

COMPARAÇÃO DE EFICIÊNCIA NA IDENTIFICAÇÃO DE EPIs UTILIZANDO DEEP LEARNING PADRÃO E TRANSFER LEARNING

Nicolas Kaneta Gasparin¹ e Wesley Candido Silva²

RESUMO

Este trabalho visa comparar o desempenho entre dois métodos de aprendizado profundo aplicados à identificação automática do uso de Equipamentos de Proteção Individual (EPIs) em ambientes industriais: o deep learning convencional e o transfer learning. Inicialmente, são abordados os conceitos fundamentais de visão computacional e redes neurais convolucionais, que fundamentam o desenvolvimento dos modelos propostos. A metodologia consiste na utilização de conjuntos de imagens categorizadas entre trabalhadores com e sem EPIs, sendo os modelos treinados com parâmetros padronizados e avaliados por meio de métricas como tempo de processamento, consumo de CPU, acurácia, função de perda, matriz de confusão e curva ROC. Diante dos testes realizados ao longo do estudo, mostrou-se que o método utilizando *transfer learning* apresentou um melhor desempenho em relação ao *deep learning* padrão.

Palavras-chave: equipamentos de proteção individuais; aprendizado profundo; industria; comparação; detecção.

EFFICIENCY COMPARISON IN PPE IDENTIFICATION USING STANDARD DEEP LEARNING AND TRANSFER LEARNING

ABSTRACT

This study aims to compare the performance of two deep learning approaches applied to the automatic identification of Personal Protective Equipment (PPE) usage in industrial environments: conventional deep learning and transfer learning. Initially, fundamental concepts of computer vision and convolutional neural networks are

¹ Graduação, Engenharia Elétrica, da Unisenai Londrina. E-mail: nicolas.gasparin00185603@sesisenaipr.org.br

² Especialista - Engenharia de Automação Industrial, da Faculdade de Tecnologia do Senai Santa Catarina. E-mail: wesley.candido@sistemafiep.org.br

discussed, which support the development of the proposed models. The methodology involves the use of image datasets categorized between workers with and without PPE, with models trained using standardized parameters and evaluated through metrics such as processing time, CPU usage, accuracy, loss function, confusion matrix, and ROC curve. Based on the tests carried out throughout the study, the transfer learning method demonstrated better performance compared to the standard deep learning approach.

Key words: personal protective equipment; deep learning; industry; comparison; detection.

1 INTRODUÇÃO

Os Equipamentos de Proteção Individual (EPIs), são acessórios, utensílios e equipamentos utilizados por trabalhadores com o intuito de proteger sua saúde contra riscos que possam surgir no ambiente de trabalho, prevenindo diversos tipos de acidentes e doenças.

De acordo com a NR-6 (Norma Regulamentadora nº 6), que trata do uso de Equipamentos de Proteção Individual (EPIs) no ambiente de trabalho, é estabelecido que o empregador é obrigado a fornecer gratuitamente aos trabalhadores os EPIs adequados aos riscos presentes na atividade, em perfeito estado de conservação e funcionamento. A norma também determina que o trabalhador deve usar corretamente, conservar e comunicar irregularidades nos equipamentos. Além disso, a NR-6 exige que os EPIs possuam Certificado de Aprovação (CA) emitido pelo Ministério do Trabalho, garantindo sua eficácia e qualidade. Caso contrário, podem causar sérios problemas, tanto físicos para o colaborador quanto judiciais e financeiros para a empresa (DOMINGUES, 2024).

Quando pensamos na identificação de pessoas e/ou objetos, logo vem à mente a utilização da inteligência artificial aplicada em câmeras, permitindo que o contratante não precise monitorar constantemente sua empresa, tornando o ambiente industrial cada vez mais seguro.

Posteriormente, surgem os estudos sobre o chamado Machine Learning, ou Aprendizado de Máquina, que utiliza algoritmos e bancos de dados previamente estabelecidos para analisá-los e aprender com eles, sendo capaz de prever ou decidir ações futuras. No entanto, caso os dados inseridos para o treinamento não sejam suficientemente precisos, o sistema pode apresentar um filtro muito amplo, gerando resultados insatisfatórios (AWARI, 2023).

Dessa forma, iniciam-se os estudos relacionados ao *Deep Learning*, ou Aprendizado Profundo, que, diferentemente do Machine Learning, utiliza centenas ou milhares de camadas e subcamadas em seu processamento, contra uma ou duas do seu antecessor, podendo assim, gerar resultados mais satisfatórios quando submetido às tarefas para as quais foi designado (HOLDSWORTH e SCAPICCHIO, 2024).

Um grande aliado do *Deep Learning* é o chamado *Transfer Learning*, ou Aprendizado por Transferência, que, quando utilizado corretamente, permite o reaproveitamento de dados de processamentos anteriores como base para a criação de novas tarefas, o que se encaixa nas necessidades do uso do *Deep Learning* (MUREL e KAVLAKOGLU, 2024). A literatura recente entre 2020 e 2025 ressalta que esse paradigma é especialmente eficaz em cenários com escassez de dados rotulados ou quando há variação entre domínios. Por exemplo, *surveys* como o de Peng Yan et al. (2023) demonstram que, em ambientes industriais com séries temporais — como manutenção preditiva e monitoramento de processos — o *Transfer Learning* permite reaproveitar modelos CNN/LSTM pré-treinados para novas máquinas ou condições, reduzindo drasticamente a necessidade de rotulação e *retrain* completo. Revisões focadas em visões gerais do campo (2024) também apontam que TL diminui o tempo de treinamento, além de mitigar o problema de *catastrophic forgetting* por meio de técnicas como congelamento seletivo de camadas.

Diante das pesquisas supracitadas, faz-se necessário o estudo para identificar qual método seria mais eficiente na identificação desses EPIs, conciliando velocidade de processamento com menor demanda de hardware, por meio da realização de testes e comparativos, os quais serão explicados e exemplificados nos tópicos a seguir.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Inteligência Artificial e Aprendizado de Máquina

A Inteligência Artificial (IA) é uma área multidisciplinar da ciência da computação que visa criar sistemas capazes de realizar tarefas que normalmente exigiriam inteligência humana, como percepção visual, reconhecimento de fala, tomada de decisão e tradução entre idiomas, sendo muito utilizada também para tarefas repetitivas. O termo "inteligência artificial" foi formalmente cunhado em 1956 na conferência de Dartmouth por John McCarthy e outros pesquisadores. Desde então, a IA passou por diferentes fases de entusiasmo e ceticismo, conhecidas como "verões" e "invernos" da IA até alcançar um novo auge nos últimos anos graças ao aumento exponencial do poder de processamento, à disponibilidade de grandes volumes de dados (big data) e aos avanços nos algoritmos de aprendizado.

A IA pode ser dividida em três categorias principais. A IA Estreita (ANI - Artificial Narrow Intelligence) é focada em uma tarefa específica, como assistentes virtuais, mecanismos de recomendação e carros autônomos. Já a IA Geral (AGI - Artificial General Intelligence) objetiva replicar a inteligência humana em múltiplas tarefas cognitivas, embora ainda esteja em estágio teórico. Por fim, a IA Superinteligente (ASI - Artificial Super Intelligence) refere-se a uma inteligência hipotética que supera amplamente a capacidade humana (RUSSELL; NORVIG, 2021).

O Aprendizado de Máquina (Machine Learning – ML) é um subconjunto da IA que permite que sistemas aprendam a partir de dados, identificando padrões e tomando decisões com mínima intervenção humana. Entre as principais abordagens de aprendizado de máquina estão o aprendizado supervisionado, em que os algoritmos aprendem com dados rotulados, como nos casos de regressão linear e árvores de decisão; o aprendizado não supervisionado, que busca padrões em dados não rotulados, como em técnicas de agrupamento (clustering), a análise de componentes principais (PCA); e o aprendizado por reforço, baseado em recompensas e punições, com aplicações em jogos e robótica (GERÓN, 2019).

2.2 Deep Learning

Deep Learning (DL) refere-se a um ramo do aprendizado de máquina que utiliza redes neurais com muitas camadas ocultas, conhecidas como redes neurais profundas (deep neural networks). Essas redes são capazes de extrair representações hierárquicas dos dados brutos, sendo particularmente eficazes em tarefas como reconhecimento de voz, visão computacional e tradução automática (CHOLLET, 2021). Conforme apontado por Chollet (2021), uma das principais vantagens do *deep learning* é a sua habilidade de aprender automaticamente características relevantes dos dados, sem a necessidade de intervenção humana direta na definição de atributos.

As arquiteturas mais comuns utilizadas em DL incluem as redes neurais convolucionais (CNNs), que são aplicadas a dados estruturados em grade, como imagens, possuindo camadas especializadas na captura de padrões espaciais locais, as redes neurais recorrentes (RNNs), que são especialmente eficazes no processamento de dados sequenciais, como séries temporais ou texto, e as redes do tipo *Long Short-Term Memory* (LSTM), que representam uma variação das RNNs e solucionam o problema do desaparecimento do gradiente, sendo ideais para modelar dependências de longo prazo em sequências (RASCHKA; MIRJALILI, 2019).

As técnicas de deep learning estão presentes em diversas inovações tecnológicas que impactam significativamente a sociedade contemporânea. Entre essas aplicações, destacam-se o diagnóstico médico por imagem, como na detecção automatizada de tumores, os sistemas de tradução automática, como o Google Translate, o reconhecimento de fala utilizado em assistentes virtuais e os veículos autônomos, que dependem de DL para interpretar o ambiente e tomar decisões em tempo real (FOSTER, 2019).

2.3 Transfer Learning

O Transfer Learning permite reaproveitar conhecimentos aprendidos por um modelo treinado em uma tarefa de origem e aplicá-lo em uma tarefa de destino.

Segundo Chollet (2021), essa abordagem é particularmente eficaz quando os dados disponíveis para o novo problema são escassos, tornando-se uma alternativa viável e eficiente em diversos contextos práticos da inteligência artificial. Entre as estratégias mais comuns de integração do modelo estão a extração de características (feature extraction), que consiste na utilização de camadas intermediárias de um modelo pré-treinado como extratoras de características relevantes dos dados; e o ajuste fino (fine-tuning), que envolve o treinamento das últimas camadas, ou até mesmo de todas as camadas, de um modelo previamente treinado para que se adequem à nova tarefa (CHOLLET, 2021).

As vantagens do uso do Transfer Learning são numerosas. Essa técnica reduz a necessidade de grandes volumes de dados rotulados, o que é especialmente importante em domínios onde a rotulagem é cara ou demorada. Além disso, acelera o tempo de desenvolvimento dos modelos e possibilita alcançar resultados superiores mesmo quando se dispõe de infraestrutura computacional limitada. Dessa forma, o Transfer Learning destaca-se como uma abordagem poderosa e acessível para a construção de soluções eficazes em aprendizado profundo, especialmente em cenários com restrições de dados ou recursos (KADDOURA, 2023).

Neste trabalho, foi utilizado um modelo baseado em Deep Learning com a técnica de Transfer Learning, adotando como topologia principal a arquitetura MobileNetV2, uma rede neural convolucional profunda (*Convolutional Neural Network* – CNN) otimizada para dispositivos com recursos computacionais limitados. A *MobileNetV2* foi originalmente treinada no banco de dados *ImageNet*, contendo milhões de imagens e milhares de categorias, o que permite que ela atue como um extrator de características genéricas altamente eficiente para tarefas de classificação visual.

A estrutura da MobileNetV2 é composta por dois blocos principais: as convoluções separáveis em profundidade (*depthwise separable convolutions*) e os blocos residuais invertidos (*inverted residuals*). A técnica de convoluções separáveis permite decompor uma convolução padrão em duas etapas: uma convolução *depthwise*, que aplica um único filtro em cada canal de entrada, e uma convolução *pointwise* (1x1), que realiza a combinação entre os canais. Essa abordagem reduz

significativamente o custo computacional e o número de parâmetros, sem perda expressiva de desempenho. Já os blocos residuais invertidos facilitam o fluxo de gradientes em redes profundas, aumentando a estabilidade durante o treinamento e melhorando a capacidade de generalização da rede.

O modelo foi aplicado em uma tarefa de classificação binária: identificar se uma imagem de uma pessoa contém ou não a presença de óculos. Para isso, as camadas convolucionais da MobileNetV2 foram congeladas, ou seja, seus pesos não foram atualizados durante o treinamento, mantendo o conhecimento prévio adquirido no ImageNet. Apenas as camadas superiores (head do modelo) foram ajustadas, incluindo uma camada de pooling global (*GlobalAveragePooling2D*), uma camada de regularização por dropout (com taxa de 30%) e uma camada densa (*Dense*) com ativação sigmoide para produzir uma saída probabilística no intervalo [0, 1].

Esse tipo de arquitetura permite a criação de um modelo eficiente e robusto, mesmo com um número relativamente pequeno de imagens, graças à reutilização do conhecimento pré-treinado, e se mostra ideal para aplicações com limitações de tempo e recursos computacionais.

2.4 Banco de Dados

A performance de qualquer sistema de Inteligência Artificial (IA) depende fortemente da qualidade e representatividade dos dados utilizados para seu treinamento. Como ressaltado por Raschka e Mirjalili (2019), dados limpos, balanceados e representativos do domínio do problema são essenciais para evitar vieses e problemas como o *overfitting*, que compromete a capacidade de generalização do modelo em dados novos. Um banco de dados bem elaborado deve apresentar algumas características fundamentais para garantir a eficácia dos modelos treinados.

Entre essas características, destaca-se a representatividade, ou seja, os dados devem cobrir adequadamente as variações existentes no problema abordado, contemplando diferentes cenários, condições ou padrões. O balanceamento também é crucial, especialmente em problemas de classificação, onde é necessário garantir

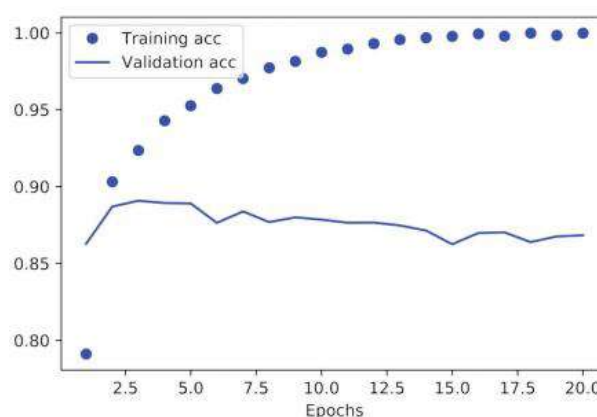
uma distribuição equitativa entre as classes para evitar que o modelo aprenda a favorecer uma classe em detrimento das outras. Além disso, a atualização contínua do banco de dados é fundamental para assegurar que o modelo permaneça relevante ao longo do tempo, acompanhando possíveis mudanças no domínio ou no comportamento dos dados.

Dentre os bancos de dados mais utilizados e citados em pesquisas acadêmicas e aplicações práticas, destacam-se o MNIST, composto por imagens de dígitos manuscritos distribuídos em 10 classes, o CIFAR-10, que contém imagens coloridas categorizadas em 10 classes distintas e o mais conhecido entre eles, ImageNet, um vasto repositório com mais de 14 milhões de imagens anotadas, amplamente utilizado em competições e *benchmarks* para tarefas de classificação de imagens.

2.5 Gráfico de Acurácia

O gráfico de acurácia é uma ferramenta fundamental na análise de desempenho de modelos de aprendizado profundo, especialmente durante o processo de treinamento e validação de redes neurais. Ele representa, visualmente, a evolução da proporção de previsões corretas realizadas pelo modelo ao longo das épocas de treinamento, permitindo avaliar sua capacidade de generalização em relação aos dados de entrada, como exemplificado na Figura 01.

Figura 01 – Exemplo de Treinamento e validação de Acurácia



Autor: Collet (2021)

A acurácia, enquanto métrica estatística, expressa a razão entre o número de classificações corretas e o total de classificações realizadas, sendo amplamente utilizada em tarefas de classificação supervisionada. De acordo com Raschka e Mirjalili (2019), esse gráfico é particularmente útil para identificar padrões de aprendizado, como melhorias consistentes ou estagnações, além de ajudar a diagnosticar problemas como *overfitting* ou *underfitting*.

Quando se observa uma curva de acurácia que cresce de maneira estável tanto no conjunto de treino quanto no de validação, é possível inferir que o modelo está aprendendo de forma adequada. Em contrapartida, grandes discrepâncias entre essas duas curvas ao longo do tempo podem indicar que o modelo está memorizando os dados de treino e falhando em generalizar, o que compromete seu desempenho em situações reais.

Chollet (2021) complementa ao afirmar que o acompanhamento da acurácia deve ser feito em conjunto com outras métricas, como a curva ROC, pois a acurácia isoladamente pode mascarar problemas em contextos com classes desbalanceadas. Isso é comum em aplicações industriais voltadas à segurança, onde há predominância de trabalhadores utilizando corretamente os Equipamentos de Proteção Individual (EPIs).

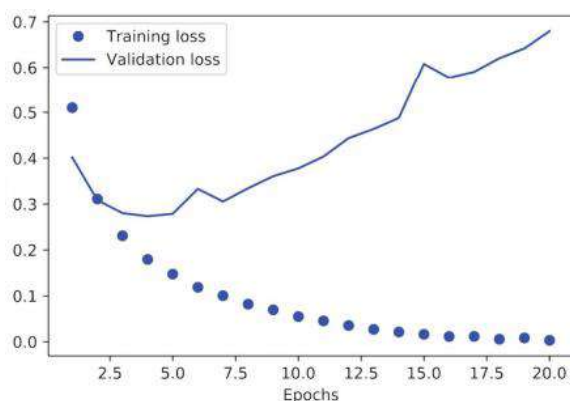
Assim, o gráfico de acurácia não apenas serve como indicador da eficácia do modelo ao longo do treinamento, mas também como instrumento essencial para ajustes estratégicos na arquitetura da rede, nos hiperparâmetros e na composição do banco de dados, contribuindo diretamente para a robustez do sistema de classificação.

2.6 Gráfico de Perda (*loss curve*)

O gráfico de perda, também conhecido como *loss curve*, é uma representação visual essencial no treinamento de modelos de aprendizado profundo, utilizado para monitorar a evolução da função de perda ao longo das épocas. A função de perda, ou *loss function*, é um componente central em redes neurais, pois quantifica

o erro entre as previsões realizadas pelo modelo e os valores reais esperados, funcionando como um guia para o processo de aprendizado, o que pode ser observado no exemplo da Figura 02.

Figura 02 – Exemplo de Treinamento e Validação de Perda (Loss)



Autor: Collet (2021)

Durante o treinamento, o objetivo do modelo é minimizar essa função de perda por meio de algoritmos de otimização, como o gradiente descendente, ajustando iterativamente os pesos da rede para melhorar a precisão das previsões. Conforme explicado por Chollet (2021), o gráfico de perda exibe essa variação em termos numéricos a cada época, permitindo observar se o modelo está, de fato, aprendendo de forma progressiva e eficiente.

Um gráfico ideal apresenta uma curva decrescente e contínua da perda no conjunto de treinamento e, preferencialmente, também no conjunto de validação. Raschka e Mirjalili (2019) ressaltam que esse acompanhamento é fundamental para identificar problemas como o overfitting, situação em que o modelo aprende excessivamente os dados de treinamento e perde capacidade de generalização.

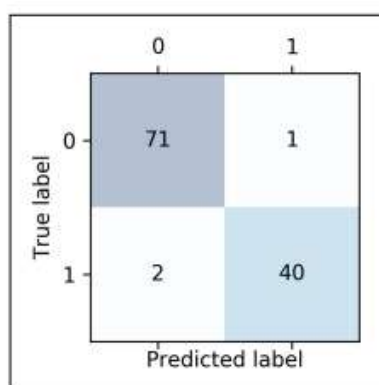
Esse comportamento pode ser identificado quando a perda de validação começa a aumentar enquanto a perda de treinamento continua diminuindo. O gráfico de perda também é utilizado para verificar a necessidade de ajustes na taxa de aprendizado, arquitetura da rede ou ainda na qualidade e equilíbrio do conjunto de dados.

Em aplicações industriais voltadas à segurança, como na identificação automatizada do uso de Equipamentos de Proteção Individual (EPIs), manter um controle rigoroso sobre a perda é essencial para garantir que o modelo esteja tomando decisões confiáveis e consistentes. Dessa forma, o gráfico de perda assume um papel crucial na fundamentação e avaliação do desempenho de modelos de deep learning, servindo como um dos principais instrumentos para validação da eficácia do processo de treinamento.

2.7 Matriz de Confusão

A matriz de confusão é uma ferramenta de avaliação amplamente utilizada na análise de desempenho de modelos de classificação supervisionada, especialmente em contextos que envolvem aprendizado profundo, como é o caso da identificação automatizada do uso de Equipamentos de Proteção Individual (EPIs). Essa ferramenta permite uma análise detalhada da performance do modelo ao comparar as previsões realizadas com os valores reais esperados, conforme exemplificado na Figura 03.

Figura 03 – Exemplo de Matriz de Confusão



	0	1
0	71	1
1	2	40

Autor: Raschka e Mirjalili (2017)

A matriz é organizada em quatro categorias principais: verdadeiros positivos (VP), quando o modelo acerta ao prever a classe positiva; verdadeiros negativos (VN), quando acerta ao prever a classe negativa; falsos positivos (FP), quando prevê incorretamente a classe positiva; e falsos negativos (FN), quando prevê erroneamente

a classe negativa. Essa estrutura fornece uma visão mais granular do comportamento do modelo, além de possibilitar o cálculo de métricas derivadas como precisão, sensibilidade (revocação) e F1-score, fundamentais para complementar a análise da acurácia (RASCHKA; MIRJALILI, 2019).

Conforme destaca Chollet (2021), o uso da matriz de confusão é especialmente útil para entender se o modelo está apenas maximizando acertos em uma classe majoritária ou se está conseguindo lidar bem com a diversidade de dados. Esse tipo de avaliação é essencial em ambientes industriais, onde os modelos podem se deparar com condições variadas de iluminação, ângulos de câmera e características físicas dos trabalhadores.

Adicionalmente, essa ferramenta ganha ainda mais relevância em cenários com desequilíbrio de classes, como ocorre quando a maioria das imagens representa trabalhadores com EPIs e apenas uma minoria sem. Nesses casos, a acurácia isolada pode ser enganosa, uma vez que o modelo pode atingir valores altos de acerto apenas por sempre prever a classe majoritária. Por isso, a análise da matriz de confusão, junto de métricas derivadas, torna-se indispensável para validar o desempenho real do modelo (PEIXEIRO, 2022; RASCHKA; MIRJALILI, 2020).

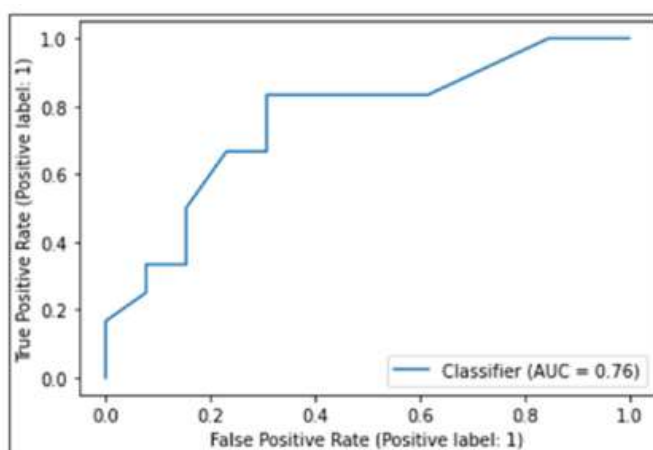
Adari e Alla (2024) ressaltam que, em sistemas críticos de segurança, como os voltados ao monitoramento automatizado em ambientes industriais, o falso negativo (FN) é o erro mais preocupante. Isso ocorre quando o sistema não identifica a ausência do EPI e classifica erroneamente o trabalhador como estando protegido. Essa falha compromete diretamente a integridade física dos colaboradores e pode gerar consequências operacionais graves.

Dessa forma, a matriz de confusão não apenas permite identificar e quantificar esse tipo de erro, como também comparar o desempenho entre diferentes abordagens, como o treinamento tradicional de deep learning e o uso de técnicas de transfer learning. A análise da matriz se mostra, portanto, um instrumento fundamental para garantir segurança, confiabilidade e precisão nos sistemas de visão computacional voltados à detecção de EPIs.

2.8 Curva ROC

A curva ROC, sigla para *Receiver Operating Characteristic*, é uma representação gráfica amplamente utilizada na avaliação do desempenho de modelos de classificação binária, sobretudo em aplicações que envolvem redes neurais profundas e aprendizado supervisionado, como no caso da identificação automática de Equipamentos de Proteção Individual (EPIs). Essa curva expressa a relação entre a taxa de verdadeiros positivos (*True Positive Rate – TPR*), também conhecida como sensibilidade ou revocação, e a taxa de falsos positivos (*False Positive Rate – FPR*), variando o limiar de decisão adotado pelo modelo, o que pode ser analisado no modelo da Figura 04.

Figura 04 – Exemplo de Curva ROC



Autor: Adari e Alla (2024)

O gráfico é construído a partir de diferentes valores de limiares, que influenciam diretamente a forma como o modelo classifica instâncias como positivas ou negativas, permitindo observar sua performance em distintos cenários. Quanto mais a curva se aproxima do canto superior esquerdo do gráfico, melhor o desempenho do modelo, indicando uma alta taxa de acertos com baixa incidência de falsos positivos.

Um aspecto essencial na análise da curva ROC é a Área Sob a Curva (*AUC – Area Under the Curve*), que quantifica numericamente a capacidade de separação entre as classes. Um valor de AUC igual a 1 indica uma classificação perfeita,

enquanto valores próximos de 0,5 representam desempenho aleatório, sem poder discriminativo.

Conforme destacado por Chollet (2021) e Raschka e Mirjalili (2019), a curva ROC é especialmente relevante em situações em que os custos dos erros tipo I e II são diferentes. Isso se aplica diretamente a sistemas de segurança ocupacional, nos quais um falso negativo (não identificar a ausência de EPI) pode acarretar consequências graves, tanto para a integridade do trabalhador quanto para a responsabilidade da empresa.

Adari e Alla (2024) reforçam que pequenas variações no limiar de decisão podem influenciar significativamente a taxa de falsos negativos, sendo preferível, em muitos casos, tolerar um leve aumento nos falsos positivos a fim de garantir maior segurança.

Dessa forma, a curva ROC se apresenta como uma ferramenta indispensável para a calibração e validação de modelos que atuam na detecção de condições críticas em ambientes industriais, contribuindo para a construção de sistemas mais confiáveis e eficazes em contextos sensíveis à segurança.

Dadas as pesquisas citadas acima, é explicado a baixo como foram utilizados os métodos estudados, de maneira à extrair os dados desejados de maneira satisfatória.

3 METODOLOGIA

O presente estudo realizou a comparação entre dois métodos de desenvolvimento de modelos de *deep learning*, sendo eles o treinamento convencional e o treinamento por meio da técnica de *transfer learning*. A avaliação dos modelos considerou métricas fundamentais de desempenho, tais como tempo de processamento, consumo de recursos computacionais, com ênfase na utilização da Unidade de Processamento Gráfico (GPU), gráfico de acurácia, função de perda (*loss*), matriz de confusão e curva ROC.

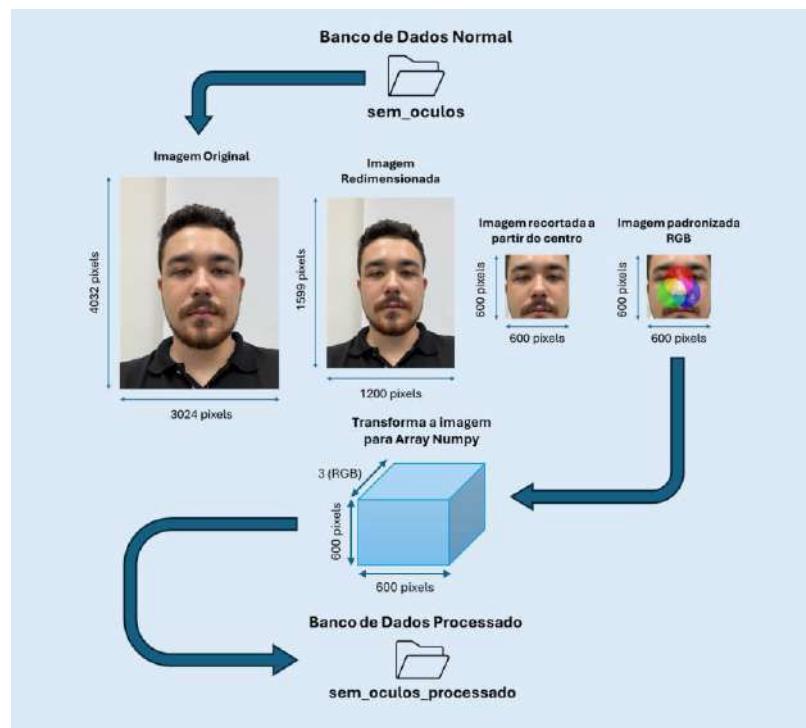
O objetivo principal deste procedimento metodológico foi identificar uma abordagem que proporcionasse uma precisão elevada, baixo custo computacional e

tempo reduzido de processamento, alinhando-se às necessidades operacionais de ambientes industriais. Ressalta-se que o sistema desenvolvido possui caráter escalável, sendo projetado para atuar de forma eficiente em diferentes setores classificados como áreas de risco. Dessa forma, busca-se garantir que a validação do uso dos Equipamentos de Proteção Individual (EPIs) seja realizada de maneira automatizada e precisa, sem a necessidade de paralisações ou interferências nas atividades produtivas, otimizando, assim, o tempo de trabalho e os recursos da organização.

Para a execução e validação dos testes do modelo, foi utilizado um notebook equipado com processador Intel Core i5 de 9ª geração, 32 GB de memória RAM e placa de vídeo dedicada NVIDIA GeForce GTX 1650, proporcionando um ambiente local com boa capacidade de processamento. As partes de desenvolvimento, treinamento do modelo e análise dos resultados, foram realizadas por meio da plataforma Google Colab, que oferece um ambiente de execução online gratuito com suporte a GPU e bibliotecas otimizadas para aprendizado profundo. Essa escolha permitiu maior flexibilidade no desenvolvimento e acelerou significativamente o treinamento do modelo, além de facilitar a integração com o Google Drive para manipulação do banco de dados de imagens.

Para a construção do banco de dados utilizado nos experimentos, foram capturadas 146 imagens envolvendo a participação de alunos e professores da instituição. O tratamento das imagens teve início com a etapa de redução da resolução, por meio da diminuição da quantidade de pixels, com o objetivo de otimizar o desempenho computacional durante o treinamento e a inferência do modelo, sem comprometer a extração das características visuais essenciais. Em seguida, já no ambiente de codificação, foi aplicada uma operação de recorte baseado no centro da imagem, com a finalidade de concentrar a área de análise na região facial do indivíduo, promovendo maior precisão na detecção do Equipamento de Proteção Individual (EPI) facial utilizado nos testes, como mostrado na figura 05.

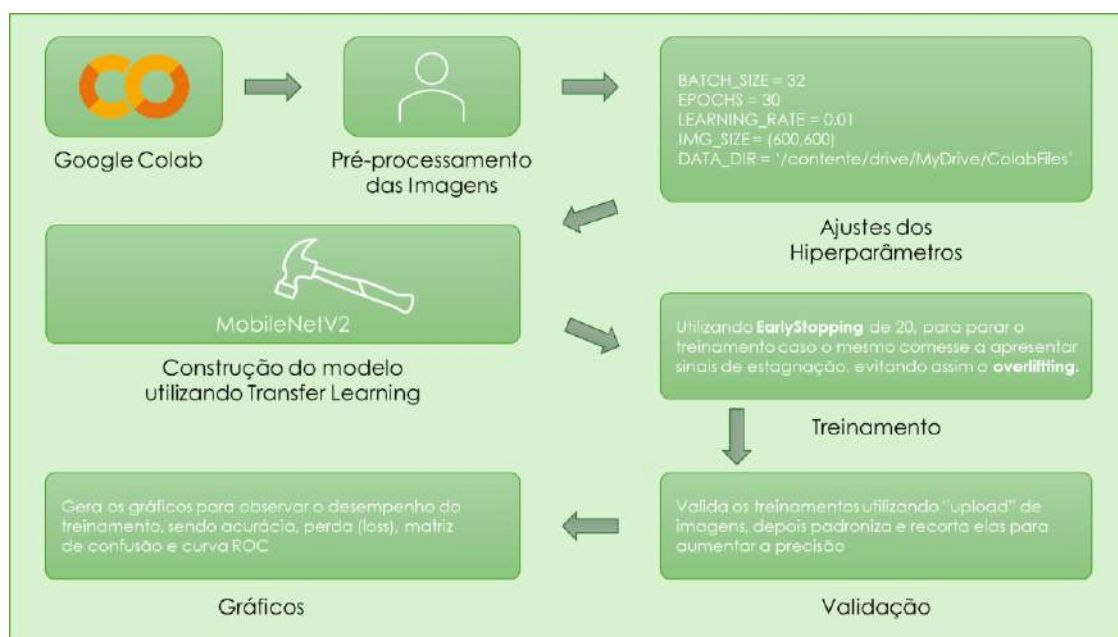
Figura 05 – Processamento das Imagens do Banco de Dados



Fonte: Autor (2025)

Esse pré-processamento contribuiu para a padronização das entradas e melhoria da eficiência do modelo na tarefa de classificação. O objetivo foi garantir diversidade entre os indivíduos fotografados, de modo a ampliar a representatividade do conjunto de dados. As imagens foram registradas em duas condições distintas: inicialmente, com os participantes utilizando óculos, e posteriormente, sem o uso do acessório, com o intuito de possibilitar a comparação e identificação da presença ou ausência de um Equipamento de Proteção Individual (EPI) facial. A estratégia de coleta foi planejada de forma a simular, dentro das possibilidades acadêmicas, um ambiente industrial real, considerando variáveis como diferentes formatos de rosto, tonalidades de pele, e outras características físicas relevantes. Essa diversidade é fundamental para assegurar a robustez e a capacidade de generalização do modelo treinado, refletindo cenários práticos em que os sistemas de detecção devem operar de maneira confiável em ambientes heterogêneos.

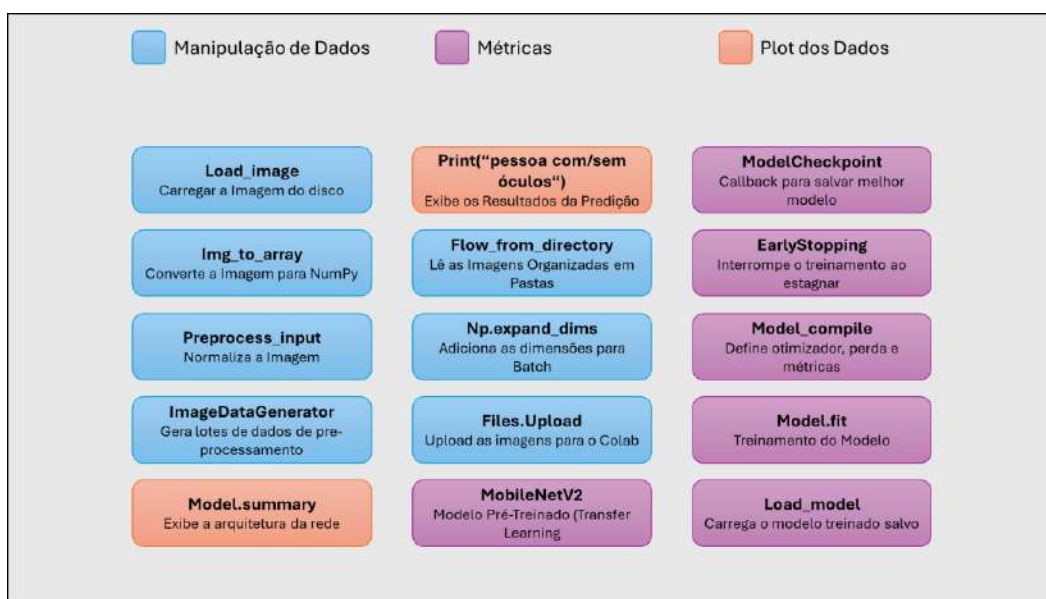
Figura 06 – Processamento das Imagens do Banco de Dados



Fonte: Autor (2025)

Para a implementação dos algoritmos e execução dos testes, foi utilizada a plataforma Google Colab, ambiente de desenvolvimento em nuvem que fornece suporte à linguagem de programação Python e recursos computacionais compatíveis com aplicações de aprendizado profundo. Nessa plataforma, foram desenvolvidos os scripts responsáveis tanto pela organização e preparação do banco de dados quanto pela construção, treinamento e avaliação dos modelos. As abordagens envolveram, inicialmente, o treinamento de um modelo de *deep learning* convencional a partir do banco de dados criado, seguido da aplicação da técnica de *transfer learning* utilizando redes pré-treinadas, onde a estrutura do código pode ser observada na figura 06. Adicionalmente, foi empregada a biblioteca *Matplotlib* para a geração dos gráficos de acurácia, perda (loss), matriz de confusão e curva ROC, os quais foram utilizados para análise e validação dos resultados obtidos ao longo deste trabalho.

Figura 07 – Estrutura de códigos



Fonte: Autor (2025)

Os treinamentos dos modelos foram conduzidos com variações na quantidade de épocas, parâmetro que determina o número de passagens completas do conjunto de dados de treinamento pela rede neural. A definição do número ideal de épocas é uma etapa crítica no processo de modelagem, uma vez que influencia diretamente na capacidade de aprendizado e generalização do sistema. Quando configurado com um número insuficiente de épocas, o modelo tende a não atingir um nível adequado de aprendizado, resultando em baixa precisão nas previsões. Por outro lado, a utilização de um número excessivo de épocas pode levar ao fenômeno conhecido como *overfitting*, no qual o modelo memoriza os padrões específicos do conjunto de treinamento, perdendo a capacidade de generalizar para novos dados. Em aplicações industriais, onde há elevada variabilidade entre os trabalhadores, com diferenças físicas, uso de EPIs e rotatividade de pessoal, esse comportamento pode comprometer significativamente a eficácia do sistema de detecção. Por esse motivo, a escolha da quantidade de épocas foi cuidadosamente ajustada com base em testes empíricos e monitoramento contínuo das curvas de acurácia e perda, de forma a equilibrar o desempenho do modelo entre precisão e generalização.

A partir da montagem do banco de dados e da subsequente etapa de treinamento do modelo, foi possível obter resultados concretos quanto à capacidade de distinção entre imagens que representavam indivíduos utilizando o Equipamento de Proteção Individual (EPI), neste caso, óculos de proteção, e aquelas em que o acessório não estava presente. Os testes demonstraram que o modelo foi capaz de identificar, com elevado grau de acurácia, a presença ou ausência do EPI nas imagens processadas, refletindo a eficácia do sistema de classificação desenvolvido. A Figura 08 ilustra um exemplo representativo do processo de detecção e categorização realizado pelo modelo treinado, evidenciando a aplicabilidade da abordagem em contextos que exigem monitoramento visual automatizado.

Figura 08 – Detecção das Imagens com e sem Óculos



Fonte: Autor (2025)

Com o objetivo de mitigar problemas recorrentes em modelos de aprendizado profundo, como overfitting, baixa velocidade de convergência e instabilidade durante o processo de treinamento, foram realizados ajustes nos principais hiperparâmetros do projeto com base em experimentação empírica (*trial and error*). Entre os hiperparâmetros ajustados, destacam-se o tamanho do lote (batch size), que determina a quantidade de amostras utilizadas antes da atualização dos pesos da rede, o número de épocas (*epochs*), que define o total de passagens completas sobre o conjunto de dados de treinamento, a taxa de aprendizado (*learning rate*), que regula

a velocidade de ajuste dos parâmetros internos do modelo e o tamanho das imagens, cuja resolução impacta diretamente no tempo de processamento e na capacidade do modelo em extrair características relevantes. Esses ajustes foram essenciais para garantir maior estabilidade no processo de aprendizado, otimizar o desempenho computacional e alcançar uma performance satisfatória na tarefa de detecção automatizada de Equipamentos de Proteção Individual (EPIs) e podem ser observados na figura 09.

Figura 09 – Ajustes dos Hiperparâmetros

```
# Hiperparâmetros
BATCH_SIZE = 32
EPOCHS = 30
LEARNING_RATE = 0.01
IMG_SIZE = (400, 400)
DATA_DIR = '/content/drive/MyDrive/ColabFiles'
```

Fonte: Autor (2025)

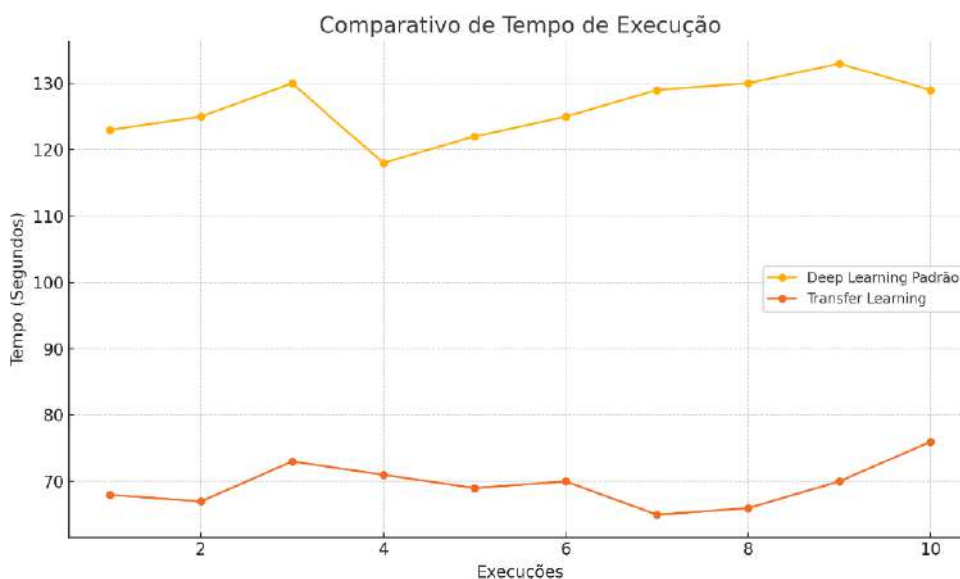
Abaixo são apresentados os resultados obtidos conforme o método explicado acima, realizando os testes de validação utilizando o método de experimentação.

4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

A estratégia adotada para a comparação do tempo de processamento entre os modelos baseou-se na realização de medições por meio de cronometragem, aplicadas a dez execuções completas do código consecutivas para cada abordagem. Visando assegurar a confiabilidade dos resultados e uma comparação direta do desempenho computacional, todos os testes foram realizados sob condições controladas de rede. Para isso, considerou-se a latência da conexão com a internet, mantendo uma média estável de aproximadamente 20 ms, conforme verificado por meio da plataforma de monitoramento NPerf. Essa padronização buscou minimizar

interferências externas relacionadas ao ambiente online do Google Colab. Os dados obtidos foram consolidados graficamente e estão representados na Figura 10.

Figura 10 – Comparativo de Tempo de Execução

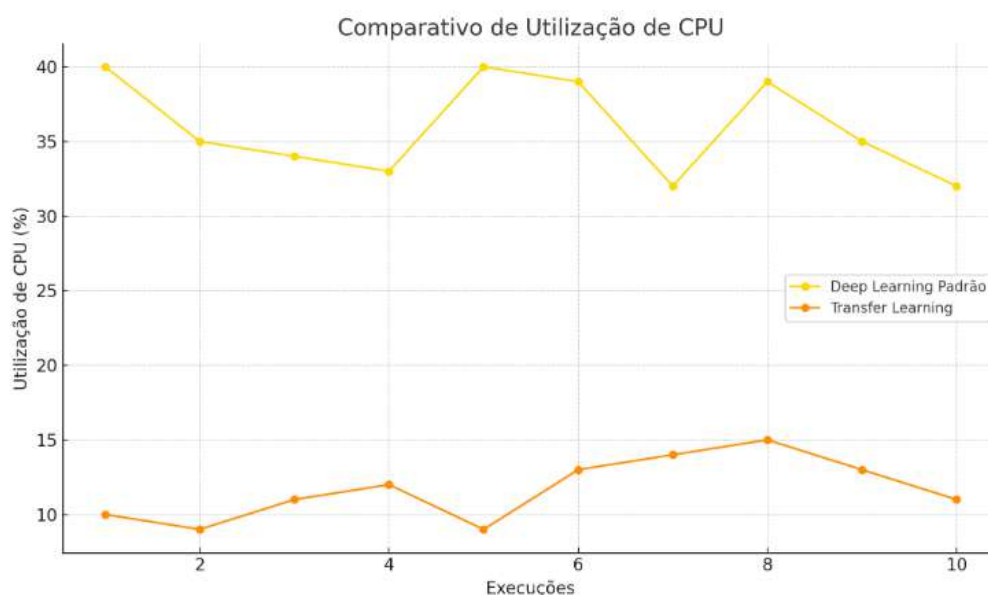


Fonte: Autor (2025)

A partir dos resultados verificados, constatou-se que o modelo treinado com *deep learning* convencional, baseado em um banco de dados estruturado de forma padrão, apresentou um tempo médio de processamento de aproximadamente 126 segundos por execução. Em contraste, o modelo desenvolvido com a técnica de *transfer learning* demonstrou desempenho consideravelmente superior, alcançando uma média de apenas 69 segundos. Essa diferença expressiva evidencia a capacidade do *transfer learning* de otimizar o tempo de execução do modelo.

Dando continuidade à análise de desempenho, foi realizada a avaliação do consumo de recursos computacionais da CPU, com o objetivo de verificar a eficiência operacional dos modelos sob o ponto de vista da carga imposta ao processador. Para isso, utilizou-se o recurso de monitoramento nativo do sistema operacional Windows, o Gerenciador de Tarefas, cujos resultados estão apresentados na Figura 11.

Figura 11 – Comparativo de Utilização da CPU

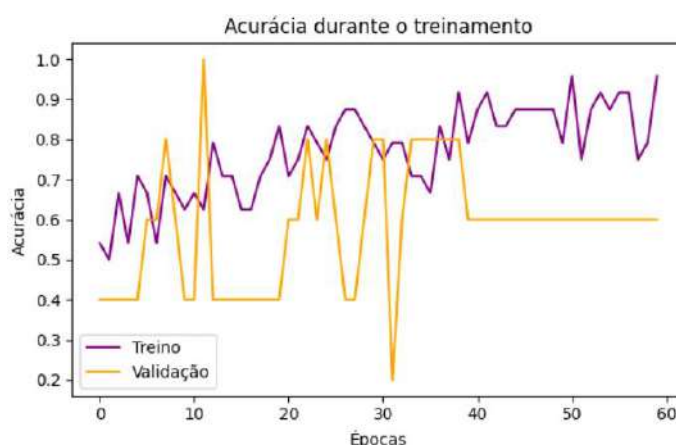


Fonte: Autor (2025)

Durante os testes, observou-se que o modelo tradicional consumiu, em média, 36% da capacidade da CPU. Por outro lado, o modelo treinado com *transfer learning* apresentou um consumo médio substancialmente menor, em torno de 12%. Essa redução no uso de recursos computacionais reforça a viabilidade do *transfer learning* para aplicações em ambientes com infraestrutura limitada, além de evidenciar seu potencial para implementação em sistemas de larga escala com menor custo operacional.

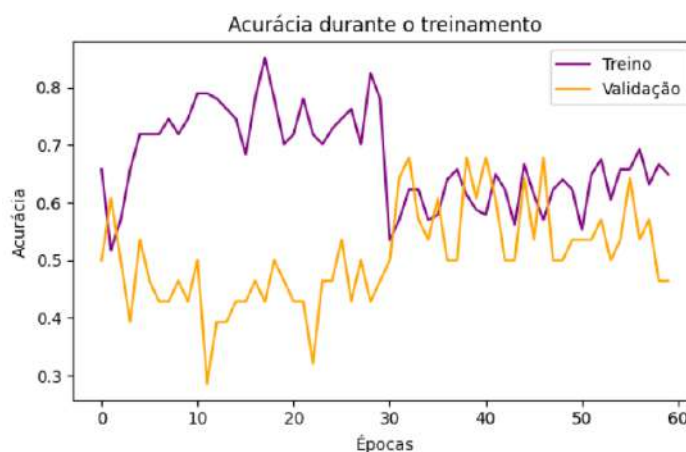
No que diz respeito à acurácia, foi conduzida uma análise detalhada da proporção de previsões corretas realizadas em relação ao total de classificações. Essa métrica permite avaliar a capacidade do modelo em generalizar a partir dos dados de treinamento e realizar inferências precisas. Para garantir a equidade na comparação, ambos os modelos foram submetidos a um regime de treinamento padronizado, com 60 épocas. A evolução da acurácia durante esse processo pode ser observada nas Figuras 12 e 13, que demonstram o desempenho dos modelos com *deep learning* convencional e com *transfer learning*, respectivamente.

Figura 12 – Gráfico de Acurácia Deep Learning Padrão



Fonte: Autor (2025)

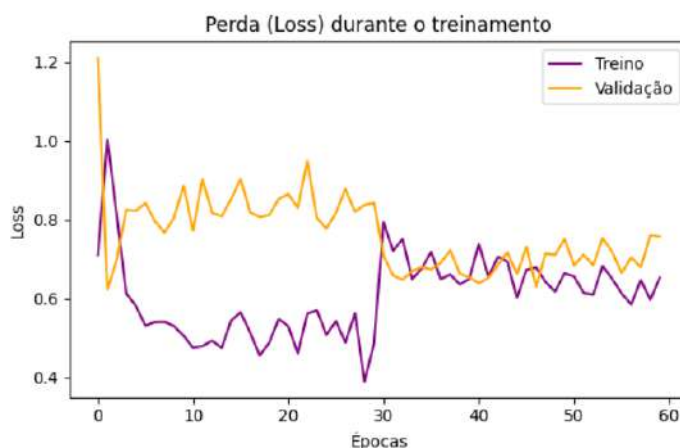
Figura 13 – Gráfico de Acurácia Deep Learning com Transfer Learnig



Fonte: Autor (2025)

De maneira complementar, foi realizada a análise da função de perda (*loss*), a qual mensura a discrepância entre as previsões realizadas pelo modelo e os valores reais. Assim como nas etapas anteriores, os testes foram conduzidos com base na mesma estrutura de treinamento, e seus resultados são apresentados nas Figuras 14 e 15. Observa-se, especialmente na Figura 14, que o modelo treinado com *transfer learning* apresentou uma trajetória mais estável de aprendizagem, com redução progressiva e contínua da perda a partir da trigésima época. Esse comportamento demonstra maior capacidade de adaptação e convergência ao longo do treinamento.

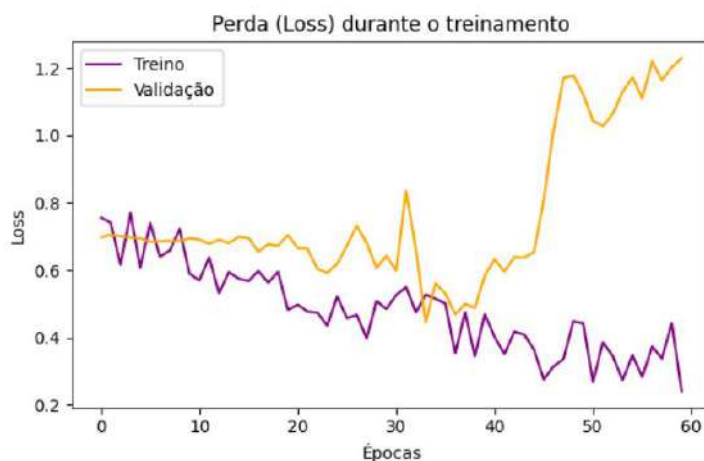
Figura 14 – Gráfico de Perda (Loss) Deep Learning com Transfer Learnig



Fonte: Autor (2025)

Em contrapartida, o modelo utilizando apenas o banco de dados padrão, apresentou um desempenho inicial insatisfatório, acompanhado por uma tendência de degradação progressiva nas épocas subsequentes. Essa instabilidade pode estar associada ao sobreajuste (*overfitting*) ou à limitação do modelo em generalizar adequadamente os padrões presentes nos dados de entrada, como pode ser observado na Figura 15. Tais aspectos reforçam a importância de técnicas mais robustas de aprendizado, como o *transfer learning*, especialmente em contextos com conjuntos de dados restritos.

Figura 15 – Gráfico de Perda (Loss) Deep Learning Padrão

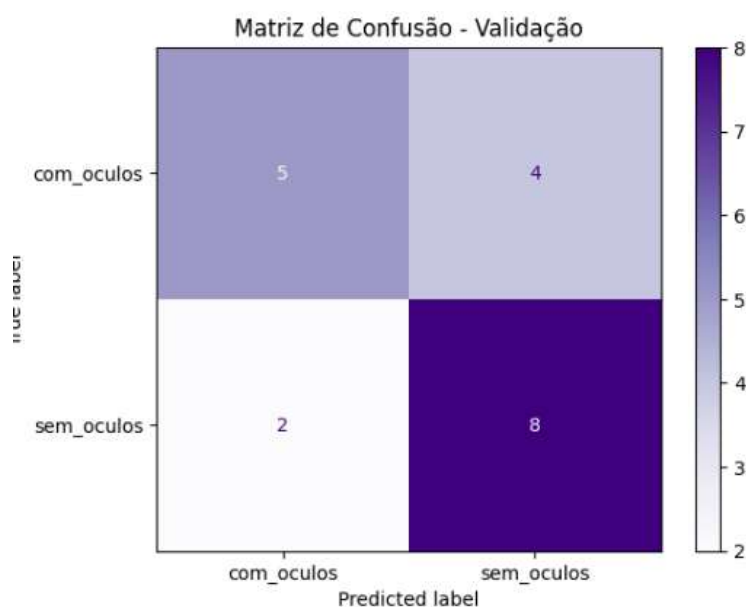


Fonte: Autor (2025)

Considerando o objetivo principal do sistema, que é a identificação automática do uso correto de Equipamentos de Proteção Individual (EPIs), destaca-se que o erro mais crítico é o falso negativo (FN), pois representa a falha em identificar um trabalhador sem EPI, classificando-o como se estivesse devidamente protegido. Tal erro compromete diretamente a segurança operacional e pode colocar em risco a integridade física dos trabalhadores, violando os propósitos preventivos do sistema.

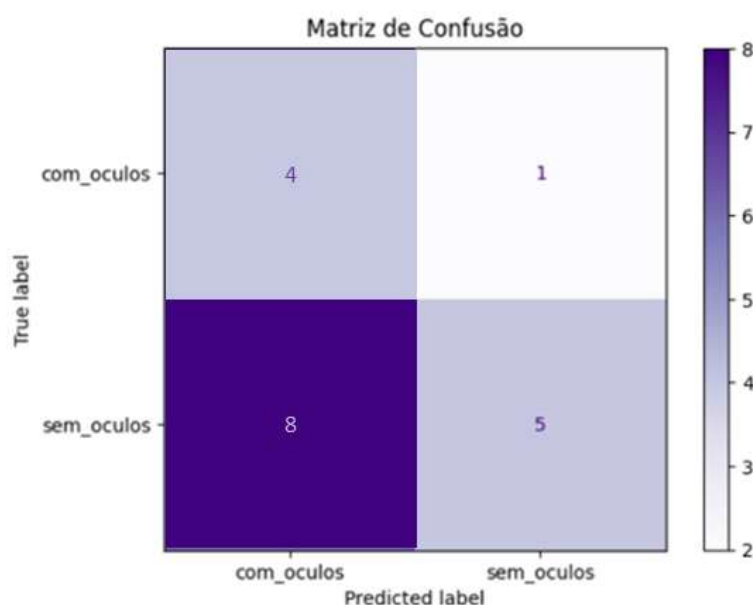
A comparação entre os dois métodos evidencia que o modelo treinado com *Transfer Learning* foi mais eficiente na mitigação de falsos negativos, resultando em apenas 2 FNs, contra 8 FNs do modelo de *Deep Learning* Padrão. Além disso, o modelo com *Transfer Learning* obteve acurácia de 68,42%, erro médio de 31,58% e erro quadrático médio de 1,05, enquanto o modelo convencional teve acurácia de apenas 50%, erro médio de 50% e EQM de 3,61. A menor variância dos erros no modelo com *Transfer Learning* também indica maior estabilidade e consistência nas previsões. Esses resultados demonstram a superioridade do *Transfer Learning* em cenários industriais onde a precisão na detecção de situações críticas é essencial para garantir a segurança e a conformidade com normas regulamentadoras.

Figura 16 – Matriz de Confusão Deep Learning com Transfer Learning



Fonte: Autor (2025)

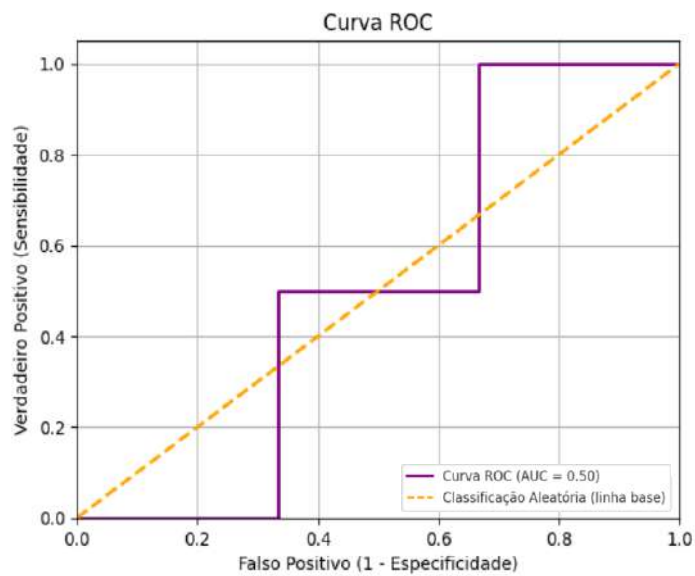
Figura 17 – Matriz de Confusão Deep Learning Padrão



Fonte: Autor (2025)

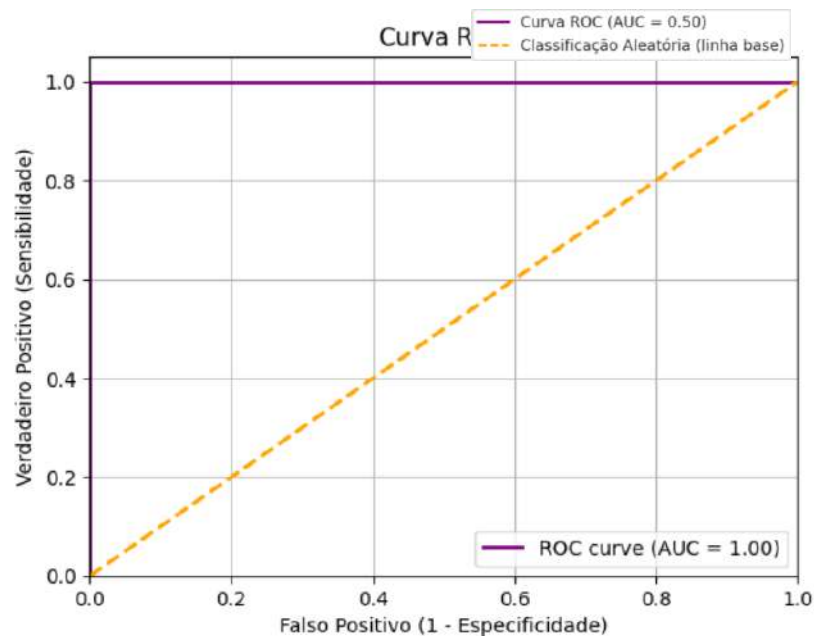
Complementando as avaliações, utilizou-se como última métrica de desempenho a Curva ROC (*Receiver Operating Characteristic*), que permite analisar a sensibilidade do modelo em diferentes limiares de decisão. A Figura 18 ilustra o comportamento da curva ROC para o modelo treinado com *deep learning* convencional, cuja Área Sob a Curva (AUC – *Area Under the Curve*) manteve-se próxima de 0,50. Tal valor indica um desempenho equivalente ao de uma escolha aleatória, o que representa uma limitação severa da abordagem tradicional no contexto de aplicação em questão.

Figura 18 – Curva ROC Deep Learning Padrão



Fonte: Autor (2025)

Figura 19 – Curva ROC Deep Learning com Transfer Learning



Fonte: Autor (2025)

Em contrapartida, a curva ROC obtida com o modelo baseado em *transfer learning* apresentou uma AUC próxima de 1,0, indicando um desempenho quase

ideal, com alta taxa de verdadeiros positivos (TPR) e baixa taxa de falsos positivos (FPR). Esses resultados confirmam a efetividade da técnica de transferência de aprendizado e reforçam sua adequação para aplicações críticas, como a verificação automatizada do uso de EPIs em ambientes industriais.

5 CONSIDERAÇÕES FINAIS

Com base nos resultados obtidos ao longo deste trabalho, foi possível concluir que a aplicação da técnica de transfer learning apresenta desempenho superior em comparação ao modelo de deep learning tradicional treinado a partir de um banco de dados padrão. A análise detalhada das métricas de desempenho, incluindo acurácia, função de perda, taxa de falsos negativos, tempo de processamento e consumo de recursos computacionais, evidenciou as vantagens significativas do transfer learning para a tarefa de identificação automatizada de Equipamentos de Proteção Individual (EPIs) em ambientes industriais. O modelo baseado em transfer learning não apenas demonstrou maior capacidade de generalização, mas também apresentou melhor estabilidade durante o treinamento, mesmo com um volume reduzido de dados, característica comum em aplicações reais. Além disso, o tempo de processamento e a utilização de CPU foram consideravelmente inferiores, tornando a solução mais viável para implementação em ambientes com restrições de hardware. A capacidade da técnica de transfer learning em reutilizar o conhecimento de modelos pré-treinados permitiu alcançar alta precisão na detecção de EPIs, reduzindo significativamente os casos de falsos negativos, o que é um fator crítico para garantir a segurança dos trabalhadores e a conformidade com as normas regulamentadoras. Dessa forma, os resultados alcançados reforçam que o transfer learning é a abordagem mais indicada para aplicações industriais que exigem monitoramento visual em tempo real, com elevada confiabilidade e baixo custo operacional.

Como continuidade deste trabalho, propõe-se a integração do sistema desenvolvido com um ambiente de hardware específico, permitindo a realização de novos testes em tempo real. Em etapas futuras, a implantação do sistema em ambiente fabril pode ser considerada, com o intuito de coletar dados reais de operação. Essa estratégia contribuirá para a ampliação do banco de dados e,

consequentemente, para o aprimoramento contínuo da acurácia e robustez do modelo.

REFERÊNCIAS

ADARI, Suman Kalyan; ALLA, Sridhar. Beginning Anomaly Detection Using Python-Based Deep Learning. 2. ed. Berkeley: Apress, 2024.

AWARI. Curso de Inteligência Artificial: do zero ao avançado. 2023. Disponível em: <<https://awari.com.br/>>. Acesso em: 22 abr. 2025.

CHOLLET, François. Deep Learning with Python. 2. ed. Shelter Island: Manning Publications, 2021.

DOMINGUES, Carlos Henrique. Segurança do Trabalho: fundamentos e práticas para ambientes industriais. São Paulo: Érica, 2024.

FAGBOHUN, Oluwole. TensorFlow Developer Certificate Guide: prepare for the certification exam. [S.l.]: Packt Publishing, 2023.

FOSTER, David. Generative Deep Learning: teaching machines to paint, write, compose, and play. Beijing: O'Reilly Media, 2019.

GHAYOUMI, Mehdi. Generative Adversarial Networks in Practice: with examples using TensorFlow 2.0. Berkeley: Apress, 2024.

HOLDSWORTH, John; SCAPICCHIO, Fabio. Inteligência Artificial na Indústria 4.0. Rio de Janeiro: Ciência Moderna, 2024.

KADDOURA, Sanaa. A Primer on Generative Adversarial Networks. Cham: Springer, 2023.

KETKAR, Nikhil; MOOLAYIL, Jojo. Deep Learning with Python. 2. ed. [S.l.]: Apress, 2021.

MUREL, Daniel; KAVLAKOGLU, Orhan. Transfer Learning: teoria e aplicação prática em Python. Porto Alegre: Bookman, 2024.

PEIXEIRO, Marco. Time Series Forecasting in Python. [S.l.]: Packt Publishing, 2022.
RASCHKA, Sebastian; MIRJALILI, Vahid. Python Machine Learning: machine learning and deep learning with Python, scikit-learn, and TensorFlow 2. 3. ed. Birmingham: Packt Publishing, 2019.

RASCHKA, Sebastian; MIRJALILI, Vahid. Python Machine Learning - Detailed and Simple. [S.l.]: Packt Publishing, 2020.

VASILEV, Ivan. Python Deep Learning. 2. ed. Birmingham: Packt Publishing, 2020.
IBM. What is Deep Learning?. 2024. Disponível em: <<https://www.ibm.com/cloud/learn/deep-learning>>. Acesso em: 10 abr. 2025.
KAGGLE. Confusion Matrix Explained. 2023. Disponível em: <<https://www.kaggle.com/general/16947>>. Acesso em: 22 abr. 2025.

MATPLOTLIB. Plotting Loss and Accuracy Curves in Keras. 2024. Disponível em: <https://matplotlib.org/stable/gallery/lines_bars_and_markers/plot_loss_accuracy.html>. Acesso em: 07 abr. 2025.

NVIDIA. What is Transfer Learning?. 2023. Disponível em: <<https://developer.nvidia.com/blog/what-is-transfer-learning/>>. Acesso em: 28 abr. 2025.

OPENCV. ROC Curve in Machine Learning. 2024. Disponível em: <https://docs.opencv.org/4.x/dc/d3d/tutorial_general_roc_curve.html>. Acesso em: 12 abr. 2025.

PORTAL DA SEGURANÇA DO TRABALHO. O que são Equipamentos de Proteção Individual (EPIs). 2023. Disponível em: <<https://segurancadotrabalho.org/o-que-sao-epis>>. Acesso em: 18 mar. 2025.

SCIKIT-LEARN DEVELOPERS. Plot Receiver Operating Characteristic (ROC) Curve. 2024. Disponível em: <https://scikit-learn.org/stable/auto_examples/model_selection/plot_roc.html>. Acesso em: 02 maio 2025.

TENSORFLOW. Using Callbacks to Track Accuracy and Loss. 2024. Disponível em: <<https://www.tensorflow.org/tutorials/keras/classification>>. Acesso em: 14 abr. 2025.

WHO. Occupational Health: Personal protective equipment. 2023. Disponível em: <https://www.who.int/occupational_health/publications/ppe_guidelines/en/>. Acesso em: 25 abr. 2025.

RUSSELL, Stuart; NORVIG, Peter. Inteligência Artificial. 4. ed. São Paulo: Pearson, 2021.

GERÓN, Aurélien. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. 2. ed. Sebastopol: O'Reilly Media, 2019.