambientes Naturais** João dos Santos Neto, Júlio Vitor Monteiro Marques, José Lindenberg Rocha Sarmiento, Romuere Rodrigues Veloso e Silva *Seminário Integrado de Software e Hardware (SEMISH)*, 2025.

2 Referencial Teórico

Este capítulo reúne os principais fundamentos teóricos que sustentam o desenvolvimento desta pesquisa. Os temas abordados servem como base para a compreensão das escolhas metodológicas adotadas, além de oferecerem o suporte conceitual necessário para a análise e interpretação dos resultados apresentados nos capítulos subsequentes.

2.1 Visão Computacional e Machine Learning

A Visão Computacional(VC) é um ramo da Inteligência Artificial(IA) que busca criar métodos para que sistemas computacionais consigam interpretar e obter dados úteis a partir de imagens ou vídeos (RAKHIMOV et al., 2021). Essa área tem ganhado destaque em várias aplicações, como reconhecimento facial (ROBERTSON; CRAW, 1994), carros autônomos (AGARWAL; CHIANG; SHARMA, 2018) e em cameras de vigilância (DICK; BROOKS, 2003). Com o avanço das tecnologias e a grande quantidade de dados disponíveis, os métodos antigos, como Máquinas de Vetores de Suporte (SVM) (KHAN; LAGHARI; AWAN, 2021), foram substituídos por técnicas mais modernas e automáticas, como o *Deep Learning*(DL) (MITTAL; SINGH; SHARMA, 2020).

O DL, por sua vez, é uma área do *Machine Learning*(ML) (KHAN; LAGHARI; AWAN, 2021) que usa redes neurais com várias camadas para aprender representações dos dados em diferentes níveis. Entre essas redes, as CNNs (ZHAO et al., 2024) tornaram-se a arquitetura principal para tarefas de VC, graças à sua habilidade de captar e aprender padrões visuais importantes das imagens. Essa técnica tem trazido grandes avanços em tarefas como detecção (GALVEZ et al., 2018) e classificação (RAWAT; WANG, 2017), mostrando ótimo desempenho mesmo em cenários difíceis, como a identificação de animais em áreas naturais (NGUYEN et al., 2017). A Figura 1 mostra uma ilustração dos subcampos que compõem a área de VC.

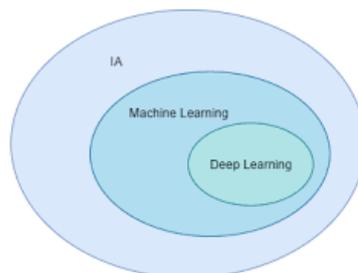


Figura 1 – Representação de subcampos entre IA, ML e DL. Fonte: (NOBRE, 2021)

2.2 Modelos de Detecção de Objetos

Os modelos de detecção de objetos são algoritmos desenvolvidos para localizar e classificar automaticamente instâncias de objetos dentro de imagens ou vídeos (DHILLON; VERMA, 2020). Ao contrário dos métodos de classificação de imagens, que apenas identificam a presença de uma classe, os detectores também informam a posição dos objetos por meio de caixas delimitadoras (*bounding boxes*) (SONG et al., 2011). Com o avanço das CNNs, surgiram modelos mais precisos e rápidos, capazes de realizar a detecção em tempo real, mesmo em cenários com múltiplos objetos e alta complexidade visual (GEUTHER et al., 2019).

Atualmente, os modelos de detecção podem ser divididos em duas grandes categorias: os de uma etapa (*one-stage*) e os de duas etapas (*two-stage*) (DU; ZHANG; WANG, 2020). Modelos como YOLO (*You Only Look Once*) (JIANG et al., 2022), SSD (*Single Shot MultiBox Detector*) (LIU et al., 2016) e RetinaNet (JETHANI et al., 2020) pertencem à primeira categoria e se destacam pela velocidade de processamento. Já modelos como o Faster R-CNN (ALBAHLI et al., 2021), da categoria de duas etapas, priorizam a precisão ao realizar primeiro a proposta de regiões e, depois, a classificação. A Figura 2 apresenta uma ilustração de modelos com uma etapa (*one-stage*) e os de duas etapas (*two-stage*).

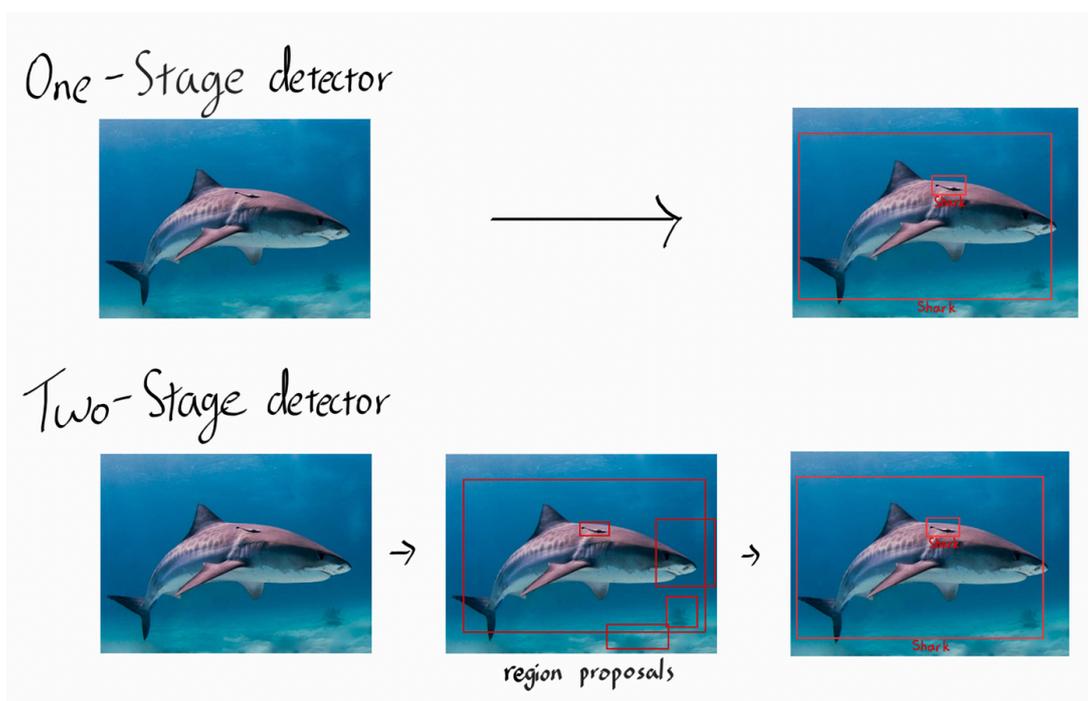


Figura 2 – Representa ilustrativa de modelos *One-Stage* e *Two-Stage* Fonte:(YUN, 2021)

2.3 Aplicações de Detecção de Animais em Ambientes Naturais

A detecção automática de animais em ambientes naturais tem sido relevante nas últimas décadas devido ao seu potencial em diversas áreas, como na vigilância por drones em tempo real de animais marinhos (SAQIB et al., 2018). Com o uso de câmeras em campo e modelos de VC, é possível identificar a presença, o comportamento e a movimentação de animais sem a necessidade de interferência humana direta (SCHINDLER; STEINHAGE, 2021). A Figura 3 apresenta um exemplo de detecção de animal em ambiente natural.



Figura 3 – Detecção de Espécies de Pássaros em Câmeras Selvagens. Fonte: (FERGUS et al., 2024)

Na agropecuária, por exemplo, a detecção automatizada de animais, como ovelhas, pode auxiliar no controle de rebanhos, na identificação de animais doentes ou perdidos e na otimização de práticas de pastoreio (VAINTRUB et al., 2021). Já na conservação ambiental, esses sistemas ajudam a monitorar populações de animais ameaçados, detectar atividades ilegais como a caça e estudar padrões migratórios (SIMÕES; BOUVEYRON; PRECIOSO, 2023). O avanço dessas aplicações depende diretamente da robustez dos modelos de detecção, que precisam lidar com variações de iluminação, camuflagem, movimento e interferências do ambiente natural (DJOLONGA et al., 2021). A Figura 4 apresenta um exemplo de detecção na agropecuária.

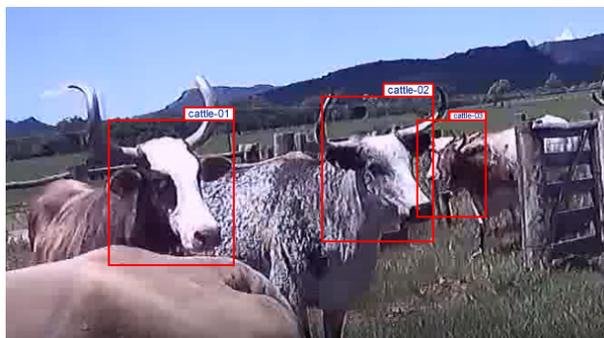


Figura 4 – Detecção de Bovinos em Pastos. Fonte: (RURAL, 2023)

3 Trabalhos Relacionados

Com o aumento do uso de técnicas de visão computacional na identificação de animais em ambientes naturais, diversos estudos têm explorado abordagens baseadas em *Deep Learning* para aprimorar a detecção e o monitoramento de diferentes espécies. Assim, nesta seção, serão apresentados e discutidos estudos relacionados ao tema, destacando as metodologias empregadas, os desafios enfrentados e as contribuições para o avanço da detecção de animais em cenários naturais.

Para realizar esse estudo foram considerados 4 critérios: I - Análise Comparativa; II - Uso de CNNs; III - Ambientes Naturais; IV - Animal de Estudo. A Tabela 1 apresenta uma comparação detalhada dos trabalhos, destacando os temas abordados com base em critérios específicos. Esta análise considera a metodologia adotada nos artigos selecionados, proporcionando uma visão abrangente das abordagens e dos resultados obtidos.

Tabela 1 – Quadro comparativo de trabalhos relacionados

Trabalhos Relacionados	I	II	III	IV
(NEUPANE, 2022)	Sim	Sim	Sim	Animais Selvagens
(SCHNEIDER; TAYLOR; KREMER, 2018)	Não	Sim	Sim	Animais Selvagens
(SILVA, 2021)	Não	Sim	Sim	Animais Selvagens
Este Trabalho	Sim	Sim	Sim	Caprinos

O trabalho de (NEUPANE, 2022) explora a aplicação de *Deep Learning* na identificação de espécies de animais em ambientes naturais. O estudo avaliou a eficácia de modelos como SSD, YOLOv4, EfficientNet, *Long Short-Term Memory* (LSTM), Mask R-CNN e Faster R-CNN na detecção de bovinos e ovinos em pastagens, contribuindo com uma abordagem de pré-processamento que melhorou a generalização dos modelos em diferentes condições climáticas e vegetativas. As limitações identificadas pelos autores envolvem a qualidade das imagens, a velocidade de processamento dos dados, o tamanho dos datasets e a movimentação dos animais durante a captura das imagens, fatores que podem afetar diretamente os resultados.

O estudo de (SCHNEIDER; TAYLOR; KREMER, 2018) trata da identificação de espécies selvagens utilizando armadilhas de câmeras ecológicas e redes neurais profundas. O autor avalia a aplicabilidade de YOLO e Faster R-CNN na detecção automática de animais em habitats naturais, com enfoque no Faster R-CNN, que apresentou a melhor capacidade de detecção. Porém, um dos desafios encontrados foi a necessidade de um grande volume de dados anotados para o treinamento eficaz dos modelos.

Por fim, a pesquisa de (SILVA, 2021) fornece um sistema de monitoramento de fauna utilizando YOLOv3 e Faster R-CNN para detecção de animais em imagens coletadas por drones. O estudo compara a precisão e a velocidade dos modelos, identificando que

o YOLOv3 apresentou melhor desempenho em tempo real, enquanto o Faster R-CNN obteve maior precisão na segmentação. O principal desafio foi a variação na iluminação das imagens aéreas, que impactou o desempenho do modelo em cenários noturnos.

Diante das limitações identificadas na literatura, este estudo propõe uma análise comparativa de redes neurais convolucionais aplicadas à detecção de ovelhas em ambientes naturais, avaliando seu desempenho por meio de métricas amplamente utilizadas, tais como mAP, precisão e recall ([PADILLA; NETTO; SILVA, 2020](#)).

4 Metodologia

Nesta seção, apresenta-se a metodologia adotada para a condução do estudo, abrangendo desde a aquisição da base de dados até a avaliação de desempenho e a execução dos experimentos. O processo foi estruturado em cinco etapas principais: Aquisição da Base de Dados, Experimentos, Detecção e Avaliação. Cada uma dessas etapas foi planejada de forma criteriosa, visando assegurar a robustez da abordagem e a precisão dos resultados obtidos. A Figura 5 ilustra o fluxograma que descreve a sequência das atividades conduzidas ao longo do estudo.

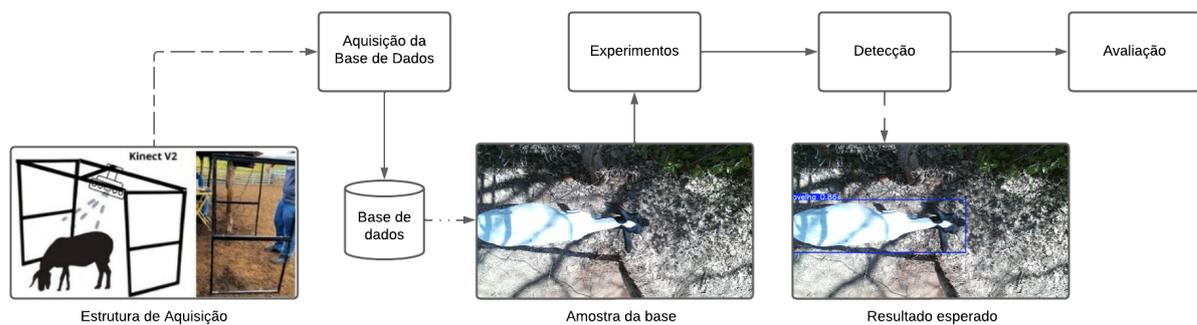


Figura 5 – Fluxograma da metodologia proposta, composta pelas etapas de: (i) aquisição da base de dados, (ii) realização dos experimentos para análise comparativa de desempenho, (iii) detecção e seleção das redes neurais convolucionais (CNNs) e (iv) avaliação dos modelos com base em métricas padronizadas.

4.1 Aquisição dos dados

O processo de aquisição da base de dados foi realizado por meio de uma estrutura metálica construída com perfis de metalon 20x20, garantindo a captação das imagens. A estrutura é composta por seis módulos idênticos e interconectáveis, que formam tanto as laterais quanto a cobertura do sistema. Cada módulo apresenta dimensões de 65 cm de altura por 36 cm de largura, garantindo um encaixe entre os componentes. Na parte superior da estrutura, foi acoplada uma câmera *Microsoft Kinect V2*, estrategicamente posicionada para capturar vídeos dos animais em movimento. A configuração adotada, semelhante a uma porteira, permite a filmagem dos animais durante sua passagem, assegurando a obtenção de imagens de qualidade para análises posteriores.

Para a condução dos experimentos, os vídeos foram coletados em duas propriedades rurais, denominadas Fazenda A e Fazenda B. No total, foram registradas imagens de 42 animais na Fazenda A e 57 animais na Fazenda B, totalizando 3175 imagens. Cada animal foi filmado individualmente em um único vídeo no formato RGB, com uma taxa de 30 quadros por segundo (FPS), resolução Full HD (1920x1080) e tamanho médio de 250 MB

por vídeo. A Figura 6 ilustra a estrutura montada em um ambiente real, destacando sua funcionalidade no processo de coleta de dados. Em seguida, A Figura 7 apresenta um exemplo da imagem utilizada neste trabalho.

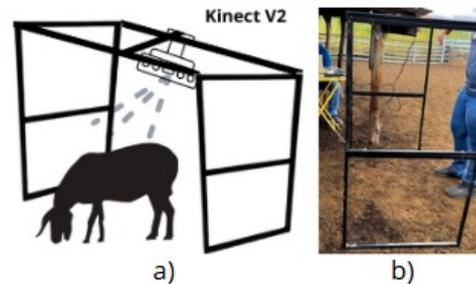


Figura 6 – Estrutura de coleta de dados: (a) representação esquemática da passagem equipada com sensor Kinect V2, posicionada acima do animal para captura de imagens; (b) implementação da estrutura em ambiente real de manejo.



Figura 7 – Imagem de exemplo da base de dados.

4.2 Experimentos

Os experimentos foram conduzidos na plataforma Google Colab (MENON et al., 2023), empregando uma estrutura organizada de diretórios para o armazenamento das imagens e suas respectivas anotações em formato XML. A divisão dos dados seguiu a abordagem de distribuir em 70% para treinamento, 20% para teste e 10% para validação. O treinamento foi realizado ao longo de cinco épocas para todos os modelos, garantindo uma avaliação dos resultados. A Taxa de Aprendizado(TA) foi ajustada de acordo com as características de cada modelo: para a rede YOLO, foi definida em 0,001 com um *batch size*(BS) de 16; para o Faster R-CNN, em 0,0001 com *batch size* de 8; para a RetinaNet, em 0,0001 com *batch size* de 16; e, por fim, para o SSD, em 0.001 com *batch size* de 4. Essa variação de parâmetros se deu por pesquisas feitas por testes e pesquisa dentro da literatura, bem como em repositórios de código aberto.

A seleção da função de perda variou conforme as especificidades de cada modelo. O Faster R-CNN empregou a função *Smooth L1 Loss* para a regressão das caixas delimitadoras (bounding boxes), sem necessidade de ajustes adicionais para classificação, visto

que o problema envolvia uma única classe. Os modelos YOLO e RetinaNet adotaram a função de perda *Focal Loss*, apropriada para cenários com desequilíbrio entre classes, contribuindo para uma melhoria na detecção de objetos de menor escala. O SSD utilizou a função de perda *MultiBox*, que integra a perda de localização (*Localization Loss*) e a perda de confiança (*Confidence Loss*), permitindo ao modelo ajustar com precisão as caixas delimitadoras e classificar corretamente os objetos detectados.

4.3 Detecção

Na etapa de Detecção, foram selecionadas quatro arquiteturas para a detecção de ovelhas em ambientes naturais: YOLOv8 (REDMON et al., 2016), Faster R-CNN (REN et al., 2015), RetinaNet (LIN et al., 2017) e SSD (LIU et al., 2016). A escolha considerou o equilíbrio entre precisão, velocidade e robustez.

A YOLOv8 foi escolhida por sua eficiência na detecção em tempo real, sendo uma arquitetura de etapa única (*Single-Stage Detector*), que realiza a identificação de objetos em uma única passagem pela rede (BAOYUAN; YITONG; KUN, 2021). A Faster R-CNN, por sua vez, foi selecionada devido à sua alta precisão em cenários complexos, utilizando uma abordagem de duas etapas (*Two-Stage Detector*), onde uma *Region Proposal Network* (RPN) sugere áreas de interesse antes da classificação e refinamento das *Bounding Boxes* (BB) (GALVEZ et al., 2018). A Figura 8 apresenta um exemplo de detecção realizada pela YOLOv8.



Figura 8 – Exemplo de Detecção - YOLOv8

A RetinaNet foi incluída por equilibrar velocidade e precisão, destacando-se pelo uso da função de perda *Focal Loss* (FL) (ROSS; DOLLÁR, 2017), que melhora a detecção de objetos em conjuntos de dados desbalanceados. Já a SSD foi escolhida por sua eficiência em inferência rápida, realizando detecções em múltiplas escalas diretamente a partir dos mapas de características, sem a necessidade de uma etapa intermediária de propostas de regiões (KUMAR; ZHANG; LYU, 2020). A Figura 8 apresenta a arquitetura do modelo RetinaNet.

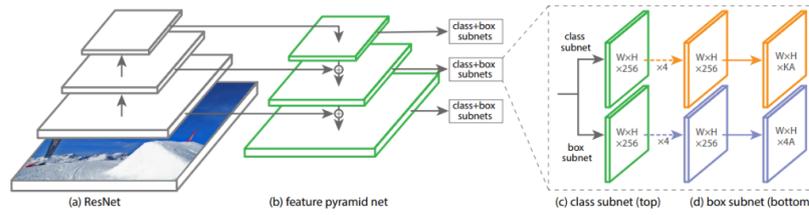


Figura 9 – Arquitetura RetinaNet

Assim, a seleção dessas arquiteturas possibilita uma análise comparativa, contemplando distintas abordagens para a detecção de objetos. Cada modelo apresenta características específicas em termos de velocidade e robustez, permitindo uma avaliação comparativa fundamentada de seu desempenho na detecção de ovelhas em ambientes naturais. A Figura 10 apresenta a arquitetura do modelo YOLOv8.

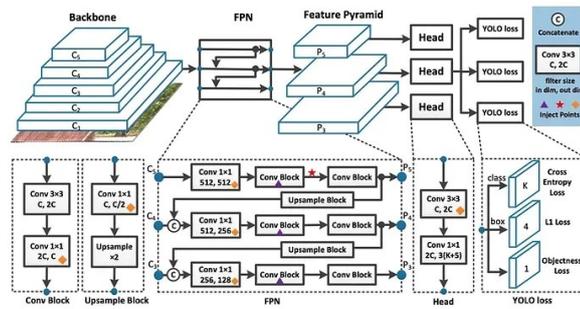


Figura 10 – Arquitetura YOLOv8.

4.4 Avaliação

Para a avaliação do desempenho dos modelos, serão utilizadas as métricas *mean Average Precision* (mAP) (HENDERSON; FERRARI, 2017), Precisão (OKSUZ et al., 2018) e Recall (BANSAL et al., 2021), amplamente empregadas na literatura para quantificar a eficácia de modelos de detecção de objetos.

A métrica Precisão mede a proporção de predições corretas em relação ao total de predições realizadas pelo modelo, sendo particularmente relevante para aplicações em que a minimização de falsos positivos é essencial. Um falso positivo ocorre quando o modelo classifica incorretamente uma amostra negativa como positiva. Por exemplo, em um sistema de detecção de fraudes em transações bancárias, um falso positivo seria uma transação legítima classificada como fraudulenta, levando a uma investigação desnecessária. A Equação (4.1) expressa matematicamente esse conceito:

$$\text{Precisão} = \frac{TP}{TP + FP}, \quad (4.1)$$

onde TP representa os verdadeiros positivos e FP os falsos positivos.

Por outro lado, o *Recall* quantifica a capacidade do modelo de identificar corretamente todas as instâncias relevantes de um objeto, podendo ser utilizada em cenários nos quais a ausência de detecção pode comprometer a análise. A Equação (4.2) define essa relação:

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (4.2)$$

onde FN representa os falsos negativos.

A métrica mAP sintetiza a precisão média em diferentes limiares de confiança e é amplamente utilizada para avaliar o desempenho global de modelos de detecção de objetos onde neste trabalho será atribuído o limiar de confiança 50. O mAP é definido como a média das precisões médias (AP) calculadas para cada classe do modelo, conforme ilustrado na Equação (4.3):

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i, \quad (4.3)$$

onde N representa o número total de classes e corresponde à precisão média da i-ésima classe, obtida a partir da curva *Precision-Recall*.

Também será levado em conta para avaliação de desempenho dos modelos a métrica de Tempo de Inferência que é uma métrica utilizada em sistemas de aprendizado de máquina e visão computacional, pois indica o tempo necessário para que um modelo processe uma entrada e forneça uma previsão. Esse tempo é especialmente relevante em aplicações em tempo real, como reconhecimento facial, carros autônomos e monitoramento de segurança, onde decisões rápidas são essenciais para o funcionamento eficiente do sistema.

A escolha dessas métricas justifica-se pelo fato de que cada uma captura um aspecto distinto do desempenho do modelo. Enquanto a Precisão reduz a incidência de falsos positivos, o *Recall* assegura que todas as instâncias relevantes sejam detectadas. O mAP, por sua vez, fornece uma visão abrangente da eficácia do modelo ao longo de diferentes limiares de confiança, permitindo uma avaliação do sistema de detecção.

5 Resultados

Nesta seção, apresentamos o desempenho comparativo de quatro modelos de detecção de ovelhas em ambientes naturais: YOLOv8, Faster R-CNN, RetinaNet e SSD. A Tabela 2 mostra os valores de mAP, precisão (Precision) e sensibilidade (Recall) obtidos nos experimentos. Em seguida, a Tabela 3 apresenta os principais parâmetros de configuração utilizados no treinamento de cada modelo, incluindo tempo médio de inferência por imagem.

Tabela 2 – Desempenho comparativo dos modelos de detecção.

Modelos	mAP	Precisão	Recall
YOLOv8	0,98	0,75	0,91
Faster R-CNN	0,38	0,42	0,54
RetinaNet	0,51	0,72	0,78
SSD	0,73	0,98	0,78

Tabela 3 – Parâmetros de configuração e tempo de inferência dos modelos.

Modelos	Camadas	Filtros	Batch Size	Learning Rate	Imagens de Teste	Tempo Médio de Inferência(s)
YOLOv8	24	Entre 32 e 1024	16	0,001	42	0,19
Faster R-CNN	50	Entre 64 e 2048	8	0,0001	42	9,86
RetinaNet	50	Entre 64 e 2028	16	0,0001	42	9,30
SSD	13	Entre 64 e 512	4	0,001	42	0,30

A análise conjunta dos resultados e dos parâmetros evidencia o impacto direto da arquitetura e das configurações de treinamento sobre o desempenho dos modelos. O YOLOv8, com arquitetura mais leve (24 camadas), detecção em etapa única e treinamento a partir de pesos pré-treinados em grandes conjuntos de dados, obteve o melhor desempenho geral — alcançando mAP de 0,98, recall de 0,91 e o menor tempo médio de inferência (0,19s por imagem). Sua taxa de aprendizado (0,001) e o tamanho de batch (16) podem ter contribuído para acelerar a convergência e garantir melhor generalização.

O SSD, embora apresente uma arquitetura menos profunda (13 camadas), obteve a maior precisão (0,98) entre os modelos testados, indicando um comportamento mais conservador e seletivo na detecção. No entanto, seu mAP (0,73) e recall (0,78) inferiores aos do YOLOv8 sugerem que o modelo pode estar deixando de detectar algumas instâncias relevantes. A Figura 11 representa duas saídas pela inferência do modelo.



Figura 11 – Exemplo de Saídas: (a) representa uma saída com acerto; (b) representa uma saída com falso positivo

Modelos mais pesados, como o Faster R-CNN e o RetinaNet (ambos com 50 camadas e filtros entre 64 e 2048), apresentaram os maiores tempos de inferência (9,86s e 9,30s, respectivamente) e desempenho inferior nas métricas de mAP e precisão. O baixo desempenho do Faster R-CNN (mAP de 0,38) pode estar relacionado à sua baixa taxa de aprendizado (0,0001), ajustada conforme alguns teste, e mesmo que tenha sido utilizado pesos pré-treinados, o que, aliado a um número limitado de imagens, comprometeu sua capacidade de generalização.

Dessa forma, os resultados indicam que modelos com menor profundidade, taxas de aprendizado mais agressivas e pré-treinamento em grandes bases, como o YOLOv8, oferecem o melhor equilíbrio entre eficiência computacional e desempenho, sendo mais adequados para aplicações em tempo real e ambientes não controlados. A Tabela 4 apresenta um comparativo dos resultados deste trabalho com os demais achados na literatura.

Tabela 4 – Comparação do desempenho do YOLOv8 com trabalhos da literatura.

Trabalhos	mAP50	Precision	Recall
(THOMAS et al.,)	0,91	0,90	0,88
(LIU et al., 2024)	0,88	0,87	0,87
(JIANG; WU, 2024)	0,78	0,85	0,72
Este Trabalho	0,98	0,75	0,91

Os resultados apresentados na Tabela 4, demonstram que o modelo treinado neste estudo atingiu o maior valor de mAP (0,98), superando todos os trabalhos analisados. O estudo de (THOMAS et al.,) obteve um mAP de 0,91, valor inferior ao atingido neste trabalho, indicando que a configuração adotada aqui permitiu uma melhor capacidade de detecção dos objetos. Similar ao estudo de (LIU et al., 2024) que apresentou um mAP de 0,88, também inferior ao obtido neste trabalho, reforçando a eficácia do modelo treinado. Já (JIANG; WU, 2024) alcançaram um mAP ainda menor (0,78), sugerindo que a abordagem adotada por esses autores pode ter limitações na identificação precisa dos objetos.

Sobre o recall, o modelo proposto nesta pesquisa também alcançou o maior resultado (0,91). (THOMAS et al.,) registrou um recall de 0,88, um valor próximo, enfatizando que o modelo treinado tem capacidade de detectar os objetos de interesse. (LIU et al., 2024) obtiveram um recall de 0,87, enquanto (JIANG; WU, 2024) apresentaram um recall de 0,72, o menor entre os trabalhos analisados. Esse resultado propõe que a metodologia adotada neste estudo possibilitou uma melhor recuperação dos objetos detectados.

Entretanto, a precisão pelo modelo treinado neste estudo atingiu (0,75), que é inferior aos demais trabalhos analisados. (THOMAS et al.,) alcançaram a maior precisão (0,90), seguidos por (LIU et al., 2024) (0,87) e (JIANG; WU, 2024) (0,85), indicando que o modelo deste trabalho desenvolvido apresenta uma maior incidência de falsos positivos. Isso enfatiza que apesar da elevada taxa de detecção, o modelo pode estar classificando alguns elementos como objetos de interesse.

Desse modo, os resultados demonstram que a configuração empregada neste trabalho favorece uma detecção ampla dos objetos, proporcionando maior recall e mAP, mas com um critério de decisão menos rigoroso, resultando em uma redução na precisão quando comparado aos trabalhos anteriores.

5.1 Limitações

Apesar dos resultados obtidos, o presente estudo ainda apresenta algumas limitações que devem ser consideradas para uma interpretação mais abrangente e crítica dos achados. A principal delas está relacionada à quantidade reduzida de épocas de treinamento — apenas cinco para todos os modelos avaliados. Essa limitação se da pela falta de recursos computacionais mais avançados para a realização dos experimentos, uma vez que todo o treinamento foi conduzido em ambiente virtual (Google Colab) utilizando processamento via CPU. No entanto, tal decisão pode ter comprometido o desempenho de arquiteturas mais complexas e profundas, como o Faster R-CNN e o RetinaNet, que, em geral, exigem um número maior de iterações para que seus parâmetros sejam devidamente ajustados e alcancem níveis superiores de desempenho.

Além disso, embora o conjunto de dados utilizado tenha totalizado 3.175 imagens da filmagem de 99 ovelhas em ambientes reais, o número relativamente reduzido de indivíduos podem restringir a capacidade de generalização dos modelos. Como solução para mitigar essa limitação, seria uma opção viável a aplicação de técnicas de aumento de dados (data augmentation), ampliando artificialmente a variedade de amostras e contribuindo para maior robustez durante o treinamento. No entanto, não foi realizada uma análise qualitativa das falhas de detecção.

6 Conclusão

Este estudo apresentou uma análise comparativa de quatro arquiteturas de redes neurais convolucionais aplicadas à detecção de ovelhas em ambientes naturais: YOLOv8, Faster R-CNN, RetinaNet e SSD. Os experimentos buscaram identificar o modelo com melhor desempenho com base nas métricas de mAP, Recall e Precision. Conforme apresentado na Tabela 2, o YOLOv8 demonstrou o melhor desempenho geral, alcançando mAP de 0.98 e Recall de 0,91, além de apresentar o menor tempo médio de inferência (0,19s), o que o torna altamente adequado para aplicações em tempo real.

Além disso, a análise dos parâmetros de cada modelo revelou que fatores como a profundidade da rede, taxa de aprendizado, uso de pré-treinamento e número de épocas influenciaram diretamente nos resultados. Como limitação do estudo, destaca-se o uso de apenas 5 épocas de treinamento por modelo, adotado para padronizar o tempo computacional entre as abordagens. Essa limitação pode ter restringido a capacidade dos modelos mais complexos de atingirem sua performance ideal, principalmente aqueles sem pré-treinamento.

Como trabalhos futuros, propõe-se a implementação de um sistema de detecção em tempo real baseado no YOLOv8, explorando sua eficiência para aplicações em campo. Além disso, pretende-se incorporar técnicas de aprendizado contínuo, permitindo que o modelo se adapte gradualmente a mudanças no comportamento, aparência e contexto ambiental dos animais ao longo do tempo. Também será conduzido um treinamento mais extenso, com um número maior de épocas e o uso de conjuntos de dados mais abrangentes, visando melhorar a capacidade de generalização dos modelos — especialmente das arquiteturas mais profundas, como o RetinaNet e o Faster R-CNN, que podem se beneficiar significativamente de configurações mais robustas de treinamento.

Em conjunto, os achados deste trabalho reforçam a viabilidade do uso do YOLOv8 como solução eficaz e eficiente para a detecção de ovelhas em ambientes naturais, representando um avanço promissor para a automação no manejo de rebanhos e a modernização da pecuária de precisão.

Referências

- AGARWAL, N.; CHIANG, C.-W.; SHARMA, A. A study on computer vision techniques for self-driving cars. In: SPRINGER. *international conference on frontier computing*. [S.l.], 2018. p. 629–634. Citado na página 13.
- ALBAHLI, S. et al. An improved faster-rcnn model for handwritten character recognition. *Arabian Journal for Science and Engineering*, Springer, v. 46, n. 9, p. 8509–8523, 2021. Citado na página 14.
- BANSAL, A. et al. Risk ranked recall: Collision safety metric for object detection systems in autonomous vehicles. In: IEEE. *2021 10th Mediterranean conference on embedded computing (MECO)*. [S.l.], 2021. p. 1–4. Citado na página 21.
- BAOYUAN, C.; YITONG, L.; KUN, S. Research on object detection method based on ff-yolo for complex scenes. *IEEE Access*, IEEE, v. 9, p. 127950–127960, 2021. Citado na página 20.
- BJERGE, K. et al. Accurate detection and identification of insects from camera trap images with deep learning. *PLOS Sustainability and Transformation*, Public Library of Science San Francisco, CA USA, v. 2, n. 3, p. e0000051, 2023. Citado na página 11.
- DHILLON, A.; VERMA, G. K. Convolutional neural network: a review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, Springer, v. 9, n. 2, p. 85–112, 2020. Citado na página 14.
- DICK, A. R.; BROOKS, M. J. Issues in automated visual surveillance. In: *International conference on digital image computing: techniques and applications*. [S.l.: s.n.], 2003. Citado na página 13.
- DJOLONGA, J. et al. On robustness and transferability of convolutional neural networks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2021. p. 16458–16468. Citado na página 15.
- DU, L.; ZHANG, R.; WANG, X. Overview of two-stage object detection algorithms. In: IOP PUBLISHING. *Journal of Physics: Conference Series*. [S.l.], 2020. v. 1544, n. 1, p. 012033. Citado na página 14.
- FERGUS, P. et al. Harnessing artificial intelligence for wildlife conservation. *Conservation*, MDPI, v. 4, n. 4, p. 685–702, 2024. Citado 2 vezes nas páginas 6 e 15.
- GALVEZ, R. L. et al. Object detection using convolutional neural networks. In: IEEE. *TENCON 2018-2018 IEEE region 10 conference*. [S.l.], 2018. p. 2023–2027. Citado 2 vezes nas páginas 13 e 20.
- GEUTHER, B. Q. et al. Robust mouse tracking in complex environments using neural networks. *Communications biology*, Nature Publishing Group UK London, v. 2, n. 1, p. 124, 2019. Citado na página 14.

- HENDERSON, P.; FERRARI, V. End-to-end training of object class detectors for mean average precision. In: SPRINGER. *Computer Vision–ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part V 13*. [S.l.], 2017. p. 198–213. Citado na página 21.
- JETHANI, S. et al. Surveillance system for monitoring social distance. In: SPRINGER. *International Advanced Computing Conference*. [S.l.], 2020. p. 100–112. Citado na página 14.
- JIANG, L.; WU, L. Enhanced yolov8 network with extended kalman filter for wildlife detection and tracking in complex environments. *Ecological Informatics*, Elsevier, v. 84, p. 102856, 2024. Citado 2 vezes nas páginas 24 e 25.
- JIANG, P. et al. A review of yolo algorithm developments. *Procedia computer science*, Elsevier, v. 199, p. 1066–1073, 2022. Citado na página 14.
- KHAN, A. A.; LAGHARI, A. A.; AWAN, S. A. Machine learning in computer vision: A review. *EAI Endorsed Transactions on Scalable Information Systems*, v. 8, n. 32, 2021. Citado na página 13.
- KUMAR, A.; ZHANG, Z. J.; LYU, H. Object detection in real time based on improved single shot multi-box detector algorithm. *EURASIP Journal on Wireless Communications and Networking*, Springer, v. 2020, n. 1, p. 204, 2020. Citado na página 20.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. Citado na página 11.
- LIN, T.-Y. et al. Focal loss for dense object detection. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, p. 2980–2988, 2017. Citado na página 20.
- LIU, W. et al. Ssd: Single shot multibox detector. *Proceedings of the European Conference on Computer Vision (ECCV)*, p. 21–37, 2016. Citado 2 vezes nas páginas 14 e 20.
- LIU, Y. et al. Camouflage detection: Optimization-based computer vision for alligator sinensis with low detectability in complex wild environments. *Ecological Informatics*, Elsevier, v. 83, p. 102802, 2024. Citado 2 vezes nas páginas 24 e 25.
- MENON, H. P. et al. A study on yolov5 for drone detection with google colab training. In: IEEE. *2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS)*. [S.l.], 2023. p. 1576–1580. Citado na página 19.
- MITTAL, P.; SINGH, R.; SHARMA, A. Deep learning-based object detection in low-altitude uav datasets: A survey. *Image and Vision computing*, Elsevier, v. 104, p. 104046, 2020. Citado na página 13.
- NEUPANE, J. e. a. A literature review on deep learning applications for livestock monitoring. *Computers and Electronics in Agriculture*, 2022. Citado na página 16.
- NEUPANE, S. B.; SATO, K.; GAUTAM, B. P. A literature review of computer vision techniques in wildlife monitoring. *IJSRP*, v. 16, p. 282–295, 2022. Citado na página 11.

- NGUYEN, H. et al. Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In: IEEE. *2017 IEEE international conference on data science and advanced Analytics (DSAA)*. [S.l.], 2017. p. 40–49. Citado 2 vezes nas páginas 11 e 13.
- NOBRE, F. L. S. Detecção e remoção automática de objetos indesejados em imagens utilizando aprendizagem profunda. 2021. Citado 2 vezes nas páginas 6 e 13.
- OKSUZ, K. et al. Localization recall precision (lrp): A new performance metric for object detection. In: *Proceedings of the European conference on computer vision (ECCV)*. [S.l.: s.n.], 2018. p. 504–519. Citado na página 21.
- PADILLA, R.; NETTO, S. L.; SILVA, E. A. D. A survey on performance metrics for object-detection algorithms. In: IEEE. *2020 international conference on systems, signals and image processing (IWSSIP)*. [S.l.], 2020. p. 237–242. Citado na página 17.
- PARK, M. J.; SACCHI, M. D. Automatic velocity analysis using convolutional neural network and transfer learning. *Geophysics*, Society of Exploration Geophysicists, v. 85, n. 1, p. V33–V43, 2020. Citado na página 11.
- RAKHIMOV, B. S. et al. Review and analysis of computer vision algorithms. *The American Journal of Applied sciences*, v. 3, n. 5, p. 245–250, 2021. Citado na página 13.
- RAWAT, W.; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation*, MIT Press, v. 29, n. 9, p. 2352–2449, 2017. Citado na página 13.
- REDMON, J. et al. You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 779–788, 2016. Citado na página 20.
- REN, S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, p. 91–99, 2015. Citado na página 20.
- ROBERTSON, G.; CRAW, I. Testing face recognition systems. *Image and Vision Computing*, Elsevier, v. 12, n. 9, p. 609–614, 1994. Citado na página 13.
- ROSS, T.-Y.; DOLLÁR, G. Focal loss for dense object detection. In: *proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 2980–2988. Citado na página 20.
- RURAL, C. *Embrapa já identifica bovinos individualmente por imagem*. 2023. Acesso em: 10 jun. 2025. Disponível em: <<https://agevolution.canalrural.com.br/embrapa-ja-identifica-bovinos-individualmente-por-imagem/>>. Citado 2 vezes nas páginas 6 e 15.
- SAQIB, M. et al. Real-time drone surveillance and population estimation of marine animals from aerial imagery. In: IEEE. *2018 International conference on image and vision computing New Zealand (IVCNZ)*. [S.l.], 2018. p. 1–6. Citado na página 15.
- SCHINDLER, F.; STEINHAGE, V. Identification of animals and recognition of their actions in wildlife videos using deep learning techniques. *Ecological Informatics*, Elsevier, v. 61, p. 101215, 2021. Citado na página 15.

- SCHNEIDER, S.; TAYLOR, G. W.; KREMER, S. Deep learning object detection methods for ecological camera trap data. In: IEEE. *2018 15th Conference on computer and robot vision (CRV)*. [S.l.], 2018. p. 321–328. Citado na página 16.
- SILVA, J. e. a. Aprendizagem profunda para monitoramento de fauna com drones. *Revista de Visão Computacional*, 2021. Citado na página 16.
- SILVA, V. J. P. d. et al. Aprendizagem de máquina aplicada ao monitoramento da presença de animais em reservas naturais de ambientes industriais. Universidade Federal de Campina Grande, 2021. Citado na página 11.
- SIMÕES, F.; BOUYEYRON, C.; PRECIOSO, F. Deepwild: Wildlife identification, localisation and estimation on camera trap videos using deep learning. *Ecological Informatics*, Elsevier, v. 75, p. 102095, 2023. Citado na página 15.
- SONG, Z. et al. Contextualizing object detection and classification. In: IEEE. *CVPR 2011*. [S.l.], 2011. p. 1585–1592. Citado na página 14.
- THOMAS, M. S. et al. Wildlife detection and recognition using yolo v8. Citado 2 vezes nas páginas 24 e 25.
- VAINTRUB, M. O. et al. Precision livestock farming, automats and new technologies: Possible applications in extensive dairy sheep farming. *Animal*, Elsevier, v. 15, n. 3, p. 100143, 2021. Citado na página 15.
- YUN, S. *Computer Vision – Object Detection: One-Stage vs Two-Stage*. 2021. Acesso em: 10 jun. 2025. Disponível em: <<https://sharkyun.medium.com/computer-vision-object-detection-one-stage-vs-two-stage-b05dbff88195>>. Citado 2 vezes nas páginas 6 e 14.
- ZHAO, X. et al. A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, Springer, v. 57, n. 4, p. 99, 2024. Citado na página 13.

Apêndices

APÊNDICE A – Apêndice



TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA “JOSÉ ALBANO DE MACEDO”

Identificação do Tipo de Documento

- Tese
- Dissertação
- Monografia
- Artigo

Eu, **João dos Santos Neto**, autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de 02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação **Análise Comparativa de Modelos de Detecção para Identificação de Pequenos Ruminantes em Ambientes Naturais** de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título de divulgação da produção científica gerada pela Universidade.

Picos-PI 26 de Junho de 2025.

Documento assinado digitalmente
gov.br JOAO DOS SANTOS NETO
Data: 26/06/2025 22:18:41-0300
Verifique em <https://validar.iti.gov.br>

Assinatura