



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS ANGICAL DO PIAUÍ
TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE SISTEMAS**

GLAUBER GIORDANO DE MORAIS LIMA

**FERRAMENTA BASEADA EM VISÃO COMPUTACIONAL PARA MONITORAMENTO
AUTOMÁTICO DO USO DE EPIS EM AMBIENTE INDUSTRIAL**

**ANGICAL DO PIAUÍ, PIAUÍ
2026**

GLAUBER GIORDANO DE MORAIS LIMA

FERRAMENTA BASEADA EM VISÃO COMPUTACIONAL PARA MONITORAMENTO
AUTOMÁTICO DO USO DE EPIS EM AMBIENTE INDUSTRIAL

Trabalho de Conclusão de Curso (artigo científico) apresentado como requisito para a obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Angical.

Orientador: Prof. M^e. Warlles Carlos Costa Machado

FERRAMENTA BASEADA EM VISÃO COMPUTACIONAL PARA MONITORAMENTO AUTOMÁTICO DO USO DE EPIS EM AMBIENTE INDUSTRIAL

Glauber Giordano De Moraes Lima¹

Warlles Carlos Costa Machado²

RESUMO

A segurança do trabalho em ambientes industriais, especialmente no setor de envase de água mineral, exige o cumprimento de normas sanitárias e o uso adequado de Equipamentos de Proteção Individual (EPIs). Em áreas críticas de produção, a ausência desses equipamentos pode comprometer a segurança dos trabalhadores e a conformidade do processo produtivo. Este trabalho desenvolveu e avaliou uma ferramenta para reconhecimento automático do uso de EPIs, utilizando visão computacional e aprendizado profundo. A ferramenta foi baseada no modelo de detecção YOLOv8n, treinado com imagens representativas do ambiente industrial e de cenários controlados. A arquitetura implementada separa processamento e atuação. As imagens são analisadas em computador dedicado, por meio de uma aplicação em Python com interface Tkinter, enquanto o módulo embarcado, baseado no Raspberry Pi Pico W, recebe os estados por requisições web e aciona um bastão luminoso colorido. Foram avaliados o desempenho do modelo de detecção e o funcionamento integrado do sistema. O YOLOv8n alcançou mAP@0.5 de 0,911, e o tempo médio entre a captura da imagem e a resposta física foi de aproximadamente 140 ms. Os resultados indicam viabilidade técnica para apoiar a verificação automática do uso de EPIs em ambiente industrial, observadas as limitações relacionadas à iluminação, oclusões, distância da câmera e diversidade dos dados de treinamento.

Palavras-chave: segurança do trabalho; equipamentos de proteção individual; visão computacional; YOLOv8n; Internet das Coisas.

ABSTRACT

Occupational safety in industrial environments, especially in the mineral water bottling sector, requires compliance with sanitary regulations and proper use of Personal Protective Equipment (PPE). In critical production areas, the absence of such equipment may compromise worker safety and process compliance. This work developed and evaluated a tool for automatic PPE recognition using computer vision and deep learning. The tool was based on the YOLOv8n detection model, trained with representative images from the industrial environment and controlled scenarios. The implemented architecture separates processing and actuation. Images are analyzed on a dedicated computer through a Python application with a Tkinter interface, while the embedded module, based on the Raspberry Pi Pico W, receives states through web

¹ Graduando do Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas. Instituto Federal de Educação, Ciência e Tecnologia do Piauí (IFPI), Campus Angical. E-mail: contatoglau-ber98@gmail.com.

² Mestre. Docente do Instituto Federal de Educação, Ciência e Tecnologia do Piauí (IFPI), Campus Angical. E-mail: warllescarlos@ifpi.edu.br.

requests and activates a colored light stick. The detection model performance and the integrated operation of the system were evaluated. YOLOv8n achieved an mAP@0.5 of 0.911, and the average time between image capture and physical response was approximately 140 ms. The results indicate technical feasibility to support automatic PPE verification in an industrial environment, considering limitations related to lighting, occlusions, camera distance, and training data diversity.

Keywords: occupational safety; personal protective equipment; computer vision; YOLOv8n; Internet of Things.

Data de aprovação: 30 de julho de 2026

1 INTRODUÇÃO

A segurança do trabalho desempenha papel fundamental em ambientes industriais, especialmente em setores do ramo alimentício, nos quais o cumprimento de normas sanitárias e operacionais é necessário para preservar a saúde dos trabalhadores e a qualidade dos produtos.³ Nesse contexto, a utilização adequada de Equipamentos de Proteção Individual (EPIs) constitui uma das principais medidas de prevenção de acidentes e de controle de riscos ocupacionais, sendo exigida pela legislação brasileira por meio da Norma Regulamentadora nº 6 (NR-6) (Brasil. Ministério do Trabalho e Previdência, 2022).

Em indústrias de envase de água mineral, a observância dessas normas torna-se ainda mais relevante devido às exigências relacionadas à segurança dos trabalhadores e à prevenção de contaminações durante o processo produtivo. Ambientes de acesso restrito, como a área da enchedora, demandam o uso contínuo de equipamentos como toucas, máscaras e botas, uma vez que a ausência desses itens pode comprometer as condições de trabalho e a conformidade sanitária da produção (Cavalcante, 2023).

Apesar da existência de procedimentos internos e treinamentos, a verificação do uso correto dos EPIs ainda depende, em muitos casos, da observação humana realizada por supervisores ou responsáveis operacionais. Esse processo tende a ser pontual, sujeito a falhas de percepção e limitado pela disponibilidade de acompanhamento contínuo. Tais limitações evidenciam a necessidade de soluções automatizadas capazes de apoiar o monitoramento das condições de conformidade em tempo real.

Técnicas de visão computacional e aprendizado profundo têm sido aplicadas em tarefas de inspeção visual automatizada, permitindo a identificação de pessoas, objetos e padrões em imagens e vídeos. Modelos de detecção de objetos da família YOLO (*You Only Look Once*) apresentam baixa latência e desempenho adequado para apli-

³ Uma versão preliminar deste trabalho foi publicada nos Anais Estendidos do XXII Simpósio Brasileiro de Sistemas de Informação (SBSI 2026), na Trilha de Indústria e Inovação em SI (Lima *et al.*, 2026).

cações que exigem resposta rápida, o que favorece sua utilização no reconhecimento automático de EPIs em ambientes industriais (Redmon *et al.*, 2016).

A integração entre visão computacional, Internet das Coisas (IoT) e sistemas embarcados amplia a aplicação prática dessas soluções, pois permite converter eventos detectados visualmente em respostas físicas no ambiente avaliado. Neste trabalho, essa integração foi implementada separando a inferência, executada em computador dedicado, da atuação física, realizada por um microcontrolador conectado a um sinalizador luminoso.

A verificação manual do uso de EPIs em ambientes industriais apresenta, no entanto, limitações relacionadas à frequência do acompanhamento, à cobertura das áreas monitoradas e à identificação imediata de inconformidades. Em setores com requisitos sanitários rigorosos, como o envase de água mineral, essas limitações podem reduzir a eficiência do acompanhamento das condições de segurança. Diante desse contexto, este trabalho responde à seguinte questão de pesquisa: *Como desenvolver e avaliar uma solução capaz de reconhecer automaticamente o uso de EPIs em ambiente industrial e sinalizar, com baixa latência, situações de inconformidade?*

A transformação digital dos ambientes industriais tem ampliado o uso de tecnologias capazes de apoiar atividades de monitoramento, controle e tomada de decisão. Nesse cenário, a aplicação de visão computacional ao acompanhamento do uso de EPIs representa uma alternativa para reduzir a dependência exclusiva da observação humana e ampliar a capacidade de identificação de inconformidades. Sob a perspectiva acadêmica, esta pesquisa contribui para a integração entre visão computacional, aprendizado profundo, Internet das Coisas (IoT) e sistemas embarcados em um problema aplicado de segurança ocupacional, fornecendo evidências experimentais sobre o uso do YOLOv8n em um cenário industrial específico, com avaliação do modelo de detecção e da arquitetura integrada.

Do ponto de vista tecnológico, a solução desenvolvida demonstra a viabilidade de integrar um modelo de detecção executado em computador dedicado a um módulo físico responsável apenas pela sinalização. Essa separação reduz a complexidade do dispositivo de atuação, preserva a capacidade de processamento do modelo e facilita a manutenção do sistema. A escolha do setor de envase de água mineral como cenário de aplicação confere relevância prática ao estudo, pois esse ambiente combina exigências de segurança ocupacional e controle sanitário. Os resultados obtidos podem apoiar novas aplicações de monitoramento automático de EPIs em outros ambientes industriais, desde que sejam consideradas as particularidades de cada contexto operacional.

Diante do exposto, o objetivo geral deste trabalho é desenvolver e avaliar uma ferramenta para reconhecimento automático do uso de EPIs em ambiente industrial, integrada a um módulo de atuação capaz de sinalizar situações de inconformidade. De

forma mais específica, os objetivos são:

- a) construir um conjunto de dados com imagens representativas do ambiente de aplicação e das classes relacionadas ao uso de EPIs;
- b) treinar e validar modelos de detecção de objetos da família YOLO para identificação automática de pessoas e EPIs;
- c) selecionar o modelo de detecção mais adequado considerando desempenho e latência;
- d) implementar uma aplicação em Python para captura, inferência, visualização e tomada de decisão;
- e) implementar a comunicação entre a aplicação e o módulo de atuação por meio de requisições HTTP (*Hypertext Transfer Protocol*);
- f) desenvolver a sinalização física das condições de conformidade utilizando microcontrolador e atuador luminoso;
- g) avaliar o desempenho do modelo de detecção e o funcionamento integrado da solução em cenários controlados e em ambiente industrial real.

Este trabalho está organizado conforme segue. Na Seção 2, são apresentados os conceitos teóricos que fundamentam a pesquisa, abordando segurança do trabalho, visão computacional, aprendizado profundo, detecção de objetos, Internet das Coisas (IoT), sistemas embarcados e trabalhos relacionados. Na Seção 3, descrevem-se os procedimentos metodológicos adotados, incluindo a caracterização da pesquisa, o ambiente de estudo, a construção do conjunto de dados, o treinamento dos modelos e os procedimentos de avaliação. Na Seção 4, apresenta-se o desenvolvimento da solução implementada, detalhando sua organização, a aplicação em Python, a comunicação HTTP, o módulo de atuação e o fluxo operacional. Na Seção 5, são apresentados e discutidos os resultados dos experimentos, contemplando o desempenho do modelo de detecção e o funcionamento integrado da solução. Por fim, na Seção 6, apresentam-se as considerações finais do trabalho, suas contribuições, limitações e possibilidades de continuidade.

2 REFERENCIAL TEÓRICO

Esta seção reúne os conceitos teóricos que fundamentam a pesquisa, abordando segurança do trabalho, visão computacional, aprendizado profundo, detecção de objetos, o modelo YOLO, Internet das Coisas (IoT), sistemas embarcados e trabalhos relacionados ao tema.

2.1 SEGURANÇA DO TRABALHO

A segurança do trabalho compreende o conjunto de medidas destinadas à prevenção de acidentes, à redução de riscos ocupacionais e à promoção da saúde dos trabalhadores durante o exercício de suas atividades. Essas medidas envolvem aspectos técnicos, organizacionais e normativos voltados a proporcionar condições adequadas de trabalho e minimizar a ocorrência de incidentes nos mais diversos ambientes produtivos (Zorzenon, Lizarelli e Moura, 2022).

Entre os recursos empregados para a proteção dos trabalhadores destacam-se os Equipamentos de Proteção Individual (EPIs), definidos como dispositivos ou produtos de uso individual destinados à proteção contra riscos capazes de ameaçar a segurança e a saúde no trabalho. No Brasil, sua utilização é regulamentada pela Norma Regulamentadora nº 6 (NR-6), que estabelece diretrizes relacionadas ao fornecimento, ao uso e às responsabilidades associadas a esses equipamentos (Brasil. Ministério do Trabalho e Previdência, 2022). Em setores alimentícios, como o envase de água mineral, os EPIs também contribuem para a prevenção de contaminações durante o processo produtivo, integrando práticas de segurança ocupacional e conformidade sanitária (Cavalcante, 2023; Zorzenon, Lizarelli e Moura, 2022).

2.2 VISÃO COMPUTACIONAL

A visão computacional é uma área da computação dedicada ao desenvolvimento de métodos capazes de extrair, processar e interpretar informações provenientes de imagens e vídeos. Em ambientes industriais, essa capacidade permite transformar registros visuais do processo produtivo em informações operacionais úteis, como presença de pessoas, objetos e condições de conformidade (Gonzalez e Woods, 2018; Zhou *et al.*, 2026).

Entre as principais tarefas associadas à visão computacional destacam-se a classificação de imagens, a detecção de objetos e a segmentação de imagens. Essas abordagens são amplamente utilizadas em aplicações que demandam reconhecimento automático de padrões visuais, abrangendo áreas como vigilância, inspeção automatizada, diagnóstico por imagem, veículos autônomos e automação industrial (Gonzalez e Woods, 2018; Redmon *et al.*, 2016).

Os avanços em poder computacional e aprendizado profundo ampliaram o uso dessa área em tarefas de inspeção automatizada. No acompanhamento visual do uso de EPIs, a análise objetiva das imagens reduz a dependência exclusiva da observação humana e permite identificar padrões de conformidade e inconformidade de forma contínua.

2.3 APRENDIZADO PROFUNDO E REDES NEURAIS CONVOLUCIONAIS

O aprendizado profundo é uma área da inteligência artificial baseada em redes neurais artificiais compostas por múltiplas camadas de processamento. Essa estrutura permite construir representações hierárquicas dos dados e identificar padrões complexos em imagens, característica essencial para distinguir pessoas e EPIs em diferentes posições, escalas e condições de iluminação (Goodfellow, Bengio e Courville, 2016; LeCun, Bengio e Hinton, 2015).

As Redes Neurais Convolucionais (CNNs) constituem uma das arquiteturas mais utilizadas em aplicações envolvendo imagens e vídeos. Essas redes utilizam operações de convolução para extrair automaticamente características relevantes dos dados visuais, como bordas, formas e texturas, reduzindo a necessidade de engenharia manual de atributos. Essa capacidade tornou as CNNs fundamentais para o desenvolvimento de sistemas modernos de classificação, reconhecimento e detecção de objetos (LeCun, Bengio e Hinton, 2015).

2.4 DETECÇÃO DE OBJETOS EM VISÃO COMPUTACIONAL

A detecção de objetos consiste na tarefa de identificar e localizar elementos de interesse em imagens ou vídeos. Diferentemente da classificação de imagens, que atribui um único rótulo à imagem como um todo, a detecção determina simultaneamente a classe do objeto e sua localização na cena, geralmente representada por meio de caixas delimitadoras (*bounding boxes*).

Essa abordagem permite identificar múltiplos objetos em uma mesma imagem, tornando-a adequada para aplicações que exigem análise detalhada do ambiente. Em razão dessa característica, a detecção de objetos é amplamente empregada em sistemas de vigilância, inspeção automatizada, controle de qualidade, veículos autônomos e monitoramento visual (Redmon *et al.*, 2016).

Para o reconhecimento do uso de EPIs, a detecção de objetos é mais adequada do que a classificação de imagens, pois não basta indicar a presença global de um equipamento na cena. É necessário localizar simultaneamente pessoas e itens de proteção, como toucas, máscaras e botas, permitindo avaliar a relação espacial entre o colaborador e os equipamentos exigidos.

Os avanços recentes em aprendizado profundo elevaram o desempenho dos modelos de detecção, inclusive em cenários com objetos de tamanhos distintos e variações de iluminação. Além disso, arquiteturas modernas são capazes de realizar inferências rápidas, característica necessária quando o resultado visual precisa ser convertido em uma ação operacional.

O desempenho desses modelos é normalmente avaliado por métricas que

consideram a qualidade da localização e da classificação dos objetos detectados. Entre as mais utilizadas destacam-se a precisão (*precision*), o *recall* e a média da precisão média (*mean Average Precision* – mAP), amplamente empregadas na comparação de modelos de detecção de objetos.

Essas métricas apoiam-se no conceito de Interseção sobre União (*Intersection over Union* – IoU), que mede a razão entre a área de sobreposição e a área de união das caixas delimitadoras predita e anotada. A notação mAP@0.5, adotada neste trabalho, indica o valor de mAP calculado considerando corretas as detecções cujo IoU em relação à anotação de referência é igual ou superior a 0,5.

2.5 O MODELO YOLO (*YOU ONLY LOOK ONCE*)

O YOLO é uma família de modelos de detecção de objetos baseada em aprendizado profundo, desenvolvida para realizar a identificação e a localização de objetos em uma única etapa de processamento. Essa abordagem permite que a rede neural produza simultaneamente as classes dos objetos e suas respectivas localizações na imagem, resultando em elevada eficiência computacional quando comparada a métodos que executam essas tarefas em etapas separadas (Redmon *et al.*, 2016).

Desde sua formulação inicial, o modelo passou por diversas evoluções voltadas ao aumento da precisão, da velocidade de inferência e da capacidade de generalização. Essas melhorias deram origem a diferentes versões da arquitetura, ampliando sua aplicação em cenários que exigem processamento em tempo real e elevada confiabilidade nas detecções (Redmon e Farhadi, 2018).

Entre as versões mais recentes destaca-se o YOLOv8, desenvolvido pela Ultralytics. Essa versão incorpora aprimoramentos arquiteturais e estratégias de treinamento que contribuem para melhor desempenho em tarefas de detecção de objetos, mantendo baixa latência e eficiência computacional. Arquiteturas leves dessa família, como o YOLOv8n, são particularmente adequadas para aplicações em tempo real, pois equilibram velocidade de inferência e precisão, requisito importante para a identificação rápida de inconformidades (Ultralytics, 2023).

Em razão de seu desempenho e versatilidade, a família YOLO consolidou-se como uma das principais referências em detecção de objetos, sendo amplamente empregada em sistemas de monitoramento, inspeção automatizada e análise visual em tempo real (Redmon *et al.*, 2016; Ultralytics, 2023).

2.6 DETECÇÃO AUTOMÁTICA DE EPIS EM AMBIENTES INDUSTRIAIS

A detecção automática de EPIs em ambientes industriais apresenta desafios específicos relacionados às condições reais de operação. Variações de iluminação,

oclusões parciais, diferentes posições dos trabalhadores e sobreposição entre pessoas e equipamentos podem dificultar a identificação correta dos objetos de interesse. Esses fatores afetam principalmente itens visualmente menores, como máscaras e toucas, cuja área na imagem tende a ser reduzida em câmeras posicionadas à distância.

Além disso, ambientes produtivos podem apresentar múltiplos trabalhadores em uma mesma cena, exigindo que o sistema diferencie pessoas, equipamentos presentes e situações de ausência de EPIs. Nesses casos, a resposta rápida do sistema torna-se relevante para que a identificação visual seja convertida em sinalização física ou alerta operacional sem atraso perceptível.

No contexto deste trabalho, esses desafios justificam a adoção de um modelo de detecção leve, executado em computador dedicado e integrado a um dispositivo de atuação. Essa combinação preserva a capacidade de inferência do modelo e permite converter o estado identificado em resposta física com rapidez. Cabe destacar que o cenário de aplicação deste trabalho foi delimitado ao fluxo individual de entrada dos colaboradores, de modo que o desafio associado a múltiplas pessoas simultâneas na cena é tratado como possibilidade de extensão futura da solução.

2.7 INTERNET DAS COISAS E SISTEMAS EMBARCADOS NA INDÚSTRIA 4.0

A Internet das Coisas (IoT) corresponde ao paradigma de interconexão entre dispositivos físicos capazes de coletar, transmitir e receber informações por meio de redes de comunicação. Essa abordagem permite a integração entre sensores, atuadores e sistemas computacionais, favorecendo a automação de processos e o compartilhamento de dados em tempo real (Bavaresco *et al.*, 2021).

No contexto da Indústria 4.0, a IoT desempenha papel fundamental na convergência entre o ambiente físico e os sistemas digitais, possibilitando maior conectividade, monitoramento e capacidade de tomada de decisão baseada em dados. Como consequência, diferentes componentes de uma aplicação podem atuar de forma integrada, permitindo que eventos identificados por sistemas computacionais desencadeiem ações automáticas em dispositivos conectados (Zorzenon, Lizarelli e Moura, 2022).

Os sistemas embarcados constituem parte essencial dessa infraestrutura, sendo projetados para executar funções específicas de controle, monitoramento e atuação. Em aplicações conectadas, esses dispositivos podem receber informações processadas por sistemas externos e convertê-las em respostas físicas, como sinalizações visuais, alarmes sonoros ou outras formas de atuação.

A comunicação entre os diferentes elementos de uma arquitetura IoT pode utilizar diferentes protocolos, escolhidos conforme os requisitos de latência, simplicidade de implementação, interoperabilidade e capacidade dos dispositivos envolvidos (Bavaresco *et al.*, 2021). Na arquitetura desenvolvida neste trabalho, a integração entre a aplicação

principal e o sistema embarcado foi implementada por meio de requisições HTTP, permitindo o envio do estado de conformidade identificado pelo módulo de visão computacional ao dispositivo responsável pela atuação física.

2.8 TRABALHOS RELACIONADOS

O monitoramento automatizado do uso de EPIs tem sido objeto de pesquisa em diferentes contextos industriais ao longo dos últimos anos. As abordagens desenvolvidas evoluíram desde sistemas baseados em sensores até soluções fundamentadas em técnicas de visão computacional e aprendizado profundo, impulsionadas pelos avanços recentes em detecção de objetos.

Barro-Torres *et al.* (2012) desenvolveram um sistema de monitoramento em tempo real baseado em sensores RFID (*Radio-Frequency Identification*) para verificar a utilização de EPIs em ambientes de construção civil. A solução reduziu a dependência de inspeções manuais, porém exigia instrumentação específica dos equipamentos monitorados e não realizava análise visual do ambiente, limitando sua aplicação em cenários onde se deseja inspecionar trabalhadores sem modificar os EPIs.

Posteriormente, Yang *et al.* (2020) desenvolveram um sistema inteligente voltado à segurança na construção civil, combinando dispositivos IoT e mecanismos de verificação automática da compatibilidade entre ferramentas e equipamentos de proteção. Embora demonstre o potencial de dispositivos conectados, o estudo não explora detecção visual baseada em aprendizado profundo, mantendo a lacuna relacionada à análise automática de imagens.

Com o avanço dos modelos de visão computacional, Protik, Rafi e Siddique (2021) apresentaram uma solução baseada no YOLOv4 para detecção automática de EPIs em imagens e vídeos. Os resultados evidenciaram a viabilidade de detectores em tempo real, mas a solução concentrou-se na detecção visual, sem integração com mecanismos de atuação física embarcada.

Ali *et al.* (2022) desenvolveram um sistema de monitoramento de segurança em laboratórios educacionais utilizando o YOLOv5 para identificação automática de equipamentos de proteção. Apesar dos resultados promissores quanto à precisão, o ambiente controlado reduz a complexidade operacional quando comparado a cenários industriais, e o estudo não aborda resposta física no ambiente monitorado.

No mesmo período, Lo, Lin e Hung (2023) exploraram modelos de aprendizado profundo para detecção de conformidade no uso de EPIs em cenários da construção civil, alcançando desempenho compatível com aplicações em tempo real. De forma semelhante, Al-Azani *et al.* (2024) apresentaram uma abordagem baseada no YOLOv7 para monitoramento de EPIs em imagens de câmeras de vigilância. Embora esses estudos reforcem a eficácia da detecção visual, permanecem concentrados em

contextos de construção civil ou vigilância, sem direcionamento específico à indústria alimentícia e sem integração completa entre visão computacional, comunicação com sistema embarcado e atuação física.

O Quadro 1 sintetiza a comparação entre os trabalhos analisados.

Quadro 1 – Comparação entre trabalhos relacionados e a solução desenvolvida

Trabalho	Tecnologia	Principal contribuição	Limitação frente a este trabalho	Visão	Atuação
Barro-Torres <i>et al.</i> (2012)	Sensores RFID	Monitoramento eletrônico do uso de EPIs	Exige instrumentação dos equipamentos e não realiza análise visual	Não	Não
Yang <i>et al.</i> (2020)	IoT e sensores	Integração entre dispositivos conectados e verificação de segurança	Não utiliza detecção visual baseada em aprendizado profundo	Não	Não
Protik, Rafi e Siddique (2021)	YOLOv4	Detecção visual de EPIs em imagens e vídeos	Não integra sistema embarcado para resposta física	Sim	Não
Ali <i>et al.</i> (2022)	YOLOv5	Identificação automática de EPIs em ambiente educacional	Aplicação em ambiente controlado e sem atuação física	Sim	Não
Lo, Lin e Hung (2023)	Aprendizado profundo	Detecção de conformidade em construção civil	Não é direcionado à indústria alimentícia nem à atuação embarcada	Sim	Não
Al-Azani <i>et al.</i> (2024)	YOLOv7	Monitoramento visual por câmeras de vigilância	Concentra-se na detecção visual, sem integração física local	Sim	Não
Este trabalho	YOLOv8n + HTTP	Detecção automática de EPIs integrada a sistema embarcado	Aplicação inicial restrita ao cenário avaliado	Sim	Sim

Fonte: elaborada pelo autor (2026), com base nos trabalhos citados.

Observa-se que os trabalhos identificados concentram-se predominantemente na detecção visual de EPIs ou na utilização de tecnologias IoT para monitoramento de condições operacionais. Entretanto, poucos estudos exploram, de forma integrada, detecção visual baseada em aprendizado profundo, comunicação entre aplicação principal e sistema embarcado, e atuação física local. Além disso, não foram identificados trabalhos direcionados especificamente ao contexto da indústria alimentícia. Dessa forma, este trabalho contribui para essa lacuna ao integrar detecção automática de EPIs, comunicação baseada em requisições HTTP e atuação embarcada com baixa latência.

3 METODOLOGIA

Esta seção descreve os procedimentos adotados na condução da pesquisa e na avaliação da solução desenvolvida. São apresentados a caracterização do estudo, o ambiente de aplicação, a construção do conjunto de dados, o treinamento dos modelos de detecção e os critérios de avaliação experimental.

3.1 CARACTERIZAÇÃO DA PESQUISA

Quanto à natureza, esta pesquisa caracteriza-se como aplicada, pois resultou no desenvolvimento de uma solução tecnológica voltada à verificação visual do uso de EPIs em ambiente industrial. Quanto aos objetivos, classifica-se como exploratória e experimental, uma vez que envolveu a construção de um conjunto de dados, o treinamento de modelos de detecção, a implementação de uma arquitetura funcional e a avaliação de seu desempenho.

A abordagem adotada foi predominantemente quantitativa. A avaliação do modelo de detecção considerou métricas usuais da área, como precisão (*precision*), *recall*, AP@0.5 e mAP@0.5, além do acompanhamento da evolução da mAP@0.5:0.95 ao longo do treinamento, como indicador complementar de qualidade de localização. A avaliação integrada considerou a latência entre a captura da imagem e a sinalização física executada pelo módulo de atuação.

3.2 AMBIENTE DE ESTUDO

O estudo foi conduzido no contexto de uma indústria de envase de água mineral, com foco na área da enchedora. Esse setor possui acesso restrito e exige o uso adequado de EPIs devido aos requisitos de segurança ocupacional e controle sanitário associados ao processo produtivo.

A coleta principal de imagens foi realizada com uma câmera IP de segurança em resolução Full HD, posicionada a aproximadamente dois metros da entrada da antessala da enchedora e em ângulo inclinado, de modo a capturar integralmente a passagem dos colaboradores. O enquadramento foi direcionado ao fluxo de entrada dos colaboradores, permitindo registrar situações de uso adequado e inadequado dos EPIs monitorados.

A Figura 1 apresenta o ambiente considerado neste trabalho. A escolha desse cenário decorre da necessidade de verificação automática do uso de equipamentos como toucas, máscaras e botas, frequentemente exigidos em áreas de produção alimentícia.

Figura 1 – Área da enchedora da indústria de envase de água mineral utilizada como referência para o estudo



Fonte: elaborada pelo autor (2026).

3.3 CONSTRUÇÃO DO CONJUNTO DE DADOS

Foi construído um conjunto de dados composto por aproximadamente 1.500 imagens, contendo pessoas em diferentes condições de utilização dos EPIs avaliados. As imagens foram obtidas a partir de três fontes: registros do ambiente industrial e coletas em ambiente controlado no Instituto Federal do Piauí (IFPI), capturadas em resolução Full HD e correspondendo, juntas, a aproximadamente 40% do total, e bases públicas compatíveis com as classes de interesse, cuja resolução variou conforme a fonte original e que são responsáveis pelos 60% restantes. Essa composição ampliou a diversidade visual das amostras, contemplando variações de iluminação, distância, enquadramento, postura e presença ou ausência dos equipamentos. O Quadro 2

resume as características do conjunto de dados.

Quadro 2 – Características gerais do conjunto de dados

Característica	Descrição
Quantidade aproximada	1.500 imagens
Resolução	Full HD (ambiente industrial e IFPI); variável, conforme a fonte original (bases públicas)
Fontes das imagens	Ambiente industrial e ambiente controlado no IFPI ($\approx 40\%$); bases públicas ($\approx 60\%$)
Método de captura	Ambiente IFPI: quadros extraídos de vídeo, com amostragem a cada dois minutos; ambiente industrial: quadros capturados pelo mecanismo de aquisição da solução (botão e câmera IP; Seção 4.7), durante a fase de coleta de dados
Formato de anotação	Compatível com modelos da família YOLO
Divisão dos dados	70% treinamento; 20% validação; 10% teste
Classes anotadas	Pessoa; com touca; sem touca; com máscara; sem máscara; com bota; sem bota

Fonte: elaborada pelo autor (2026).

As imagens obtidas no ambiente industrial e no ambiente controlado foram anotadas manualmente no formato YOLO, por meio de uma ferramenta própria desenvolvida pelo autor para essa finalidade. Das bases públicas do Roboflow Universe foram selecionadas apenas as imagens compatíveis com as classes de interesse deste trabalho; suas anotações, disponíveis no mesmo formato, foram revisadas e harmonizadas com a taxonomia de sete classes adotada, ajustando-se os rótulos quando necessário. Em ambos os casos, cada anotação registrou a classe do objeto e as coordenadas normalizadas da caixa delimitadora correspondente. Foram consideradas sete classes: *pessoa, com touca, sem touca, com máscara, sem máscara, com bota e sem bota*. A inclusão de classes positivas e negativas permitiu representar explicitamente situações de conformidade e inconformidade durante o treinamento. Essas classes foram distribuídas de forma aproximadamente equilibrada no conjunto de dados, reduzindo o risco de viés do modelo em favor de classes majoritárias.

Após a anotação, o conjunto de dados foi dividido de forma aleatória simples em subconjuntos de treinamento, validação e teste, correspondendo a 70%, 20% e 10% das imagens, respectivamente. Essa divisão permitiu treinar os modelos, acompanhar o desempenho durante o ajuste dos parâmetros e avaliar a capacidade de generalização em amostras não utilizadas no treinamento.

Foram aplicadas técnicas de aumento de dados (*data augmentation*) para reduzir o risco de sobreajuste e ampliar a variabilidade das amostras. As transformações

incluiram rotações, variações de brilho e alterações de escala, preservando a coerência visual das classes anotadas.

3.3.1 Aspectos éticos da coleta

A utilização das imagens coletadas no ambiente industrial e no ambiente controlado ocorreu mediante autorização dos participantes. Todos os colaboradores que participaram diretamente da aquisição das imagens assinaram Termo de Consentimento Livre e Esclarecido (TCLE), autorizando o uso de imagem exclusivamente para fins acadêmicos, restringindo sua aplicação ao desenvolvimento e à avaliação da solução descrita neste trabalho. A autorização para a captação de imagens no ambiente produtivo foi concedida verbalmente pelo gerente responsável pela unidade industrial, não havendo documento formal de autorização institucional; em razão disso, e por solicitação da empresa, sua identidade é mantida anonimizada ao longo deste trabalho.

O tratamento das imagens observou as diretrizes da Lei Geral de Proteção de Dados Pessoais (LGPD), Lei nº 13.709/2018, contemplando finalidade específica e acadêmica do tratamento e acesso restrito aos pesquisadores envolvidos. A divulgação limita-se às imagens reproduzidas neste documento, cuja utilização está amparada pelos termos de autorização assinados ou pela licença de uso livre das bases públicas. Este trabalho não foi submetido a Comitê de Ética em Pesquisa (CEP), por não envolver coleta de dados sensíveis, procedimentos invasivos ou risco direto aos participantes, restringindo-se ao registro de imagens em ambiente de trabalho mediante consentimento individual.

3.4 TREINAMENTO E SELEÇÃO DO MODELO

O reconhecimento automático dos EPIs foi realizado por modelos de detecção de objetos da família YOLO. Foram avaliadas as arquiteturas YOLOv5n e YOLOv8n, escolhidas por apresentarem versões leves e compatíveis com aplicações sensíveis ao tempo de resposta.

O treinamento utilizou o conjunto de dados previamente descrito e manteve a divisão entre treinamento, validação e teste. Durante essa etapa, os modelos aprenderam a localizar pessoas e EPIs por meio de caixas delimitadoras e a classificar cada objeto segundo as classes anotadas. A configuração principal utilizada no treinamento do YOLOv8n é apresentada na Tabela 1.

Tabela 1 – Configuração do treinamento do modelo YOLOv8n

Parâmetro	Configuração
Modelo	YOLOv8n
Framework	Ultralytics
Épocas	150
<i>Batch size</i>	8
Resolução de entrada	640 × 640 pixels
Hardware	NVIDIA RTX 3060 Ti

Fonte: elaborada pelo autor (2026).

Os demais hiperparâmetros de treinamento, não listados na Tabela 1, seguiram a configuração padrão do framework Ultralytics para o YOLOv8n.

A seleção do modelo final considerou simultaneamente desempenho de detecção e custo computacional. Foram analisadas precisão (*precision*), *recall*, AP@0.5 e mAP@0.5, além da latência média de inferência. Esse critério foi adotado porque a aplicação depende de resposta rápida para converter a detecção visual em sinalização física no ambiente avaliado. Ambos os modelos avaliados foram treinados sob a mesma configuração de épocas, *batch size*, resolução de entrada e hardware.

3.5 PROCEDIMENTOS DE AVALIAÇÃO

A avaliação experimental foi conduzida em duas etapas. A primeira avaliou o desempenho dos modelos de detecção no conjunto de teste, considerando as métricas de precisão (*precision*), *recall*, AP@0.5 e mAP@0.5. Essa etapa permitiu comparar as arquiteturas avaliadas e selecionar o YOLOv8n como modelo final da aplicação.

A segunda etapa avaliou o funcionamento integrado do sistema. Foram executados testes envolvendo captura de imagem, inferência na aplicação, envio do estado por requisição HTTP e acionamento da sinalização física. A latência total foi medida considerando o intervalo entre a captura da imagem e a atualização do atuador visual.

Os testes foram realizados em cenários controlados com situações de conformidade e inconformidade no uso dos EPIs. Além disso, a solução encontra-se implantada e em operação contínua na indústria de envase de água mineral. Nesse ambiente, foi realizada uma janela formal de observação instrumentada de aproximadamente vinte minutos, destinada ao registro do comportamento integrado entre detecção automática, comunicação HTTP e atuação embarcada, cujos resultados são relatados na Seção 5.

4 DESENVOLVIMENTO DA SOLUÇÃO

Esta seção descreve a implementação da solução. Diferentemente da metodologia, que apresenta os procedimentos da pesquisa, esta parte detalha as decisões de projeto, os componentes de software e hardware, a integração entre os módulos e o fluxo operacional.

4.1 VISÃO GERAL DA ARQUITETURA DA SOLUÇÃO

A ferramenta foi implementada em três módulos principais: aquisição de imagens, processamento e atuação. Essa organização separa a captura do ambiente, a execução do modelo de detecção e a sinalização física das condições identificadas.

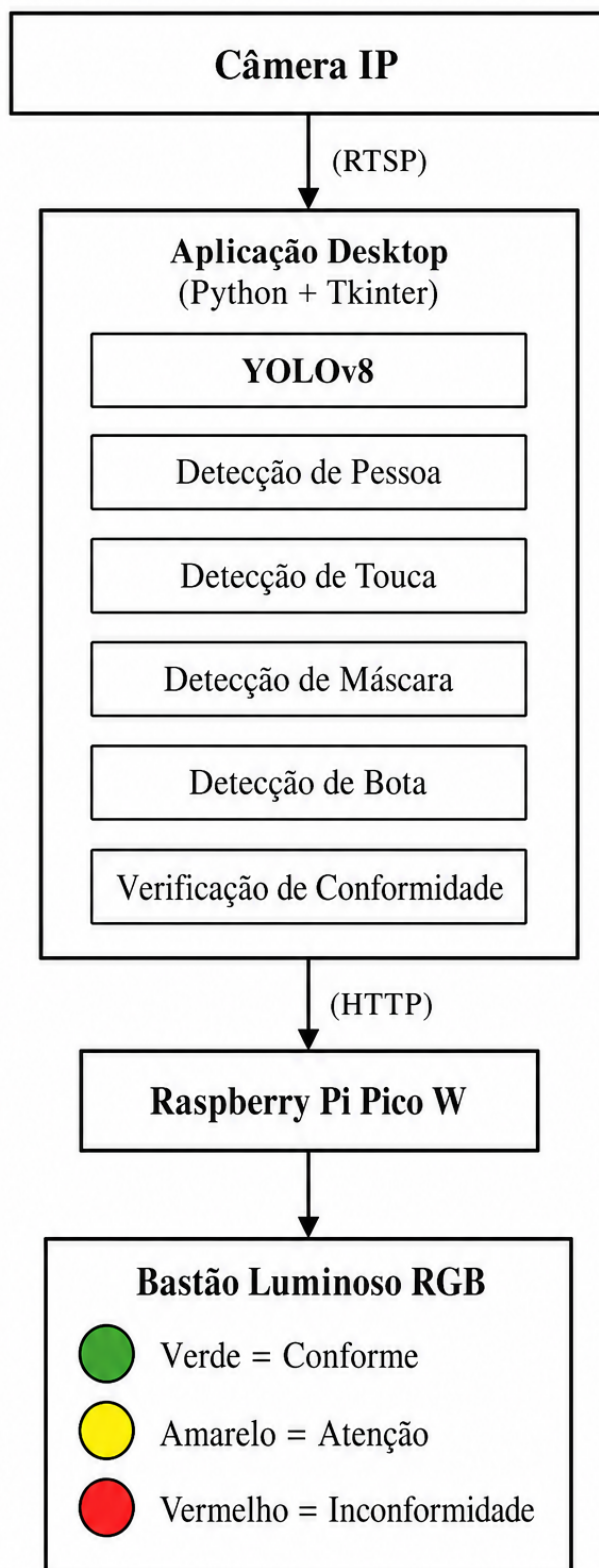
O módulo de aquisição utiliza uma câmera IP direcionada ao fluxo de entrada dos colaboradores, cujo fluxo de vídeo é transmitido continuamente por meio do protocolo RTSP (*Real Time Streaming Protocol*). A captura do quadro utilizado na inferência, no entanto, é acionada por um botão físico pressionado pelo colaborador ao adentrar o ambiente monitorado. O quadro correspondente é encaminhado ao computador dedicado, responsável pelo processamento visual, pela execução do YOLOv8n e pela definição do estado de conformidade com base nas classes detectadas.

O módulo de atuação é formado por um Raspberry Pi Pico W conectado a um bastão luminoso RGB (*Red, Green, Blue*).⁴ O dispositivo recebe o estado definido pela aplicação por meio de requisições HTTP e aciona a cor correspondente. A inferência não foi executada no microcontrolador porque essa tarefa exige maior capacidade de processamento; por isso, ele foi mantido apenas nas funções de comunicação e acionamento.

A separação entre processamento e atuação reduz a carga computacional do módulo embarcado, facilita a manutenção e permite substituir ou atualizar o modelo de detecção sem alterar a lógica de sinalização. A Figura 2 apresenta a organização geral implementada.

⁴ Versões preliminares do protótipo, apresentadas em publicação anterior derivada deste trabalho, utilizavam comunicação assíncrona via MQTT e um atuador baseado em matriz de LEDs 5x5. Ao longo do desenvolvimento, a arquitetura foi consolidada na configuração aqui descrita – requisições HTTP e bastão luminoso RGB –, sendo essa a versão avaliada nos experimentos relatados nas Seções 3 e 5.

Figura 2 – Arquitetura geral da solução desenvolvida



Fonte: elaborada pelo autor (2026).

4.2 TECNOLOGIAS UTILIZADAS

A implementação utilizou tecnologias compatíveis com os requisitos de baixa latência, simplicidade de integração e execução local. O Quadro 3 apresenta os principais recursos empregados e suas funções no sistema.

Quadro 3 – Principais tecnologias utilizadas na solução

Tecnologia	Função na solução
Python	Linguagem utilizada no desenvolvimento da aplicação principal
Botão físico	Acionamento do início da captura pelo colaborador
OpenCV	Captura e manipulação do fluxo de vídeo
Ultralytics	Execução do modelo YOLOv8n treinado
Tkinter	Construção da interface gráfica da aplicação
HTTP	Comunicação entre aplicação principal e módulo embarcado
Raspberry Pi Pico W	Recepção dos estados e acionamento do bastão luminoso RGB

Fonte: elaborada pelo autor (2026).

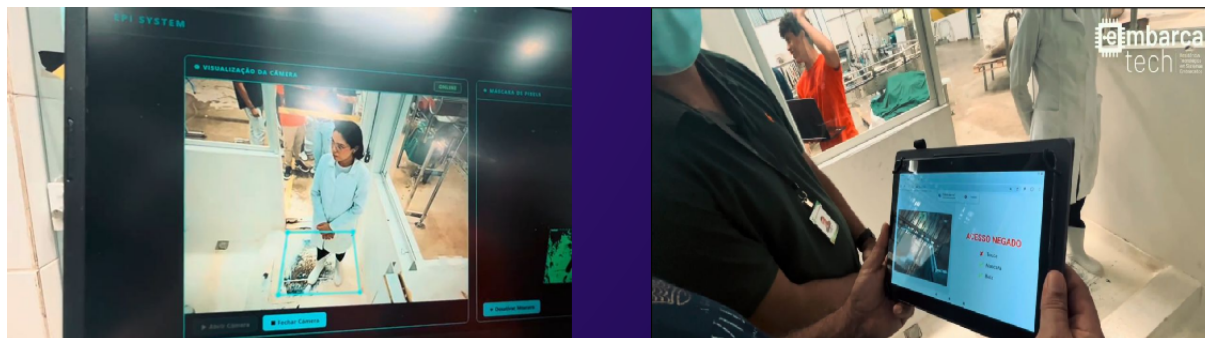
4.3 DESENVOLVIMENTO DA APLICAÇÃO DESKTOP

A interface foi desenvolvida em Python com recursos gráficos implementados em Tkinter. A escolha de uma interface local foi adequada ao contexto do experimento porque permite operar o fluxo de vídeo, visualizar as detecções e acionar a comunicação com o módulo físico sem depender de infraestrutura web externa.

O aplicativo integra a captura das imagens, a execução do YOLOv8n, a visualização das detecções, a classificação do estado de conformidade e o envio de requisições HTTP. A biblioteca OpenCV foi utilizada para tratar o fluxo de vídeo, enquanto a Ultralytics permitiu carregar os pesos treinados e executar as inferências. O Tkinter foi adotado por ser nativo ao ecossistema Python, leve e suficiente para a construção da interface necessária ao acompanhamento operacional.

A Figura 3 apresenta a interface desenvolvida para o acompanhamento das detecções e das condições de conformidade dos EPIs.

Figura 3 – Interfaces da aplicação desenvolvida para acompanhamento dos EPIs



Fonte: elaborada pelo autor (2026).

Além da visualização, o software centraliza a tomada de decisão. A partir das classes detectadas, define o estado lógico enviado ao módulo de atuação. Essa escolha evita transferir regras de interpretação para o microcontrolador e mantém a lógica principal no equipamento com maior capacidade computacional.

4.4 CRITÉRIO DE DECISÃO DA APLICAÇÃO

A lógica de decisão foi definida com base nas classes detectadas pelo modelo. Quando os EPIs exigidos são identificados em conformidade, o estado correspondente é associado à sinalização verde. Quando há ausência de item obrigatório, como touca, máscara ou bota, o estado é classificado como inconforme e associado à sinalização vermelha.

Também foi previsto um estado intermediário, associado à cor amarela, correspondente à condição de espera do sistema, quando nenhum colaborador está sendo processado pelo fluxo de captura e inferência. Essa definição simplifica a comunicação visual e mantém o módulo embarcado restrito ao acionamento da sinalização recebida.

A associação entre as detecções de EPI e a pessoa correspondente segue um critério espacial baseado nas caixas delimitadoras (*bounding boxes*) fornecidas pelo modelo. Apenas detecções com confiança mínima de 70% são consideradas válidas pela aplicação. O sistema foi projetado para a detecção de um colaborador por vez, condição compatível com o fluxo de entrada monitorado na antessala da enchedora. Nesse cenário, as detecções de EPI cuja sobreposição com a caixa delimitadora da classe *pessoa* seja igual ou superior a 70% são associadas a esse indivíduo, compondo o conjunto de classes utilizado na definição do estado de conformidade. Essa abordagem não contempla a ocorrência de múltiplas pessoas simultâneas no campo de visão da câmera. Nesse cenário, a associação entre EPIs e indivíduos exigiria critérios adicionais de correspondência, não implementados na versão atual da solução.

4.5 INTEGRAÇÃO ENTRE APLICAÇÃO E SISTEMA EMBARCADO

A comunicação entre a aplicação e o módulo embarcado foi implementada por meio de requisições HTTP. Quando há alteração no estado de conformidade, a informação necessária para atualização do atuador visual é enviada ao dispositivo conectado à rede local.

A adoção do HTTP simplificou a integração por dispensar componentes intermediários, como serviços de mensageria ou servidores adicionais. O protocolo também favoreceu a interoperabilidade entre o software em Python e o dispositivo conectado à rede local sem fio, mantendo baixo esforço de implementação.

Após receber a requisição, o dispositivo interpreta o estado e executa a sinalização física correspondente. A troca de informações ocorre em tempo reduzido, permitindo que a resposta visual seja acionada logo após a identificação das condições avaliadas.

4.6 DESENVOLVIMENTO DO SISTEMA EMBARCADO

O módulo físico foi implementado com Raspberry Pi Pico W. A escolha desse dispositivo deve-se à conectividade sem fio integrada, ao baixo custo e à capacidade de acionar saídas digitais suficientes para controlar a sinalização visual.

O comportamento do bastão luminoso foi organizado em três estados:

- a) verde: todos os EPIs exigidos foram identificados em conformidade;
- b) amarelo: estado de espera, sem colaborador sendo processado pelo fluxo;
- c) vermelho: ausência de um ou mais EPIs obrigatórios.

Essa estratégia fornece comunicação visual direta para operadores e colaboradores, sem exigir processamento de imagem ou regras complexas no microcontrolador.

4.7 FLUXO OPERACIONAL DA SOLUÇÃO

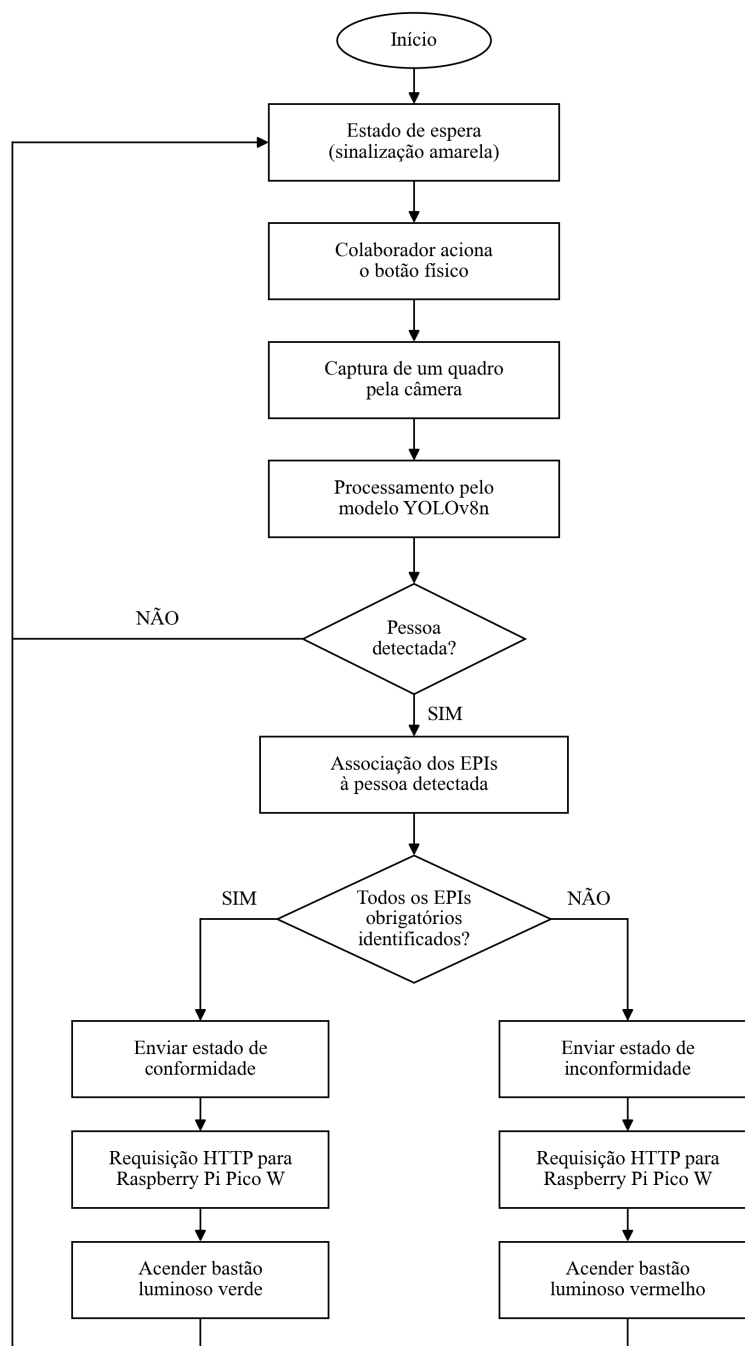
O funcionamento inicia-se quando o colaborador aciona um botão físico ao adentrar o ambiente monitorado, disparando a captura de um quadro pela câmera. Em seguida, a aplicação processa o quadro imediatamente após a captura, executa o YOLOv8n e identifica as classes relacionadas aos EPIs.

Com base nas detecções, o software determina o estado de conformidade e envia a informação por requisição HTTP. Ao receber o estado, o módulo de atuação atualiza a cor do sinalizador, indicando visualmente a condição identificada, e o colaborador segue para a área de produção somente após essa sinalização.

A Figura 4 apresenta o fluxo operacional da solução, desde o estado de espera e

o acionamento do botão físico até a sinalização física e o retorno ao estado de espera.

Figura 4 – Fluxo operacional da solução desenvolvida



Fonte: elaborada pelo autor (2026).

5 RESULTADOS E DISCUSSÃO

Esta seção apresenta os resultados obtidos nos experimentos. A análise contempla o desempenho dos modelos de detecção, a evolução do treinamento, o comportamento do YOLOv8n por classe e o funcionamento integrado entre software e módulo de atuação.

5.1 DESEMPENHO DOS MODELOS DE DETECÇÃO

Os experimentos avaliaram a capacidade dos modelos de detecção em reconhecer automaticamente os EPIs no ambiente industrial. Inicialmente, foram comparadas as arquiteturas YOLOv5n e YOLOv8n, considerando desempenho de detecção e latência média de inferência, medida no mesmo hardware utilizado no treinamento (Tabela 1).

A Tabela 2 apresenta os valores globais de mAP@0.5 e latência média obtidos por cada modelo durante os experimentos.

Tabela 2 – Comparação entre os modelos avaliados

Modelo	mAP@0.5	Latência média (ms)
YOLOv5n	0,785	7,2
YOLOv8n	0,911	8,6

Fonte: elaborada pelo autor (2026).

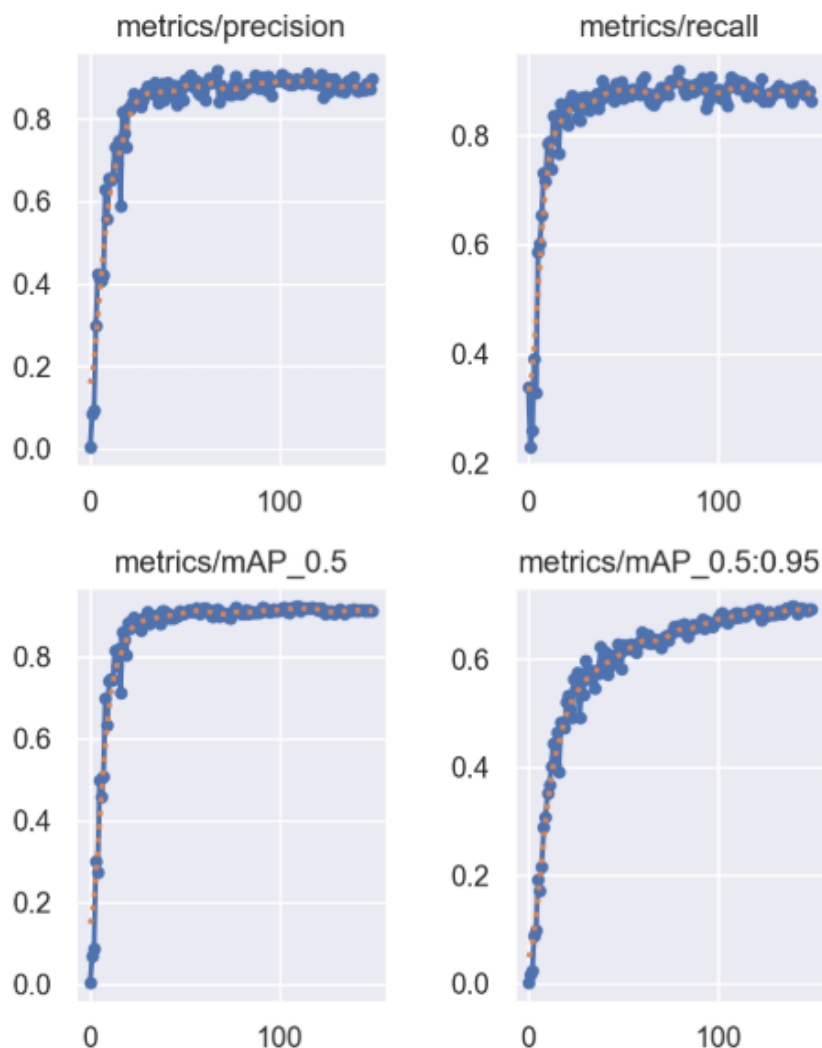
O YOLOv8n apresentou desempenho superior ao YOLOv5n. A diferença de mAP@0.5 indica maior capacidade de localizar e classificar os EPIs nas condições avaliadas. Embora sua latência média tenha sido 1,4 ms maior, esse acréscimo é pequeno diante do fluxo operacional do sistema e não comprometeu a resposta da sinalização. Por esse motivo, o modelo mais recente foi adotado na versão final da aplicação.

A superioridade observada é coerente com o uso de uma arquitetura mais recente e com o fluxo de treinamento da biblioteca Ultralytics, que favorece melhor ajuste em bases com múltiplas classes e objetos de tamanhos distintos. Esse aspecto é relevante porque alguns EPIs, como máscaras e toucas, ocupam regiões pequenas da imagem e são mais sensíveis a variações de pose, iluminação e distância da câmera.

5.2 EVOLUÇÃO DO TREINAMENTO

Além da comparação entre arquiteturas, foi analisada a evolução das principais métricas durante o treinamento do modelo selecionado. A Figura 5 apresenta as curvas de precisão (*precision*), *recall*, mAP@0.5 e mAP@0.5:0.95 registradas ao longo das épocas.

Figura 5 – Evolução das métricas de treinamento do modelo YOLOv8n



Fonte: elaborada pelo autor (2026).

As curvas indicam crescimento acentuado nas primeiras épocas, seguido de estabilização progressiva. Esse comportamento sugere convergência consistente, sem oscilações expressivas ao final do treinamento. A $mAP@0.5$ atingiu patamar próximo ao valor consolidado apresentado na Tabela 2. Já a $mAP@0.5:0.95$ evoluiu de forma mais gradual, comportamento esperado por exigir maior precisão espacial entre as caixas preditas e as anotações de referência.

5.3 AVALIAÇÃO POR CLASSE

Após a seleção do modelo final, foi realizada uma análise do desempenho por classe. A Tabela 3 apresenta os resultados obtidos no conjunto de teste.

Tabela 3 – Desempenho por classe do modelo YOLOv8n

Classe	Precisão	Recall	AP@0.5
Pessoa	0,97	0,96	0,96
Com touca	0,91	0,88	0,89
Sem touca	0,94	0,92	0,92
Com máscara	0,90	0,87	0,88
Sem máscara	0,84	0,81	0,83
Com bota	0,96	0,94	0,95
Sem bota	0,95	0,93	0,94

Fonte: elaborada pelo autor (2026).

Os melhores resultados ocorreram nas classes *pessoa*, *com bota* e *sem bota*. Essas categorias tendem a ocupar regiões maiores da imagem e apresentam contornos mais distinguíveis, o que favorece a detecção. A classe *pessoa*, em particular, apresenta área visual mais ampla e padrões corporais mais estáveis em comparação aos EPIs menores.

A classe *sem máscara* apresentou o menor desempenho relativo. Esse resultado pode estar associado a oclusões parciais, variações de iluminação, diferentes posicionamentos faciais e menor área visual da região analisada. Em imagens capturadas à distância, pequenas mudanças de pose ou sombra podem dificultar a distinção entre ausência de máscara, uso incorreto e baixa visibilidade do rosto. Na prática, esse comportamento indica que a posição da câmera e a qualidade das imagens são fatores críticos para aplicações que dependem da análise de EPIs localizados na região facial.

A Figura 6 apresenta exemplos de detecções realizadas pelo modelo em diferentes condições visuais, incluindo ambientes internos, áreas externas e situações com múltiplas pessoas.

Figura 6 – Exemplos de detecções de EPIs realizadas pelo modelo YOLOv8n



Fonte: elaborada pelo autor (2026), com imagens de registro próprio e das bases públicas do Roboflow Universe (TEST, 2024; EPP Internship, 2025; Detect, 2026).

As amostras mostram que o modelo identificou EPIs em cenários com variação de escala, distância e iluminação. Também aparecem predições com menor confiança, principalmente em classes associadas à máscara e à touca. Esses casos reforçam a influência do tamanho reduzido dos objetos, da oclusão e da qualidade do enquadramento. Assim, embora o desempenho global seja adequado, a aplicação prática exige cuidado na instalação da câmera e na atualização do conjunto de dados quando houver mudanças no ambiente.

5.4 AVALIAÇÃO DA ARQUITETURA INTEGRADA

Além do desempenho do modelo de detecção, foi avaliado o comportamento do sistema completo, considerando captura, inferência, tomada de decisão, comunicação HTTP e atuação embarcada.

Os testes foram realizados em ambiente controlado, utilizando rede local sem fio para comunicação entre a aplicação e o dispositivo de atuação. A análise considerou o tempo entre a captura da imagem e a atualização do sinalizador.

O tempo médio observado entre a captura da imagem e a execução da resposta física foi de aproximadamente 140 ms, com desvio padrão inferior a 12 ms. Esse valor inclui aquisição da imagem, inferência do modelo, envio da requisição HTTP e proces-

samento pelo módulo embarcado. A baixa variação temporal indica comportamento estável nos testes realizados, aspecto importante para uso operacional, pois respostas muito irregulares poderiam dificultar a interpretação da sinalização pelos usuários.

A integração permitiu converter a saída do modelo em sinalização física objetiva. Essa característica reduz a dependência de acompanhamento manual da interface e facilita a percepção das condições de conformidade no ambiente avaliado.

Além dos testes em ambiente controlado, o funcionamento da solução foi verificado em ambiente industrial real. O sistema foi implantado em junho de 2026 e opera diariamente no local desde então, encontrando-se em fase de consolidação do uso pelos colaboradores da empresa. Durante a janela formal de observação instrumentada, de aproximadamente vinte minutos, o fluxo completo de captura, inferência, comunicação HTTP e sinalização funcionou de forma estável e coerente com o comportamento registrado nos testes controlados. Nesse período, não foram observadas falhas de comunicação nem inconsistências na atualização do sinalizador luminoso.

5.5 DISCUSSÃO DOS RESULTADOS

Os resultados demonstram a viabilidade técnica da ferramenta desenvolvida para verificação automática do uso de EPIs em ambiente industrial. O YOLOv8n apresentou desempenho superior ao YOLOv5n no conjunto avaliado e manteve latência compatível com o fluxo operacional do aplicativo.

A análise por classe indica, entretanto, que o desempenho não é uniforme para todos os EPIs. Objetos menores, como máscaras e toucas, são mais afetados por distância da câmera, iluminação, pose e oclusões parciais. Esses fatores explicam a redução observada principalmente na classe *sem máscara* e indicam que o posicionamento da câmera e a diversidade do conjunto de dados são determinantes para a robustez do sistema.

A avaliação integrada confirmou a adequação da separação entre inferência e atuação: o módulo físico permaneceu simples, e o tempo de resposta manteve-se dentro do necessário para a sinalização operacional.

Apesar dos resultados obtidos, a generalização deve ser analisada com cautela. Os testes instrumentados foram conduzidos em cenários controlados e, no ambiente industrial, restringiram-se a uma janela de observação instrumentada, embora o sistema permaneça em operação contínua nesse local. Cabe registrar, ainda, três ressalvas adicionais. O conjunto de teste é composto por aproximadamente 150 imagens, tamanho relativamente reduzido que recomenda cautela na extrapolação das métricas obtidas. A divisão dos dados foi realizada de forma aleatória simples, sem estratificação por origem das imagens. Por fim, as latências reportadas referem-se ao hardware utilizado nos experimentos, não tendo sido caracterizadas no equipamento em operação

na indústria. Variações de iluminação mais intensas, maior circulação simultânea de pessoas, diferentes uniformes, mudanças na posição da câmera e novos tipos de EPIs podem exigir ampliação do conjunto de dados e novo ajuste do modelo.

De modo geral, a ferramenta apresenta potencial de adaptação para outros contextos industriais, desde que sejam revistos o conjunto de classes, as regras de conformidade e as condições de coleta de imagens. Essa necessidade de adaptação é esperada em sistemas de visão computacional aplicados, nos quais o desempenho depende diretamente da representatividade dos dados utilizados no treinamento.

6 CONSIDERAÇÕES FINAIS

Este trabalho apresentou o desenvolvimento e a avaliação de uma ferramenta para monitoramento automático do uso de Equipamentos de Proteção Individual em ambiente industrial. A solução combinou visão computacional, aprendizado profundo, Internet das Coisas (IoT) e sistema embarcado, tendo como cenário de aplicação uma indústria de envase de água mineral.

A arquitetura desenvolvida separou processamento e atuação. As imagens foram analisadas em computador dedicado por meio de uma aplicação em Python, enquanto o módulo embarcado permaneceu responsável pela recepção dos estados e pelo acionamento luminoso. Essa decisão permitiu utilizar o YOLOv8n para inferência visual sem transferir ao microcontrolador uma carga computacional incompatível com sua função de sinalização.

Os objetivos definidos foram alcançados, nos limites descritos na Seção 5. Foi construído um conjunto de dados com imagens representativas do cenário avaliado, foram treinados e comparados modelos da família YOLO, foi selecionado o YOLOv8n como modelo final e foi implementada a integração entre aplicação, comunicação HTTP e atuação embarcada. O sistema também foi avaliado em cenários controlados e verificado em ambiente industrial real, onde permanece em operação contínua.

Os resultados responderam ao problema de pesquisa ao demonstrar que é possível reconhecer automaticamente o uso de EPIs e sinalizar situações de inconformidade com baixa latência no contexto avaliado. O YOLOv8n alcançou mAP@0.5 de 0,911 e apresentou desempenho superior ao YOLOv5n. A integração entre inferência, comunicação HTTP e atuação física apresentou tempo médio de aproximadamente 140 ms entre a captura da imagem e a resposta visual, valor adequado para sinalização operacional das condições analisadas.

Como contribuição científica, o trabalho apresenta uma avaliação aplicada do YOLOv8n para detecção de EPIs em cenário industrial alimentício, discutindo desempenho, limitações e fatores que afetam as classes analisadas. Como contribuição

tecnológica, demonstra a integração entre modelo de detecção, interface local, comunicação HTTP e sinalização física, sem restringir o resultado à visualização em tela.

Entre as limitações, destacam-se a realização de testes majoritariamente em ambiente controlado, a janela restrita de observação instrumentada na indústria (embora o sistema permaneça em operação contínua no ambiente), a sensibilidade de algumas classes à iluminação, oclusões, distância da câmera e tamanho reduzido dos EPIs, e o critério de associação espacial entre EPIs e pessoa, restrito à detecção de um colaborador por vez e não estendido a cenários com múltiplas pessoas simultâneas no campo de visão. Essas restrições não invalidam os resultados, mas indicam que a implantação em ambientes com maior variabilidade operacional exige novas coletas, ajustes no modelo e validação prolongada.

Como trabalhos futuros, recomenda-se ampliar o conjunto de dados, incluir novas condições de iluminação e enquadramento, avaliar outros tipos de EPIs, realizar testes prolongados em diferentes ambientes industriais e integrar a solução a sistemas corporativos de gestão de segurança. Recomenda-se também estender o critério de associação entre EPIs e pessoa para cenários com múltiplas pessoas simultâneas no campo de visão, o que demandaria estratégias de rastreamento individual e critérios de correspondência mais robustos que a simples sobreposição de caixas delimitadoras. Também podem ser investigados registro automático de eventos e modelos de detecção mais recentes, desde que preservada a compatibilidade com os requisitos de latência da aplicação.

Conclui-se que a integração entre detecção visual e atuação embarcada constitui uma alternativa viável para apoiar a avaliação das condições de conformidade no uso de EPIs. No contexto estudado, a ferramenta desenvolvida reconheceu situações de conformidade e inconformidade e as converteu em sinalização física imediata, encerrando a pesquisa com evidências de viabilidade técnica e com limites claramente identificados para sua ampliação.

Por fim, o autor agradece a Deus, pela força, sabedoria e perseverança concedidas ao longo de toda a trajetória acadêmica; à família, em especial aos pais, pelo apoio incondicional, incentivo constante e por acreditarem na importância da educação; aos professores, pelo conhecimento compartilhado, orientação e dedicação ao longo do processo de formação; e ao programa EmbarcaTech, pelo apoio, incentivo e pelas oportunidades de aprendizado que contribuíram significativamente para o desenvolvimento deste projeto.

REFERÊNCIAS

- AL-AZANI, S. *et al.* Real-time monitoring of personal protective equipment compliance in surveillance cameras. **IEEE Access**, v. 12, p. 121882–121895, 2024.
- ALI, L. *et al.* Development of YOLOv5-based real-time smart monitoring system for increasing lab safety awareness in educational institutions. **Sensors**, v. 22, n. 22, p. 8820, 2022.
- BARRO-TORRES, S. *et al.* Real-time personal protective equipment monitoring system. **Computer Communications**, v. 36, n. 1, p. 42–50, 2012.
- BAVARESCO, R. *et al.* Internet of things and occupational well-being in industry 4.0: A systematic mapping study and taxonomy. **Computers & Industrial Engineering**, v. 161, p. 107670, 2021.
- Brasil. Ministério do Trabalho e Previdência. **Norma Regulamentadora nº 6 (NR-6): Equipamento de Proteção Individual (EPI)**. Brasília, DF: [s.n.], 2022. Diário Oficial da União. Disponível em: <<https://www.gov.br/trabalho-e-emprego/pt-br/assuntos/securanca-e-saude-no-trabalho/normas-regulamentadoras>>. Acesso em: 17 nov. 2025.
- CAVALCANTE, J. M. Segurança do trabalho em serviços de alimentação. **Revista Brasileira de Medicina do Trabalho**, v. 21, n. 3, p. e20231060, 2023.
- Detect. **ppe Dataset**. 2026. Roboflow Universe. Licença CC BY 4.0. Disponível em: <<https://universe.roboflow.com/detect-sjigm/ppe-yezzu>>. Acesso em: 17 jul. 2026.
- EPP Internship. **PPE 2 - Reduced Dataset**. 2025. Roboflow Universe. Licença CC BY 4.0. Disponível em: <<https://universe.roboflow.com/epp-internship-0rrmj/ppe-2-reduced>>. Acesso em: 17 jul. 2026.
- GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 4. ed. New York: Pearson, 2018.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge: MIT Press, 2016.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015.
- LIMA, G. G. D. M. *et al.* Um sistema integrado para monitoramento automático do uso de EPIs na indústria alimentícia com detecção visual e atuação embarcada em tempo real. In: **Anais Estendidos do XXII Simpósio Brasileiro de Sistemas de Informação (SBSI) – Trilha de Indústria e Inovação em SI**. Porto Alegre: Sociedade Brasileira de Computação (SBC), 2026. p. 86–90.
- LO, J.-H.; LIN, L.-K.; HUNG, C.-C. Real-time personal protective equipment compliance detection based on deep learning algorithm. **Sustainability**, v. 15, n. 1, p. 391, 2023.

PROTIK, A. A.; RAFI, A. H.; SIDDIQUE, S. Real-time personal protective equipment (PPE) detection using YOLOv4 and TensorFlow. In: **2021 IEEE Region 10 Symposium (TENSYP)**. Jeju: IEEE, 2021. p. 1–6.

REDMON, J. *et al.* You only look once: Unified, real-time object detection. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Las Vegas: IEEE, 2016. p. 779–788.

REDMON, J.; FARHADI, A. YOLOv3: An incremental improvement. **arXiv preprint arXiv:1804.02767**, 2018.

TEST. **Hairnet Dataset**. 2024. Roboflow Universe. Licença Public Domain. Disponível em: <<https://universe.roboflow.com/test-icuyg/hairnet-v6o20>>. Acesso em: 17 jul. 2026.

Ultralytics. **YOLOv8: Real-Time Object Detection**. 2023. Repositório GitHub. Disponível em: <<https://github.com/ultralytics/ultralytics>>. Acesso em: 02 fev. 2026.

YANG, X. *et al.* Automated PPE-tool pair check system for construction safety using smart IoT. **Journal of Building Engineering**, v. 32, p. 101721, 2020.

ZHOU, H. A. *et al.* Generative AI in industrial machine vision: A review. **Journal of Intelligent Manufacturing**, v. 37, p. 1447–1470, 2026.

ZORZENON, R.; LIZARELLI, F. L.; MOURA, D. B. A. A. What is the potential impact of industry 4.0 on health and safety at work? **Safety Science**, v. 153, p. 105802, 2022.