

Detecção Automática de Bolhas em Grãos Propelentes de Motores Foguete Utilizando a Arquitetura YOLOv8

Ana Flávia Simões Pires¹, Lucas Henrique Amorim da Silva², Paulo Roberto Sakai³, Rodrigo Roversi Rapozo³, Wilson Francisco Moreira de Souza Ceron⁴, Glauco da Silva^{1,3}

¹Faculdade de Tecnologia de Guaratinguetá, Guaratinguetá/SP - Brasil

²Universidade Federal de São Paulo, São José dos Campos/SP - Brasil

³Instituto de Aeronáutica e Espaço, São José dos Campos/SP - Brasil

⁴Instituto Nacional de Pesquisas Espaciais, São José dos Campos/SP - Brasil

Resumo—A inspeção radiográfica é amplamente utilizada para identificação de defeitos internos em grãos propelentes empregados em motores foguetes a propelente sólido. Entre esses defeitos, destacam-se as bolhas, cuja presença pode comprometer a integridade estrutural do material e influenciar o desempenho do sistema propulsivo. Este trabalho apresenta a aplicação da arquitetura YOLOv8 na detecção automática de bolhas em radiografias industriais de propelentes. Para isso, foi desenvolvido um conjunto de dados a partir de radiografias reais, submetido a etapas de preparação, anotação e treinamento do modelo. Também foram avaliadas diferentes estratégias de processamento e treinamento com o objetivo de aprimorar o desempenho da detecção. Os resultados obtidos demonstram o potencial da utilização de técnicas de visão computacional como ferramenta de apoio à inspeção radiográfica de propelentes sólidos, contribuindo para a automatização do processo de identificação de defeitos.

Palavras-Chave—Grão propelente, Inteligência artificial, Radiografia

Abstract—Radiographic inspection is widely used to identify internal defects in solid rocket propellant grains. Among these defects, bubbles are particularly important because their presence may compromise the structural integrity of the material and affect the performance of the propulsion system. This work presents the application of the YOLOv8 architecture for the automatic detection of bubbles in industrial radiographic images of solid propellants. A dataset was developed from real radiographic images and subjected to preparation, annotation, and model training stages. Different processing and training strategies were also evaluated to improve detection performance. The results demonstrate the potential of computer vision techniques as a support tool for the radiographic inspection of solid propellants, contributing to the automation of defect identification processes.

Keywords—Propellant grain, Artificial intelligence, Radiography

I. INTRODUÇÃO

Motores foguete a propelente sólido são amplamente empregados em aplicações aeroespaciais devido à sua elevada confiabilidade, simplicidade construtiva e capacidade de armazenamento por longos períodos [1]. O desempenho desses

motores depende diretamente da integridade estrutural do grão propelente, tornando necessária a realização de inspeções capazes de identificar defeitos internos que possam comprometer sua operação [2].

Entre os defeitos observados em propelentes sólidos destacam-se trincas, descolamentos e vazios internos, frequentemente denominados bolhas. Esses defeitos podem alterar o comportamento mecânico do material e impactar o desempenho do sistema propulsivo, tornando sua identificação uma etapa relevante dos processos de controle de qualidade e avaliação estrutural [2].

A radiografia por raios X é uma das principais técnicas de ensaio não destrutivo empregadas para inspeção de componentes industriais, permitindo a visualização de estruturas internas sem causar danos ao material analisado. Entretanto, a interpretação dessas imagens depende da experiência dos especialistas e pode demandar tempo significativo. Nos últimos anos, técnicas de visão computacional e aprendizado profundo têm sido amplamente aplicadas na detecção automática de defeitos em diferentes contextos industriais, incluindo soldagem, materiais compósitos e ensaios não destrutivos, apresentando resultados promissores em termos de precisão e automação do processo de inspeção [3], [4].

Entre as arquiteturas mais utilizadas para detecção de objetos destaca-se a família YOLO (*You Only Look Once*), inicialmente proposta por Redmon *et al.* [5], desenvolvida para realizar detecção em tempo real com elevado desempenho. Nesse contexto, o presente trabalho tem como objetivo avaliar a utilização da arquitetura YOLOv8 na detecção automática de bolhas em radiografias de grãos propelentes, investigando especialmente o impacto da preparação dos dados, da revisão das anotações e de diferentes estratégias de treinamento sobre o desempenho do modelo.

O restante do artigo está distribuído da seguinte forma: os trabalhos relacionados são apresentados na Seção II; na Seção III são apresentados os materiais e métodos utilizados na pesquisa; os resultados obtidos e a discussão são apresentados na Seção IV; e, por fim, a Seção V apresenta as considerações finais da pesquisa.

II. TRABALHOS RELACIONADOS

O uso de técnicas de visão computacional e aprendizado profundo tem crescido significativamente em aplicações de

A. Pires, ana.pires@aluno.cps.sp.gov.br; L. Silva, lucas.henrique@unifesp.br; P. Sakai, sakaiprs@fab.mil.br; R. Rapozo, rapozorr@fab.mil.br; W. Ceron, wilson.ceron@inpe.br; G. Silva, glaucogs@fab.mil.br. Este trabalho foi parcialmente financiado pelo CNPq, por meio do Processo 196632/2025-0.

inspeção não destrutiva (*Non-Destructive Testing* – NDT). Modelos baseados em redes neurais convolucionais (CNNs) têm demonstrado elevado desempenho na detecção automática de defeitos em imagens industriais, reduzindo a dependência de inspeções exclusivamente humanas.

A família de modelos YOLO introduziu uma abordagem unificada para detecção de objetos em tempo real, permitindo simultaneamente localização e classificação em uma única etapa de processamento. Versões mais recentes, como o YOLOv8 [6], apresentam melhorias em eficiência computacional e precisão, sendo amplamente utilizadas em aplicações industriais.

Na área de inspeção radiográfica, estudos recentes têm aplicado aprendizado profundo para detecção automática de falhas em soldas, componentes metálicos e materiais compósitos [4], [7]–[11], demonstrando ganhos significativos em precisão e velocidade de análise quando comparados a métodos tradicionais baseados em regras. Revisões recentes [3], [11] indicam que técnicas de *deep learning* são particularmente eficazes na identificação de padrões complexos e sutilezas presentes em imagens de raios X.

Entretanto, a aplicação dessas técnicas em propelentes sólidos ainda é limitada na literatura, principalmente devido à escassez de bases de dados públicas e às restrições associadas a materiais aeroespaciais. Nesse contexto, o presente trabalho contribui ao investigar a viabilidade do uso do YOLOv8 para detecção automática de defeitos em radiografias de grãos propelentes, mesmo em cenários caracterizados por conjuntos de dados reduzidos.

III. MATERIAS E MÉTODOS

A. Conjunto de dados

O conjunto de dados utilizado neste trabalho foi composto inicialmente por seis radiografias de grãos propelentes obtidas durante atividades de inspeção radiográfica. As imagens apresentavam resolução aproximada de 2100 x 1250 *pixels* e continham regiões com e sem defeitos internos.

Considerando a elevada resolução das radiografias e o reduzido tamanho visual das bolhas em relação à imagem completa, foi empregada a técnica de *image patching* (*tiling*) para subdividir cada radiografia em regiões menores. Para isso, foi utilizada uma grade fixa de 4 x 2 regiões, resultando em oito *patches* por imagem e totalizando 48 imagens derivadas. Cada *patch* apresentou dimensões aproximadas de 525 x 625 *pixels*.

A utilização dessa estratégia teve como objetivo aumentar a quantidade de amostras disponíveis para treinamento e preservar detalhes locais dos defeitos. Caso as radiografias originais fossem redimensionadas diretamente para a entrada da rede neural, parte das características associadas às bolhas poderia ser perdida devido à redução da resolução espacial. O particionamento das imagens permitiu aumentar a representatividade visual dos defeitos durante o treinamento, facilitando sua identificação pelo modelo.

Após a geração dos *patches*, o conjunto de dados foi dividido em treinamento, validação e teste. O conjunto de treinamento foi composto inicialmente por 32 imagens, enquanto os conjuntos de validação e teste receberam 8 imagens cada. Com o objetivo de ampliar a variabilidade dos dados e reduzir os efeitos da quantidade limitada de radiografias originais,

técnicas de aumento de dados foram aplicadas exclusivamente ao conjunto de treinamento, elevando o número de imagens dessa etapa para 128. Os conjuntos de validação e teste permaneceram inalterados durante todos os experimentos. A Tabela I apresenta a caracterização geral do conjunto de dados utilizado ao longo do projeto.

TABELA I
CARACTERIZAÇÃO DO CONJUNTO DE DADOS

Item	Quantidade
Radiografias originais	6
Patches gerados	48
Treinamento (após <i>data augmentation</i>)	128
Imagens de validação	8
Imagens de teste	8

B. Rotulação das imagens

O processo de rotulação foi realizado utilizando a ferramenta *LabelImg*¹, amplamente empregada na construção de conjuntos de dados para tarefas de detecção de objetos. As anotações foram realizadas por meio de caixas delimitadoras (*bounding boxes*), definindo a posição e a classe de cada objeto presente nas radiografias.

O conjunto de dados foi composto por quatro classes: bolha, elástico, porta de passagem e IQI (*Image Quality Indicator*). A classe bolha representa o principal defeito de interesse deste trabalho, enquanto as demais classes correspondem a elementos presentes nas radiografias que auxiliam na contextualização da imagem e na diferenciação entre defeitos e estruturas legítimas do processo de inspeção.

Durante o desenvolvimento do projeto, foi realizada uma revisão completa das anotações com o objetivo de identificar inconsistências e defeitos não rotulados anteriormente. Essa etapa foi motivada pela análise dos resultados obtidos nos primeiros ciclos experimentais, que indicaram a existência de bolhas presentes nas radiografias, mas ausentes no conjunto de referência utilizado para treinamento e avaliação.

A revisão concentrou-se principalmente em regiões contendo bolhas pequenas e áreas de baixo contraste, considerando os cenários mais desafiadores para a detecção automática. Como resultado, o número total de instâncias anotadas aumentou de 3.865 para 6.072, representando um crescimento de aproximadamente 57% em relação ao conjunto original. Esse refinamento contribuiu não apenas para aumentar a quantidade de exemplos disponíveis para treinamento, mas também para tornar as avaliações mais consistentes, reduzindo situações em que detecções corretas eram contabilizadas como erros devido à ausência de anotações correspondentes no conjunto de referência.

A Fig. 1 apresenta um exemplo de radiografia após o processo de rotulação manual, destacando as diferentes classes utilizadas durante o treinamento do modelo.

C. Arquitetura utilizada

Os experimentos foram conduzidos utilizando inicialmente a arquitetura YOLOv8n, uma versão compacta da família YOLO, desenvolvida para tarefas de detecção de objetos em

¹<https://labelimg.org/>

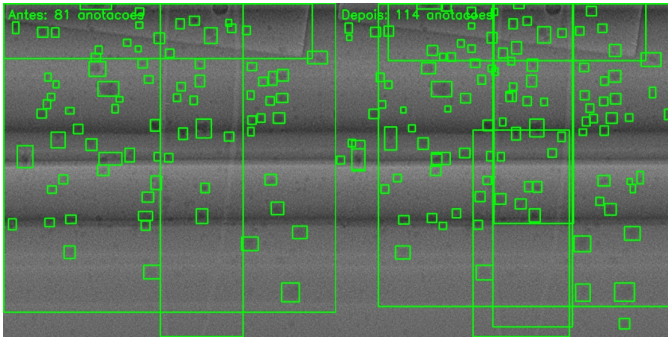


Fig. 1. Exemplo de radiografia após o processo de anotação manual.

tempo real [6]. A escolha dessa arquitetura foi motivada pelo equilíbrio entre desempenho computacional e capacidade de detecção, característica particularmente importante em cenários com recursos computacionais limitados e conjuntos de dados reduzidos.

O treinamento foi realizado empregando a estratégia de *transfer learning*, utilizando pesos pré-treinados disponibilizados pela *Ultralytics*². Esses pesos foram originalmente obtidos a partir do conjunto de dados COCO (*Common Objects in Context*), permitindo que o modelo iniciasse o treinamento já possuindo conhecimento prévio sobre padrões visuais gerais e reduzindo a necessidade de grandes volumes de dados específicos para alcançar resultados satisfatórios [12], [13].

Além da YOLOv8n, foram avaliadas as arquiteturas YOLOv8s, YOLOv10n e RT-DETR, com o objetivo de investigar a influência da arquitetura no desempenho da detecção de defeitos em radiografias de propelentes sólidos. A comparação entre diferentes modelos foi motivada pela necessidade de verificar se as limitações observadas ao longo dos experimentos estavam associadas principalmente às características do conjunto de dados ou à capacidade de representação da arquitetura utilizada.

Todos os modelos foram treinados utilizando as mesmas condições experimentais, incluindo conjunto de dados, estratégia de divisão, resolução de entrada e parâmetros de treinamento, permitindo uma comparação direta e mais justa entre as arquiteturas avaliadas. Os resultados dessa comparação são apresentados e discutidos na Subseção IV-C.

D. Estratégia experimental

O desenvolvimento do modelo foi conduzido por meio de ciclos sucessivos de treinamento, avaliação e refinamento da metodologia. Inicialmente, os experimentos foram realizados utilizando o conjunto de dados original e a arquitetura YOLOv8n como modelo de referência. A partir dos resultados obtidos, foram avaliadas diferentes estratégias envolvendo preparação dos dados, resolução de entrada, pré-processamento das imagens e revisão das anotações.

Para ampliar a variabilidade do conjunto de treinamento, foram aplicadas técnicas de aumento de dados (*data augmentation*) exclusivamente nas imagens destinadas ao treinamento. As transformações utilizadas incluíram espelhamento horizontal, espelhamento vertical e a combinação simultânea de ambas, contribuindo para aumentar a diversidade visual das

amostras sem alterar as características dos defeitos presentes nas radiografias.

Também foram investigadas diferentes técnicas de pré-processamento com o objetivo de aumentar a visibilidade das bolhas nas imagens radiográficas. Entre os métodos avaliados estão ajustes lineares de contraste com diferentes fatores de amplificação, a técnica CLAHE (*Contrast Limited Adaptive Histogram Equalization*) [14] e métodos de realce de bordas. A Fig. 2 apresenta um exemplo de imagem submetida ao processo de pré-processamento.

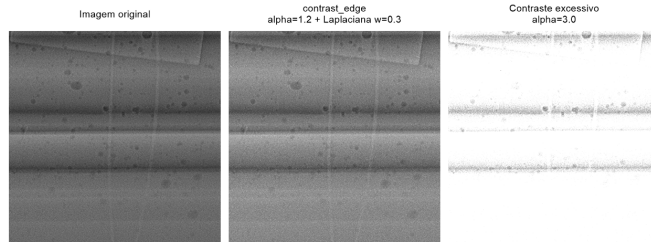


Fig. 2. Comparação entre uma radiografia original e a imagem obtida após a aplicação do pré-processamento *contrast_edge*.

Além das estratégias de pré-processamento, foram avaliadas diferentes resoluções de entrada para a rede neural, incluindo 640 x 640, 768 x 768 e 1024 x 1024 *pixels*. Os experimentos buscaram investigar a influência da resolução na capacidade de detecção de defeitos pequenos e de baixo contraste.

Ao longo do desenvolvimento do projeto, também foi realizada uma revisão das anotações do conjunto de dados com o objetivo de identificar inconsistências e instâncias não rotuladas anteriormente. Essa etapa foi conduzida após a análise dos resultados obtidos nos ciclos iniciais e passou a integrar os experimentos subsequentes.

As diferentes combinações de resolução, pré-processamento e conjunto de dados foram avaliadas por meio de ciclos experimentais sucessivos. Os resultados obtidos em cada etapa são apresentados e discutidos na Seção IV.

E. Métricas de avaliação

O desempenho dos modelos foi avaliado utilizando métricas amplamente empregadas em tarefas de detecção de objetos: *Precision*, *Recall*, mAP50 (*mean Average Precision*) e mAP50-95. A métrica *Precision* quantifica a proporção de detecções corretas em relação ao total de detecções realizadas pelo modelo, enquanto *Recall* mede a capacidade de identificar os objetos realmente presentes nas imagens.

As métricas mAP50 e mAP50-95 avaliam simultaneamente a capacidade de localização e classificação dos objetos detectados. O mAP50 considera uma detecção correta quando a sobreposição entre a caixa prevista e a anotação de referência atinge pelo menos 50% de interseção sobre união (*Intersection over Union – IoU*), enquanto o mAP50-95 utiliza múltiplos limiares de IoU entre 0,50 e 0,95, fornecendo uma avaliação mais rigorosa da qualidade das detecções.

Além das métricas quantitativas, também foram analisadas as matrizes de confusão geradas durante a validação dos modelos. Essas matrizes permitiram identificar padrões de acerto e erro entre as diferentes classes presentes no conjunto de dados, auxiliando na interpretação dos resultados obtidos ao longo dos ciclos experimentais.

²<https://www.ultralytics.com/>

F. Validação cruzada

Considerando a quantidade limitada de radiografias originais disponíveis, foi realizada uma validação cruzada do tipo *K-Fold* com o objetivo de avaliar a capacidade de generalização do modelo e reduzir possíveis vieses associados a uma única divisão dos dados.

Foi adotada uma estratégia de 4 *folds*, na qual cada iteração utilizou uma imagem original distinta para validação e as demais para treinamento. Dessa forma, todas as radiografias foram utilizadas ao menos uma vez como conjunto de validação ao longo do processo experimental.

Em cada *fold*, o modelo foi treinado utilizando os mesmos parâmetros empregados nos experimentos principais, permitindo avaliar a estabilidade das métricas em diferentes distribuições dos dados. As métricas *Precision*, *Recall*, *mAP50* e *mAP50-95* foram calculadas individualmente para cada execução e posteriormente consolidadas por meio da média e do desvio padrão.

A utilização da validação cruzada permitiu obter uma estimativa mais robusta do desempenho do modelo, especialmente em função da quantidade reduzida de imagens originais disponíveis para treinamento e validação.

IV. RESULTADOS E DISCUSSÃO

A. Evolução dos ciclos experimentais

O desenvolvimento do modelo foi realizado de forma incremental, por meio de ciclos sucessivos de treinamento, avaliação e refinamento da metodologia. O Ciclo 1 correspondeu ao treinamento inicial utilizando o conjunto de dados original. A partir dos resultados obtidos, foram identificadas limitações relacionadas à estrutura das classes e à qualidade das anotações disponíveis.

Os ciclos 2 e 3 foram agrupados por fazerem parte da mesma etapa de refinamento do projeto. O Ciclo 2 concentrou-se na análise dos resultados obtidos anteriormente, na identificação de erros do modelo e na definição de novas estratégias de melhoria. Como não houve um novo treinamento ao final dessa etapa, não foram geradas métricas independentes para o Ciclo 2.

As modificações propostas foram implementadas e avaliadas durante o Ciclo 3, que incluiu a investigação de diferentes técnicas de pré-processamento, como ajustes lineares de contraste e aplicação de CLAHE, além da execução dos treinamentos correspondentes. Os experimentos permitiram observar o impacto dessas transformações na visibilidade dos defeitos e no comportamento das métricas de detecção. Dessa forma, as métricas apresentadas na Tabela II representam o resultado consolidado das atividades desenvolvidas ao longo dos ciclos 2 e 3.

TABELA II
EVOLUÇÃO DAS MÉTRICAS AO LONGO DOS PRINCIPAIS CICLOS DO PROJETO

Ciclo	Precision	Recall	mAP50	mAP50-95
Ciclo 1	0,465	0,596	0,493	0,178
Ciclos 2-3	0,671	0,824	0,791	0,487
Ciclo 4 (640)	0,867	0,622	0,823	0,522
Ciclo 4 (768)	0,757	0,824	0,857	0,626
Ciclo 5	0,836	0,829	0,864	0,641
Ciclo 6	0,872	0,882	0,879	0,560

Posteriormente, no Ciclo 4, foram comparadas diferentes resoluções de entrada para a rede neural, permitindo avaliar o compromisso entre preservação de detalhes e estabilidade do treinamento.

Nos ciclos mais recentes, a análise dos resultados evidenciou a necessidade de uma revisão mais detalhada das anotações. Esse processo resultou na identificação de novas bolhas, principalmente em regiões de baixo contraste e defeitos de pequenas dimensões, elevando o número total de instâncias anotadas de 3.865 para 6.072. A revisão tornou o conjunto de dados mais representativo e reduziu inconsistências que poderiam impactar tanto o treinamento quanto a avaliação dos modelos.

Observa-se uma evolução significativa em relação ao modelo inicial. O *mAP50* aumentou de 0,493 no Ciclo 1 para 0,879 no Ciclo 6, enquanto o *Recall* evoluiu de 0,596 para 0,882. Esses resultados indicam que o modelo passou a identificar uma parcela significativamente maior dos defeitos presentes nas radiografias ao longo dos ciclos de refinamento.

Os resultados também evidenciam que diferentes estratégias impactaram as métricas de maneiras distintas. A configuração utilizando resolução de 640 x 640 *pixels* apresentou elevada *Precision*, porém com redução do *Recall*. Por outro lado, a resolução de 768 x 768 *pixels* demonstrou melhor equilíbrio entre capacidade de detecção e qualidade geral das previsões, apresentando os melhores resultados para *mAP50* e *mAP50-95*.

A comparação entre os ciclos também sugere que as melhorias associadas à qualidade do conjunto de dados exerceram influência significativa sobre o desempenho final do modelo. O refinamento das anotações e a inclusão de novas instâncias permitiram aumentar a representatividade dos defeitos presentes nas radiografias, contribuindo para a obtenção dos melhores resultados observados ao longo do projeto.

B. Resultados da validação cruzada

Além da divisão convencional entre treinamento, validação e teste, foi realizada uma validação cruzada do tipo *K-Fold* com o objetivo de avaliar a capacidade de generalização do modelo em diferentes distribuições dos dados. Essa etapa é particularmente importante em conjuntos de dados reduzidos, pois permite verificar se o desempenho observado depende excessivamente de uma configuração específica de treinamento e validação.

Foi adotada uma estratégia de 4 *folds*, na qual cada radiografia original foi utilizada individualmente como conjunto de validação em uma das execuções, enquanto as demais foram empregadas para treinamento. Dessa forma, todas as imagens participaram do processo de validação ao menos uma vez. Os resultados obtidos em cada *fold* podem ser vistos na Tabela III.

Os resultados demonstram comportamento consistente entre os diferentes *folds* avaliados. A métrica *mAP50* apresentou média de 0,691 e desvio padrão de apenas 0,020, indicando baixa variação entre as diferentes imagens utilizadas na validação. De forma semelhante, o *Recall* apresentou média de 0,619 e desvio padrão de 0,027, sugerindo que a capacidade de detecção do modelo permaneceu relativamente estável mesmo quando avaliado em imagens completamente inéditas.

Observa-se que o *Fold* 1 apresentou os menores resultados de *Precision* e *mAP50-95*. Esse comportamento está associado

TABELA III
RESULTADOS DA VALIDAÇÃO CRUZADA K-FOLD

Fold	Imagem de Validação	Precision	Recall	mAP50	mAP50-95
0	img_bolhas_01	0,725	0,613	0,693	0,456
1	img_bolhas_02	0,479	0,663	0,674	0,395
2	img_bolhas_03	0,755	0,589	0,673	0,466
3	img_limpa_01	0,853	0,613	0,722	0,477
Média		0,703	0,619	0,691	0,448
±		±	±	±	±
DP		0,138	0,027	0,020	0,032

à imagem “img_bolhas_02”, considerada uma das mais desafiadoras do conjunto de dados devido à elevada concentração de bolhas pequenas e regiões de baixo contraste, características que aumentam a dificuldade da tarefa de detecção.

Embora os resultados médios obtidos no *K-Fold* sejam inferiores aos melhores resultados observados nos ciclos experimentais, essa diferença é esperada. Enquanto os experimentos convencionais utilizam uma divisão fixa dos dados, a validação cruzada impõe um cenário mais rigoroso, avaliando o modelo em radiografias originais que não participaram do treinamento. Dessa forma, os resultados apresentados fornecem uma estimativa mais conservadora e realista da capacidade de generalização do modelo.

C. Comparação entre arquiteturas

Com o objetivo de investigar a influência do modelo sobre o desempenho da tarefa, foram realizados experimentos comparativos utilizando diferentes arquiteturas de detecção de objetos.

Todos os modelos foram treinados utilizando as mesmas condições experimentais, incluindo conjunto de dados, estratégia de divisão, resolução de entrada e parâmetros de treinamento. Dessa forma, as diferenças observadas podem ser atribuídas principalmente às características de cada arquitetura. A Tabela IV apresenta os resultados obtidos para as arquiteturas avaliadas.

TABELA IV
COMPARAÇÃO ENTRE ARQUITETURAS DE DETECÇÃO

Modelo	Precision	Recall	mAP50	mAP50-95
YOLOv8n	0,872	0,882	0,879	0,560
YOLOv8s	0,847	0,855	0,869	0,598
YOLOv10n	0,525	0,795	0,765	0,478

Observa-se que as arquiteturas YOLOv8n e YOLOv8s apresentaram desempenhos bastante próximos, alcançando os melhores resultados entre os modelos avaliados. A YOLOv8n obteve os maiores valores de *Precision*, *Recall* e mAP50, enquanto a YOLOv8s apresentou o melhor resultado para mAP50-95.

A arquitetura YOLOv10n apresentou desempenho inferior em todas as métricas analisadas, especialmente em *Precision*, indicando menor capacidade de identificação correta dos defeitos presentes nas radiografias quando comparada às demais arquiteturas avaliadas.

Adicionalmente, avaliou-se a viabilidade do modelo RT-DETR-L. Contudo, devido ao elevado custo computacional

inerente a essa arquitetura e à ausência de aceleração por *hardware* (GPU) no ambiente de desenvolvimento utilizado, o treinamento demonstrou-se inviável em CPU, sendo interrompido precocemente nas primeiras etapas. Diante dessa limitação de infraestrutura, a arquitetura foi desconsiderada do escopo experimental para salvaguardar a integridade e a confiabilidade da comparação quantitativa com os demais modelos.

Considerando os resultados obtidos, a arquitetura YOLOv8n foi mantida como modelo principal do projeto, apresentando o melhor desempenho global entre os modelos avaliados.

D. Avaliação do modelo final

Com base nos resultados obtidos ao longo dos ciclos experimentais e da comparação entre arquiteturas, a configuração final selecionada foi composta pela arquitetura YOLOv8n, resolução de entrada de 768 x 768 *pixels*, pré-processamento *contrast_edge* e conjunto de dados revisado.

Essa configuração apresentou os melhores resultados globais do projeto, alcançando *Precision* de 0,872, *Recall* de 0,882, mAP50 de 0,879 e mAP50-95 de 0,560. Os resultados indicam que o modelo foi capaz de identificar corretamente a maior parte das bolhas presentes nas radiografias, mantendo simultaneamente baixos níveis de detecções incorretas.

A Fig. 3 apresenta a matriz de confusão obtida para o modelo final. Observa-se predominância de classificações corretas para todas as classes avaliadas, indicando que o modelo foi capaz de distinguir adequadamente as bolhas dos demais elementos presentes nas radiografias.

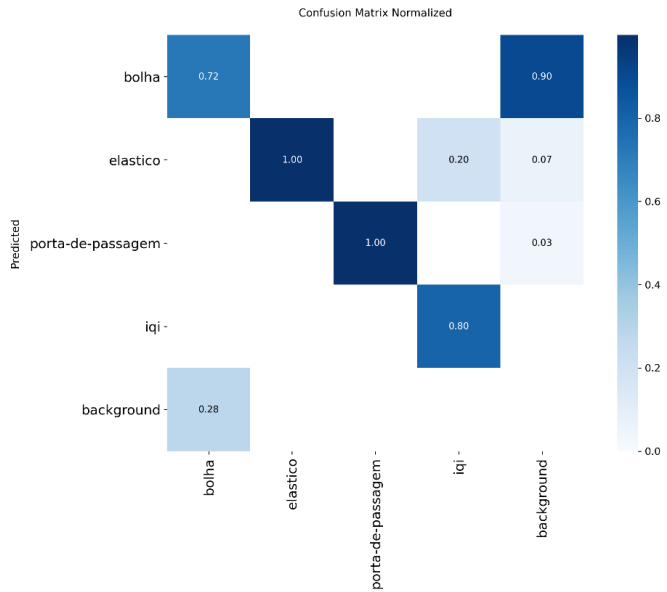


Fig. 3. Matriz de confusão normalizada do modelo final.

Complementarmente, a Fig. 4 apresenta exemplos de predições realizadas pelo modelo. As caixas delimitadoras demonstram a capacidade da rede em localizar defeitos em diferentes regiões das radiografias, incluindo cenários contendo múltiplas bolhas distribuídas ao longo da imagem.

Os resultados obtidos evidenciam a evolução do desempenho ao longo do projeto e demonstram a viabilidade da aplicação de técnicas de aprendizado profundo para auxiliar

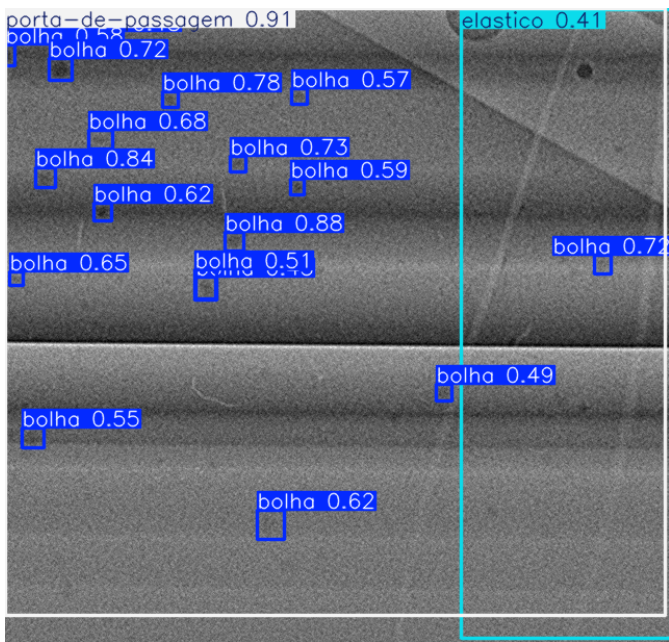


Fig. 4. Predição realizada pelo modelo final em imagens do conjunto de validação.

a identificação automática de defeitos em radiografias de propelentes sólidos.

V. CONSIDERAÇÕES FINAIS

Este trabalho apresentou a aplicação de técnicas de aprendizado profundo para a detecção automática de bolhas em radiografias de grãos propelentes. Ao longo do desenvolvimento foram conduzidos ciclos sucessivos de treinamento e refinamento, envolvendo preparação dos dados, aumento de dados, pré-processamento de imagens, revisão das anotações e avaliação de diferentes arquiteturas de detecção.

Os resultados obtidos demonstraram uma evolução consistente do desempenho do modelo ao longo dos experimentos. A configuração baseada na arquitetura YOLOv8n, associada à resolução de entrada de 768 x 768 pixels e ao pré-processamento *contrast_edge*, apresentou os melhores resultados entre as abordagens avaliadas, alcançando métricas superiores às obtidas nas etapas iniciais do projeto.

Além das melhorias observadas nas métricas de detecção, o trabalho permitiu compreender de forma mais aprofundada os desafios associados à identificação automática de defeitos em radiografias de propelentes, especialmente em regiões contendo bolhas pequenas e padrões de baixo contraste. A análise dos resultados também evidenciou a importância da qualidade das anotações e do refinamento contínuo do conjunto de dados para o desempenho dos modelos de aprendizado profundo.

Os resultados obtidos demonstram o potencial da utilização de técnicas de visão computacional como ferramenta de apoio à inspeção radiográfica de propelentes sólidos, contribuindo para o avanço de métodos automatizados de detecção de defeitos em aplicações aeroespaciais.

Como continuidade desta pesquisa, pretende-se realizar novos experimentos visando melhorar as métricas de detecção, com foco principalmente na identificação de bolhas pequenas e regiões de baixo contraste. Também serão conduzidas novas etapas de refinamento do conjunto de dados e das estratégias de treinamento.

Após a consolidação dos resultados para a classe bolha, a metodologia será expandida para a detecção de outros defeitos presentes em propelentes sólidos, como trincas e descolamentos.

AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo auxílio financeiro por meio do Programa Institucional de Bolsas de Iniciação Científica (PIBIC), Processo 196632/2025-0.

Ao Instituto de Aeronáutica e Espaço (IAE) pela oportunidade de desenvolvimento da pesquisa.

REFERÊNCIAS

- [1] G. P. Sutton, O. Biblarz, J. H. Morehart, *Rocket propulsion elements*. John Wiley & Sons, 2026.
- [2] R. R. Rapozo, “Influência de defeitos no grão propelente e interfaces no desempenho de motores foguete a propelente sólido composto,” Tese (Doutorado em Ciências e Tecnologias Espaciais), Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, Brasil, 2025.
- [3] M. Amarnath, N. Sudharshan, P. Srinivas, “Automatic detection of defects in welding using deep learning - a systematic review,” *Materials Today: Proceedings*, 2023, DOI: 10.1016/j.matpr.2023.03.268.
- [4] J. Cheng, W. Tan, Y. Yuan, Z. Zhao, Y. Cheng, “Research on defect detection method for composite materials based on deep learning networks,” *Applied Sciences*, vol. 14, no. 10, 2024, DOI: 10.3390/app14104161.
- [5] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, “You only look once: Unified, real-time object detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788, DOI: 10.1109/CVPR.2016.91.
- [6] M. Sohan, T. S. Ram, C. V. R. Reddy, “A review on yolov8 and its advancements,” in *Data Intelligence and Cognitive Informatics*, I. J. Jacob, S. Piramuthu, P. Falkowski-Gilski, Eds. Singapore: Springer Nature Singapore, 2024, pp. 529–545, DOI: 10.1007/978-981-99-7962-2_39.
- [7] S. Yoon, A. (Song-Kyoo) Kim, W. J. Cantwell, C. Y. Yeun, C.-S. Cho, Y.-J. Byon, T.-Y. Kim, “Defect detection in composites by deep learning using solitary waves,” *International Journal of Mechanical Sciences*, vol. 239, p. 107882, 2023, DOI: 10.1016/j.ijmecsci.2022.107882.
- [8] S. S. da Cruz, S. B. Mafra, E. H. Teixeira, “Detecção automática de defeitos de superfície e componentes em placas de circuito impresso usando modelos de aprendizado profundo,” *Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT2024)*, vol. 1, p. 4, 2024, DOI: 10.14209/sbrt.2024.1571032822.
- [9] X. Zhang, H. Chen, Z. Liu, S. Anli, “Lightweight detr algorithm for x-ray weld defect detection,” *Scientific Reports*, vol. 15, no. 1, p. 41873, 2025, DOI: 10.1038/s41598-025-25916-2.
- [10] M. Zhang, M. Feng, C. Chen, X. Yu, G. Lian, “Weld defect detection: deep learning-based image processing and the mechanisms of defect formation,” *Archives of Computational Methods in Engineering*, vol. 33, no. 3, pp. 3525–3563, 2026, DOI: 10.1007/s11831-025-10407-4.
- [11] B. Eren, “Deep learning approaches for weld defect detection: A comprehensive review of models, applications, and future directions,” *Computers & Industrial Engineering*, vol. 212, p. 111725, 2026, DOI: 10.1016/j.cie.2025.111725.
- [12] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick, “Microsoft coco: Common objects in context,” in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 740–755, DOI: 10.1007/978-3-319-10602-1_48.
- [13] S. J. Pan Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010, DOI: 10.1109/TKDE.2009.191.
- [14] K. Zuiderveld, *Contrast limited adaptive histogram equalization*. USA: Academic Press Professional, Inc., 1994, p. 474–485.