

CLASSIFICAÇÃO DE DEFEITOS EM COURO UTILIZANDO O MODELO YOLOV11: UMA ABORDAGEM BASEADA EM APRENDIZADO POR TRANSFERÊNCIA

Amanda Caroline Soares – d2021013766@unifei.edu.br
Matheus Brendon Francisco – matheus_brendon@unifei.edu.br

Palavras-chave: Classificação de defeitos, YOLOv11, Aprendizado por transferência

1. INTRODUÇÃO

A transição para uma produção mais inteligente, eficiente e sustentável é essencial para um futuro renovável, e tecnologias de Inteligência Artificial (IA) têm se mostrado estratégicas na modernização de processos industriais, especialmente onde a inspeção de qualidade ainda depende de avaliação humana. Na indústria do couro, a variabilidade do material e a subjetividade na detecção de defeitos afetam a produtividade e o aproveitamento sustentável da matéria-prima (ATAÇ et al., 2024). Este estudo segue os princípios da produção inteligente e da norma ISO 9001:2015, promovendo decisões baseadas em dados para aumentar a qualidade dos produtos e reduzir desperdícios.

Nos últimos anos, técnicas de IA têm possibilitado a classificação automática de defeitos, aumentando a precisão e a rapidez da inspeção de qualidade em diferentes setores industriais. Esses métodos permitem identificar padrões visuais complexos e reduzir a dependência de avaliação manual, contribuindo para processos mais padronizados e eficientes.

Este estudo expande a abordagem de Banduka et al. (2024), focando em seis categorias de defeitos no couro (folding marks, growth marks, grain off, loose grain, pinhole e non defective), mostrando que mesmo com recursos computacionais limitados e sem técnicas avançadas de aumento de dados, é possível obter excelentes resultados, explorando o YOLO para classificação, além de detecção, aprimorando a inspeção automatizada de defeitos.

2. DESENVOLVIMENTO

2.1 Referencial teórico

Embora os modelos tradicionais de aprendizado de máquina tenham grande potencial, seu treinamento do zero exige muitos dados e alto poder computacional. Uma alternativa são os modelos pré-treinados, redes já treinadas com grandes volumes de dados que podem ser ajustadas para tarefas específicas, reduzindo tempo e esforço no treinamento (SHARMA; GULERIA, 2024) e servindo como base para aprendizado em novos conjuntos de dados com poucos ajustes (HE et al., 2025). Esses modelos oferecem representações visuais detalhadas, capazes de identificar padrões complexos mesmo em conjuntos de dados reduzidos (EHTISHAM et al., 2023), atuando como extratores de características primárias (CHEN et al., 2020) e permitindo inicialização de pesos com grande flexibilidade na rede (SHORTEN; KHOSHGOFTAAR, 2019).

Entre os modelos pré-treinados, o YOLO (You Only Look Once) se destaca na detecção de objetos em tempo real. Introduzido por Redmon et al. (2016), ele divide a imagem em uma grade, prevendo simultaneamente caixas delimitadoras e probabilidades de classe, utilizando camadas convolucionais e caixas de ancoragem em múltiplas escalas. O funcionamento envolve três etapas: divisão em células que predizem posição e classe, geração de caixas delimitadoras com localização, dimensão e confiança, e detecção final calculando uma pontuação combinada (BADGUJAR; POULOSE; GAN, 2024).

Apesar de amplamente utilizado para detecção, seu potencial para classificação é pouco explorado; neste estudo, será empregado o YOLO11-cls para identificar e categorizar defeitos no couro.

2.2 Metodologia

O estudo adotou a metodologia Knowledge Discovery in Databases (KDD). O dataset utilizado, obtido no Kaggle e criado por Praveen Kumar Moganam e Denis Ashok Sathia, contém 3.600 imagens distribuídas em cinco categorias de defeitos no couro e uma categoria “non defective”. As imagens foram verificadas quanto à

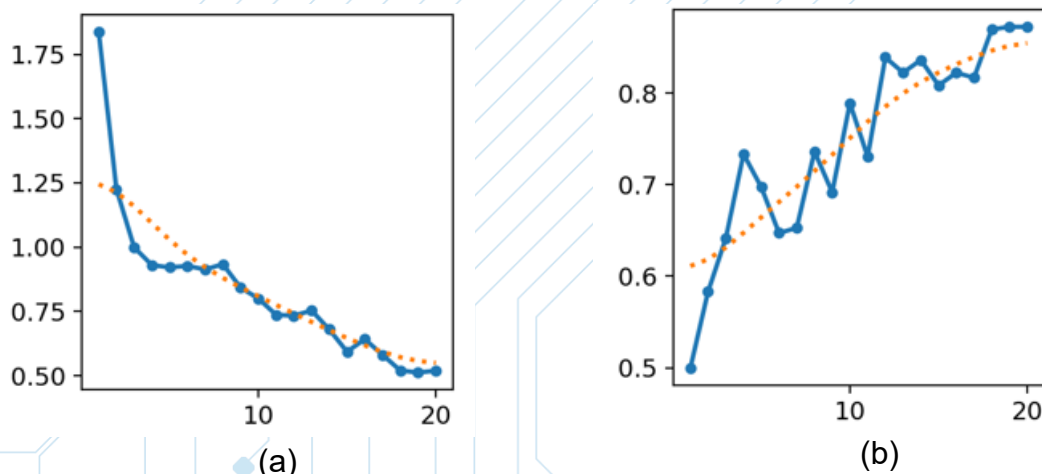
integridade e divididas em conjuntos de treinamento, validação e teste, garantindo o balanceamento entre as classes.

O treinamento foi realizado no Kaggle com uma GPU T4, utilizando o modelo YOLOv11n-cls pré-treinado, em 20 épocas, com imagens de 224×224 pixels e batch size 32. Técnicas de regularização e aumento de dados foram aplicadas para evitar overfitting e melhorar a robustez, e o processo durou aproximadamente 7 minutos e 26 segundos, evidenciando eficiência computacional. A avaliação do modelo foi realizada por meio da métrica de acurácia, permitindo analisar seu desempenho na classificação dos defeitos.

3. RESULTADOS E DISCUSSÃO

Nesta seção, apresentam-se os resultados do treinamento do modelo YOLOv11 para detecção de defeitos em couro. O desempenho foi avaliado ao longo de 20 épocas com base na função de perda (loss) e na acurácia, métricas comuns em tarefas de classificação, conforme ilustrado na Figura 1.

Figura 1 - Evolução do (a) erro e da (b) acurácia ao longo do treinamento

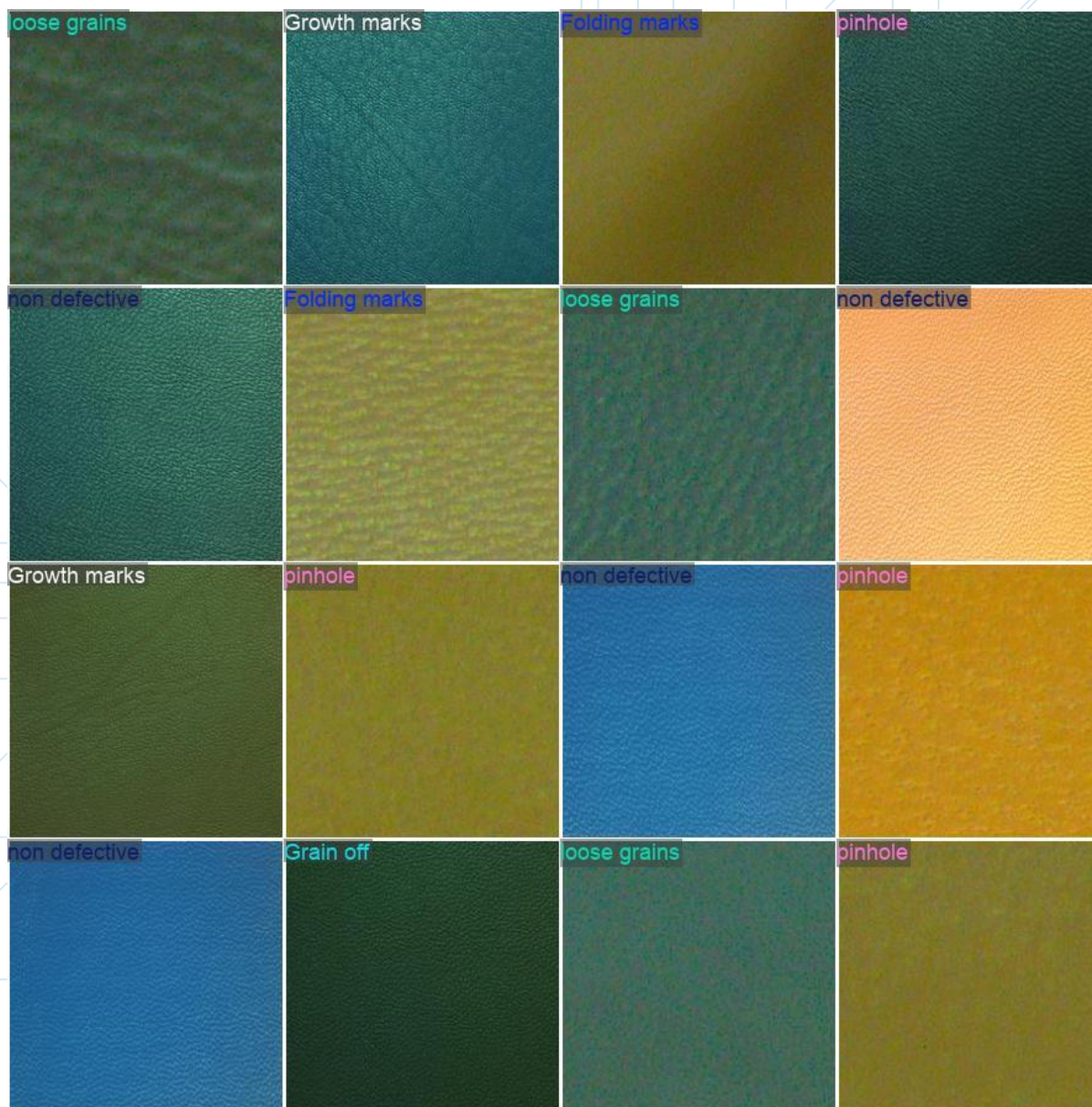


Fonte: yolo11n-cls.pt (2025)

A subfigura (a) mostra uma queda consistente da função de perda, de cerca de 1,8 para valores inferiores a 0,5, indicando que o modelo aprendeu progressivamente os padrões das imagens. Na subfigura (b), a acurácia evoluiu de aproximadamente 50% para mais de 88% após 20 épocas, apesar de pequenas

oscilações, demonstrando melhora na previsão correta das classes de defeitos. Esses resultados evidenciam o bom potencial do YOLOv11 e sua capacidade de generalização crescente, ilustrada na Figura 2, que mostra as detecções com rótulos das classes previstas sobre as imagens.

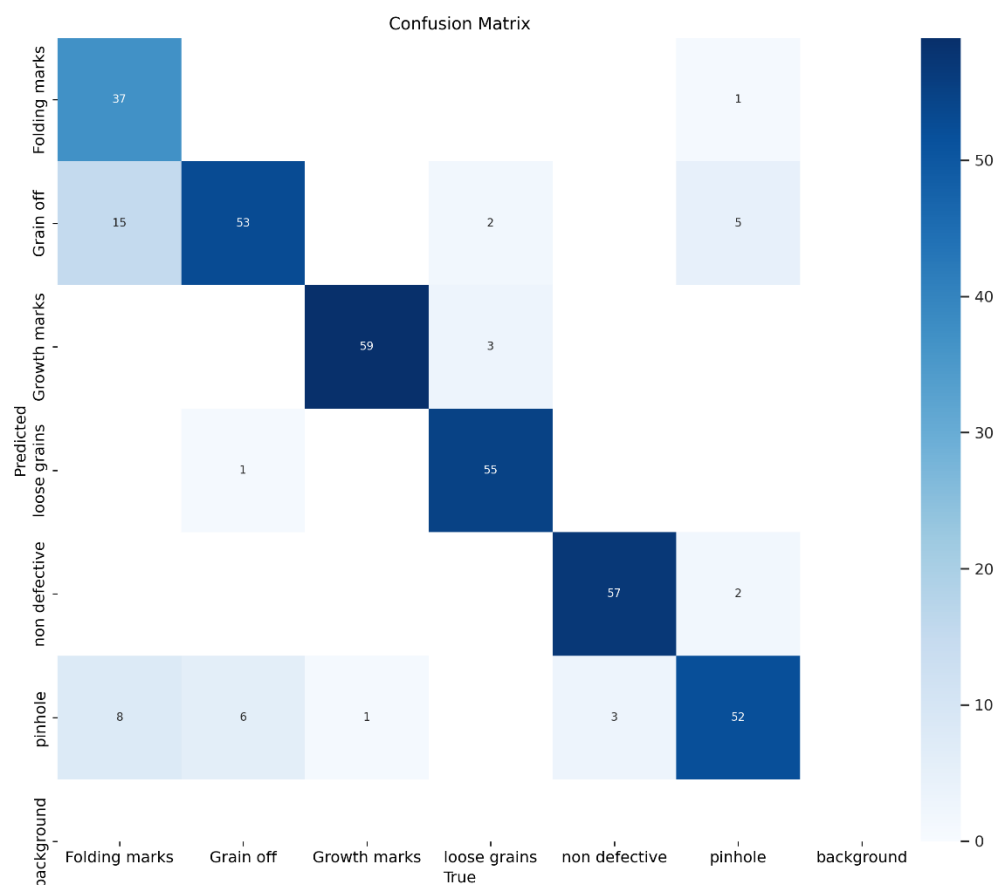
Figura 2 - Exemplos de classificação de defeitos



Fonte: yolo11n-cls.pt (2025)

A amostragem demonstra que o modelo identificou corretamente todas as classes (folding marks, growth marks, loose grains, grain off, pinhole e non defective), com caixas delimitadoras e nomes das classes sobre as imagens. A segmentação e classificação indicam que o modelo capturou padrões relevantes na textura e coloração, mesmo com a variação natural das amostras, evidenciando seu potencial para inspeção automatizada de qualidade na indústria de couro, conforme ilustrado na Figura 3.

Figura 3 - Matriz de confusão



Fonte: yolo11n-cls.pt (2025)

A Figura 3 apresenta a matriz de confusão do YOLOv11, mostrando o desempenho na classificação dos defeitos no couro. Classes como Growth marks, Loose grains, Non defective e Grain off tiveram altos acertos, com 59, 55 e 57 amostras corretamente classificadas, respectivamente. Algumas confusões ocorreram, como Grain off classificada como Folding marks em 15 casos e Pinhole

com leves erros em relação a Non defective e Folding marks, mas ainda com 52 acertos. A análise evidencia que classes com características visuais semelhantes geram erros, indicando oportunidades de melhorias no treinamento, base de dados ou balanceamento das classes, conforme resumido na Tabela 2.

Tabela 2 - Métricas obtidas

Classe	Precisão	Recall	F1-Score
0	0.6333	0.9744	0.7677
1	0.8333	0.7067	0.7852
2	0.9833	0.9516	0.9672
3	0.9167	0.9821	0.9483
4	0.9500	0.9661	0.9580
5	0.8667	0.7536	0.8062

Fonte: Elaborado pelos autores (2025)

A Tabela 2 apresenta as métricas de avaliação do YOLOv11 por classe, incluindo precisão, recall e F1-score. A classe Growth marks (2) obteve os melhores resultados, com precisão 0,9833, recall 0,9516 e F1-score 0,9672, seguida pela classe Non defective (4) com F1-score 0,9580. A classe Folding marks (0) teve menor precisão (0,6333) apesar de alto recall (0,9744), indicando muitos falsos positivos, principalmente em relação à classe Grain off (1). As classes Grain off (1) e Pinhole (5) apresentaram F1-scores mais baixos (0,7852 e 0,8062), refletindo dificuldades do modelo em diferenciá-las devido à semelhança visual entre os padrões de textura.

4. CONSIDERAÇÕES FINAIS

Os resultados obtidos demonstram a eficácia do modelo YOLOv11 na detecção automática de defeitos em couro, apresentando alto desempenho nas principais métricas de avaliação, especialmente nas classes *Growth marks*, *Loose grains* e *Non defective*. A matriz de confusão e os índices de precisão, recall e F1-score revelam que o modelo é robusto, mas ainda encontra desafios em distinguir algumas classes com características visuais similares, como *Folding marks* e *Grain off*.

O uso de técnicas de inteligência artificial como o YOLO se mostra promissor para aplicações industriais, contribuindo significativamente para o controle de qualidade automatizado. Futuros trabalhos podem explorar o balanceamento do conjunto de dados, o aumento de amostras com *data augmentation* e o ajuste fino de hiperparâmetros para melhorar ainda mais o desempenho nas classes com menor acurácia.

AGRADECIMENTOS

Os autores agradecem à Universidade Federal de Itajubá pelo suporte institucional e acadêmico, e à FAPEMIG pelo apoio financeiro.

REFERÊNCIAS

ATAÇ, Hasan Onur; KAYABAŞI, Ahmet; ASLAN, M. Fatih. **The study on multi-defect detection for leather using object detection techniques.** *Collagen and Leather*, v. 6, n. 1, p. 1-12, 2024.

BADGUJAR, C. M.; POULOSE, A.; GAN, H. **Agricultural object detection with You Only Look Once (YOLO) algorithm: a bibliometric and systematic literature review.** *Computers and Electronics in Agriculture*, v. 223, p. 109090, ago. 2024.

BANDUKA, N. *et al.* **Automated dual-side leather defect detection and classification using YOLOv11: a case study in the finished leather industry.** *Processes*, v. 12, n. 12, p. 2892, 2024.

CHEN, J. *et al.* **Using deep transfer learning for image-based plant disease identification.** *Computers and Electronics in Agriculture*, v. 173, p. 105393, jun. 2020.

HE, A. *et al.* **DVPT: Dynamic visual prompt tuning of large pre-trained models for medical image analysis.** *Neural Networks*, v. 185, p. 107168, maio 2025.

REDMON, J. *et al.* **You only look once: unified, real-time object detection.** In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. Proceedings [...]. [S. l.: s. n.], 2016. p. 779-788.

SHARMA, S.; GULERIA, K. **A systematic literature review on deep learning approaches for pneumonia detection using chest X-ray images.** *Multimedia Tools and Applications*, v. 83, p. 24101–24151, mar. 2024.

SHORTEN, C.; KHOSHGOFTAAR, T. M. **A survey on Image Data Augmentation for Deep Learning.** *Journal of Big Data*, v. 6, n. 60, 2019.