



# Influência de Hiperparâmetros no Treinamento do YOLO

Uma investigação sobre como decisões de configuração afetam o  
aprendizado e o desempenho final

---

*Relatório técnico de pesquisa*

---

**Autor:** Miguel Batista Pinotti

**R.A.:** G839684

**Disciplina:** Processamento de imagem

**Orientador:** Jonathan Gustavo Rogéri

São José do Rio Preto  
17 de setembro de 2025

## Resumo

Este trabalho investiga, de forma empírica, como diferentes hiperparâmetros de treinamento e pós-processamento influenciam o desempenho e o custo operacional de modelos da família YOLO, com foco prático na variante YOLOv8m. São avaliados taxa de aprendizado, tamanho do *batch* e *gradient accumulation*, escolha de otimizador e *scheduler*, regularização (por exemplo, *weight decay* e EMA), conjunto de *augmentations* (mosaic, mixup, HSV, perspectiva), tamanho de entrada, número de épocas e critérios de NMS. O delineamento experimental inicia com variações de fator único para triagem de sensibilidade, seguido de busca reduzida sobre os fatores mais influentes. As métricas principais são mAP@0.5 e mAP@[0.5:0.95], além de precisão, revocação e F1; também são medidos tempo por época, latência de inferência e uso de memória. O objetivo é produzir recomendações objetivas para equilibrar desempenho e custo em cenários com restrição de recursos.

## Introdução

A detecção de objetos consolidou-se como uma tarefa central em visão computacional, sustentando aplicações que vão de inspeção industrial a sistemas embarcados de segurança viária. Entre as famílias de detectores em tempo real, a linha *You Only Look Once* (YOLO) evoluiu para arquiteturas mais leves, rápidas e com melhor precisão. O YOLOv8, em particular na variante m (“*medium*”), equilibra custo computacional e desempenho, tornando-se uma escolha frequente quando há restrições de latência e memória sem abrir mão de acurácia competitiva.

Apesar da popularidade, o comportamento do YOLOv8m é fortemente condicionado por hiperparâmetros e escolhas de pré-processamento e *data augmentation*. Taxa de aprendizado, tamanho do *batch*, política de *scheduler*, *weight decay*, *momentum*, composição de *augmentations* (por exemplo, *mosaic*, *mixup*, rotação, *shear*, *hsv*), tamanho de entrada (*img size*), número de épocas, critério de parada antecipada e estratégia de *fine-tuning* (camadas congeladas, pesos pré-treinados) podem alterar não apenas as métricas finais, mas também a estabilidade do treino, a velocidade de convergência e a capacidade de generalização para dados fora da amostra.

Nesta pesquisa, o objetivo é investigar de forma sistemática como diferentes hiperparâmetros influenciam o processo de treinamento e o resultado final do YOLOv8m. Em linhas gerais, buscamos responder: (i) quais hiperparâmetros exercem maior efeito nas métricas de avaliação em detecção de objetos; (ii) como esses efeitos interagem quando combinados; e (iii) qual é o *trade-off* entre desempenho (por exemplo, mAP) e custos operacionais relevantes, como tempo por época e latência de inferência.

**Contexto do modelo.** O YOLOv8 introduz aperfeiçoamentos de arquitetura e treinamento que priorizam eficiência, com cabeça de detecção desacoplada e blocos internos que favorecem o fluxo de gradiente. Na prática, a sensibilidade a hiperparâmetros permanece alta por envolver um pipeline complexo: normalização e *augmentations* na entrada, estratégia de *optimizer* (por exemplo, SGD/AdamW), agenda de aprendizado (*schedulers* como *cosine* ou *step*), regularização explícita (*weight decay*, *label smoothing*) e mecanismos de estabilização (EMA de pesos, *warmup*).

**Variáveis de interesse.** Nesta investigação, os principais fatores considerados incluem:

- **Taxa de aprendizado** inicial e *warmup*: controla a velocidade de atualização dos pesos e influencia diretamente a estabilidade da curva de perda.
- **Tamanho do *batch*** e *accumulation*: impactam a estimativa do gradiente, uso de memória e o ruído do processo de otimização.
- **Otimizador** e *scheduler*: diferentes combinações (por exemplo, AdamW com *cosine annealing*) podem alterar o ponto de convergência e a resistência a *overfitting*.
- **Regularização**: *weight decay*, *label smoothing* e congelamento parcial do *backbone* afetam a capacidade de generalização e o aproveitamento de pesos pré-treinados.
- **Data augmentation** e **tamanho da imagem**: *mosaic*, *mixup*, *hsv*, rotação e *perspective* podem melhorar robustez, mas interferem no regime de aprendizado e no custo computacional; o *img size* altera tanto o detalhamento das *features* quanto o tempo de processamento.
- **Épocas** e **parada antecipada**: definem o regime de treino efetivo; insuficiente resulta em subtreino, excessivo pode induzir *overfitting*.
- **CrITÉRIOS de pós-processamento**: *confidence threshold* e *IoU* do NMS influenciam diretamente as métricas de detecção na avaliação, mesmo sem alterar os pesos do modelo.

**Métricas e desfechos.** A avaliação reportará, prioritariamente, mAP@0.5 e mAP@[0.5:0.95], além de precisão, revocação e F1. Também serão observados indicadores operacionais: tempo médio por época, tempo médio de inferência por imagem e uso de memória. Curvas de perda (box, cls e dfl), junto de *learning curves*, apoiarão a análise de estabilidade e convergência.

**Desenho experimental resumido.** Para isolar efeitos e interações relevantes, o estudo adotará um *screening* inicial por fator único controlado sobre um conjunto base de hiperparâmetros razoáveis, seguido de uma busca mais informada (por exemplo, *grid* reduzido ou abordagem bayesiana) nos fatores mais sensíveis. A reprodutibilidade será favorecida por semente fixa de aleatoriedade, divisão consistente de treino/validação/teste e reaproveitamento dos mesmos pesos iniciais em todas as corridas comparáveis. Sempre que pertinente, serão usadas três repetições por configuração para mitigar variância estocástica.

**Escopo e limitações.** O foco é a variante m do YOLOv8; resultados podem não se transferir linearmente para variantes menores ou maiores. Diferenças de hardware, versão de *framework* e particularidades do *dataset* também constituem ameaças à validade externa. Ainda assim, a análise pretende oferecer diretrizes práticas para seleção de hiperparâmetros que equilibram desempenho e custo, informando experimentos futuros e implantação em ambientes com restrições reais.

Os próximos capítulos (não incluídos aqui) apresentarão a metodologia detalhada, especificação do conjunto de dados, matriz experimental, resultados e discussão.