

MONITORAMENTO DO USO DO CAPACETE DE SEGURANÇA POR VISÃO COMPUTACIONAL

MONITORING SAFETY HELMET USE BY COMPUTER VISION

RODRIGUES.J { juliano.sales@fatec.sp.gov.br }

SILVA.M { maria.silva545@fatec.sp.gov.br }

RESUMO

Este artigo apresenta um sistema de monitoramento baseado em visão computacional e redes neurais, voltado à prevenção de acidentes de trabalho causados pela ausência do capacete de segurança em ambientes laborais. O sistema integra câmeras de segurança (DVR) em tempo real a uma arquitetura de visão computacional YOLO, treinada a partir de uma rede neural convolucional (CNN), e a uma plataforma web que disponibiliza os dados capturados de forma organizada, auxiliando o responsável pela segurança do trabalho. Os resultados obtidos demonstraram um desempenho satisfatório na detecção do uso e da ausência de capacetes, apresentando potencial positivo para aplicações em monitoramento de segurança industrial com uma taxa de identificação de 100% em ambientes iluminados e 92.2% para capacetes com cores variadas. Entretanto, verificou-se a necessidade de aprimoramento do modelo em ambientes com baixa luminosidade, onde a detecção apresentou limitações com uma taxa de identificação de 36.3%.

PALAVRAS-CHAVE: Equipamento de Segurança; Segurança e Saúde do Trabalho; Deep Learning; Visão Computacional; Redes Neurais.

ABSTRACT

This paper presents a monitoring system based on computer vision and neural networks, aimed at preventing workplace accidents caused by the absence of safety helmets in

workplaces. The system integrates real-time security cameras (DVRs) with a YOLO computer vision architecture, trained using a convolutional neural network (CNN), and a web platform that makes the captured data available in an organized manner, assisting those responsible for workplace safety. The results obtained demonstrated satisfactory performance in detecting the use and absence of helmets, showing positive potential for industrial safety monitoring applications, with a 100% identification rate in well-lit environments and 92.2% for helmets of varying colors. However, the model needed to be improved in low-light environments, where detection showed limitations, with an identification rate of 36.3%.

KEYWORDS: Safety Equipment; Occupational Health and Safety; Deep Learning; Computer Vision; Neural Networks.

INTRODUÇÃO

Em 2024, o Brasil registrou aproximadamente 180.292 concessões de Benefícios Previdenciários Acidentários (B91), das quais 52.333 ocorreram no estado de São Paulo. Nesse mesmo estado, foram concedidas 1.027 aposentadorias por invalidez acidentária (B92) e 15.386 auxílios-acidente (B94) decorrentes de acidentes de trabalho.

No município de Registro (SP), foram notificadas 181 ocorrências de acidentes de trabalho, enquanto em Cajati (SP) foram registradas 99 notificações e um óbito, conforme dados da Previdência Social, com base no tratamento e análise do SmartLab (2023).

Esses números demonstram a urgência de ações voltadas à saúde e segurança do trabalho, principalmente em atividades de maior risco. Entre as estratégias, destaca-se o uso adequado de Equipamentos de Proteção Individual (EPI), sendo o capacete de segurança um dos equipamentos essenciais para a proteção contra impactos e quedas de objetos. Segundo Zocchio, o uso de EPI's evita lesões ou atenua sua gravidade, reduzindo ou impedindo o contato direto entre o agente causador e a cabeça do trabalhador.

Com o objetivo de investigar a conformidade e os desafios da SST, realizou-se uma pesquisa de campo voltada exclusivamente ao uso do capacete de segurança na empresa Pirâmide, localizada em Registro, no estado de São Paulo, especializada na extração de areia. A escolha se justifica pela natureza da atividade, que envolve operação de maquinários pesados em ambientes de alto risco, nos quais o uso do capacete é obrigatório. Atualmente, a organização conta com cerca de 140 colaboradores, distribuídos em aproximadamente 40 por unidade, atuando em seis setores distintos, todos classificados como de grau de risco 4, o que torna o uso constante do capacete indispensável em todas as áreas operacionais.

Embora o gestor entrevistado não registre acidentes diretamente ligados à ausência do capacete, a área da usina foi identificada como o ponto de maior risco devido à presença de estruturas suspensas, o que eleva a probabilidade de incidentes em caso de negligência.

O levantamento revelou uma divergência clara entre a percepção formal e a prática: a maioria dos colaboradores afirmou nunca ter sido advertido pela falta do capacete,

entretanto, a resposta predominante sobre ter avistado colegas sem o equipamento foi "Às vezes". Este dado sugere a ocorrência de desvios ou descumprimento ocasional da norma.

Diante da persistência de falhas humanas na fiscalização e no cumprimento das normas de segurança, este estudo propõe o desenvolvimento e a aplicação de um sistema de monitoramento automatizado para a verificação do uso do capacete em áreas cobertas. A solução utiliza um DVR (Digital Video Recorder), tecnologia de vigilância comumente empregada, integrada à análise em tempo real por meio de técnicas de Aprendizagem Profunda (Deep Learning).

Esta pesquisa busca contribuir para o fortalecimento das práticas de segurança do trabalho, automatizando a identificação de riscos, reduzindo a subjetividade na fiscalização e prevenindo acidentes, de forma a promover um ambiente laboral mais seguro e saudável.

OBJETIVO

Este projeto tem como objetivo auxiliar o responsável pela saúde e segurança do trabalho na vigilância contínua do ambiente, por meio do desenvolvimento de um sistema de monitoramento, utilizando técnicas de Visão Computacional e Rede Neural Convolucional, com capacidade para identificar a ausência do uso do capacete de segurança em ambiente laboral, onde o uso desse EPI é indispensável.

OBJETIVOS ESPECÍFICOS

- a) Aplicar técnicas de Visão Computacional junto a Rede Neural Convolucional para identificar a ausência do capacete de segurança.
- b) Integrar o sistema DVR à plataforma web para processamento, armazenamento e análise dos registros.

ESTADO DA ARTE

Nesta seção, serão explorados trabalhos correlatos para identificar a atual fronteira científica na área. Para isso, foram selecionados artigos recentes que, de forma semelhante a este projeto, estuda a aplicação da Visão Computacional como meio de prevenir acidentes ocupacionais e garantir a integridade dos colaboradores.

O estudo de (DELHI; SANKARLAL; THOMAS, 2020) utiliza locais de construção como área de pesquisa e tem como principal objetivo detectar a adesão dos colaboradores a dois EPIs essenciais: o capacete de segurança e o colete reflexivo. Utilizando técnicas de Aprendizagem Profunda (Deep Learning) baseadas em Visão Computacional, os autores desenvolveram um algoritmo capaz de classificar as imagens de acordo com o nível de segurança. Para isso, empregaram Redes Neurais Convolucionais (CNNs), que são redes neurais artificiais altamente eficazes no processamento de dados estruturados em grade, permitindo a classificação em categorias como "seguro", "não seguro", "sem capacete" ou "sem jaqueta".

O projeto de (Balakreshnan et al., 2020), por sua vez, foca na detecção do uso de óculos de

proteção em ambientes laboratoriais. Com a finalidade de obter um sistema rápido e de baixo custo de infraestrutura, os autores adotaram uma câmera inteligente de 9 megapixels operando em Linux. Essa arquitetura permitiu a conexão com o Microsoft Azure, criando um sistema híbrido. Tal sistema une a captura local da câmera à plataforma de serviços em nuvem utilizando dispositivos de Internet das Coisas (IoT), por meio do IoT Hub da Azure. Essa abordagem demonstra uma estratégia eficiente para evitar o processamento pesado na própria câmera.

Em uma abordagem mais ampla, (Liu et al., 2021) discute a relação entre a Inteligência Artificial (IA) e a segurança do trabalho. Diferentemente dos trabalhos anteriores, o artigo apresenta uma pesquisa que comenta e explica quais tecnologias são mais adaptáveis à tarefa de análise de comportamentos perigosos na construção civil. No contexto específico da utilização de EPIs, o autor descreve que o processo de Aprendizagem de Máquina (Machine Learning) pode ser dividido em quatro etapas: detecção do corpo humano, detecção do EPI, reconhecimento da pessoa equipada e, por fim, a avaliação da eficácia e precisão do sistema. Além da Aprendizagem de Máquina, o artigo faz referência à Aprendizagem Profunda (Deep Learning), que possibilita o desenvolvimento de sistemas capazes de aprender de forma autônoma e executar tarefas análogas às capacidades cognitivas humanas.

A partir da síntese entre a área científica apresentada nos artigos e os resultados da pesquisa de campo, observa-se a aplicabilidade da visão computacional utilizando a aprendizagem profunda e Redes Neurais Convolucionais para colaborar com a Segurança e Saúde do Trabalho, de forma eficiente e com baixo custo.

METODOLOGIA

A metodologia dessa pesquisa é baseada no desenvolvimento e na aplicação de uma plataforma focada na detecção do uso do capacete de segurança. A arquitetura do sistema compreende dois componentes principais: o módulo de análise e uma Plataforma Web para gestão.

Visão Computacional

A visão computacional foi testada para identificar se o capacete está corretamente posicionado na cabeça do colaborador. O treinamento foi realizado utilizando a YOLO e a CNN. A YOLO é um modelo de detecção de objetos em tempo real, ela divide a imagem em uma rede de grade $S \times S$, onde em cada uma dessas redes são retiradas caixas delimitadoras (REDMON et al., 2016) verificando, em cada espaço, se há um objeto classificado. Um algoritmo de perda (*loss function*) compara se a imagem analisada corresponde ao rótulo correto, ajustando o modelo para reduzir os erros. A CNN é a rede neural convolucional responsável por reconhecer os aspectos (brilho, cor, tamanho, entre outros) e aprender os padrões presentes nas imagens (LI et al., 2021). As características identificadas pela CNN são enviadas às camadas finais da YOLO, permitindo que o modelo verifique se a classificação existe em cada espaço da grade.

O pré-processamento inicial envolveu a extração de frames de vídeos com taxa de 1 frame por segundo e a normalização dos valores de pixel que variam entre 0 e 1. Para o treinamento do modelo, foi construído um dataset com 531 imagens anotadas na plataforma

Roboflow, distribuídas em conjuntos de treino (371 imagens), validação (107 imagens) e teste (53 imagens).

O modelo desenvolvido processa as imagens aplicando múltiplas camadas convolucionais, cada uma seguida por uma camada de *max pooling*. Inicialmente, a imagem passa por camadas convolucionais com 32 filtros de tamanho 3x3, que extraem características básicas, como bordas e texturas, ativadas pela função ReLU (Rectified Linear Unit). A ReLU contribui para a eficiência, pois desativa neurônios cuja saída da transformação linear é zero (SHARMA; SHARMA, 2020). Em seguida, ocorre a primeira etapa de *max pooling* (2x2). O *max pooling* ocorre a partir da definição de uma camada de subamostragem (pooling) e uma área AxA, qual percorre toda a saída da camada anterior resumindo a informação daquela área em um único valor. O *max pooling* é o método de ativação máxima, onde pega-se apenas o valor máximo dentro da vizinhança. (PINHEIRO JUNIOR, 2024)

Na sequência, a imagem é processada por camadas convolucionais intermediárias com 64 filtros de 3x3, extraíndo características profundas, seguida por uma nova etapa de *max pooling* (2x2). Posteriormente, o fluxo passa pelas camadas convolucionais avançadas, que empregam 128 filtros. Estas camadas são capazes de identificar características de nível mais alto, essenciais para a detecção precisa de capacetes. Para aumentar a acurácia na localização, o modelo também incorpora mecanismos de atenção (ou Attention Mechanism), que direcionam o foco da rede neural para as áreas da imagem que são mais relevantes para a tarefa de detecção

Na camada de saída, é utilizada a função de ativação *sigmoid* para a detecção e classificação, gerando *bounding boxes* com coordenadas e níveis de confiança para cada capacete identificado. Após essa etapa, aplica-se a supressão não máxima (*Non-Maximum Suppression – NMS*), que elimina detecções redundantes, e os resultados são armazenados em um banco de dados MongoDB por meio da biblioteca PyMongo, garantindo a persistência dos dados.

Por fim, o modelo entrega as previsões em tempo real, utilizando-se um conjunto de métricas de desempenho. As métricas são coletadas ao longo das épocas, que representam o número de vezes que o algoritmo percorre todo o conjunto de dados de treinamento. As principais métricas utilizadas são: Recall (Revocação), Precision (Precisão), F1-Score e mAP (mean Average Precision). O Recall é fundamental, pois mede sua capacidade de encontrar todos os objetos relevantes do modelo, resultando na taxa de detecção. A Precision demonstra a confiabilidade das detecções e está intimamente ligada à confiança média das previsões, mostrando o quanto o modelo acerta quando faz uma detecção. O F1-Score oferece um equilíbrio entre a Precision e o Recall. O mAP combina a precisão média com o Intersection over Union (IoU). A base para todas essas avaliações é o Intersection over Union (IoU), também conhecido como coeficiente de similaridade de Jaccard, uma estatística para estimar a similaridade entre a predição e a área real (Von Wangenheim, 2022).

Em testes práticos, o sistema demonstrou alta eficácia, processando um vídeo com 142 quadros e obtendo detecção em 100% deles, com confiança média de 0,84, validando a abordagem proposta para o monitoramento de segurança em ambientes industriais.

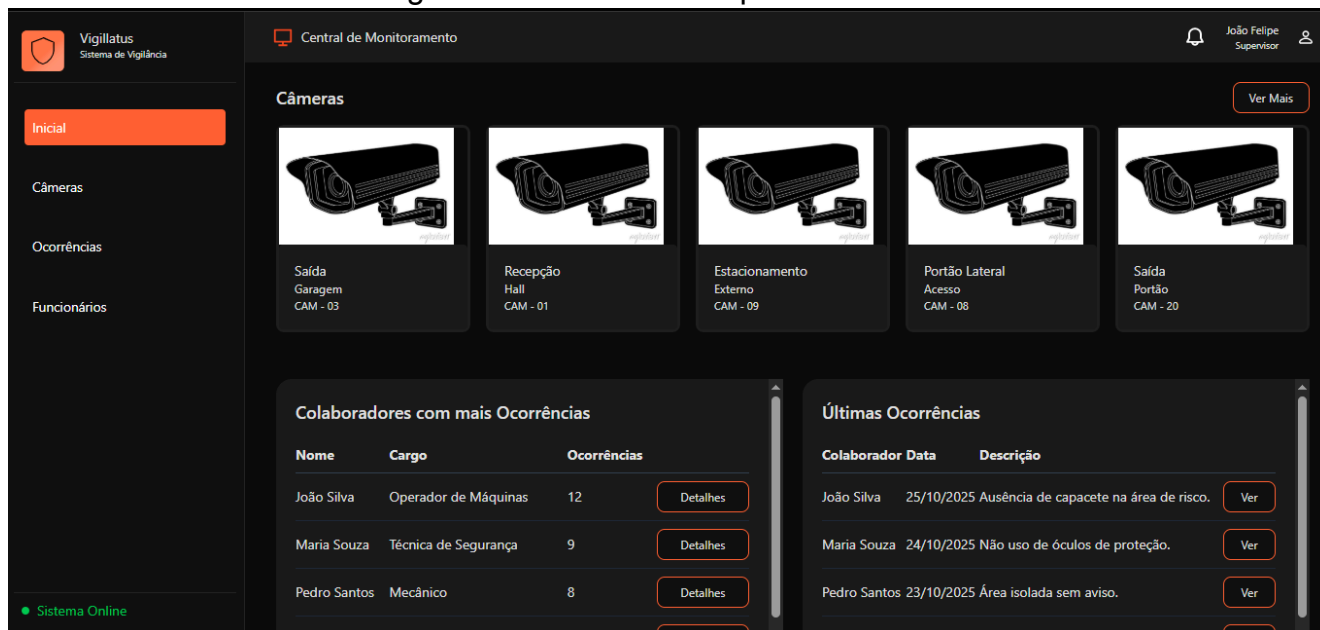
Aplicação Web

A fim de facilitar a visualização dos dados captados pela aplicação, desenvolveu-se uma

plataforma Web que apresenta as informações relevantes de maneira organizada ao gestor. A interface foi desenvolvida utilizando React, e os dados exibidos são enviados por meio de uma API em Node.js, que centraliza a comunicação entre o front-end, a base de dados e as rotas do sistema.

A figura 1 apresenta a interface principal do sistema, onde é possível observar as câmeras em tempo real, além de verificar as últimas análises feitas pelo sistema, quais são distribuídas nas listas: últimas ocorrências, número de ocorrências por colaborador.

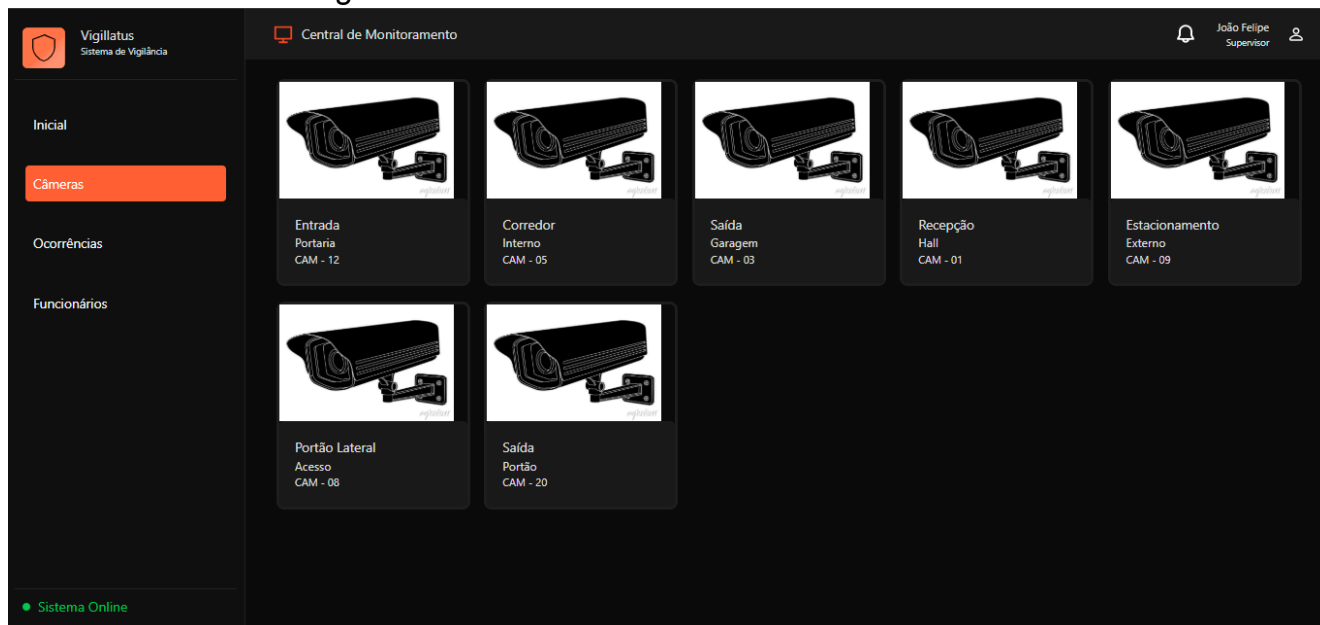
Figura 1 - Interface Principal do Sistema



Fonte: Autoria Própria (2025)

A Figura 2 apresenta a interface de monitoramento das câmeras do sistema. Nela, é possível visualizar todas as câmeras interligadas, as quais já integram o módulo de Visão Computacional.

Figura 2 - Interface de Monitoramento do Sistema



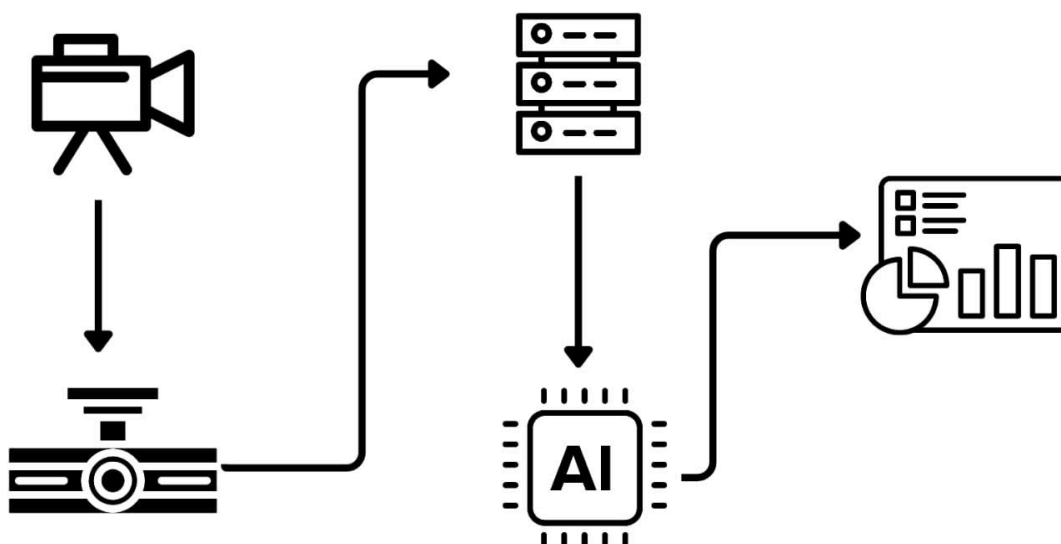
Fonte: Autoria Própria (2025)

Comunicação de Redes

A estrutura de comunicação tem início com as imagens capturadas pelo sistema DVR, que são enviadas ao servidor. Em seguida, a visão computacional analisa essas imagens e, ao detectar a ausência de capacete, realiza o envio das informações ao sistema web por meio de requisições da API. Esse processo permite que os dados sejam gerados e as informações sejam apresentadas de forma organizada na plataforma.

A figura 3 ilustra o fluxo de maneira a entender a comunicação do sistema.

Figura 3 - Fluxo de Redes



Fonte: Autoria Própria (2025)

RESULTADOS PRELIMINARES

Os testes realizados buscaram avaliar o desempenho do modelo de visão computacional em diferentes condições de iluminação e variação de cores dos capacetes. O treinamento foi considerado concluído ao atingir 80% de desempenho nas épocas, momento em que o modelo apresentou resultados satisfatórios.

Nos testes realizados em ambientes com pouca iluminação, foram analisados quatro vídeos, totalizando 1.589 frames. O modelo detectou 577 frames, alcançando uma taxa de detecção média de 36,3% e confiança média de 0,758. Observou-se que vídeos com menor visibilidade apresentaram pior desempenho, indicando que a ausência de luz adequada dificulta a identificação dos capacetes.

Em condições de alta iluminação, também foram testados quatro vídeos, com um total de

1.734 frames. Nessa situação, o modelo obteve 100% de taxa de detecção e confiança média de 0,829, demonstrando excelente desempenho e maior estabilidade nos resultados, uma vez que níveis de iluminação satisfatórios favoreceram a análise das imagens.

Por fim, na avaliação de capacetes com diferentes cores, foram processados quatro vídeos, somando 1.280 frames. O modelo alcançou 92,2% de taxa de detecção e confiança média de 0,838, mostrando boa capacidade de adaptação às variações cromáticas dos objetos. Ainda assim, observou-se leve variação no desempenho conforme a tonalidade do capacete, especialmente em cores menos contrastantes com o ambiente, como o cinza e o laranja.

Tabela 1 - Resultados apresentados por condição de teste

Condição do Teste	Total de Frames	Frames com Detecção	Taxa de Detecção (%)	Confiança Média
Baixa Iluminação	1.589	577	36,3%	0,758
Alta Iluminação	1.734	1.734	100,0%	0,829
Capacetes com cores diferentes	1.280	1.180	92,2%	0,838

De modo geral, os resultados preliminares indicam que o modelo apresenta alto desempenho em ambientes iluminados e em capacetes de cores variadas, mas reduz sua eficiência em situações de baixa visibilidade.

CONCLUSÃO

Diante do grande número de acidentes de trabalho, o presente estudo demonstra o potencial da aplicação da visão computacional em ambientes laborais, contribuindo para o monitoramento e a prevenção de riscos. A integração das redes neurais com a visão computacional apresentou resultados positivos na identificação correta do uso e da falta de capacetes, oferecendo uma ferramenta que pode auxiliar de forma otimizada os responsáveis pela segurança do trabalho.

Entretanto, observou-se a necessidade de aprimoramento do modelo em condições de baixa iluminação, uma vez que a maioria dos ambientes industriais apresenta variações de luminosidade que afetam a precisão da detecção. Assim, como continuidade desta pesquisa, serão realizados novos testes em ambientes com baixa iluminação, bem como o aperfeiçoamento do sistema de visão computacional, com o objetivo de tornar a solução mais robusta e eficaz em diferentes contextos de operação.

REFERÊNCIAS

BALAKRESHNAN, Balamurugan et al. PPE Compliance Detection using Artificial Intelligence in Learning Factories. *Procedia Manufacturing*, v. 45, p. 277–282, 2020. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2351978920310556>.

Acesso em: [Inserir data de acesso]. DOI: <https://doi.org/10.1016/j.promfg.2020.04.017>.

BISHOP, C.M. Pattern Recognition and Machine Learning. [S. l.]: Springer, 2006.

BOOCH, G.; RUMBAUGH, J.; JACOBSON, I. The Unified Modeling Language User Guide. [S. l.]: Addison-Wesley, 1999.

CANDIDO, Carlos Henrique; SANTOS MELLO, Ronaldo dos. Ferramenta de Modelagem de Banco de Dados Relacionais brModelo v3. In: ESCOLA REGIONAL DE BANCO DE DADOS (ERBD), 13., 2017. Anais... [S. l.]: ERBD, 2017.

CARVALHO, V. MySQL: Comece com o principal banco de dados open source do mercado. [S. l.]: Casa do Código, 2015.

DELHI, Venkata Santosh Kumar; SANKARLAL, R.; THOMAS, Albert. Detecção da conformidade dos equipamentos de proteção individual (EPI) em canteiros de obras utilizando técnicas de aprendizado profundo baseadas em visão computacional. *Frontiers in Built Environment*, v. 6, p. 136, 2020. Disponível em: <https://www.frontiersin.org/journals/built-environment/articles/10.3389/fbuil.2020.00136/full>. Acesso em: [Inserir data de acesso].

EBAC. O que é figma e como usar. [S. l.: s. n.], 2023. Disponível em: [Inserir link, se houver, pois o original não o continha]. Acesso em: [Inserir data de acesso].

FORBES. The Evolution Of Technology In Outsmarting Industrial Safety Hazards. [S. l.: s. n.], 2022. Disponível em: <https://www.forbes.com/sites/forbestechcouncil/2022/06/17/the-evolution-o-f-technology-in-outsmarting-industrial-safety-hazards/?sh=2e5bcc374fb3>. Acesso em: 12 de setembro 2023.

GARCIA-MOLINA, H.; ULLMAN, J.D.; WIDOM, J. Database Systems: The Complete Book. [S. l.]: Pearson Prentice Hall, 2009.

GOVERNO FEDERAL DO BRASIL. Consolidação das Leis do Trabalho. [S. l.: s. n.], 1943. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto-lei/del5452.htm. Acesso em: 12 de setembro 2023.

HAYERBEKE, M. Eloquent JavaScript: A Modern Introduction to Programming. [S. l.]: No Starch Press, 2011.

HORTON, John et al. Workplace safety futures: the impact of emerging technologies and platforms on work health and safety and workers' compensation over the next 20 years. Canberra, ACT, Australia: CSIRO, 2018.

IBERDROLA. Deep learning': um conceito-chave para levar a inteligência artificial a um nível superior. [S. l.: s. n.]. Disponível em: <https://www.iberdrola.com/inovacao/deep-learning>. Acesso em: 12 de setembro 2023.

KHAN, S. et al. A Guide to Convolutional Neural Networks for Computer Vision. [S. l.]: Morgan & Claypool Publishers, 2018.

LI, Z. et al. A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems, IEEE, 2021.

LIU, Wenyao et al. Applications of Computer Vision in Monitoring the Unsafe Behavior of Construction Workers: Current Status and Challenges. Buildings, v. 11, n. 9, 2021. Disponível em: <https://www.mdpi.com/2075-5309/11/9/409>. Acesso em: [Inserir data de acesso].

LOELIGER, J.; MCCULLOUGH, M. Version Control with Git: Powerful tools and techniques for collaborative software development. [S. l.]: O'Reilly Media, 2012.

MATTHES, E. Python Crash Course, 3rd Edition: A Hands-On, Project-Based Introduction to Programming. [S. l.]: No Starch Press, 2023.

MPT. Mortalidade no trabalho cresce em 2022 e acidentes notificados ao SUS batem recorde. [S. l.: s. n.], 2023. Disponível em: <https://www.prt20.mpt.mp.br/informe-se/noticias-do-mpt-se/1130-mortalidade-no-trabalho-cresce-em-2022-e-acidentes-notificados-ao-sus-batem-recorde>. Acesso em: 12 de setembro 2023.

ONU. Sobre o nosso trabalho para alcançar os Objetivos de Desenvolvimento Sustentável no Brasil. [S. l.: s. n.], 2015. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 8 de setembro 2023.

PILONE, D.; PITMAN, N. UML 2.0 in a Nutshell. [S. l.]: O'Reilly Media, 2005.

PINHEIRO JUNIOR, Uemerson. Registro não-rígido de imagens médicas usando Block-Based Principal Component Analysis como camada de pooling. 2023. Dissertação (Mestrado em Ciência da Computação) – Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Rio Claro, 2023.

PISHGAR, Maryam et al. REDECA: a novel framework to review artificial intelligence and its applications in occupational safety and health. International journal of environmental research and public health, MDPI, v. 18, n. 13, p. 6705, 2021.

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. [S.l.: s.n.], 2016. p. 779–788.

ROBSON, E.; FREEMAN, E. Head First HTML and CSS: A Learner's Guide to Creating Standards-Based Web Pages. [S. l.]: O'Reilly Media, 2012.

RUSSEL, S.; NORVIG, P. Artificial Intelligence: A Modern Approach. [S. l.]: Pearson, 2016.

SHARMA, S.; SHARMA, S. Activation functions in neural networks. International Journal of

Engineering Applied Sciences and Technology, v. 4, n. 12, p. 310–316, 2020.

SZELISKI, R. Computer Vision: Algorithms and Applications (Texts in Computer Science. [S. l.]: Springer, 2010