

INSPEÇÃO INTELIGENTE 360 PARA MANUFATURA: SEGMENTAÇÃO, CLASSIFICAÇÃO, DIAGNÓSTICO E PROGNÓSTICO

Luiz Fernando Ramos Lemos – luiz.lemos@ifsuldeminas.edu.br

José Arnaldo Barra Montevechi – montevechi@unifei.edu.br

Matheus Brendon Francisco – matheus_brendon@unifei.edu.br

Palavras-chave: Inspeção inteligente, Segmentação, Classificação, Deep Learning
IA generativa, LLM

1. INTRODUÇÃO

Apesar dos avanços, a aplicação de IA em ambientes industriais ainda enfrenta desafios significativos. A escassez de *datasets* que contenham informações técnicas além das máscaras de segmentação limita a capacidade de generalização dos modelos e a sua interpretabilidade. Ademais, observa-se a utilização recorrente de métodos de *deep learning* sem uma definição clara das *features* principais do problema, resultando em sistemas de funcionamento opaco e de baixa explicabilidade (Moganam; Seelan, 2022). Diante disso, evidencia-se a necessidade de frameworks estruturados que orientem desde a caracterização do processo fabril e suas falhas potenciais até a seleção dos parâmetros de extração, classificação e segmentação. Um framework cientificamente fundamentado permite a padronização das análises e o uso combinado de abordagens estatísticas, de aprendizado profundo e de IA generativa, garantindo reprodutibilidade e confiabilidade aos resultados obtidos.

Neste contexto, propõe-se a formulação do framework Inspeção Inteligente 360, concebido para integrar técnicas de *Machine Learning* (ML), *Deep Learning* (DL), IA generativa e modelos de linguagem de larga escala (LLMs) em um fluxo unificado de inspeção visual. O framework compreende: (i) a definição do processo de manufatura e das falhas envolvidas; (ii) a identificação das *features* representativas dos defeitos (como histogramas de cor, entropia e angulação); (iii) a escolha dos métodos de classificação e segmentação mais adequados (ML, DL ou *ensembles*

híbridos); (iv) a geração de relatórios quantitativos e qualitativos que incluam classe, grau de confiança e artefatos encontrados; e (v) o uso de LLMs para analisar a coerência das previsões e elaborar relatórios inteligíveis com diagnóstico e prognóstico. Essa abordagem amplia a capacidade de explicação dos modelos e facilita a tomada de decisão baseada em evidências visuais e métricas objetivas.

2. DESENVOLVIMENTO

Nesta seção, apresentamos o referencial teórico que situa os principais marcos da IA em classificação e segmentação de imagens — do Perceptron e k-NN às CNNs e arquiteturas baseadas em Transformers (ex.: SegFormer) — com foco em aplicações em couro; em seguida, descrevemos a metodologia da pesquisa aplicada: definição de objetivos e métricas do framework Inspeção Inteligente 360.

2.1 Referencial teórico

O desenvolvimento da Inteligência Artificial (IA) aplicada à inspeção fabril passou por um processo evolutivo que acompanha o avanço das tecnologias de aprendizado de máquina e visão computacional. Desde a formulação do perceptron por Rosenblatt (1958), passando pelo *k*-Nearest Neighbors (Cover; Hart, 1967), pelo algoritmo de retropropagação (Rumelhart; Hinton; Williams, 1986) e pelos classificadores de margem máxima (Cortes; Vapnik, 1995), as bases matemáticas foram lançadas para a criação dos modelos de aprendizado profundo, que, posteriormente, revolucionaram a área com arquiteturas convolucionais como a LeNet-5 (LeCun; Bengio; Haffner, 1998) e a AlexNet (Krizhevsky; Sutskever; Hinton, 2017). A partir de 2020, com os *Transformers* (Dosovitskiy et al., 2020) e segmentadores como o SegFormer (Xie et al., 2021), consolidou-se uma nova geração de modelos capazes de representar contextos visuais com maior eficiência. A Tabela 1 sintetiza os principais marcos dessa trajetória, demonstrando o progresso dos modelos de classificação e segmentação de imagens e suas aplicações em manufatura e detecção de defeitos em couro (Bergmann et al., 2019).



IEPG SUMMIT'25

Pensando o futuro com inteligência
artificial e consciência social

Quadro 1 - Marcos da IA em classificação/segmentação de imagens (fundamentos e aplicações em couro).

Ano	Referência (título)	Aplicação/Contribuição
1958	Rosenblatt – The perceptron: a probabilistic model...	Perceptron; aprendizado linear supervisionado.
1967	Cover & Hart – Nearest neighbor pattern classification	Formulação clássica do k-NN; base teórica para classificadores por vizinhança.
1986	Rumelhart et al. – Learning representations by back-propagating errors	Retropropagação de erro; permite o treinamento de MLPs profundas.
1995	Cortes & Vapnik – Support-vector networks	Definição das SVM; princípio da margem máxima e métodos de kernel.
1998	LeCun et al. – Gradient-based learning applied to document recognition	LeNet-5; CNN para reconhecimento de padrões e documentos.
2012	Krizhevsky et al. – ImageNet classification with deep convolutional neural networks	AlexNet; salto de desempenho com GPU e ReLU.
2015	Simonyan & Zisserman – Very deep convolutional networks for large-scale image recognition	VGG16/19; profundidade simples (3x3) aprimora desempenho.
2020	Dosovitskiy et al. – An image is worth 16x16 words: Transformers for image recognition at scale	Vision Transformer (ViT); aplicação de Transformers em visão computacional.
2021	Xie et al. – SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers	SegFormer; segmentador leve com atenção e decodificador eficiente.
2012–2024	Asha; Aslam; Liong; Kumar & Ashok; Jawahar; Moganam & Seelan; Smith; Iqbal; Jawahar-FARP	Aplicações em couro: detecção de defeitos, segmentação multiclases e uso de ViT em *small data*.

Fonte: Elaborado pelos autores (2025) com base em Rosenblatt (1958); Cover e Hart (1967); Rumelhart, Hinton e Williams (1986); Cortes e Vapnik (1995); LeCun, Bengio e Haffner (1998); Krizhevsky, Sutskever e Hinton (2017); Simonyan e Zisserman (2015); Dosovitskiy et al. (2020); Xie et al. (2021); Bergmann et al. (2019).

2.2 Metodologia

A metodologia adotada fundamenta-se em uma pesquisa aplicada de caráter exploratório e experimental, estruturada a partir de uma revisão bibliográfica aprofundada sobre o estado da arte em Machine Learning (ML), Deep Learning (DL) e Modelos de Linguagem de Larga Escala (LLMs). Essa etapa permitiu mapear as principais abordagens empregadas na inspeção visual automatizada, bem como identificar lacunas teóricas e técnicas relacionadas à explicabilidade, à escassez de dados técnicos e à integração entre visão computacional e IA generativa. Foram analisados trabalhos clássicos e contemporâneos — desde os precursores baseados em aprendizado supervisionado até arquiteturas modernas de Transformers e segmentadores leves — de forma a embasar cientificamente as escolhas de métodos e métricas a serem aplicadas no framework Inspeção Inteligente 360.

A segunda etapa consiste na implementação e experimentação de métodos representativos dos paradigmas atuais da área. Foram desenvolvidos e comparados modelos de classificação e segmentação que contemplam tanto abordagens baseadas em features interpretáveis, como o Cosine–Histogram Association (CHA), quanto arquiteturas de aprendizado profundo inspiradas em SegFormer e Vision Transformers, adaptadas para o contexto de manufatura. Além disso, foram realizadas implementações de LLMs como o DeepSeek adaptadas via LoRA para o domínio fabril, com o objetivo de traduzir as saídas numéricas e geométricas dos modelos em laudos técnicos compreensíveis, contendo diagnóstico e prognóstico das não conformidades detectadas.

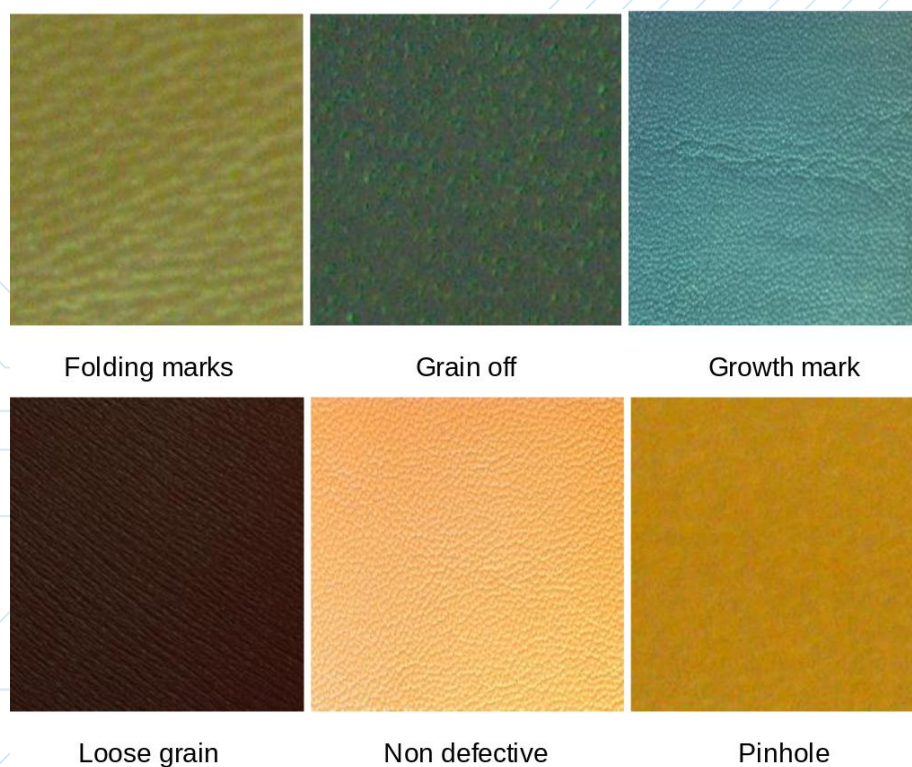
A terceira etapa envolve o contato com representantes do setor manufatureiro, de modo a compreender as demandas práticas relacionadas à automatização da inspeção de qualidade, bem como os requisitos de segurança e custo computacional. Essa interação norteou a implantação e a validação dos experimentos, permitindo confrontar os resultados obtidos com as condições reais de produção. Por fim, a aplicação da proposta no setor fabril avaliará o framework e poderá sugerir alterações.

3. RESULTADOS E DISCUSSÃO

Esta seção apresenta métodos explorados no decorrer da pesquisa, um classificador interpretável (CHA) e uma segmentação leve adaptada ao padrão de textura do couro (SegFormerLite-Leather) trabalhando em conjunção com um módulo de linguagem ajustado ao domínio (DeepSeek com treinamento LoRA).

O conjunto de dados testados com o método CHA contempla seis classes representativas do processo (folding marks, grain off, growth mark, loose grain, non defective, pinhole).

Figura 1 – Exemplo de amostras de couro.



Fonte: Formulação dos autores (2025).

No método CHA (Cosine–Histogram Association) as imagens são vetorizadas por histogramas RGB com número de bins selecionado por otimização Bayesiana, e a decisão é obtida pelo máximo do cosseno entre o vetor da amostra e os protótipos de classe, processo representado na Figura 2. A ausência de treinamento reduz substancialmente a latência de implantação e de inferência e mantém a interpretabilidade do traço decisório, permitindo associar diretamente as

características cromáticas às classes. A discussão comparativa de eficiência e efetividade, apresentada nas Tabelas 1 e 2, indica que o CHA supera baselines convolucionais em acurácia de teste e tempo de execução.

Tabela 1 – Acurácia de teste (CNNs vs CHA).

Modelo	Acurácia (teste)
AlexNet	0,876
VGG16	0,858
ResNet-50	0,836
Inception v3	0,800
CNN (sem otimização)	0,924
CNN (com otimização)	0,916
CHA (otimizado)	0,9708

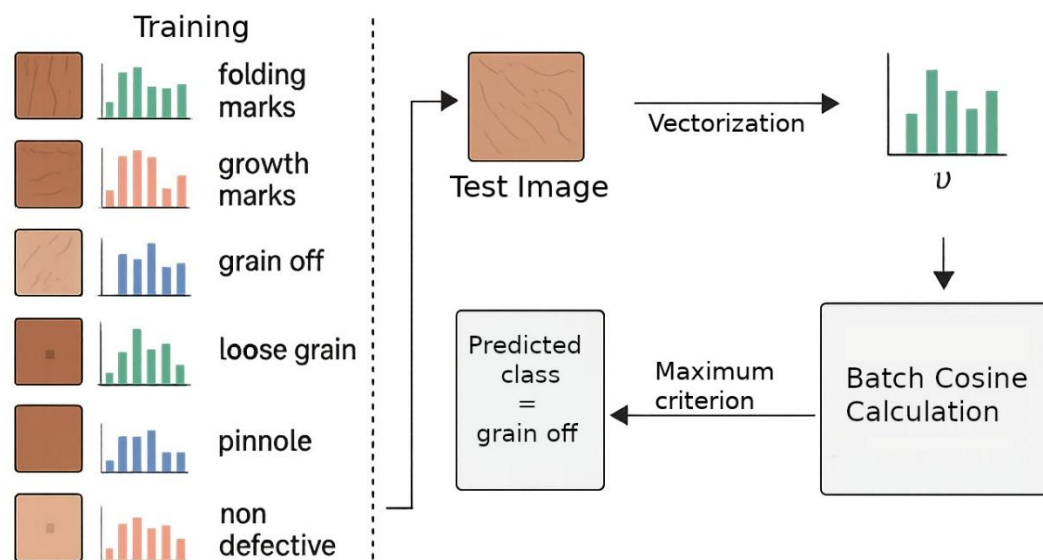
Fonte: Formulação dos autores (2025).

Tabela 2 – Tempo de execução (treinamento e inferência).

Método	Treinamento (s)	Inferência (s)
CNN (sem otimização)	105,3220	1,8184
CNN (com otimização)	140,9328	0,8354
CHA (otimizado)	— (vetorização 0,7949)	0,4048

Fonte: Formulação dos autores (2025).

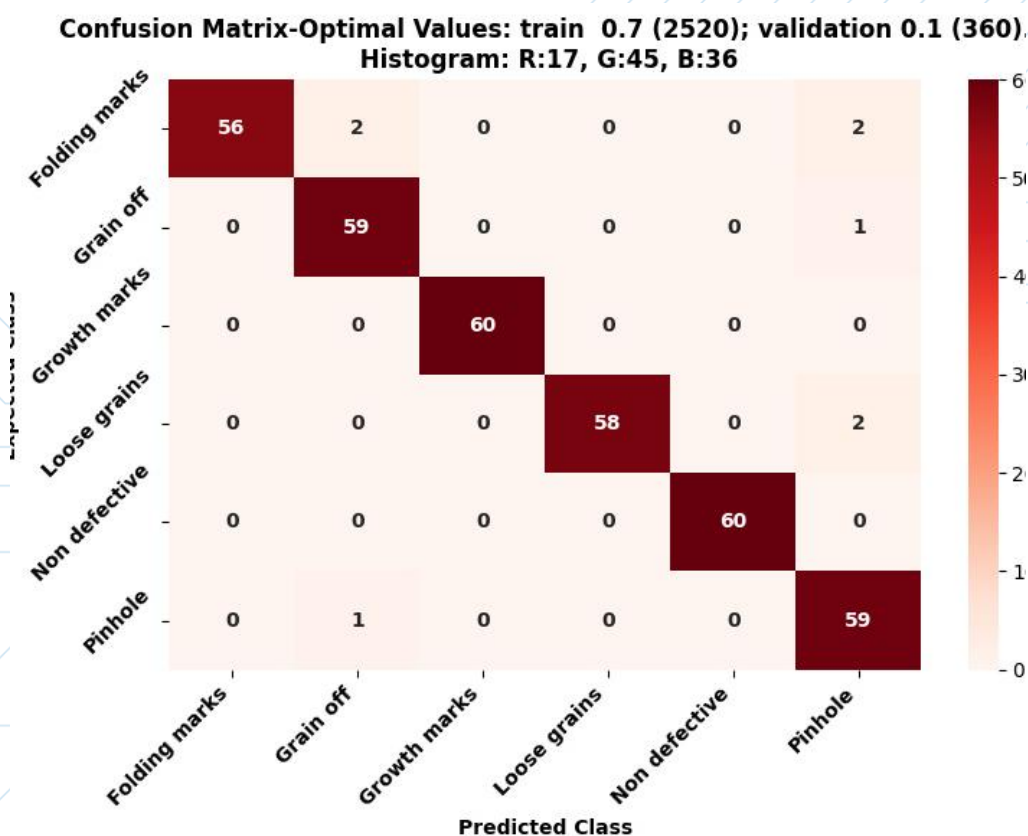
Figura 2– Representação esquemática do método CHA.



Fonte: Formulação dos autores (2025).

A avaliação global do classificador é sintetizada pela matriz de confusão do CHA otimizado, calculada com parametrização de histogramas R:17, G:45, B:36 e particionamento 0,7/0,1/0,2 para treino, validação e teste. Observa-se predominância da diagonal principal e poucos erros específicos em pares de classes de textura/croma próximos, o que é coerente com a acurácia de teste superior a 97% e com a estabilidade do método. Esses achados explicam o diferencial de velocidade e robustez frente a arquiteturas profundas, discutidos nas Tabelas 1 e 2.

Figura 3— Matriz de confusão do método CHA.



Fonte: Formulação dos autores (2025).

Para demonstrar o ciclo 360° (segmentação → classificação → checagem por LLM → diagnóstico e prognóstico), utilizou-se a SegFormerLite-Leather, rede neural de elaboração própria inspirada na SegFormer, que diferentemente do usual se selecionou features previamente, RGB, Sobel magnitude/ângulo e histogramas locais, realça ranhuras e bordas sutis típicas do material. O modelo LLM, base

DeepSeek-R1-Distill-Qwen-7B adaptada por LoRA, lê a classe e a confiança, interpreta o inventário de artefatos e emite parecer técnico. No primeiro exemplo, a LLM identifica incompatibilidade entre rótulo e evidência segmentada e recomenda inspeção manual, atuando como salvaguarda de coerência do processo.

Figura 4– Exemplo de segmentação em que a LLM diverge da classificação pela existência de artefatos.



Fonte: Formulação dos autores (2025).

O Quadro 2 apresenta o laudo gerado automaticamente para a Figura 4, estruturado em resumo quantitativo, diagnóstico e ação recomendada.

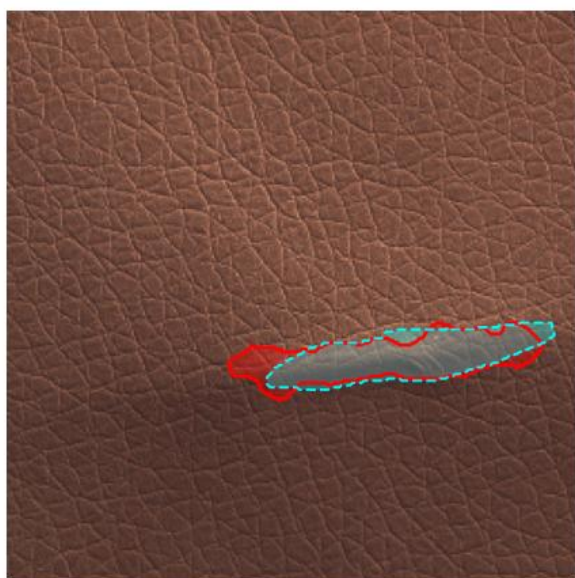
Quadro 2- Resposta associada à Figura 4

Segmentação e Classificação incompatíveis, problema identificado pela LLM que pediu por medidas.
Setting `pad\_token\_id` to `eos\_token\_id`:151643 for open-end generation.
[G][color_018.png] class_pred='good' (conf=0.9373) | comps_total=1 | IoU=0.330 Dice=0.496 | cores:
PRED=red(000000) GT=cyan(00ffff)
- comp=001 tipo=mancha area=40 bbox[x=182,y=183,w=5,h=9] centro=(184.0,186.7)
norm=(0.722,0.732) len≈9.6 wid≈5.4 ang≈83.2°
[G][LLM][DEBUG] Sem tags válidas; aplicando fallback heurístico enriquecido.
Diagnóstico
A classe prevista é 'good' (material dentro do padrão, sem falhas funcionais aparentes) (conf≈0.94),
mas foram detectados 1 mancha(s). Maior ocorrência pequeno no faixa central, no centro.
Prognóstico
Encaminhar para **inspeção manual** para confirmar os artefatos e revisar a classificação; aplicar
retrabalho local/segregação conforme severidade.

Fonte: Formulação dos autores (2025).

No segundo exemplo, o pipeline opera de forma consistente: a segmentação destaca o defeito linear, a classe prevista apresenta alta confiança e as métricas de sobreposição sustentam o parecer do LLM. O caso ilustra a rastreabilidade obtida quando máscara e inventário de artefatos corroboram a classificação e o encaminhamento.

Figura 5— Exemplo onde o processo 360, segmentação, classificação, diagnóstico e prognóstico funcionou de forma adequada.



Fonte: Formulação dos autores (2025).

O Quadro 3 lista as medidas e o parecer correspondente à Figura 5, incluindo geometria do componente principal e orientação prática de contenção e mitigação de recorrência.

Quadro 3- Resposta associada à Figura 5

Funcionamento adequado de todo o processo de classificação, segmentação, relatório quantitativo, diagnóstico e prognóstico.

Setting `pad\_token\_id` to `eos\_token\_id`:151643 for open-end generation.

[G][fold_015.png] class_pred='fold' (conf=0.9997) | comps_total=1 | IoU=0.693 Dice=0.819 | cores: PRED=red(#ff0000) GT=cyan(#00ffff)

- comp=001 tipo=linha area=2390 bbox[x=100,y=140,w=141,h=35] centro=(171.1,155.5) norm=(0.671,0.610) len≈155.4 wid≈21.7 ang≈174.5°

[G][LLM][DEBUG] Sem tags válidas; aplicando fallback heurístico enriquecido.

Diagnóstico

A classe prevista é 'fold' (vinco/dobra que altera o relevo) (conf≈1.00); foram identificados 1 linha(s).
Maior ocorrência grande no faixa central, à direita.

Prognóstico

Conter a área afetada e avaliar impacto funcional/estético; recomenda-se segregar para peças menores ou recorte seletivo e ajuste de parâmetros para reduzir reincidência.

Fonte: Formulação dos autores (2025).

Por fim, o terceiro exemplo documenta o comportamento esperado em amostras conformes. A ausência de componentes segmentados confirma a classe good, e o LLM emite um parecer objetivo, adequado para aprovação de lotes e manutenção de rastreabilidade sem onerar o processo com revisões desnecessárias.

Figura 6– Exemplo de amostra sem defeitos qualificada adequadamente.



Fonte: Formulação dos autores (2025).

O Quadro 4 sintetiza o laudo do LLM para a Figura 6, destacando a ausência de artefatos e a recomendação de rotina.

Quadro 4- Resposta associada à Figura 5

Classificação perfeita, peça sem problemas.

Setting `pad\_token\_id` to `eos\_token\_id`:151643 for open-end generation.

[G][good_006.png] class_pred='good' (conf=0.9249) | comps_total=0 | IoU=0.000 Dice=0.000 | cores: PRED=red(#ff0000) GT=cyan(#00ffff)

[G][LLM][DEBUG] Sem tags válidas; aplicando fallback heurístico enriquecido.

Diagnóstico

A classe prevista é 'good' (material dentro do padrão, sem falhas funcionais aparentes) (conf≈0.92). Não foram segmentados artefatos visíveis nesta tira.

Prognóstico

Aproveitamento integral e inspeção visual de rotina para rastreabilidade.

Fonte: Formulação dos autores (2025).

4. CONSIDERAÇÕES FINAIS

Os resultados parciais obtidos até o momento indicaram a importância da explicabilidade e de métodos simples baseados no contexto específico da produção em questão, particularmente na produção de couro o método CHA (Cosine–Histogram Association) alcançou resultados expressivos na classificação de defeitos em couro, com acurácia de teste superior a 97%, superando redes neurais convolucionais de referência (AlexNet, VGG16 e ResNet-50) e apresentando tempo de inferência significativamente menor, conforme demonstrado nas Tabelas 2 e 3 do artigo-base. O CHA mostrou-se um método interpretável, rápido e eficiente, demonstrando que técnicas baseadas em *features* bem definidas podem ser competitivas com modelos profundos, especialmente em cenários industriais onde há restrições computacionais. Os experimentos preliminares têm sido conduzidos com o *dataset* MVTec Leather mostraram que além de facilitar o processo de tomada de decisão por meio de diagnósticos e prognósticos humanamente inteligíveis o uso de LLM é capaz de corrigir classificações feitas pelos métodos de segmentação e classificação.

Ademais, o presente trabalho busca desenvolver tecnologias de baixo custo computacional e operação local (*on-premise*), dispensando o uso de nuvem e assegurando a confidencialidade dos dados industriais. Assim, pretendemos com a inspeção inteligente 360 definir parâmetros científicos para inspeção autônoma do sensoramento ao prognóstico melhorando a produtividade, precisão técnica, interpretabilidade e segurança da informação na inspeção fabril.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro do Instituto de Engenharia de Produção e Gestão da Universidade Federal de Itajubá (IEPG–UNIFEI) e o apoio institucional do Instituto Federal do Sul de Minas Gerais (IFSULDEMINAS), incluindo a concessão de afastamento para dedicação integral ao doutorado. Este trabalho também foi apoiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

REFERÊNCIAS

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, v. 65, n. 6, p. 386–408, 1958. DOI: <https://doi.org/10.1037/h0042519>

COVER, T.; HART, P. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, v. 13, n. 1, p. 21–27, 1967. DOI: <https://doi.org/10.1109/TIT.1967.1053964>

RUMELHART, D.; HINTON, G.; WILLIAMS, R. Learning representations by back-propagating errors. *Nature*, v. 323, p. 533–536, 1986. DOI: <https://doi.org/10.1038/323533a0>

CORTES, C.; VAPNIK, V. Support-vector networks. *Machine Learning*, v. 20, p. 273–297, 1995. DOI: <https://doi.org/10.1007/BF00994018>

LECUN, Y.; BENGIO, Y.; HAFNER, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, v. 86, n. 11, p. 2278–2324, 1998. DOI: <https://doi.org/10.1109/5.726791>

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, v. 60, n. 6, p. 84–90, 2017. DOI: <https://doi.org/10.1145/3065386>

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2015. DOI: <https://doi.org/10.48550/arXiv.1409.1556>

DOSOVITSKIY, A. et al. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale. *arXiv preprint arXiv:2010.11929*, 2020. DOI: <https://doi.org/10.48550/arXiv.2010.11929>



IEPG SUMMIT'25

*Pensando o futuro com inteligência
artificial e consciência social*

XIE, E. et al. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. arXiv preprint arXiv:2105.15203, 2021. DOI: <https://doi.org/10.48550/arXiv.2105.15203>

BERGMANN, P.; FAUSER, M.; SATTLEGGGER, D.; STEGER, C. MVTec AD – A comprehensive real-world dataset for unsupervised anomaly detection. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. p. 9592–9600. DOI: <https://doi.org/10.1109/CVPR.2019.00982>