André Amaro Pereira

# Impacto do Data Augmentation na Detecção de Bovinos em Bezerreiros Tropicais com Modelos Baseados na Arquitetura YOLO: Uma Análise Comparativa, via Roboflow, em Cenários com Restrição de Dados.

Recife

2026

André Amaro Pereira

# Impacto do Data Augmentation na Detecção de Bovinos em Bezerreiros Tropicais com Modelos Baseados na Arquitetura YOLO: Uma Análise Comparativa, via Roboflow, em Cenários com Restrição de Dados.

Artigo Científico apresentado ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Universidade Federal Rural de Pernambuco – UFRPE

Departamento de Estatística e Informática

Curso de Bacharelado em Sistemas de Informação

Orientador: Profº Victor Wanderley Costa de Medeiros

Banca Examinadora - Profº Cícero Garrozi

Recife

2026

A Deus, aos meus pais, por serem a base da minha vida e por me ensinarem, com amor e dedicação, valores que me trouxeram até aqui.

À minha esposa e aos meus filhos, pelo amor, paciência, compreensão e apoio nos momentos em que precisei me ausentar para seguir em busca deste objetivo.

À minha família, especialmente à minha irmã Catarina Alexsandra, e a todos que, direta ou indiretamente, contribuíram para que esta conquista se tornasse possível.

# Agradecimentos

Agradeço, primeiramente, a Deus, por me conceder força, saúde, sabedoria e perseverança para enfrentar os desafios, que não foram fáceis, nesta caminhada.

Aos meus pais, minha eterna gratidão por todo amor, cuidados, apoio e ensinamentos, que me proporcionaram ser a pessoa que sou. Sem a base que recebi deles, certamente eu não teria chegado até aqui.

À minha esposa e aos meus filhos, agradeço pelo amor, pela paciência e pela compreensão nos momentos em que precisei me ausentar, dedicar noites, fins de semana e parte do meu tempo à construção deste trabalho. O apoio de vocês foi fundamental para que eu pudesse seguir firme até a conclusão desta etapa.

Aos professores que fizeram parte da minha trajetória acadêmica, deixo meu sincero agradecimento. Cada um, à sua maneira, contribuiu para que eu enxergasse a tecnologia, a informática e a pesquisa de uma forma mais ampla, crítica e transformadora. Em especial, agradeço aos professores Victor Medeiros, Cícero Garrozi, Cleviton Monteiro, Cleyton Magalhães, Roberta Gouveia, Silvana Bocanegra, Gabriel Alves, Rodrigo Assad, Maria da Conceição Moares, Gilberto Cysneiros, Guillherme Vilar, Rodrigo Nonamor, Marcelo Gama e a todos os professores do Departamento de Estatística e Informática, Departamento de Administração e Departamento de Computação por suas imensuráveis e incansáveis contribuições ao longo da minha formação acadêmica.

Agradeço também aos colegas de sala Keila, Raquel, Caio, Leonardo, Danielly José Francisco e todos os outros, pessoas admiráveis, inteligentes, generosas e de grande coração. Cada conversa, ajuda, incentivo e gesto de companheirismo tornou esta jornada mais leve e significativa. Sou grato por terem caminhado comigo, segurado a minha mão, e por contribuírem, de forma tão espontânea e carinhosa, para que eu pudesse avançar em cada etapa do curso.

Por fim, agradeço a todos que, direta ou indiretamente, fizeram parte desta caminhada. Cada apoio, palavra de incentivo, orientação e gesto de colaboração teve importância na realização deste trabalho e na concretização deste sonho.

*“Ninguém educa ninguém, ninguém educa a si mesmo, os homens se educam entre si, mediatizados pelo mundo.”*

*(Paulo Freire)*

## RESUMO

A utilização de técnicas de Visão Computacional na bovinocultura tem se mostrado uma alternativa promissora para automatizar atividades de monitoramento animal em ambientes de produção. No entanto, o desempenho de modelos de detecção de objetos depende diretamente da qualidade, da quantidade e da diversidade das imagens disponíveis para treinamento, o que representa um desafio em cenários com restrição de dados. Neste contexto, este trabalho investiga o impacto de diferentes técnicas de *Data Augmentation* no desempenho de modelos baseados na arquitetura YOLO para a detecção de bovinos em bezerreiros tropicais. A pesquisa utilizou conjuntos de dados compostos por 50, 100 e 228 imagens coletadas em ambiente real de produção leiteira, originário de um projeto de pesquisa denominado “Criação de bezerras sob diferentes métodos de fornecimento de água de dessedentação durante a fase de aleitamento”, coordenado pelo departamento de Ciências Agrárias da UFRPE, cujas imagens apresentavam variações de iluminação, oclusões e movimentação de animais. Essas imagens foram organizadas, anotadas e processadas na plataforma Roboflow, onde foram aplicadas transformações geométricas e fotométricas, incluindo espelhamento, rotação, ajustes de brilho e saturação, conversão para escala de cinza e inserção de ruído. Os experimentos foram organizados em grupos de teste e avaliados por meio das métricas mAP@0.5, Precision e Recall. Os resultados demonstraram que o impacto do *Data Augmentation* varia conforme o volume de dados disponível e a natureza das transformações aplicadas. Transformações fotométricas moderadas apresentaram os resultados mais consistentes, especialmente na base de 100 imagens, contribuindo para melhorias nas métricas de detecção. Em contrapartida, combinações excessivas de transformações geométricas e fotométricas produziram degradação do desempenho, evidenciando que o aumento artificial de dados não garante ganhos automáticos de generalização. Conclui-se, portanto, que a eficácia do *Data Augmentation* depende diretamente da coerência entre as transformações aplicadas e as características reais do domínio investigado, sendo mais eficiente quando empregado de forma controlada e compatível com as condições do ambiente de criação animal.



**Palavras-chave:** Visão Computacional; Detecção de Objetos; Data Augmentation; YOLO; Roboflow; Bovinocultura de Precisão.

## ABSTRACT

The use of Computer Vision techniques in cattle farming has emerged as a promising alternative for automating animal monitoring activities in production environments. However, the performance of object detection models depends directly on the quality, quantity, and diversity of the images available for training, which represents a challenge in data-constrained scenarios. In this context, this study investigates the impact of different Data Augmentation techniques on the performance of YOLO-based models for cattle detection in tropical calf-rearing facilities. The research employed datasets composed of 50, 100, and 228 images collected in a real dairy production environment, originating from a research project entitled “Raising Dairy Calves Under Different Drinking Water Supply Methods During the Milk Feeding Phase”, coordinated by the Department of Agricultural Sciences at the Federal Rural University of Pernambuco (UFRPE), whose images exhibited variations in lighting conditions, occlusions, and animal movement. The images were organized, annotated, and processed using the Roboflow platform, where geometric and photometric transformations were applied, including image flipping, rotation, brightness and saturation adjustments, grayscale conversion, and noise insertion. The experiments were organized into test groups and evaluated using the mAP@0.5, Precision, and Recall metrics. The results demonstrated that the impact of Data Augmentation varies according to the amount of data available and the nature of the transformations applied. Moderate photometric transformations produced the most consistent results, particularly in the 100-image dataset, contributing to improvements in object detection performance. Conversely, excessive combinations of geometric and photometric transformations led to performance degradation, demonstrating that artificial data expansion does not automatically guarantee gains in model generalization. Therefore, it is concluded that the effectiveness of Data Augmentation depends directly on the consistency between the applied transformations and the real characteristics of the investigated domain, being more effective when employed in a controlled manner and aligned with the conditions of the animal production environment.



**Keywords:** Computer Vision; Object Detection; Data Augmentation; YOLO; Roboflow; Precision Livestock Farming.

**SUMÁRIO**

## 1. INTRODUÇÃO

O monitoramento de rebanhos bovinos em ambientes de bezerreiros tropicais representa um desafio crescente para a agropecuária de precisão, sobretudo ao tentar realizar esse processo de maneira eficiente. Variações de iluminação ao longo do dia, movimentação constante de animais e oclusão parcial entre os indivíduos são alguns dos fatores que implicam em falhas e aumento de complexidade no processo de supervisão manual, reforçando a necessidade de sistemas automatizados baseados em Visão Computacional.

Diante deste cenário, os modelos de detecção de objetos baseados em Deep Learning, com destaque para modelos da família YOLO (*You Only Look Once*) (REDMON et al., 2016) têm conquistado mais espaço dentre as ferramentas utilizadas para a identificação automática de animais em imagens e vídeos. A capacidade de operar em tempo real, com alto desempenho, mesmo diante de variações visuais adversas, contribui para a adequação de arquiteturas semelhantes no contexto de produção animal (BOCHKOVSKIY; WANG; LIAO, 2020).

No entanto, a confiança nos modelos construídos depende diretamente da qualidade, da quantidade e da diversidade do conjunto de dados disponível, uma vez que bases de dados limitadas tendem a impactar negativamente nos resultados obtidos. Uma das abordagens mais utilizadas para lidar com restrições de dados é o *Data Augmentation*, que consiste na aplicação de transformações controladas sobre os dados de treinamento, como, por exemplo, alterações geométricas e fotométricas capazes de elevar a diversidade do conjunto de dados (HANEL, 2021).

Ainda assim, essas técnicas não asseguram a eficácia do modelo, uma vez que transformações que não representam adequadamente o ambiente real podem introduzir ruídos indesejados e comprometer o bom desempenho do modelo.

Este trabalho está inserido em um projeto de pesquisa mais amplo denominado “Criação de bezerras sob diferentes métodos de fornecimento de água de dessedentação durante a fase de aleitamento”, coordenado pelo departamento de Ciências Agrárias da UFRPE, no contexto da agropecuária digital, voltado à bovinocultura, que tem como objetivo avaliar as respostas fisiológicas, comportamento ingestivo e desempenho de bezerras mestiças criadas em bezerreiros tropicais. O projeto contempla a coleta de dados em ambiente real de produção, bem como o uso

de tecnologias não invasivas, tais como: sensores embarcados e processamento de imagens para estimar peso e monitorar o comportamento animal.

Nesse contexto, o presente trabalho se propõe a investigar o impacto de diferentes técnicas de *Data Augmentation* no desempenho de modelos de Visão Computacional baseados na arquitetura YOLO, aplicados à detecção e análise de comportamentos de bezerros (bovinos) em ambientes reais de bezerreiros tropicais. Para isso, utiliza-se a plataforma Roboflow como pipeline experimental, permitindo a criação e comparação sistemática de diferentes configurações de aumento de dados.

A relevância deste estudo está na avaliação prática e comparativa das técnicas de *Data Augmentation* em um contexto real, contribuindo para a compreensão de sua viabilidade e limitações em cenários com restrição de dados. Além disso, o trabalho busca identificar quais estratégias são mais adequadas para melhorar métricas de desempenho do modelo, tais como *mAP, Precision* e *Recall*, sem comprometer a representatividade do domínio analisado.

## 2. OBJETIVOS

Esta seção descreve o objetivo geral deste trabalho, detalhando os seus objetivos específicos.

### 2.1 OBJETIVO GERAL

O presente trabalho tem como objetivo investigar o impacto de diferentes técnicas de *Data Augmentation* no desempenho de modelos de visão computacional baseados na arquitetura YOLO para a detecção e classificação de comportamentos de bovinos em bezerreiros tropicais.

### 2.2 OBJETIVOS ESPECÍFICOS

Como estratégia para alcançar o objetivo geral definido, foram definidos seis objetivos específicos, listados a seguir:

- Mapear os principais conceitos acerca dos temas relacionados com Visão Computacional, Detecção de Objetos e *Data Augmentation*, visando a construção do referencial teórico;
- Organizar o conjunto de dados (imagens), realizando anotações e rotulagens na plataforma Roboflow;
- Selecionar e aplicar técnicas de *Data Augmentation* disponíveis na versão gratuita da plataforma Roboflow, incluindo transformações geométricas e fotométricas;
- Realizar um conjunto de experimentos para comparar cenários com e sem aumento de dados;
- Avaliar o desempenho do modelo de detecção de objetos a partir das diferentes transformações aplicadas no pré-processamento de dados;
- Analisar o impacto da utilização de técnicas de modelagem de dados baseadas nas métricas mAP, Precision e Recall.

## 3. TRABALHOS RELACIONADOS

No contexto da agropecuária de precisão, técnicas de Visão Computacional têm sido utilizadas de forma ampla para investigar a eficiência da automação do monitoramento animal como estratégia em sistemas de produção bovina. Estudos têm apresentado avanços relevantes na aplicação de modelos de detecção de objetos, com destaque para arquiteturas baseadas em Deep Learning, no processo de identificação e acompanhamento de animais em ambientes reais de bezerreiros tropicais.

Nolêto et al. (2023), apontam um forte crescimento no uso de redes neurais artificiais na pecuária e demonstram que aplicações com foco em identificar e monitorar o comportamento de animais têm impactado positivamente na previsão do desempenho produtivo. Na sua revisão sistemática, os autores ressaltam que essas abordagens apresentam resultados positivos, mesmo com limitações ligadas às variações adversas nos ambientes, em especial ao tratar da aplicação em dados visuais.

Na mesma direção, Moreira (2022) discute sobre os desafios que envolvem o uso de visão computacional em imagens de animais obtidas em campo e enfatiza o impacto que aspectos como variação de iluminação, ângulo, movimentação e oclusão parcial podem causar. Tais fatores interferem no processo de treinamento, tornando-o mais complexo e prejudicando a capacidade de generalização dos modelos, sobretudo quando se trata de cenários com restrição na disponibilidade de dados.

Em linhas gerais, a aplicação de técnicas de *Data Augmentation* costuma ser explorada para mitigar o problema gerado pelo baixo volume de dados disponíveis, sendo uma alternativa capaz de ampliar o conjunto de dados artificialmente e, consequentemente, melhorando o desempenho do modelo. Em seu estudo, Rici, Santos e Ottoni (2021) defendem que a aplicação de transformações geométricas e fotométricas contribui consideravelmente para a precisão e acurácia dos modelos de classificação. Contudo, os autores alertam que essa técnica só é eficiente se as transformações forem aplicadas de maneira coerente e controlada.

Hanel (2021) e Moreira (2022) apontam ainda, que ao aplicar transformações de forma descontrolada há o risco de introduzir ruídos no conjunto de dados, gerando amostras irreais e prejudicando o desempenho do modelo. Ao se trabalhar em contextos reais, tais como a bovinocultura, esse fator possui alta relevância, uma vez que as condições ambientais tendem a apresentar características específicas que devem ser respeitadas durante o processo de *Data Augmentation*.

Em relação à escolha da plataforma, Da Silva et al (2023) apresentam a viabilidade no uso do Roboflow como ambiente de treinamento de modelos de redes neurais focados na detecção de objetos em tempo real. A plataforma utiliza de modelos pré-treinados, que segundo os autores, colaboram para a simplificação no processo de construção de novos modelos. Isso torna a implementação de visão computacional algo mais acessível, em especial em cenários de baixa infraestrutura especializada.

É possível observar que boa parte dos estudos estão concentrados em aplicações com bases de dados padronizados, com foco em problemas de classificação de imagens. A quantidade de trabalhos voltados para a detecção em ambientes reais e com restrição de dados é bem menor. Além disso, as pesquisas

geralmente realizam comparações sistemáticas entre diferentes estratégias de aumento de dados, com aplicação em modelos da família YOLO.

Dessa forma, foi identificada uma lacuna na literatura em relação à avaliação prática e comparativa do impacto gerado por diferentes técnicas de *Data Augmentation* em modelos de detecção de objetos aplicados à bovinocultura em ambientes tropicais. Deste modo, essa pesquisa tem como finalidade contribuir para o avanço dessa área de estudo, a partir da investigação acerca da eficácia das técnicas de *Data Augmentation* como forma de melhorar o desempenho de modelos baseados na arquitetura YOLO, em cenários reais com restrição de dados.

## 4. FUNDAMENTAÇÃO TEÓRICA

Esta seção tem como objetivo apresentar os principais conceitos relacionados com o presente trabalho, nela são apresentados: Visão Computacional, Detecção de Objetos, *Data Augmentation*, Arquitetura YOLO e Roboflow.

### 4.1. VISÃO COMPUTACIONAL

A detecção de objetos é a base da área de Visão Computacional, com foco em localizar e identificar objetos de estudo em uma imagem ou vídeos. Segundo Marim (2019), a detecção de objetos pode ser definida como a combinação da Localização e Classificação de um objeto a partir de imagens, isto é: a junção de dois subproblemas.

Esse processo faz parte de um conjunto de tarefas executáveis sobre uma imagem, a partir da Visão Computacional. A Classificação, por exemplo, é o nível mais básico, nela o modelo identifica qual o objeto principal da imagem observada. Indo além, a Localização é responsável por determinar em que local da imagem esse objeto está (Marim, 2019).

O processo de detecção tem como elemento fundamental o *bounding box* (caixa delimitadora), que consiste em um retângulo que cerca o objeto detectado, definido por coordenadas que indicam sua posição e dimensão na imagem (WANGENHEIM, 2018). O *bounding box* é o maior fruto de uma detecção na prática,

pois é a partir dele que o sistema identifica o local exato onde intervir, recortar ou rastrear um objeto em uma sequência de *frames*.

Vale salientar que, apesar do avanço das técnicas de detecção, ainda existem uma série de desafios para o treinamento de modelos que executam essa tarefa, a mais comum é o *overfitting* (sobreajuste). Por definição, o *overfitting* ocorre quando um modelo se apega aos detalhes e ruídos dos dados utilizados durante o treinamento, impactando negativamente o seu desempenho diante de dados diferentes. Como resultado, o modelo gera previsões pouco confiáveis, sendo incapaz de generalizar para novos cenários.

No contexto das redes neurais, esse comportamento se manifesta quando há um elevado número de camadas e neurônios, sendo agravado em cenários com baixo volume de dados, excesso no tempo de treinamento e arquiteturas muito profundas (Mota, 2025). Uma técnica comumente aplicada para reduzir o risco de *overfitting* é o aumento artificial de dados, conhecida popularmente como *Data Augmentation*. O objetivo desta técnica é justamente combater o *overfitting*, pois força o modelo a entender que um determinado objeto mantém a sua identidade independentemente de estar invertido verticalmente (ponta-cabeça), no escuro ou posicionado nos cantos da fotografia.

A Visão Computacional é uma área da Inteligência Artificial (IA) que, segundo Savekar e Kumar (2021), é dedicada a estudar a capacitação, interpretação e compreensão de informações a partir de imagens e vídeos do mundo real. Essa tecnologia se originou a partir do estudo de Lawrence Roberts, um cientista da Computação que realizou uma das primeiras publicações no campo em 1963, no MIT (Massachusetts Institute of Technology).

Para Oliveira e Melo (2023), é uma área promissora dado o seu potencial em demandas de inspeção e automação, sendo uma ótima ferramenta para tomada de decisão. A visão computacional opera através de um processo de análise visual, a Fundação CERTI (2026) utiliza uma analogia para descrever o processo: máquinas, câmeras e sensores fazem o papel dos olhos, de forma que as redes neurais recebam a imagem e extraiam as informações através de algoritmos de inteligência artificial. Na prática, isso inclui desde aspectos básicos como cores e pixels, até análises mais complexas, tais como contornos e objetos.

Como possibilidades de aplicação prática, se destacam: serviços ligados à saúde e ao agronegócio. O segmento da pecuária, em específico, apresenta ampla

diversidade de aplicações, uma vez que a capacidade de ver, analisar e compreender imagens aumenta consideravelmente com o uso de programas computacionais (Nolêto et al., 2023). Na bovinocultura, essas ferramentas costumam ser utilizadas para contagem de animais, rastreamento de postura, monitoramento de comportamento e avaliação de saúde.

As redes neurais artificiais são as principais responsáveis pelo avanço da visão computacional no contexto da pecuária. Por definição, Redes Neurais Artificiais são sistemas computacionais formados por unidades interconectadas, semelhante aos neurônios do cérebro biológico, cujo objetivo é a transmissão de sinais (Angarita-Zapata et al., 2021). Há, ainda, um subconjunto de redes neurais aplicados em contextos mais complexos, são conhecidas como *Deep Learning* (redes neurais profundas) e atuam em processamento de altos volumes de dados.

O trabalho de Gomes (2025) relata que tecnologias envolvendo redes neurais e visão computacional apresentam resultados positivos quando usadas para prever ganho de peso, monitorar comportamento e saúde animal, sendo aliadas para detectar distúrbios diversos e minimizar emissão de gases de efeito estufa. Nolêto et al. (2023), através de uma revisão sistemática sobre o tema, apontaram que o uso das redes neurais na pecuária vem crescendo devido a sua eficácia na previsão da produção e identificação de animais por meio de detecção de imagens.

## 4.2. DETECÇÃO DE OBJETOS

A detecção de objetos é a base da área de Visão Computacional, com foco em localizar e identificar objetos de estudo em uma imagem ou vídeos. Segundo Marim (2019), a detecção de objetos pode ser definida como a combinação da Localização e Classificação de um objeto a partir de imagens, isto é: a junção de dois subproblemas.

Esse processo faz parte de um conjunto de tarefas executáveis sobre uma imagem, a partir da visão computacional. A Classificação, por exemplo, é o nível mais básico, nela o modelo identifica qual o objeto principal da imagem observada. Indo além, a Localização é responsável por determinar em que local da imagem esse objeto está (Marim, 2019).

O processo de detecção tem como elemento fundamental o *bounding box* (caixa delimitadora), que se trata de um retângulo que cerca o objeto detectado, e é definido por coordenadas que indicam sua posição e dimensão na imagem (Wangenheim, 2018). O *bounding box* é o maior fruto de uma detecção na prática, pois é a partir dele que o sistema sabe o local exato onde intervir, recortar ou rastrear um objeto em uma sequência de frames.

Vale salientar que, apesar do avanço das técnicas de detecção, ainda existem uma série de desafios para o treinamento de modelos que executam essa tarefa, a mais comum é o *overfitting* (sobreajuste). Por definição, o *overfitting* ocorre quando um modelo se apega aos detalhes e ruídos dos dados utilizados durante o treinamento, de modo a impactar o seu desempenho com dados diferentes. Como resultado, o modelo gera previsões pouco confiáveis, sendo incapaz de generalizar para novos dados.

No contexto das redes neurais, esse comportamento se manifesta quando há um grande número de camadas e neurônios, e é agravado em cenários com baixo volume de dados, excesso no tempo de treinamento e arquiteturas muito profundas (Mota, 2025). Uma técnica comumente aplicada para reduzir o risco de *overfitting* é o aumento artificial de dados, mais conhecida como *Data Augmentation*. O objetivo desta técnica é justamente combater o *overfitting*, pois força o modelo a entender que um determinado objeto continua sendo um determinado objeto, esteja ele de ponta-cabeça, no escuro ou no canto da imagem.

## 4.3. DATA AUGMENTATION

*Data Augmentation* (*DA*) é uma técnica aplicada em cenários com baixo volume de dados para treinamento dos modelos de detecção de objetos. É comum que sistemas que utilizam aprendizagem de máquina apresentem dificuldades em coletar novos dados para o treinamento de uma rede neural, e segundo Hanel (2021), essa estratégia se apresenta como uma solução para esse tipo de problema.

A técnica de *DA* parte do princípio de que é possível utilizar um conjunto de dados existente para gerar dados sintéticos, de maneira artificial, gerando um aumento na base utilizada para treinamento (HANEL, 2021). Na prática, o *DA* aplica transformações controladas sobre as imagens pré-existentes, gerando novas

amostras com características visuais variadas, sem comprometer a classe do objeto presente no frame.

Dentre as transformações mais comuns, se destacam: rotação, espelhamento, *zoom*, recorte aleatório e ajustes de brilho e contraste. Essas práticas podem ser aplicadas tanto de maneira individual quanto combinadas, e isso aumenta as possibilidades de variação nos dados de treinamento, reduzindo o risco de *overfitting* e aumentando a robustez do modelo (Distrito, 2023).

É importante ter em mente que as vantagens de aplicar técnicas de *Data Augmentation* não se limitam à expansão do volume de dados. Segundo o estudo realizado por Rici, Santos e Ottoni (2021), o uso de imagens sintéticas, geradas por *Data Augmentation* no treinamento de modelos, podem gerar boas proporções nas taxas de classificação. O trabalho aponta a necessidade de utilizar combinações de transformações adequadas, isto é, bem calibradas. Algumas das transformações aplicadas no estudo como forma de melhorar a capacidade de generalização do modelo, são: *zoom*, rotação, intensidade de brilho e espelhamento horizontal (RICI; SANTOS; OTTONI, 2021).

No contexto da detecção de objetos em imagens de animais, o *Data Augmentation* apresenta um alto potencial dada as condições de captura das imagens, que raramente são padronizadas. Além disso, as imagens tendem a apresentar alta variação de iluminação, ângulo, movimento e oclusão parcial (MOREIRA, 2022). Desse modo, as transformações geométricas (translação, rotação e escala) e fotográficas (brilho e contraste) são aliadas no processo de tornar os modelos invariantes a essas variações, viabilizando a aprendizagem de características relevantes.

## 4.4. ARQUITETURA YOLO

O YOLO é uma das principais arquiteturas de modelos utilizados para detecção de objetos dentro da área de Visão Computacional. Sua sigla, “you only look once” (na tradução livre “você só olha uma vez”), foi proposta por Joseph Redmon et al. (2016), e trata a detecção de objetos como um problema de regressão, diferente das abordagens até então conhecidas, que tratavam a detecção como um processo de múltiplas etapas separadas.

Uma das características do YOLO é a sua rapidez, especialmente no contexto de aplicações em tempo real. Isso se deve ao fato de que o YOLO realiza todo o processo de detecção passando uma única vez pela rede neural, nomeado por Redmon et al (2016) como “*single-stage detector*”.

Segundo Redmon et al. (2016), o funcionamento da arquitetura consiste em dividir a imagem de entrada em uma *grid* (grade), de modo que o modelo seja responsável por prever as *bounding boxes* (caixas delimitadoras) e as classes dos objetos, baseadas em probabilidade, para cada célula da grade. Cada *bounding box* passa a ser definida por: coordenadas espaciais (largura e altura) e uma média de confiança, isto é, a probabilidade de conter um objeto e a precisão da localização indicada.

Além disso, o YOLO utiliza redes neurais convolucionais (*Convolutional Neural Networks*, da sigla CNNs) no processo de extração das características relevantes da imagem. As CNNs identificam os padrões visuais e permitem que o modelo seja capaz de reconhecer objetos mesmo diante de variações visuais adversas, tais como iluminação ou oclusão (BOCHKOVSKIY; WANG; LIAO, 2020).

No que tange detecção de objetos em ambientes não controlados, a arquitetura YOLO ganha destaque dada a sua capacidade de lidar com múltiplos objetos de forma simultânea, sendo sua aplicação viável para a identificação e monitoramento de animais em tempo real, independentemente das variações de iluminação, movimentação e sobreposição que podem afetar os cenários de entrada.

Naturalmente, o YOLO apresenta algumas limitações que se acentuam a depender da versão do modelo utilizada, que podem gerar um trade-off entre velocidade e precisão. Dentre elas se destaca: a dificuldade de detectar objetos muito pequenos ou muito próximos entre si, que na maioria das vezes esse comportamento está relacionado com a divisão fixa das grids definidas (REDMON et al., 2016).

Em linhas gerais, o YOLO se apresenta como uma abordagem eficiente para detecção de objetos, com ampla possibilidade de aplicação em sistemas que exigem velocidade e confiança no processamento, tais como os que envolvem sistemas de vigilância ou monitoramento e produção animal.

### 4.5. ROBOFLOW

O Roboflow é uma plataforma integrada de Visão Computacional desenvolvida para simplificar o ciclo completo de criação de modelos baseados em *deep learning*. Com o Roboflow, é possível treinar uma rede neural para reconhecer qualquer objeto em tempo real, podendo ser utilizado em diferentes aplicações do setor produtivo (DA SILVA et al., 2023).

A plataforma atua como um pipeline integrado para Visão Computacional, unindo etapas como organização de *datasets*, anotação de imagens, aplicação de *Data Augmentation*, treinamento e implantação de modelos em diferentes ambientes computacionais (SOLAWETZ, 2020b). Internamente, a plataforma Roboflow utiliza bibliotecas e frameworks de Deep Learning desenvolvidos predominantemente em Python, além de operações de processamento digital de imagens compatíveis com arquiteturas YOLO.

São disponibilizados, ainda, recursos avançados de gerenciamento e versionamento de dados, que permitem o armazenamento, estruturação e acompanhamento da evolução de conjuntos de imagens ao longo do processo experimental. O sistema de anotação do Roboflow combina rotulagem manual com mecanismos assistidos por inteligência artificial, o que possibilita a criação eficiente de *bounding boxes* e outros tipos de rótulos, garantindo padronização e qualidade — fatores determinantes para o desempenho de modelos de detecção de objetos (ROBOFLOW, 2025).

Outro componente central da ferramenta é o módulo de *Data Augmentation*, que oferece uma ampla variedade de transformações geométricas e fotométricas. A documentação oficial destaca que essas técnicas ampliam artificialmente a diversidade do dataset, contribuindo para a robustez e capacidade de generalização dos modelos, especialmente em cenários com restrição de dados (SOLAWETZ, 2020a).

Além disso, o Roboflow disponibiliza *pipelines* de treinamento compatíveis com arquiteturas contemporâneas, como YOLOv5, YOLOv8 e YOLO11, fornecendo relatórios detalhados e avaliação automática do modelo. Durante o treinamento dos experimentos com modelo YOLOv11, disponibilizado pelo Roboflow, a plataforma calcula automaticamente métricas como *mAP@0.5, Precision* e Recall, conforme descrito por Joseph e Johnson (2023), a partir de comparação entre as predições realizadas pelos modelos e as anotações reais do dataset.

No contexto deste estudo, o Roboflow foi utilizado para organizar o banco de imagens, coletado em um bezerreiro tropical; realizar a anotação dos objetos (bezerros e outros animais que constem na imagem); aplicar técnicas de aumento de dados e treinar modelos baseados na arquitetura YOLO. A plataforma desempenhou papel fundamental na condução dos experimentos comparativos, permitindo avaliar de forma sistemática o impacto do *Data Augmentation* na detecção de bovinos em cenários com limitações amostrais.

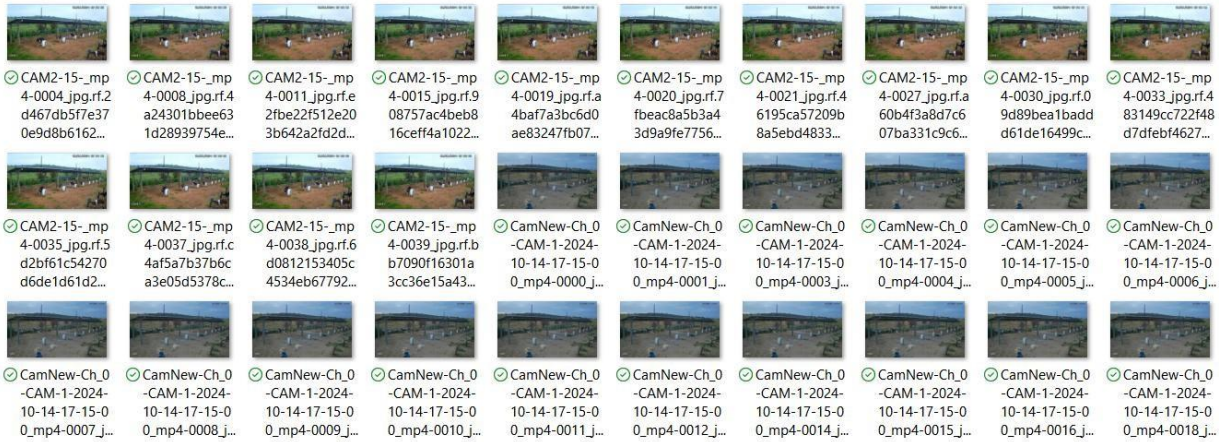
## 5. METODOLOGIA

O presente trabalho busca avaliar o impacto de diferentes técnicas de *Data Augmentation* no desempenho de modelos de detecção de objetos e, portanto, se caracteriza como uma pesquisa de abordagem quantitativa e natureza experimental aplicada. O estudo apresenta um caráter comparativo, uma vez que analisa diferentes configurações de pré-processamento de dados em cenários controlados.

Como insumo para o estudo, foi utilizado um conjunto de dados originário de um projeto de pesquisa mais amplo, denominado “Criação de bezerras sob diferentes métodos de fornecimento de água de dessedentação durante a fase de aleitamento”, coordenado pelo departamento de Ciências Agrárias da UFRPE. As atividades de campo do referido projeto foram conduzidas em uma propriedade comercial de criação de bovinos de leite, localizada no município de Capoeiras, Região Agreste do Estado de Pernambuco, com 30 bezerras mestiças (Holandês/Gir), criadas em instalação tipo bezerreiros tropicais - construções rurais e ambiência, no contexto de bovinocultura, com foco na agropecuária digital.

As imagens foram coletadas em momentos anteriores ao início da execução deste trabalho, por integrantes (professores, colaboradores e discentes) do projeto. Para o presente estudo, foram utilizadas bases de dados compostas por 50, 100 e 228 imagens, todas apresentando o ambiente real de criação de bezerros (Figura 1).

**FIGURA 1: VISÃO GERAL DA BASE DE DADOS**




**FONTE: O AUTOR (2026)**

Por se tratar de imagens de ambientes não controlados, os dados apresentam ampla variação de iluminação, oclusões parciais e movimento de animais, conforme exemplificado na Figura 2.

**FIGURA 2: EXEMPLO DE IMAGENS DA BASE DE DADOS**


**FONTE: O AUTOR (2026)**

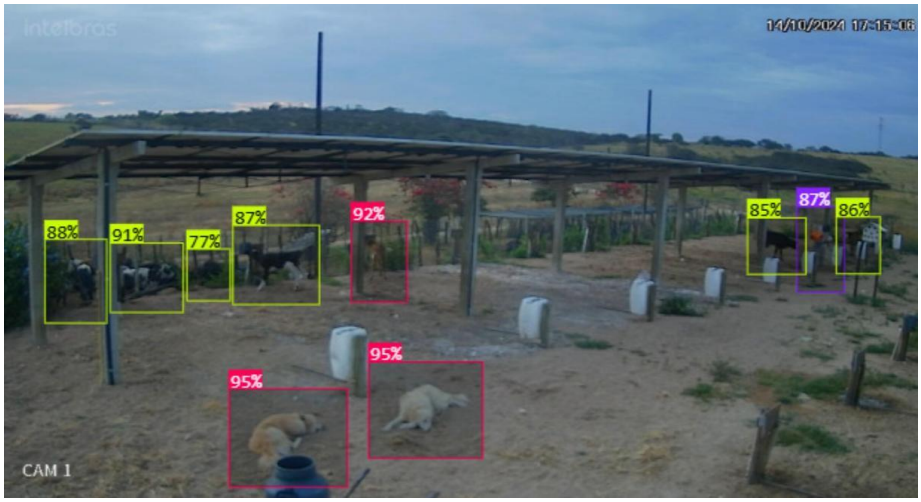
Após o recebimento das bases de imagens, os dados foram inseridos na plataforma Roboflow, utilizando a sua versão de plano gratuito público, onde grande parte das etapas deste estudo foi realizada. Dentre elas, a etapa de organização e anotação das imagens, que consistiu em identificar e rotular integralmente de forma manual os objetos presentes nas imagens por meio de *bounding boxes* (Figura 3). Não foram utilizadas ferramentas de auto anotações baseadas em Inteligência Artificial fornecidas pelo Roboflow, assegurando que o processo de marcação fosse totalmente supervisionado por humanos.

**FIGURA 3. PROCESSO DE ANOTAÇÃO MANUAL DE OBJETOS**


**FONTE: O AUTOR (2026)**

Nesse processo foram consideradas três classes de interesse - bezerro, cachorro e humano. A precisão das anotações influencia diretamente o desempenho dos modelos de detecção de objetos, tendo em vista que marcações incorretas podem degradar métricas como *mAP, Precision e Recall,* essa etapa foi crucial para garantir a qualidade dos dados de treinamento, assegurando a consistência das delimitações (Figura 4).

**FIGURA 4. RESULTADO DA DETECÇÃO DE OBJETOS**




**FONTE: O AUTOR (2026)**

As imagens utilizadas neste estudo foram extraídas diretamente de frames dos vídeos coletados no projeto original. Após a extração, os frames foram misturados de forma homogênea para desconectar a sequência temporal contínua. Concluída a rotulagem, o conjunto de imagens foi particionado de forma aleatória em aproximadamente 80% para treinamento, 10% para validação e 10% para teste. Embora a *Ultralytics* (s.d.) recomende o particionamento 70/15/15 como *baseline* para

modelos YOLO, a própria documentação enfatiza que a disciplina do processo é mais relevante do que os percentuais exatos.

Dessa forma, o particionamento aproximado 80/10/10 adotado neste estudo segue as diretrizes da *Ultralytics* ao assegurar: ausência de vazamento entre conjuntos, estratificação das classes, e preservação de um conjunto de teste intocado, utilizado apenas para a avaliação final do modelo.

Desta forma, foram estruturados diversos cenários de experimentos com variações sistêmicas de transformações geométricas e fotométricas (Quadro 1), permitindo avaliar de forma comparativa o impacto dessas técnicas no desempenho dos modelos em um ambiente caracterizado por restrições de dados.

**QUADRO 1: CENÁRIOS DE EXPERIMENTOS**

| ID DO GRUPO | NOME | CONFIGURAÇÃO DOS PARÂMETROS |
|---|---|---|
| Grupo 0 | Baseline | No augmentation |
| Grupo 1 | Geométrico | Flip H, V |
| Grupo 2 | Cor e Luz | Saturation ±30%; Brightness ±30% |
| Grupo 3 | Combinação | Flip H, Rotation: 15°, Ruído: 2.51% of pixels |
| Grupo 4 | Combinações Extremas | Flip H, Grayscale: 25%, Saturation ±30%, Brightness ±30%, Exposure ±15% e Noise: up to 1.99% of pixels |

**FONTE: O AUTOR (2026)**

A inclusão do espelhamento vertical (*Flip V*) no Grupo 1 foi motivada pela experimentação: o propósito foi avaliar e quantificar o nível de degradação que transformações irreais causariam no modelo em cenário real. Contudo, durante as simulações da base de 50 imagens foram identificados sinais de *overfitting*. Diante disso, optou-se por ajustar os parâmetros deste grupo para essa base em específico, visando preservar a consistência mínima discutida na literatura (como o estudo de Rici et al., 2023).

Além disso, os experimentos foram conduzidos utilizando diferentes perfis de treinamento disponibilizados, à época dos experimentos, pela plataforma Roboflow, originalmente denominados “*Fast*” e “*Accurate*”. Essas configurações representam arquiteturas focadas em atender necessidades distintas, entre velocidade e precisão,

um princípio amplamente documentado na literatura de detecção de objetos (Redmon et al., 2016; Bochkovskiy et al., 2020).

Modelos do tipo *Fast* empregam *datasets* menores, com menor número de parâmetros, priorizando rapidez de treinamento e inferência, enquanto modelos *Accurate* utilizam, geralmente, arquiteturas mais profundas e largas, otimizadas para maximizar métricas como *mAP, Precision* e *Recall*. A adoção de múltiplas configurações é metodologicamente relevante, pois permite avaliar como diferentes capacidades de modelagem respondem às técnicas de *Data Augmentation*, ampliando a robustez e a generalização dos resultados. Assim, cada configuração foi aplicada aos conjuntos de dados, possibilitando uma análise comparativa do impacto das transformações geométricas e fotométricas sobre modelos de diferentes escalas.

Como forma de nomear os experimentos para melhor descrição e discussão dos resultados, foi adotado como o padrão de nomenclatura: “ID do Grupo + Nome do grupo de teste + tamanho da base de dados + configuração do treinamento aplicada”.

A arquitetura YOLOv11 é integrada nativamente à plataforma Roboflow, e foi utilizada como modelo base para detecção de objetos em todos os experimentos. Essa versão mantém o equilíbrio característico da família YOLO entre velocidade de inferência e precisão, apresentando desempenho robusto mesmo em condições desfavoráveis, como alta variação de iluminação, oclusões parciais e movimento de animais.

Como estratégia para garantir a consistência das comparações, foi adotado um padrão na inicialização dos treinamentos no qual foram utilizados o mesmo ponto de partida (*checkpoint*). Esse procedimento utilizou o modelo treinado sem aplicação de técnicas de *Data Augmentation*, denominado “*Baseline*”, partindo-se em seguida para experimentos, *com Data Augmentation*, variando os parâmetros dos experimentos e, consequentemente, as opções de argumentos possíveis que possam estar contidas nas imagens.

Neste sentido, a avaliação dos modelos treinados neste estudo utiliza três métricas principais: *Precision*, *Recall* e mAP (*mean Average Precision*). Essas métricas são calculadas a partir da lógica da matriz de confusão, que classifica as predições do modelo em quatro grupos: verdadeiros positivos (TP), falsos positivos (FP), falsos negativos (FN) e verdadeiros negativos (TN). Na detecção de objetos

geralmente não se considera os verdadeiros negativos, pois representam apenas regiões de fundo da imagem (SOLAWETZ, 2026).

A *Precision* indica quantas das detecções realizadas estavam corretas (TP / (TP + FP)), enquanto o *Recall* indica quantos dos bovinos presentes nas imagens foram de fato detectados (TP / (TP + FN)). Segundo Rezende (2018), essas métricas expressam tanto a capacidade de acerto, quanto a cobertura do modelo, sendo complementares na interpretação do desempenho.

O *mAP* sintetiza essas duas dimensões em um único valor. No presente estudo, foi adotada especificamente a métrica mAP@0.5, que define que uma detecção é considerada um Verdadeiro Positivo (TP) se a sobreposição espacial entre a caixa delimitadora predita (pelo modelo) e a caixa real (anotada manualmente) for de no mínimo 50%. Ele é calculado a partir da curva Precisão–Revocação, que mostra como *Precision* e *Recall* se comportam conforme o limiar de confiança que o modelo varia. A área sob essa curva define a *Average Precision (AP)* de cada classe, e o *mAP* é a média das *APs* de todas as classes. Conforme Solawetz (2025), quanto maior essa área, mais equilibrado e robusto é o modelo.

No presente estudo, essas métricas foram a base quantitativa para comparar os modelos YOLO treinados com e sem *Data Augmentation*, via Roboflow, permitindo identificar padrões de erro — como falsos negativos em bezerros parcialmente ocultos ou falsos positivos provocados por sombras e estruturas do bezerreiro — em cenários típicos de ambientes tropicais quentes e úmidos, onde as condições de iluminação e o comportamento dos animais dificultam consideravelmente a detecção (PADILLA et al., 2020).

Cada experimento foi executado uma única vez e o processo de avaliação dos modelos foi baseado em métricas comumente utilizadas em detecção de objetos, que foram calculadas e fornecidas nativamente pela plataforma Roboflow ao final de cada treinamento realizado.

Os resultados obtidos em cada experimento foram registrados em uma planilha, disponível para acesso integral no Apêndice A, com o objetivo de organizar e facilitar a análise comparativa dos modelos. Para identificar variações de desempenho causadas pela aplicação das técnicas de *Data Augmentation*, foram verificados os valores das métricas em cada experimento, comparando os resultados obtidos em relação ao modelo Baseline.

A partir dessa comparação, foram calculadas as variações (positivas ou negativas) em cada métrica, permitindo identificar alterações de desempenho associadas às diferentes técnicas de transformação aplicadas. Dentre os parâmetros avaliados, se destacam: espelhamento horizontal e vertical (*flip H, V*), saturação (*saturation*), brilho (*brightness*), rotação (*rotation*), escala de cinza (*grayscale*), exposição (*exposure*), e ruído (*noise*). Esse processo foi aplicado em todas as bases analisadas.

Os resultados foram organizados por grupos de experimentos e por tamanho de base, possibilitando a comparação do comportamento das métricas diante das diferentes configurações de *Data Augmentation*. Essa organização permitiu identificar os padrões de desempenho além de verificar quais técnicas trouxeram efeitos mais significativos, apresentando resultados mais consistentes.

A partir dessa comparação, foi possível avaliar a adequação das diferentes técnicas de *Data Augmentation* ao contexto deste trabalho, verificando o desempenho dos modelos gerados a partir do aumento gradativo dos parâmetros de transformação. Essas modificações possibilitaram a criação das diversas versões de *dataset* na plataforma Roboflow, permitindo a análise comparativa da variação das métricas de avaliação em cada cenário experimental.

## 6. RESULTADOS E DISCUSSÃO

Esta seção apresenta os resultados e discussões acerca dos experimentos realizados para avaliar o desempenho do modelo diante de cenários de testes e técnicas de transformações diferentes.

### 6.1. GRUPO 0 – BASELINE

O Grupo 0 representa o ponto de referência dos experimentos realizados neste estudo, sendo composto por modelos treinados sem aplicação de técnicas de *Data Augmentation*. Os resultados foram usados para definir uma linha base (*baseline*), utilizada para avaliação de todas as variações subsequentes. Desta forma, os

experimentos foram conduzidos nas bases disponíveis, compostas por 50, 100 e 228 imagens, com treinamentos configurados em *Fast* e *Accurate*. Os resultados dos experimentos desse grupo estão apresentados na Tabela 1.

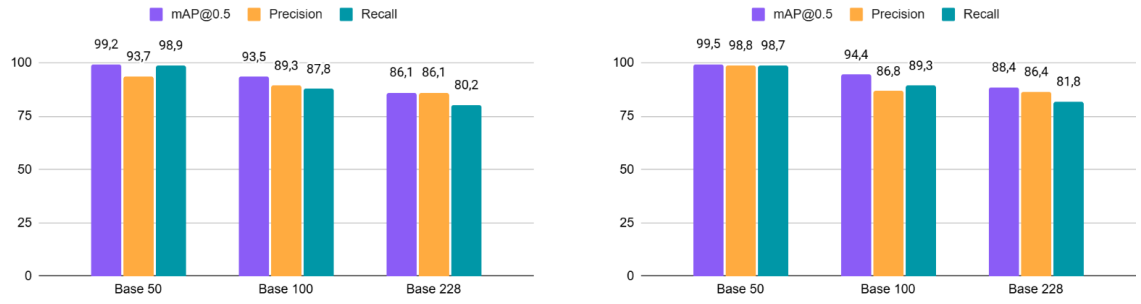
**TABELA 1. RESULTADO DOS EXPERIMENTOS DO GRUPO 0 - BASELINE**

| Base | Configuração | mAP@0.5 | Precision | Recall |
|---|---|---|---|---|
| 50 | Fast | 99,2 | 93,7 | 98,9 |
| 50 | Accurate | 99,5 | 98,8 | 98,7 |
| 100 | Fast | 93,5 | 89,3 | 87,8 |
| 100 | Accurate | 94,4 | 86,8 | 89,3 |
| 228 | Fast | 86,1 | 86,1 | 80,2 |
| 228 | Accurate | 88,4 | 86,4 | 81,8 |

**FONTE: O AUTOR (2026)**

Os resultados do Grupo 0 revelam um padrão que exige atenção ao se discutir o impacto do Data Augmentation: o número de imagens disponíveis interfere de forma não linear nas métricas obtidas, e devem ser interpretadas com cautela (Figura 5).

**FIGURA 5. RESULTADOS DO GRUPO 0 - BASELINE**




**FONTE: O AUTOR (2026).**

Inicialmente, a base composta por 50 imagens apresentou os maiores valores absolutos de *mAP* (99,2% e 99,5%), *Precision* (93,7% e 98,8%) e *Recall* (98,9% e 98,7%) para as configurações *Fast* e *Accurate*, respectivamente. Entretanto, é importante considerar que o conjunto de imagens composto por esta base de dados é de apenas 50 imagens, divididas, respectivamente, em 40 (80%) referente a treinamento, 5 (10%) referente à validação e 5 (10%) referente a teste, o que compromete a representatividade estatística dos valores obtidos.

Nesse cenário, há a probabilidade de o modelo ter alcançado um desempenho alto devido ao overfitting, ou seja: a limitação do conjunto de testes pode ter ocultado

algumas limitações do modelo. Esse comportamento é razoável e está de acordo com os trabalhos relacionados, que alerta sobre o risco do modelo ser superestimado em avaliações com poucas amostras.

Já a base de dados correspondente a 100 imagens alcançou resultados intermediários, com *mAP* entre 93,5% e 94,4%, *Precision* entre 89,3% e 86,8% e *Recall* entre 87,8% e 89,3%. Apesar do conjunto de dados ser um pouco maior que o modelo anterior, este volume de imagens ainda é considerado baixo para se chegar a conclusões definitivas a respeito do modelo em comento. Ainda assim, é possível observar que a configuração *Accurate* gerou *mAP* levemente superior ao *mAP Fast* (94,4% vs. 93,5%), agregando, por sua vez, mais ganho no *Recall* (89,3% vs. 87,8%). Isso sugere que o tempo adicional de treinamento contribuiu para um desempenho melhor na detecção dos objetos.

Em relação à base composta por 228 imagens, percebe-se que as condições de treinamento foram mais equilibradas. A base foi distribuída em 182 imagens (80%) para treinamento, 23 (10%) para validação e 23 (10%) para teste, este modelo apresentou menores valores absolutos do grupo: *mAP* de 86.1% e 88,4%, Precision de 86.1% e 86.4% e Recall de 80.2% e 81.8%, respectivamente para *Fast* e *Accurate*.

Essa redução é relativa e demonstra limitações naturais do conjunto de dados decorrente da baixa diversidade visual disponível para treinamento. Embora o modelo tenha conseguido detectar parte significativa dos bovinos, o cenário de 228 imagens trouxe desafios reais do ambiente de bezerreiros tropicais, tais como: diferentes condições de iluminação, movimentação e oclusão.

Em síntese, os resultados do Grupo 0 apresentam evidências concretas de que o tamanho e qualidade do conjunto de dados exercem muita influência sobre o desempenho dos modelos, mesmo não havendo estratégias de aumento de dados adotadas no treinamento dos experimentos a princípio.

## 6.2. GRUPO 1 – GEOMÉTRICO

O Grupo 1 testa a aplicação de técnicas geométricas, como o *flip* (espelhamento) horizontal e vertical. Neste grupo, a base de 50 imagens utilizou apenas o espelhamento horizontal (Flip H), enquanto as bases de 100 e 228 imagens utilizaram o espelhamento horizontal e vertical (Flip H + V). Apesar dessa variação

limitar a comparação absoluta entre as bases, os resultados contribuem com a discussão de forma individual. Esses resultados estão apresentados na Tabela 2.

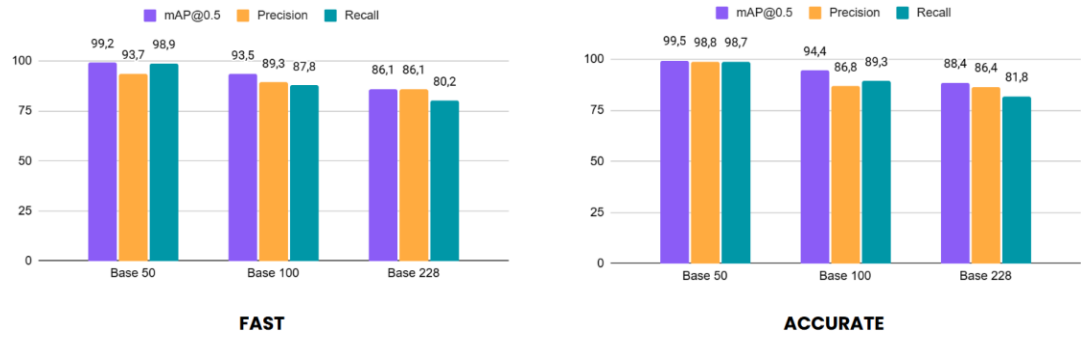
**TABELA 2. RESULTADO DOS EXPERIMENTOS DO GRUPO 1 – GEOMÉTRICO**

| Base | Configuração | mAP@0.5 | Precision | Recall |
|---|---|---|---|---|
| 50 | Fast | 99,2 | 93,7 | 98,9 |
| 50 | Accurate | 99,5 | 98,8 | 98,7 |
| 100 | Fast | 93,5 | 89,3 | 87,8 |
| 100 | Accurate | 94,4 | 86,8 | 89,3 |
| 228 | Fast | 86,1 | 86,1 | 80,2 |
| 228 | Accurate | 88,4 | 86,4 | 81,8 |

**FONTE: O AUTOR (2026)**

O comportamento heterogêneo revelado pelos resultados do Grupo 1, de certa forma, não se deve apenas à técnica de *Data Augmentation* aplicada, é necessário, também, considerar a variação do tipo de *Flip* nos experimentos, que aumentou a diversidade espacial do dataset, permitindo que o modelo YOLOv11 aprendesse diferente posições corporais dos bovinos (Figura 6).

**FIGURA 6. RESULTADOS DO GRUPO 1 – GEOMÉTRICO**




**FONTE: O AUTOR (2026)**

Na base de dados de 50 imagens, o espelhamento horizontal gerou variações positivas nas métricas *mAP* que apresentou ganho de +0,2 pontos percentuais (p.p.) na configuração *Fast* e permaneceu estável na *Accurate*. Em relação à Precision, observou-se aumento de +3,3 p.p. na configuração *Fast* e leve redução de -0,2 p.p. na *Accurate*. Já o *Recall* apresentou pequenas quedas em ambas as configurações, com variações de -0,8 p.p. no Fast e -0,6 p.p. no *Accurate*.

Cabe destacar que, conforme discutido no Grupo 0, os resultados da base 50 tendem a apresentar resultados mais precisos com valores percentuais acima de 99% para *mAP,* devido ao baixo número de imagens no conjunto de testes aplicado ao experimento. Ainda assim, é possível observar um ganho nas métricas *Precision e Recall,* indicando a redução de falsos positivos. Esses valores aumentam a sensibilidade das métricas e mostra que o *Data Augmentation* teve pouca influência no experimento.

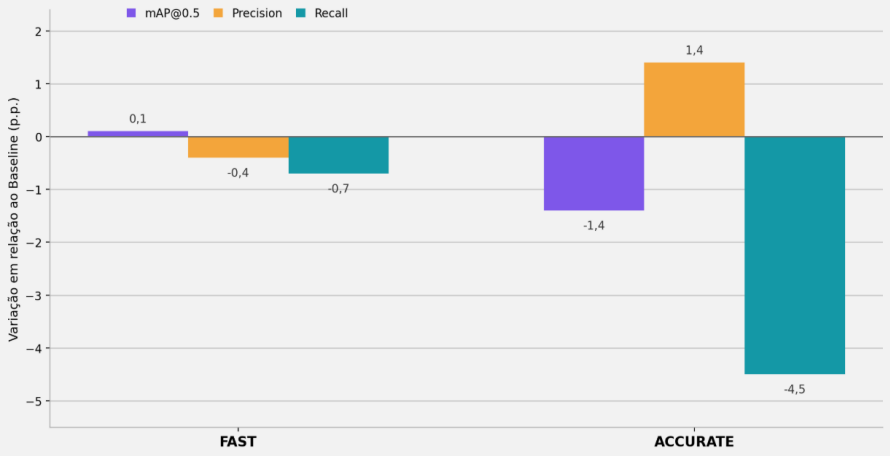
Na base de dados de 100 imagens, o espelhamento “Horizontal + Vertical” produziu variações discretas: na configuração *Fast* houve ganhos de +1,5 p.p. em *Precision* e *Recall*, apesar da leve queda de -0,3 p.p. no *mAP*. Já na configuração Accurate, o *Precision* permaneceu estável (+0,1 p.p.), mas houve pequenas reduções de -0,7 p.p. tanto no *mAP* quanto no *Recall*. Assim, é possível concluir que os resultados da diversificação geométrica, via espelhamento, apresentaram efeitos limitados sobre a capacidade de generalização do modelo em si.

Em relação à base de dados de 228 imagens, o espelhamento “Horizontal + Vertical” gerou resultados negativos em relação ao Baseline, especialmente na configuração *Accurate*, que apresentou queda de -1,4 p.p. no *mAP* e -4,5 p.p. no *Recall*. Já na configuração Fast, o Precision e Recall apresentaram pequenas quedas de -0,4 p.p. e -0,7 p.p., respectivamente, enquanto o mAP se manteve quase estável.

Uma hipótese razoável para esse resultado é que o espelhamento vertical trouxe amostras pouco representativas no ambiente real do bezerreiro, uma vez que os animais não costumam aparecer invertidos verticalmente. Ainda assim, esse resultado é consistente e confirma o alerta de Hanel (2021) e Moreira (2022) sobre o desempenho do modelo ser comprometido por transformações que não se aplicam ao contexto real.

A Figura 7 apresenta as variações das métricas em relação ao Grupo 0 para cada uma das bases e configurações.

**FIGURA 7. VARIAÇÕES DAS MÉTRICAS DO GRUPO 1 – GEOMÉTRICO**




**FONTE: O AUTOR (2026).**

Em síntese, os resultados do Grupo 1 conduzem a duas conclusões principais: I - ficou evidente o aumento moderado na *Precision* em vários modelos, indicando redução de falsos positivos; II - em diversos experimentos ocorreu redução do Recall, sugerindo que o modelo se tornou mais seletivo nas detecções realizadas, reduzindo erros de classificação, mas deixando de detectar alguns bovinos presentes nas imagens. Diante desse *trade-off*, é possível concluir que a capacidade das transformações geométricas simples é relativamente limitada, especialmente quando aplicadas de forma isolada.

## 6.3. GRUPO 2 – FOTOMÉTRICO

O Grupo 2 teve como foco testar transformações fotométricas manipulando saturação e brilho, simulando condições de iluminação. Esse conjunto de experimentos possui alta relevância para o contexto de bezerreiros tropicais, uma vez que variações de iluminação natural são um fator constante na realidade. Como parâmetros foram aplicados ±30% de saturação e brilho nas bases de 50, 100 e 228 imagens, os resultados estão organizados na Tabela 3.

**TABELA 3. RESULTADO DOS EXPERIMENTOS DO GRUPO 2 – FOTOMÉTRICO**

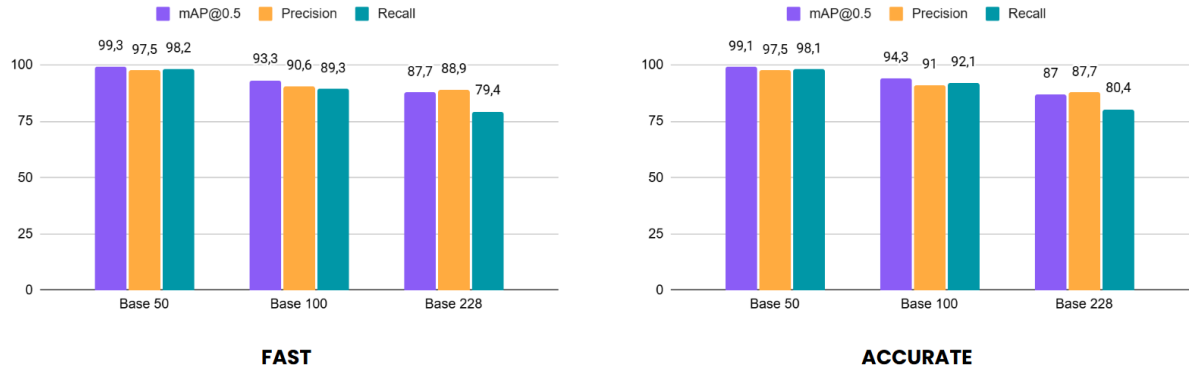
| Base | Configuração | mAP@0.5 | Precision | Recall |
|---|---|---|---|---|
| 50 | Fast | 99,3 | 97,5 | 98,2 |
| 50 | Accurate | 99,1 | 97,5 | 98,1 |

| | | | | |
|---|---|---|---|---|
| 100 | Fast | 93,3 | 90,6 | 89,3 |
| 100 | Accurate | 94,3 | 91,0 | 92,1 |
| 228 | Fast | 87,7 | 88,9 | 79,4 |
| 228 | Accurate | 87,0 | 87,7 | 80,4 |

**FONTE: O AUTOR (2026).**

Os resultados do Grupo 2 apresentam um padrão diferente do Grupo 1: as variações em relação ao *Baseline* são mais próximas de zero ou levemente positivas. Isso demonstra que ajustes fotométricos têm mais influência no desempenho dos modelos testados neste estudo (Figura 8).

**FIGURA 8. RESULTADOS DO GRUPO 2 – FOTOMÉTRICO**




**FONTE: O AUTOR (2026)**

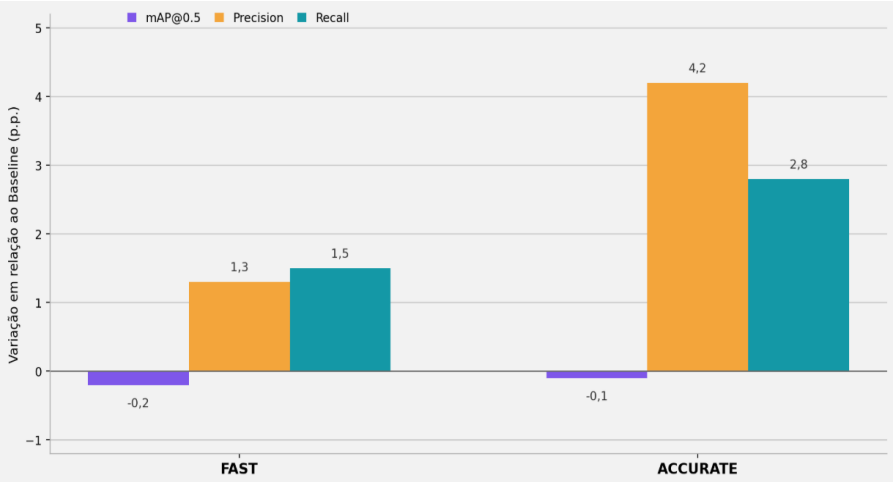
Na base de 50 imagens, mais uma vez as variações foram baixas devido ao volume do conjunto de testes. O *Precision* apresentou um ganho de 3.8 p.p. *(Fast)*, quando comparado com o *Baseline*, vindo acompanhado de uma leve redução no *Recall* (-0,7 p.p.). Esses resultados reforçam que o tamanho reduzido da base oferece pouco espaço para diferenciar estratégias de aumento de dados.

Na base de dados de 100 imagens, o *mAP* se manteve estável e apresentou pequenos ganhos em *Precision* (+ 2,5 p.p. e +4,2 p.p.), no *Fast* e *Accurate*, respectivamente. O *Recall* também teve ganhos modestos trazendo + 0,8 p.p. no *Fast* e + 2,7 p.p. no *Accurate*. A partir da configuração *Accurate* o modelo foi capaz de alcançar um *Recall* de 92,1%, isso representa um ganho de + 2,7 p.p. em relação ao *Baseline* dessa configuração. Esse resultado indica que, em volumes de dados moderados, a aplicação de transformações fotométricas contribui positivamente na cobertura e detecções sem gerar prejuízos na precisão do modelo de forma geral.

Na base de dados de 228 imagens, a configuração *Fast* gerou um ganho de +1,6 p.p. no *mAP*, somado à um aumento de +2,8 p.p. na *Precision*, enquanto o Recall teve uma ligeira redução de -0,8 p.p. Já na configuração *Accurate*, tanto o mAP quanto o *Recall* apresentaram queda de -1,4 p.p., enquanto o *Precision* apresentou ganho de 1,3 p.p., sugerindo que mesmo o aumento de tempo de treinamento não foi suficiente para compensar as transformações fotométricas dessa configuração.

A Figura 9 apresenta as variações das métricas em relação ao Grupo 2 para cada uma das bases e configurações.

**FIGURA 9. VARIAÇÕES DAS MÉTRICAS DO GRUPO 2 – FOTOMÉTRICO**




**FONTE: O AUTOR (2026).**

De modo geral, os resultados do Grupo 2 indicam que as transformações de cor e luz mostram melhoria tanto na qualidade das detecções, quanto na capacidade geral de localização dos objetos, mesmo alguns modelos apresentando pequenas oscilações negativas no Recall, indicando que determinadas alterações fotométricas podem distorcer os padrões visuais reais dos bovinos.

## 6.4. GRUPO 3 – COMBINAÇÃO

O Grupo 3 testa a aplicação de técnicas de *Data Augmentation* de forma combinada, unindo transformações geométricas (espelhamento horizontal e rotação) com adição de ruído, com o objetivo de representar estratégias mais robustas em relação aos grupos anteriores. Deste modo, é possível simular simultaneamente

fatores como orientação, ângulo de captura e imperfeições visuais comuns em ambientes de campo.

Os parâmetros aplicados foram *Flip* H, *Rotation* ±15° e *Ruído* de aproximadamente 1,5% de pixels na base de dados de 50 imagens e 2,5% nas bases de 100 e 228 imagens. Os resultados estão apresentados na Tabela 4.

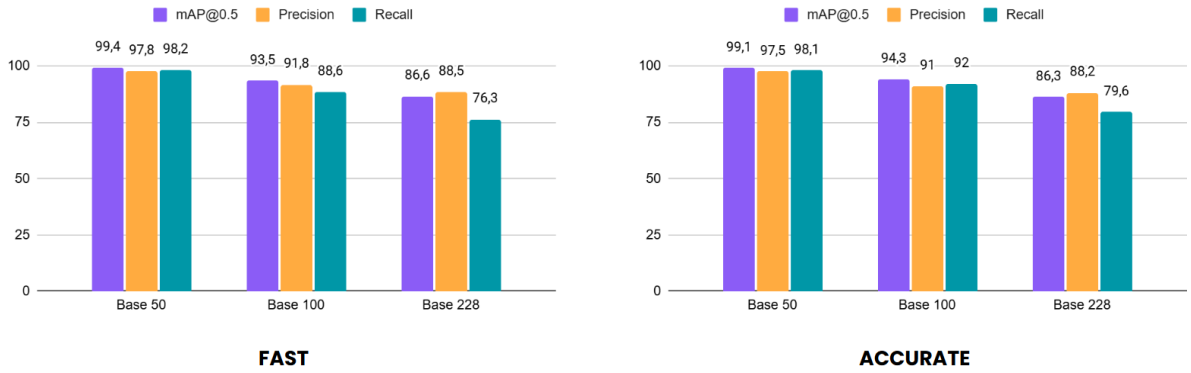
**TABELA 4. RESULTADO DOS EXPERIMENTOS DO GRUPO 3 – COMBINAÇÃO**

| Base | Configuração | mAP@0.5 | Precision | Recall |
|---|---|---|---|---|
| 50 | Fast | 99,4 | 97,8 | 98,2 |
| 50 | Accurate | 99,1 | 97,5 | 98,1 |
| 100 | Fast | 93,5 | 91,8 | 88,6 |
| 100 | Accurate | 94,3 | 91,0 | 92,0 |
| 228 | Fast | 86,6 | 88,5 | 76,3 |
| 228 | Accurate | 86,3 | 88,2 | 79,6 |

**FONTE: O AUTOR (2026).**

Os resultados do Grupo 3 apresentaram padrões mistos, com variações diversas que dependem das bases e configurações aplicadas, um comportamento que reforça o impacto do volume de dados em relação à estratégia de transformação aplicada (Figura 10).

**FIGURA 10. RESULTADOS DO GRUPO 3 – COMBINAÇÕES**




**FONTE: O AUTOR (2026)**

Na base de dados de 50 imagens, a combinação de transformações aplicadas gerou ganho de +0,2 p.p. no *mAP*, acompanhado de aumento expressivo de +4,1 p.p. na *Precision* para a configuração *Fast*. Já na configuração Accurate, o *mAP* apresentou pequena queda de -0,4 p.p., e redução de -1,3 p.p. na *Precision*. Nas duas

configurações, o *Recall* apresentou queda, -0,7 p.p. (*Fast*) e -0,6 p.p. (*Accurate*). Esses resultados evidenciam as limitações metodológicas na base de 50 imagens, que tendem a gerar inconsistências diante de configurações diversificadas.

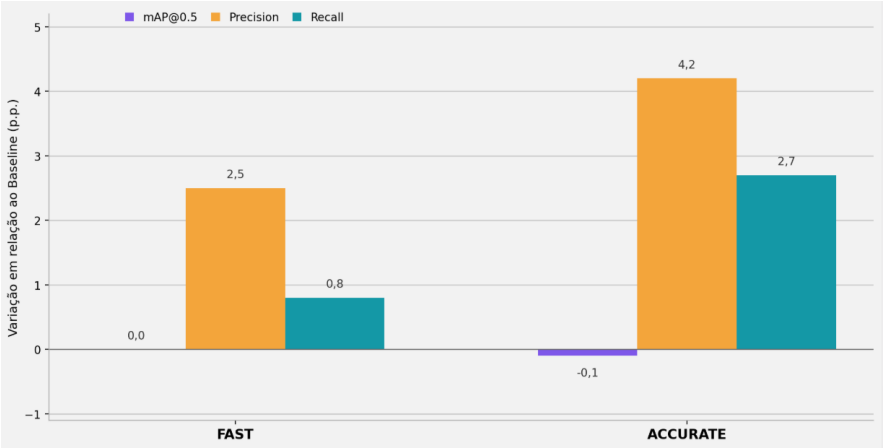
Na base de dados de 100 imagens, os resultados se mantiveram estáveis: na configuração *Fast*, as métricas *Precision* e *Recall* apresentaram variações positivas (+2,5 p.p. e +0,8 p.p., respectivamente), enquanto o *mAP* permaneceu igual. Na configuração *Accurate*, o *mAP* teve uma leve queda de -0,1 p.p, mas manteve os ganhos positivos de 4,2 p.p. no *Precision* e 2,7 p.p. no *Recall*. Tais resultados apontam que, ao aplicar combinação de transformações em volumes moderados de dados, o desempenho do modelo não é comprometido e ainda apresenta ganhos pontuais.

Em relação à base de dados de 228 imagens, foi identificado um comportamento mais complexo. Na configuração *Fast*, houve um aumento sutil de +0,5p.p. no *mAP*, além de um aumento de +2,4p.p. no *Precision*; entretanto o *Recall* trouxe uma perda expressiva de -3,9p.p. Por definição, esse *trade-off* sugere que o modelo se tornou mais conservador nas detecções, reduzindo os falsos positivos, mas perdendo cobertura.

A configuração *Accurate* também chamou atenção: o *mAP* sofreu redução de -2,1 p.p., o *Recall* também trouxe perdas (-2,2 p.p.), e apenas o *Precision* trouxe ganhos, com +1,8 p.p. Vale destacar que, mesmo com a degradação, as reduções observadas nas métricas do Grupo 3 mantêm-se em níveis mais moderados do que os do Grupo 1, indicando que a adição de ruído e rotação foi menos prejudicial ao aprendizado do que a aplicação anterior do *flip* vertical isolado.

A Figura 11 apresenta as variações das métricas em relação ao Grupo 3 para cada uma das bases e configurações.

**FIGURA 11. VARIAÇÕES DAS MÉTRICAS DO GRUPO 3 – COMBINAÇÕES**




**FONTE: O AUTOR (2026).**

## 6.5. GRUPO 4 – COMBINAÇÕES EXTREMAS

O Grupo 4 se refere ao conjunto de experimentos de abordagem mais radical, aplicados apenas na base de 228 imagens, com a finalidade de investigar o impacto de transformações fotométricas e geométricas mais incisivas. Os parâmetros aplicados incluíram: *Flip* Horizontal, *Grayscale* (25%), *Saturation* ±30%, *Brightness* ±30%, *Exposure* ±15% e *Noise* (até 1,99% de pixels), com algumas variações de *Rotation* ±15°. Os resultados estão apresentados na Tabela 5.

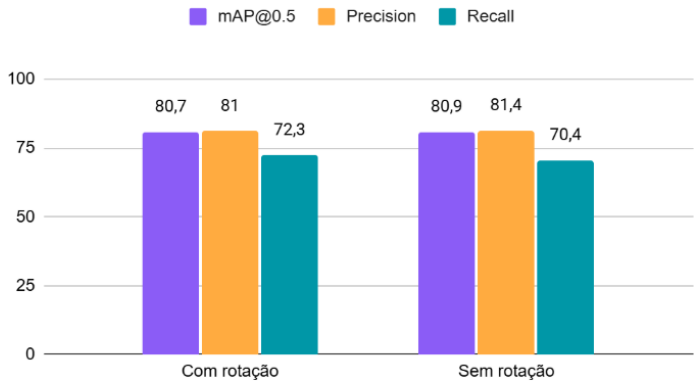
**TABELA 5. RESULTADO DOS EXPERIMENTOS DO GRUPO 4 – COMBINAÇÕES EXTREMAS**

| Base | Configuração | mAP@0.5 | Precision | Recall |
|---|---|---|---|---|
| 228 | Fast - sem rotação | 80,9 | 81,4 | 70,4 |
| 228 | Fast - com rotação | 80,7 | 81,0 | 72,3 |

**FONTE: O AUTOR (2026).**

Os dados revelam um padrão consistente: a aplicação de transformações extremas e combinadas causam degradações no desempenho em relação ao *Baseline* (Figura 12).

**FIGURA 12. RESULTADOS DO GRUPO 4 (FAST) – COMBINAÇÕES EXTREMAS**

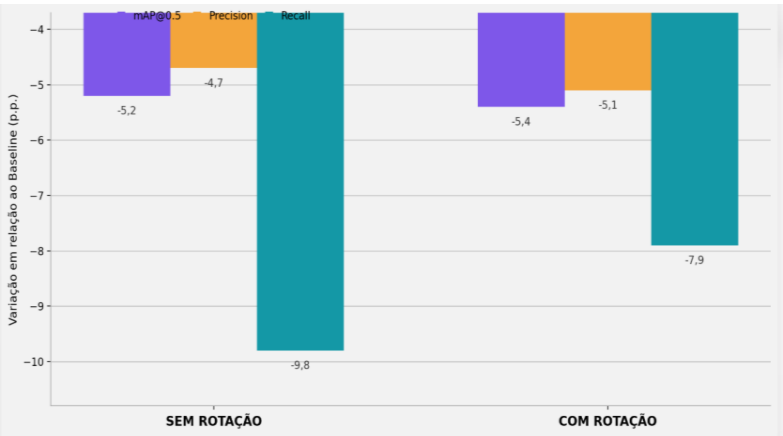



**FONTE: O AUTOR (2026)**

O experimento sem rotação apresentou uma queda de -5,2 p.p. no mAP, acompanhada de redução de -4,7 p.p. na *Precision* e -9,8 p.p. no *Recall*. Esses números representam uma das maiores degradações observadas em qualquer grupo do estudo até o presente momento. O experimento com adição de rotação segue no mesmo sentido, mantendo a queda nas métricas: -5,4 p.p. no *mAP*, -5,1 p.p. na *Precision* e -7,9 p.p. no *Recall*.

Os resultados deste grupo (Figura 13) reforçam que transformações extremas e combinadas em excesso prejudicam o desempenho de modelos de detecção na base 228, mesmo que o objetivo inicial seja aumentar a diversidade de dados. A partir do Grupo 4 foi possível gerar evidências sobre o ponto de *overfitting* crítico causado pela aplicação extrema de *Data Augmentation*, demonstrando que em contextos de bases limitadas, o simples aumento quantitativo de dados não compensa a degradação qualitativa do modelo.

**FIGURA 13. VARIAÇÕES DAS MÉTRICAS DO GRUPO 4 – COMBINAÇÕES EXTREMAS**



**FONTE: O AUTOR (2026).**

Mais uma vez, os alertas de Hanel (2021) e Moreira (2022) sobre os riscos de *Data Augmentation* descontrolado se mostram consistentes. Uma interpretação possível para esses resultados é que, a combinação excessiva de transformações gera amostras muito distantes da configuração real dos dados de bezerreiros tropicais, impactando na capacidade do modelo de aprender características realmente relevantes para o domínio. Ou seja, o modelo é forçado a aprender a partir de dados que não refletem a realidade, o que prejudica a generalização e aumenta o *overfitting* a padrões falsos introduzidos artificialmente pelas transformações.

## 7. CONCLUSÃO E TRABALHOS FUTUROS

O desenvolvimento deste estudo possibilitou a análise do impacto de diferentes técnicas de *Data Augmentation* na detecção de objetos em ambientes de bezerreiros tropicais, por meio do uso de diferentes bases de dados e configurações de treinamento (*Fast* e *Accurate*). Os resultados confirmam que o tamanho do conjunto de dados influencia de forma crítica e não linear na interpretação das métricas de desempenho.

A base de 50 imagens apresentou valores de *mAP*, *Precision* e *Recall* altos em relação a quase todos os experimentos. Entretanto, esse comportamento não implica necessariamente que o modelo teve um desempenho superior, pois se trata de um sobreajuste (*overfitting*) e limitação do conjunto de testes que tende a superestimar resultados reais.

Por outro lado, a base de 228 se mostrou mais realista para avaliar a capacidade de generalização dos modelos, justamente por apresentar maior complexidade e semelhanças com cenários reais (variações de iluminação, oclusões parciais e movimentação de animais), enfatizando que *Data Augmentation* leves e moderados apresentam desempenho significativamente superior às transformações excessivamente agressivas.

A análise comparativa entre os grupos de experimento revela, ainda, que nem todas as transformações trazem benefícios para o modelo. As transformações Geométricas (Grupo 1) causaram efeitos irreais e limitados, especialmente com a

técnica de Flip Horizontal + Vertical, prejudicando as amostras geradas e comprometendo o aprendizado do modelo, nas bases de 100 e 228 imagens.

Já as transformações de Cor e Luz (Grupo 2), foram as que mais agregaram impacto positivo em termos de consistência, sobretudo na base de 100 imagens. Os resultados indicaram um ganho significativo de *Precision* e *Recall*, quando o modelo investiu mais tempo no treinamento (configuração *Accurate*). Dessa forma, conclui-se que a simulação de variações de brilho e saturação possui alta aderência prática à realidade de bezerreiros tropicais, expostos a oscilações constantes de luz natural, oclusões e movimentações de animais.

Ao utilizar técnicas extremas combinadas (Grupos 3 e 4), houve um cenário de sobreajuste crítico do método de transformações excessivas. As características visuais foram altamente distorcidas, gerando a maior degradação do desempenho observada em todo estudo e reduzindo a capacidade de detecção de objetos.

Desta forma, o estudo demonstra que, em Visão Computacional, o aumento meramente quantitativo de dados a partir de técnicas de *Data Augmentation* não substitui a qualidade e a coerência dessas transformações, ou seja, não cria novas informações semânticas reais. A aplicação de *Data Augmentation* descontrolada força o modelo a aprender padrões falsos que não refletem a realidade e comprometem a capacidade de generalização do modelo.

Por outro lado, os resultados obtidos neste estudo demonstram que o *Data Augmentation* pode contribuir significativamente para melhoria da capacidade de generalização de modelos YOLO em cenários com limitação de dados, desde que as modificações geométricas e alterações fotométricas sejam leves. Portanto, conclui-se que a eficácia do *Data Augmentation* depende diretamente da coerência das transformações sintéticas com as características reais do domínio investigado, sendo capaz de produzir maior aderência ao cenário real, quando empregado de forma controlada.

Como limitação do estudo, destaca-se a velocidade evolutiva acelerada da plataforma Roboflow, um fator que impactou a repetição dos experimentos em momentos distintos do estudo, uma vez que para alguns experimentos as configurações de testes não estão mais disponíveis na plataforma. Além disso, reconhece-se como limitação a ausência de validação cruzada, ideal para mitigar a instabilidade estatística e o risco de *overfitting* em bases de dados menores, tais como a de 50 imagens.

Para trabalhos futuros, é recomendada a expansão do volume de imagens reais como estratégia de mitigar o efeito de *overfitting* visto em bases menores. Além disso, propõe-se a execução de experimentos que combinem as técnicas e parâmetros que apresentaram os melhores resultados neste estudo, priorizando técnicas fotométricas no lugar de rotações ou espelhamentos espaciais agressivos, avaliando o ganho dessa união.

## USO DE TECNOLOGIAS DE IA GENERATIVA

Na condução deste estudo foi utilizada a ferramenta “Chat GPT”, da OpenAI, unicamente para busca e formatação e referências bibliográficas conforme as normas da ABNT, atuando como ferramenta de curadoria. Toda formulação teórica, condução e análise dos experimentos e conclusões foram escritas, revisadas e editadas pelo autor.

**REFERÊNCIAS**

ANGARITA-ZAPATA, Juan S. et al. A taxonomy of food supply chain problems from a computational intelligence perspective. Sensors, v. 21, n. 20, p. 6910, 2021.

BOCHKOVSKIY, Alexey; WANG, Chien-Yao; LIAO, Hong-Yuan Mark. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934, 2020.

DA SILVA, Lázaro Eduardo et al. Detecção de Objetos em Tempo Real com a Plataforma ROBOFLOW. In: 19ª Semana de Ciência e Tecnologia do CEFET-MG-2023.

DE OLIVEIRA, Bruno Vicente Nunes; DE MELO, Filipe Torres. Fundamentos da visão computacional: arcabouço teórico do reconhecimento artificial de imagens e vídeos. *Humanidades & Inovação, v. 10, n. 17, p. 312-327, 2023*. Disponível em: https://revista.unitins.br/index.php/humanidadeseinovacao/article/view/9078. Acesso em: 30 maio 2026.

DISTRITO. Data augmentation: o que é e como usar essa técnica? Distrito, 2023. Disponível em: https://www.distrito.me/blog/data-augmentation-o-que-e-e-como-usar-essa-tecnica. Acesso em: 30 maio 2026.

FUNDAÇÃO CERTI. Visão computacional: o que é, sua importância e aplicações práticas. Florianópolis: Fundação CERTI, 2026. Disponível em: https://certi.org.br/blog/visao-computacional/. Acesso em: 30 maio 2026.

GOMES, Maysa Feitosa. Aplicação de métodos de visão computacional na nutrição animal e na bovinocultura. 2025.

HANEL, Mateus Schoenardie. Análise de métodos de data augmentation para melhoria do desempenho de uma rede neural de detecção de defeitos em superfícies metálicas. 2021. Trabalho de Conclusão de Curso (Engenharia de Controle e Automação) — UFRGS – Universidade Federal do Rio Grande do Sul, Escola de

Engenharia, Porto Alegre, 2021. Disponível em: https://lume.ufrgs.br/handle/10183/235162. Acesso 10 maio 2026

JOSEPH, N.; JOHNSON, B. Building Computer Vision Applications with Roboflow. Roboflow Research Publications, 2023.

MARIM, Yan Victor Ribeiro. Detecção de Objetos: Estudo e Aplicação da Arquitetura R-CNN. Projeto de Graduação — Universidade Federal do Espírito Santo, Vitória, 2019. Disponível em: https://engenhariaeletrica.ufes.br/sites/engenhariaeletrica.ufes.br/files/field/anexo/projeto_de_graduacao_ii_-_yan_victor_ribeiro_marim.pdf. Acesso em: 20 maio 2026.

MOREIRA, Andressa Gomes. Análise do desempenho de Redes Adversárias Generativas como um método de aumento de dados de imagens da mucosa ocular de ovinos. Trabalho de Conclusão de Curso (Graduação em Engenharia da Computação) — Universidade Federal do Ceará, Sobral, 2022. Disponível em: https://repositorio.ufc.br/bitstream/riufc/70525/1/2022_tcc_agmoreira.pdf. Acesso em 25 maio 2026.

MOTA, Guilherme. Dominando o overfitting: por que suas redes neurais e SVMs precisam ser treinadas com cuidado? *DIO — Digital Innovation One*, *26 abr 2025*. Disponível em: https://www.dio.me/articles/dominando-o-overfitting-por-que-suas-redes-neurais-e-svms-precisam-ser-treinadas-com-cuidado-54e7103003b6. Acesso em 30 maio 2026.

NOLÊTO, Raquel M. A.; NOLÊTO, Carleandro; SANTOS, Natanael P. S.; MADEIRA, Aline M. A.. Inovações no reconhecimento e detecção de animais: ama análise da literatura com ênfase em redes neurais e aprendizado de máquina. In: ENCONTRO UNIFICADO DE COMPUTAÇÃO DO PIAUÍ (ENUCOMPI), 16. , 2023, Piripiri/PI. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2023 . p. 33-40. DOI: https://doi.org/10.5753/enucompi.2023.26614.

PADILLA, Rafael et al. A Comparative Analysis of Object Detection Metrics. 2020.

REDMON, Joseph et al. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 779-788.

REZENDE, Solange O. Sistemas Inteligentes: Fundamentos e Aplicações. Manole, 2003.

RICI, Pedro; SANTOS, Samara Oliveira Silva; OTTONI, André Luiz Carvalho. Análise da seleção de hiperparâmetros de Data Augmentation na detecção de Covid-19 em imagens de raio-x com Deep Learning. In: **CONGRESSO BRASILEIRO DE INTELIGÊNCIA COMPUTACIONAL, 15., 2021, Joinville, SC**. *Anais [...]* Joinville, SC: SBIC, 2021. p. 1-7. DOI: 10.21528/CBIC2021-22.

ROBOFLOW. Data Annotation for High-Performing Computer Vision Models. Roboflow Blog, 22 maio 2025. Disponível em: https://blog.roboflow.com/data-annotation/. Acesso em: 5 jun. 2026.

SAVEKAR, Avinash; KUMAR, Vineet. AI in Computer Vision Market: global opportunity analysis and industry forecast, 2021-2030. Portland: Allied Market Research, 2021.

SOLAWETZ, Jacob. Data Augmentation in YOLOv4. Roboflow Blog, 13 maio 2020a. Disponível em: https://blog.roboflow.com/yolov4-data-augmentation/. Acesso em: 5 jun. 2026.

SOLAWETZ, Jacob. Getting Started with Data Augmentation in Computer Vision. Roboflow Blog, 2 jun. 2020b. Disponível em: https://blog.roboflow.com/boosting-image-detection-performance-with-data-augmentation/. Acesso em: 5 jun. 2026.

SOLAWETZ, Jacob. What is Mean Average Precision (mAP) in Object Detection? Roboflow Blog, 1 maio 2026. Disponível em: https://blog.roboflow.com/mean-average-precision/. Acesso em: 5 jun. 2026.

ULTRALYTICS. Split the Dataset Correctly. Disponível em: https://academy.ultralytics.com/courses/dataset-readiness-for-yolo/split-the-dataset-correctly. Acesso em 18/05/2026.

WANGENHEIM, Aldo von. Deep Learning: Detecção de Objetos em Imagens. LAPIX/UFSC, 2018. Disponível em: https://lapix.ufsc.br/ensino/visao/visao-computacionaldeep-learning/deteccao-de-objetos-em-imagens/. Acesso 30 maio 2026.

# APÊNDICE A

**FIGURA 14. VISÃO GERAL DOS RESULTADOS DOS EXPERIMENTOS**

| ID | Base | Experimento | Tamanho_Dataset | Treino | Validação | Teste | Modelo_Base | Aug_Params_Resumo | mAP@0.5 | Precision | Recall | Δ mAP@0.5 vs Baseline | Δ Precision vs Baseline | Δ Recall vs Baseline |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Bezerreiros 50 | Grupo 0 Baseline Fast | 50 | 40 | 5 | 5 | YOLOv11 | No augmentations | 99,2 | 93,7 | 98,9 | 0,0 | 0,0 | 0,0 |
| 2 | Bezerreiros 50 | Grupo 0 Baseline Accurate | 50 | 40 | 5 | 5 | YOLOv11 | No augmentations | 99,5 | 98,8 | 98,7 | 0,0 | 0,0 | 0,0 |
| 3 | Bezerreiros 50 | Grupo 1 Geométrico Fast | 83 | 73 | 5 | 5 | YOLOv11 | Flip: H | 99,4 | 97,0 | 98,1 | 0,2 | 3,3 | -0,8 |
| 4 | Bezerreiros 50 | Grupo 1 Geométrico Accurate | 83 | 73 | 5 | 5 | YOLOv11 | Flip: H | 99,5 | 98,6 | 98,1 | 0,0 | -0,2 | -0,6 |
| 5 | Bezerreiros 50 | Grupo 2 Cor, Luz Fast | 130 | 120 | 5 | 5 | YOLOv11 | Saturation ±30%; Brightness ±30% | 99,3 | 97,5 | 98,2 | 0,1 | 3,8 | -0,7 |
| 6 | Bezerreiros 50 | Grupo 2 Cor, Luz Accurate | 130 | 120 | 5 | 5 | YOLOv11 | Saturation ±30%; Brightness ±30% | 99,1 | 97,5 | 98,1 | -0,4 | -1,3 | -0,6 |
| 7 | Bezerreiros 100 | Grupo 0 Baseline Fast | 100 | 80 | 10 | 10 | YOLOv11 | No augmentations | 93,5 | 89,3 | 87,8 | 0,0 | 0,0 | 0,0 |
| 8 | Bezerreiros 100 | Grupo 0 Baseline Accurate | 100 | 80 | 10 | 10 | YOLOv11 | No augmentations | 94,4 | 86,8 | 89,3 | 0,0 | 0,0 | 0,0 |
| 9 | Bezerreiros 100 | Grupo 1 Geométrico Fast | 204 | 184 | 10 | 10 | YOLOv11 | Flip H, V | 93,2 | 90,8 | 89,3 | -0,3 | 1,5 | 1,5 |
| 10 | Bezerreiros 100 | Grupo 1 Geométrico Accurate | 214 | 194 | 10 | 10 | YOLOv11 | Flip H, V | 93,7 | 86,9 | 88,6 | -0,7 | 0,1 | -0,7 |
| 11 | Bezerreiros 100 | Grupo 2 Cor, Luz Fast | 260 | 240 | 10 | 10 | YOLOv11 | Saturation ±30%; Brightness ±30% | 93,3 | 90,6 | 89,3 | -0,2 | 1,3 | 1,5 |
| 12 | Bezerreiros 100 | Grupo 2 Cor, Luz Accurate | 260 | 240 | 10 | 10 | YOLOv11 | Saturation ±30%; Brightness ±30% | 94,3 | 91,0 | 92,1 | -0,1 | 4,2 | 2,8 |
| 13 | Bezerreiros 100 | Grupo 3 Combinação Fast | 260 | 240 | 10 | 10 | YOLOv11 | Flip H, Rotation: ±15°, Ruído:2.51% of pixels | 93,5 | 91,8 | 88,6 | 0,0 | 2,5 | 0,8 |
| 14 | Bezerreiros 100 | Grupo 3 Combinação Accurate | 260 | 240 | 10 | 10 | YOLOv11 | Flip H, Rotation: ±15°, Ruído:2.51% of pixels | 94,3 | 91,0 | 92,0 | -0,1 | 4,2 | 2,7 |
| 15 | Bezerreiros 228 | Grupo 0 Baseline Fast | 228 | 182 | 23 | 23 | YOLOv11 | No augmentations | 86,1 | 86,1 | 80,2 | 0,0 | 0,0 | 0,0 |
| 16 | Bezerreiros 228 | Grupo 0 Baseline Accurate | 228 | 182 | 23 | 23 | YOLOv11 | No augmentations | 88,4 | 86,4 | 81,8 | 0,0 | 0,0 | 0,0 |
| 17 | Bezerreiros 228 | Grupo 1 Geométrico Fast | 459 | 413 | 23 | 23 | YOLOv11 | Flip H, V | 86,2 | 85,7 | 79,5 | 0,1 | -0,4 | -0,7 |
| 18 | Bezerreiros 228 | Grupo 1 Geométrico Accurate | 463 | 417 | 23 | 23 | YOLOv11 | Flip H, V | 87,0 | 87,8 | 77,3 | -1,4 | 1,4 | -4,5 |
| 19 | Bezerreiros 228 | Grupo 2 Cor, Luz Fast | 592 | 546 | 23 | 23 | YOLOv11 | Saturation ±30%; Brightness ±30% | 87,7 | 88,9 | 79,4 | 1,6 | 2,8 | -0,8 |
| 20 | Bezerreiros 228 | Grupo 2 Cor, Luz Accurate | 592 | 546 | 23 | 23 | YOLOv11 | Saturation ±30%; Brightness ±30% | 87,0 | 87,7 | 80,4 | -1,4 | 1,3 | -1,4 |
| 21 | Bezerreiros 228 | Grupo 3 Combinação Fast | 592 | 546 | 23 | 23 | YOLOv11 | Flip H, Rotation: 15°, Ruído: 2.51% of pixels | 86,6 | 88,5 | 76,3 | 0,5 | 2,4 | -3,9 |
| 22 | Bezerreiros 228 | Grupo 3 Combinação Accurate | 592 | 546 | 23 | 23 | YOLOv11 | Flip H, Rotation: 15°, Ruído: 2.51% of pixels | 86,3 | 88,2 | 79,6 | -2,1 | 1,8 | -2,2 |
| 23 | Bezerreiros 228 | Grupo 4 Diversos Fast | 546 | 477 | 46 | 23 | YOLOv11 | Flip H, Grayscale: 25%, Saturation ±30%, Brightness ±30%, Exposure ±15% e Noise: up to 1.99% of pixels | 80,9 | 81,4 | 70,4 | -5,2 | -4,7 | -9,8 |
| 24 | Bezerreiros 228 | Grupo 4 Diversos Fast | 546 | 477 | 46 | 23 | YOLOv11 | Flip H, Rotation: ±15°, Grayscale: 25%, Saturation ±30%, Brightness ±30%, Exposure ±15% e Noise: up to 1.99% of pixels | 80,7 | 81,0 | 72,3 | -5,4 | -5,1 | -7,9 |

**FONTE: O AUTOR (2026).**