

Definição e Identificação de Cenários de Variação

Contexto

O conjunto de dados disponível é reduzido (125 imagens para treino e 120 para teste). Em situações assim, o modelo tende a sofrer com baixa generalização, exigindo maior dependência de técnicas de data augmentation. As arquiteturas avaliadas foram: YOLOv8, YOLOv11 e Faster R-CNN.

Variações Relevantes no Domínio

Com base na análise do dataset e nas técnicas aplicadas, os principais cenários de variação que afetam o desempenho dos detectores foram:

Variações de posição e escala do objeto

O alvo (broca) pode aparecer centralizado, parcialmente cortado ou em escalas diferentes. A configuração `crop_fraction(0.8)` e `translate(0.2)` foi fundamental para expor o modelo a esse tipo de variação.

Variações de orientação

O uso de `fliplr(0.7)`, `flipud(0.1)` e `degrees(10°)` simulou a ocorrência de orientações não canônicas da praga. Esses cenários representam bem a realidade no campo, onde a praga pode estar de cabeça para baixo ou inclinada.

Variações de iluminação e textura

A aplicação de `auto_augment`, `mixup` e `erasing` buscou compensar a ausência de diversidade luminosa no dataset, ampliando o espaço de variação visual.

Variações de composição de cena

O uso de `mosaic`, `copy_paste` e `mixup` gerou contextos artificiais onde múltiplas ocorrências ou ruídos visuais coexistem. Isso simula situações em que diferentes elementos da plantação ou outros insetos podem estar presentes.

Observações por Arquitetura

YOLOv8 e YOLOv11: Mostraram melhor adaptação aos cenários de variação. Apesar de instabilidades no início do treino (oscilações fortes no mAP), a convergência final foi superior ao FasterR-CNN. → YOLOv11 apresentou melhor equilíbrio entre *precision* (0.9285) e *recall* (0.905), sugerindo maior robustez a variações espaciais e de composição.

Faster R-CNN: Apesar de maior estabilidade no treino (curva suave), teve pior desempenho em *mAP50-95* (0.6519). A limitação decorre do **pré-processamento mais restrito** e da maior dependência de dados anotados em volume, já que não absorveu bem cenários de variação com dataset pequeno.

Análise de Degradação (Pontos Críticos)

Contexto

A degradação refere-se ao impacto do aumento da complexidade do cenário ou da insuficiência de dados na performance do modelo. Foram considerados indicadores como **overfitting**, **estabilidade**, **eficiência** e **generalização**.

Evidências de Degradação

- **Oscilações de treino (instabilidade inicial)**
Nos YOLOs, houve forte oscilação no mAP50-95 até cerca de 40% do progresso do treino. Isso reflete dificuldade inicial em aprender padrões consistentes, típica em datasets pequenos. O Faster R-CNN foi estável, mas com teto de performance inferior.
- **Overfitting (gap treino/val)**
YOLOv8: gap = 0.0277 → baixo risco.
YOLOv11: gap = 0.0652 → indício maior de sobreajuste, possivelmente pelo maior poder de representação.
Faster R-CNN: gap ≈ 0 (0.003 negativo), mas com performance global inferior.
→ **Ponto crítico:** YOLOv11 pode saturar rápido e memorizar variações artificiais, perdendo generalização real.
- **Generalização (avaliação no conjunto nunca visto)**
YOLOv8: 0.2628
YOLOv11: 0.244
FRCNN: 0.1864
→ Todos os valores são baixos, confirmando a limitação causada pelo dataset pequeno. Há **forte degradação de performance** fora do conjunto de treino/validação.
- **Eficiência e Custo Computacional**
YOLOv11: melhor trade-off entre acurácia e tempo (211 ms/img).
YOLOv8: próximo em velocidade (230 ms/img).
FRCNN: mais lento (590 ms/img), inviável para cenários em tempo real.
→ **Ponto crítico:** degradação da aplicabilidade do FRCNN devido à latência elevada.

Pontos Críticos Identificados

- **Generalização insuficiente:** todos os modelos sofrem degradação ao avaliar no conjunto exclusivo (nunca visto).
- **Oscilações de treino (instabilidade):** mais pronunciada em YOLOv8 e YOLOv11, exigindo regularização e mais dados.
- **Overfitting em YOLOv11:** melhor desempenho interno, mas tendência a memorizar artefatos do dataset.
- **Custo computacional do FRCNN:** inviável em cenários críticos.

Recomendações

- O maior ponto crítico é a **baixa generalização**, diretamente associada ao tamanho do dataset.
- A degradação no desempenho é mais forte quando o modelo encontra **objetos em posições e contextos não vistos**.
- Estratégias necessárias: Aumento real do dataset (coleta e anotação).