

Uso de Inteligência Artificial para Reconhecimento de Imagens na Triagem de Serviços de Nova Ligação

Tema: Atendimento Digital

Autores: Rafael Perez Pagan

Co-Autores: Felipe Resek Mohallem, Carlos Henrique Cruz Xavier, Fernando Costa Lasmar, Gustavo Siqueira Marquardt

Empresa: DDMX Inteligência em Análise de Dados

Resumo

O reconhecimento de imagens avançou significativamente devido à aplicação de métodos computacionais modernos, especialmente em áreas com grande geração de dados, como os serviços de triagem de concessionárias de energia elétrica. Este estudo avalia a eficácia de dois algoritmos de visão computacional na melhoria da precisão e eficiência do reconhecimento de imagens e busca automatizar o processo manual de validação de imagens enviadas durante a contratação de um serviço de nova ligação. Por meio de um modelo treinado, o sistema pode ser capaz de classificá-las sem a necessidade de um especialista, tornando o serviço mais ágil e eficiente. Os resultados indicam que os modelos são capazes de identificar corretamente as imagens enviadas pelos clientes, com baixo custo computacional e a adaptabilidade do sistema, indicando viabilidade para serviços de triagem em tempo real.

1. Introdução

O setor de energia elétrica, especialmente em empresas subsidiárias que fornecem serviços como ligações e novas instalações, enfrenta desafios crescentes para atender às demandas diárias. Com o aumento do número de atendimentos, os funcionários precisam constantemente avaliar situações de campo e classificá-las de acordo com o tipo de atendimento necessário. Contudo, esses atendimentos podem resultar em deslocamentos ineficientes, aumentando custos operacionais e gerando atrasos no atendimento aos clientes.

A falta da instalação correta do padrão de energia pelo cliente que solicita uma nova ligação é uma das principais causas de deslocamentos desnecessários de equipes. E triar os serviços de nova ligação que estão prontos para receber o serviço é um dos principais desafios das operações de despacho.

Nos últimos anos, o uso de tecnologias de automação, como *chatbots*, mostrou-se eficaz em tarefas repetitivas e manuais. Por meio de um *chatbot*, o cliente deve enviar fotos como da fachada da casa ou do padrão de energia, por exemplo, para auxiliar na tomada de decisão do atendente. Nesse cenário, a avaliação de uma grande quantidade de imagens – essencial para identificar a natureza dos problemas – continua sendo um processo manual, demorado e sujeito a erros. Essa limitação impede que as empresas alcancem maior eficiência operacional e ofereçam respostas mais rápidas e precisas aos clientes.

Nesse contexto, a aplicação de modelos de inteligência artificial (IA) surge como uma solução promissora. A automação da classificação de imagens possibilita a identificação rápida e precisa dos problemas, acelerando o processo de tomada de decisão e otimizando o atendimento ao cliente. Este artigo explora como a IA pode automatizar o processo de avaliação de imagens, promovendo um ganho em eficiência e reduzindo custos operacionais, enquanto melhora a experiência dos clientes.

2. Desenvolvimento

2.1 Trabalhos Relacionados

Nessa seção são discutidos alguns artigos do estado da arte para compreensão dos últimos avanços na detecção de imagens, especialmente vinculada ao atendimento de clientes e serviços na área de energia elétrica.

O artigo (Zou et al. 2023) discute o avanço dos últimos vinte anos no reconhecimento de imagens. A detecção de objetos serve de base para tarefas de Visão Computacional como segmentação e rastreamento de objeto, que podem ser aplicadas em atividades como direção autônoma e vigilância. Segundo o artigo, os principais pontos de avanço na tecnologia foram os seguintes: (i) detecção leve, voltada para dispositivos de baixo consumo energético, crucial para aplicações como realidade aumentada e cidades inteligentes; (ii) detecção de objetos de ponta a ponta busca melhorar a precisão e eficiência eliminando etapas separadas; (iii) a detecção de pequenos objetos, útil para contagem de pessoas e alvos militares; (iv) a detecção em 3D para as aplicações para direção autônoma; (v) a detecção em vídeos explora correlações espaciais e temporais entre quadros; e (vi) a detecção multimodal combina fontes diversas, como imagens Red Green Blue (RGB) e sons, para maior precisão.

O artigo (Lin et al. 2019) discute um sistema desenvolvido para auxiliar usuários de cadeiras de rodas no acesso a ônibus, integrando reconhecimento de imagens baseado em IA com o *chatbot* LINE e a plataforma em nuvem Heroku. O modelo YOLOv3 detecta usuários em pontos de ônibus, enviando notificações para motoristas via Flask e LINE. Usando 400 imagens rotuladas de cadeiras de rodas, o modelo foi treinado, refinado e implantado em um Raspberry Pi 3 B+ para detecção em tempo real. A solução melhora a acessibilidade no transporte público, promovendo inclusão, porém ainda carece de eficiência. Os autores prometem incluir métodos de Computação de Borda visando essa melhoria.

O artigo (Jalal, Khalil, e Haq 2024) traz uma revisão sobre como os sistemas fotovoltaicos são afetados por condições ambientais e discute a necessidade de inspeção visual para detectar falhas utilizando técnicas como imagem termografia infravermelha. Modelos de aprendizado profundo (*deep learning*), como redes neurais convolucionais (CNNs) e redes neurais recorrentes (RNNs), além de algoritmos de clusterização, mostram potencial para detecção de falhas.

As aplicações de visão computacional no setor de auxílio ao público e vinculadas a problemas do setor energético têm evoluído constantemente. Contudo, conforme as revisões de literatura desenvolvidas, ainda não foi publicada uma solução de identificação de imagens enviadas por clientes de redes energéticas urbanas. Esse modelo pode reduzir a intervenção humana, reduzir custos com logística e otimizar atividades de nova ligação energia elétrica.

2.2 Metodologia

Visando automatizar o reconhecimento das imagens enviadas pelos clientes, foram necessárias diversas etapas de tratamento dos dados. Como primeiro passo, a análise e compreensão do banco de imagens disponível para o projeto, compreendendo as características do *dataset*, incluindo o formato das imagens, o balanceamento entre as classes e o tipo de imagens.

As imagens utilizadas foram fotos reais enviadas pelos clientes através da solução DDMX Bot Callback implantada na área de concessão da atendida pela NCSR Energia, prestadora de serviços da CEMIG, na região metropolitana de Belo Horizonte e Betim, em Minas Gerais. O DDMX Bot Callback é uma ferramenta de triagem automática que utiliza um fluxo conversacional inteligente para realizar o *callback* de serviços de campo através de um *chatbot* via *whatsapp*. Para a triagem em serviços de nova ligação, o DDMX Bot Callback solicita ao cliente, além de outras informações, fotos do padrão de energia inteiro, da caixa sem tampa e da fachada do imóvel que foi solicitada a nova ligação. No total, foram utilizadas 10142 fotos, coletadas entre os dias 01 de janeiro de 2024 e 15 de maio de 2024.

Tratando-se de um problema de classificação de rótulos em imagens, o modelo escolhido para a pesquisa foi o *YOLOv8x-cls*, devido a sua alta capacidade de generalização e eficiência em tarefas específicas de classificação (-cls) (Ultralytics, 2024). Para o desenvolvimento da segmentação foi utilizado um modelo Mask R-CNN pela eficiência e integração com o detectron2.

Os treinamentos foram realizados utilizando a biblioteca *Ultralytics YOLO* (Ultralytics, 2024) no ambiente do Google Colab. O *dataset* foi dividido em proporções para aprendizado supervisionado. 70% para treinamento, 20% para validação e 10% para teste. Os parâmetros dos treinamentos incluíram o uso de 30 épocas e taxa de aprendizado ajustada automaticamente pelo otimizador Adam (Ultralytics, 2024).

O treinamento foi realizado utilizando uma GPU NVIDIA T4, disponibilizada pela plataforma de desenvolvimento, o que garantiu alta capacidade computacional para o processamento do modelo. O processo teve uma duração aproximada de quatro horas, abrangendo ajustes no número de épocas, visando melhorar a generalização do modelo e refinar a extração de características das imagens.

O processo de classificação e segmentação foi realizado para separar e identificar objetos de interesse em imagens relacionadas a padrões de energia. Após a classificação inicial, o processo de segmentação foi conduzido utilizando o modelo Mask R-CNN (He et al., 2017), reconhecido por sua capacidade de realizar segmentação de instâncias com alta precisão. Este modelo, uma extensão do Faster R-CNN, foi configurado para prever máscaras de segmentação para cada objeto detectado, além de realizar a classificação e sua localização na imagem. A abordagem envolveu a preparação de um dataset composto por 184 imagens, em que, de forma manual, identificou-se os perímetros dos segmentos das imagens que representavam as classes “caixa”, “disjuntor” e “fios”. A Figura 1 apresenta um exemplo de uma foto com os segmentos identificados manualmente.

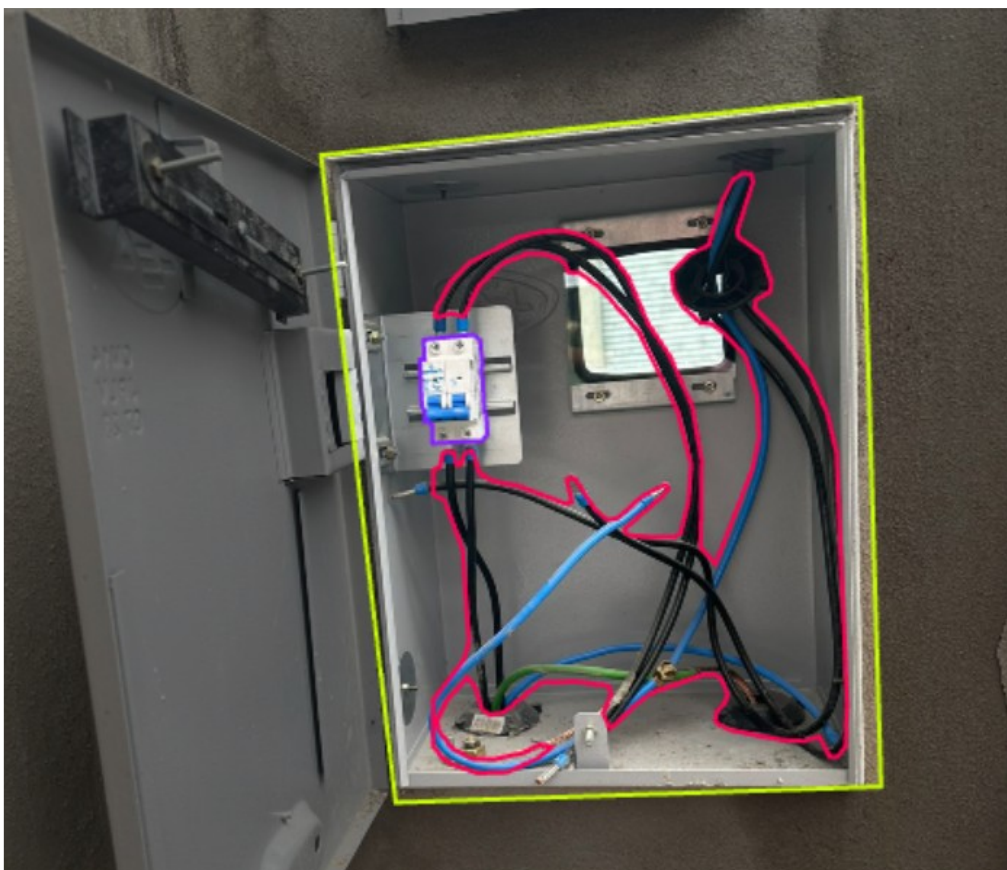


Figura 1 - Exemplo de segmentação aplicada manualmente para preparação do dataset de treinamento. Este dataset foi então utilizado como base para o treinamento do modelo de segmentação. O treinamento foi conduzido em um ambiente CPU, respeitando as limitações computacionais e otimizando o desempenho para atender aos objetivos do estudo.

2.3 Resultados

2.3.1 Classificação

O modelo classificador YOLOv8X-CLS foi avaliado para determinar o desempenho na classificação correta de imagens. Durante os experimentos, métricas como acurácia e precisão foram analisadas para medir a eficácia do modelo.

Os dados foram representados graficamente, proporcionando uma visualização mais clara do desempenho do modelo. A Figura 2 ilustra a evolução da acurácia durante o treinamento, destacando que a acurácia permaneceu consistentemente elevada ao longo do processo, com variações mínimas. Este comportamento indica a estabilidade do modelo.

Adicionalmente, a Figura 2 também apresenta as curvas de perda para treinamento e validação, que mostraram uma redução progressiva, estabilizando-se em valores baixos após certo período. Esse padrão reflete a capacidade do modelo de se ajustar gradualmente aos dados de treinamento, aprimorando sua taxa de aprendizado e eficácia na tarefa de classificação.

Tratando-se de um problema de classificação de rótulos em imagens, o modelo escolhido para a pesquisa foi o *YOLOv8x-cls*, devido a sua alta capacidade de generalização e eficiência em tarefas específicas de classificação (-cls) (Ultralytics, 2024). Para o desenvolvimento da segmentação foi utilizado um modelo Mask R-CNN pela eficiência e integração com o detectron2.

Os treinamentos foram realizados utilizando a biblioteca *Ultralytics YOLO* (Ultralytics, 2024) no ambiente do Google Colab. O *dataset* foi dividido em proporções para aprendizado supervisionado. 70% para treinamen-

to, 20% para validação e 10% para teste. Os parâmetros dos treinamentos incluíram o uso de 30 épocas e taxa de aprendizado ajustada automaticamente pelo otimizador Adam (Ultralytics, 2024).

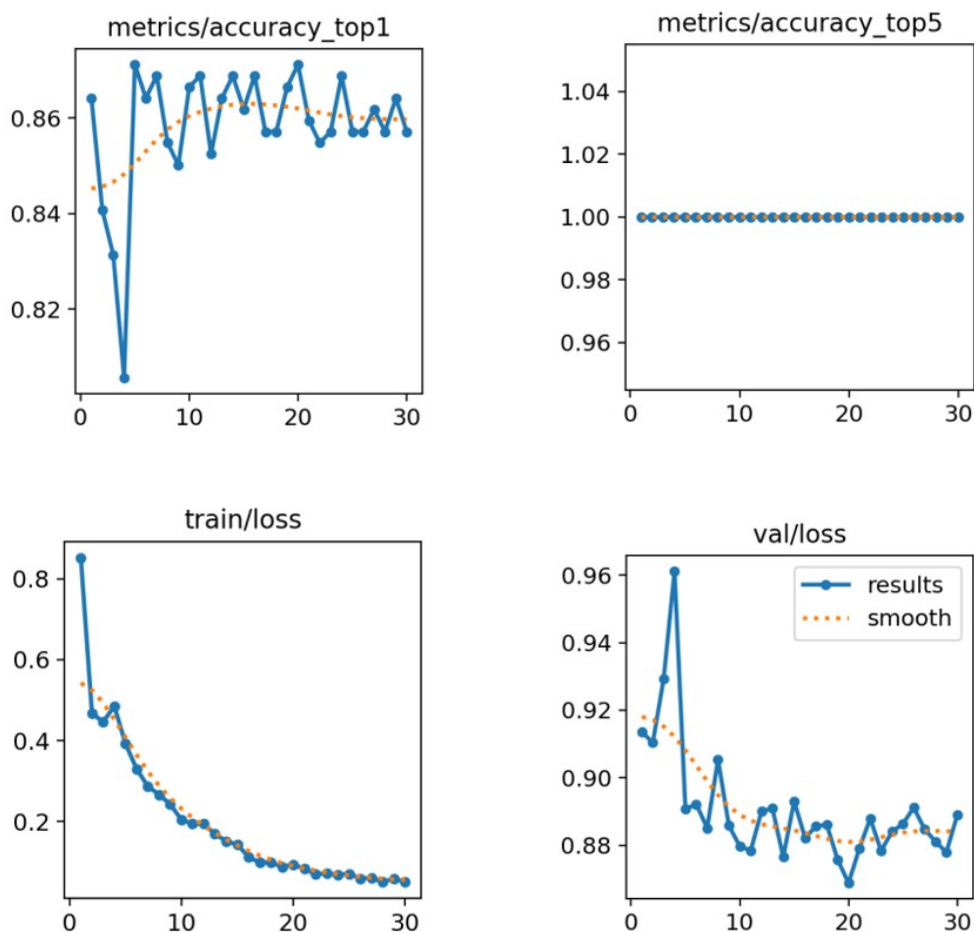


Figura 2 - Gráficos de acurácia, perdas e validação do modelo de classificação

Adicionalmente, foi gerada uma matriz de confusão normalizada, que possibilitou identificar possíveis pontos fracos do modelo. Como visto na Figura 3, o eixo X representa os rótulos verdadeiros, enquanto o eixo Y mostra as predições realizadas pelo modelo. A diagonal principal evidencia os acertos do modelo, destacando maior precisão nas classes "caixa sem tampa" e "padrão inteiro".

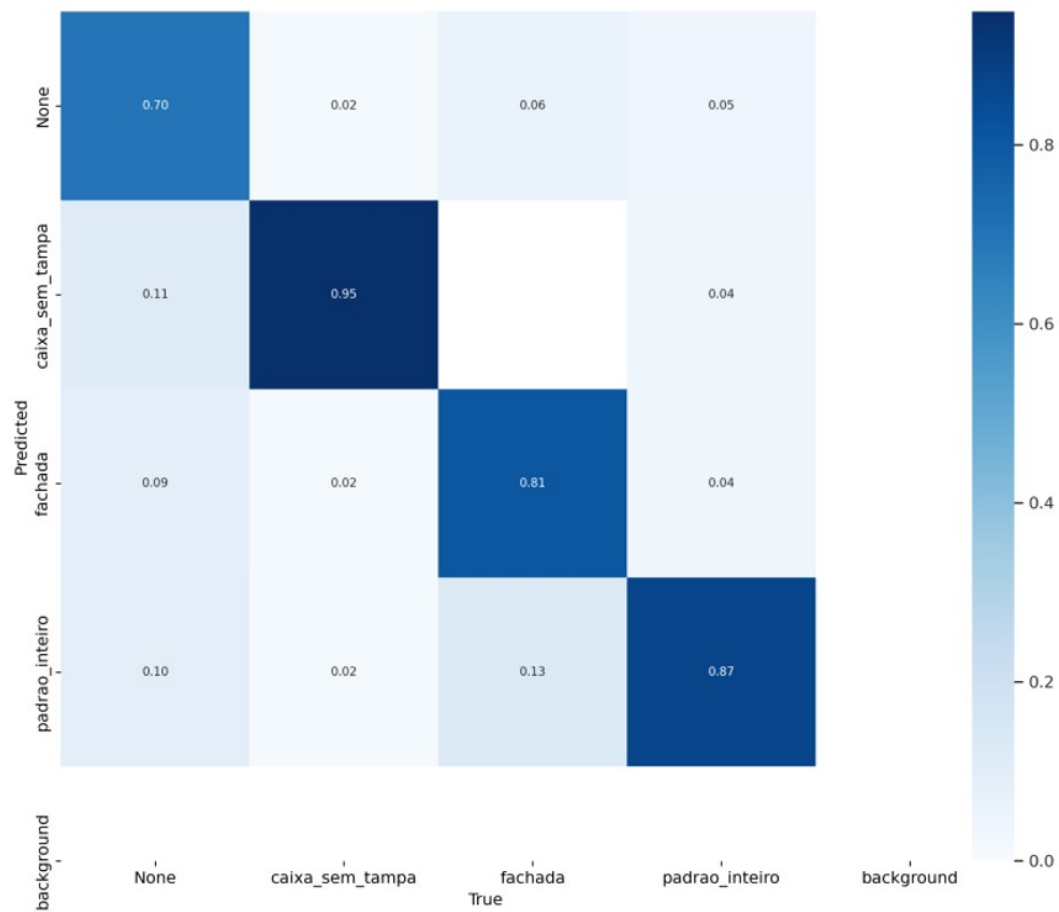


Figura 3 - Matriz de confusão do entre real e previsto do modelo de classificação

Por fim, o modelo também forneceu exemplos práticos de validação por meio de caixas preditivas. Nesses exemplos, a Figura 4 representa a predição realizada pelo modelo após o treinamento enquanto a Figura 5 indica a classe verdadeira da imagem.



Figura 4 - Lote de imagens classificadas pelo modelo



Figura 5 - Lote de imagens com classificação verdadeira

2.3.2 Segmentação

Após a classificação, separando as imagens condizentes com o objetivo, inicia-se a segmentação. Para avaliar o desempenho do modelo treinado, realizamos testes em um conjunto de dados composto por 1551 imagens contendo padrões associados às classes "disjuntor", "fios" e "caixa-sem-tampa", além de um conjunto adicional de 1600 imagens que não apresentavam nenhum desses elementos, totalizando 3151 imagens. Esse cenário foi projetado para simular situações reais, permitindo analisar a capacidade do modelo de identificar corretamente os padrões desejados enquanto evita falsos positivos em imagens irrelevantes. O objetivo era identificar se o modelo de segmentação era capaz de identificar as classes "caixa-sem-tampa", "disjuntor" e "fios" e com qual nível de confiança.

Para cada imagem, o modelo devolve uma resposta com uma lista de percentuais de confiança (score) que um determinado segmento foi encontrado. A Figura 6 apresenta um exemplo de como o modelo identifica a "caixa sem tampa", "fios" e "disjuntor" assim como o nível de confiança (score) de cada identificação.

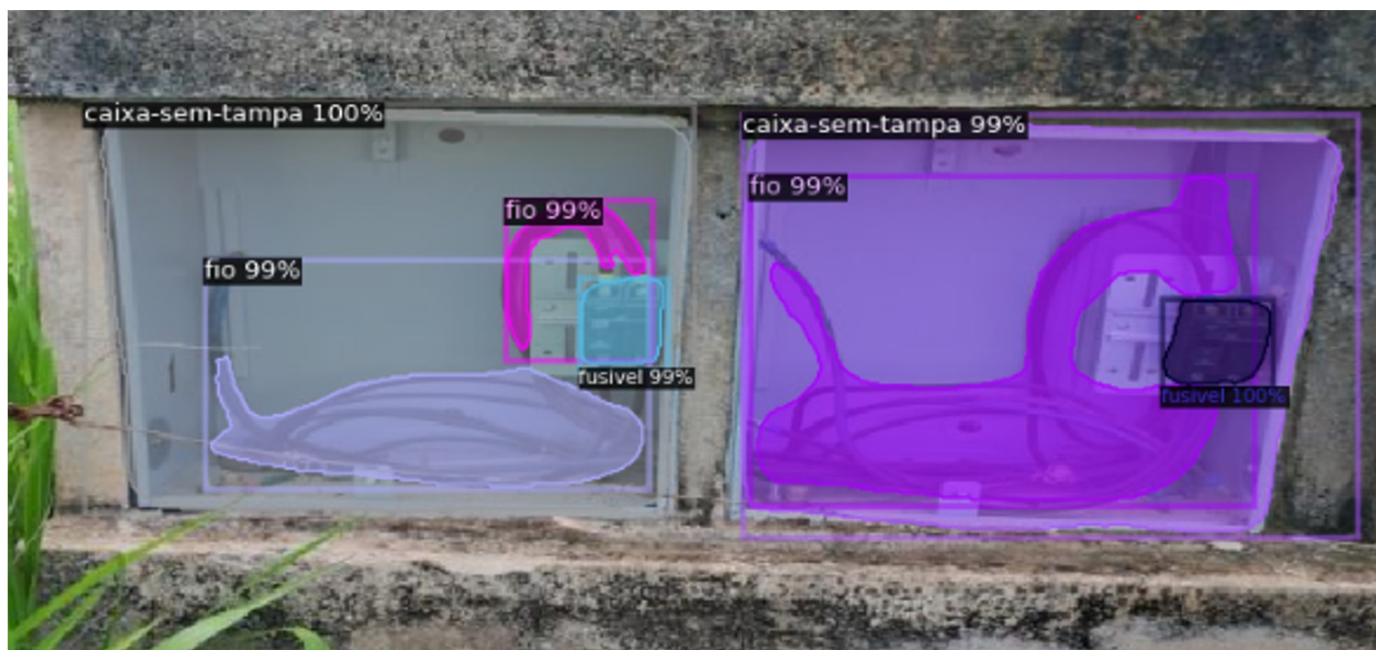


Figura 6 - Resposta do modelo de segmentação

A tabela 1 apresenta um resumo do resultado da avaliação das 3151 imagens. Pode-se observar a comparação entre o previsto e a realidade da imagem. A previsão é considerada “apta” quando o modelo aponta que possui as três classes com nível de confiança mínimo de 90%.

		Realidade		Total
Análise		Apto	Inapto	
Previsto (Score >90%)	Apto	1550	16	1566
	Inapto	1	1584	1585
	Total	1551	1600	3151

Tabela 1 - Matriz de confusão entre real e previsto do modelo de segmentação

Portanto, das 1551 imagens que já se sabia previamente que a foto possui as classes de “caixa sem tampa”, “fios” e “fusível”, em 1550 o modelo acertou (99,9% de assertividade) dizendo que possuía. Nas 1600 imagens que já se sabia previamente que a foto não possui as classes, o modelo também acertou em 1584 casos (99,0% de assertividade) não possuía.

3. Conclusão

O artigo apresentou o desenvolvimento de uma solução baseada em inteligência artificial para automatizar a classificação e segmentação de imagens em serviços do setor de energia elétrica. O uso dos modelos YOLOv8x-cls para classificação e Mask R-CNN para segmentação demonstrou resultados promissores, com precisão elevada na identificação de padrões específicos. A abordagem propõe reduzir custos opera-

cionais, otimizar atendimentos e minimizar a necessidade de intervenção humana. Apesar do desempenho satisfatório, ajustes no limiar de confiança e maior diversificação do conjunto de dados podem aprimorar ainda mais a aplicação prática do modelo. O estudo reforça o potencial da visão computacional em resolver desafios específicos do setor energético.

Outrossim, novos incrementos no modelo poderão aperfeiçoar cada vez mais suas funcionalidades, como a segmentação para diferenciar tipos de disjuntores, possibilitando identificar se o disjuntor está adequado para a carga solicitada.

Neste momento, pretende-se agregar a tecnologia desenvolvida neste trabalho com a ferramenta *DDMX Bot Callback*. Desta maneira, o operador do sistema terá a classificação das fotos e a sugestão de que o serviço está ou não pronto para execução de maneira automatizada. Esta integração permitirá também que o próprio cliente seja alertado quando o modelo identificar que há uma alguma não conformidade na foto enviada. Agilizando a resposta ao atendimento.

Agradecimentos

Os autores agradecem à NCSR Energia, por permitir a utilização das fotos coletadas através da plataforma *DDMX Bot Callback* e por acreditar que este desenvolvimento poderá melhorar a produtividade de seu processo de triagem.

4. Referências bibliográficas

Chang, Zhonghao, e Te Han. 2024. “Prognostics and health management of photovoltaic systems based on deep learning: A state-of-the-art review and future perspectives”. *Renewable and Sustainable Energy Reviews* 205 (novembro):114861. <https://doi.org/10.1016/j.rser.2024.114861>.

He, Kaiming, Georgia Gkioxari, Piotr Dollár, e Ross Girshick. 2017. “Mask R-CNN”. Em *2017 IEEE International Conference on Computer Vision (ICCV)*, 2980–88. <https://doi.org/10.1109/ICCV.2017.322>.

Jalal, Marium, Ihsan Ullah Khalil, e Azhar ul Haq. 2024. “Deep learning approaches for visual faults diagnosis of photovoltaic systems: State-of-the-Art review”. *Results in Engineering* 23 (setembro):102622. <https://doi.org/10.1016/j.rineng.2024.102622>.

Lin, Pei-Jung, Stephen Hung, Shing Fai Steven Lam, e Bo Ching Chen. 2019. “Object Recognition with Machine Learning: Case Study of Demand-Responsive Service”. Em *2019 IEEE International Conference on Internet of Things and Intelligence System (IoT&IS)*, 129–34. <https://doi.org/10.1109/IoT&IS47347.2019.8980440>.

NVIDIA T4. NVIDIA. Disponível em: <https://www.nvidia.com/pt-br/data-center/tesla-t4/>. Acesso em: 7 dez. 2024.

Ultralytics YOLOv8 Documentation. Ultralytics. Disponível em: <https://docs.ultralytics.com/pt/models/yolov8/>. Acesso em: 5 dez. 2024.

Ultralytics YOLOv8. Ultralytics. Disponível em: <https://huggingface.co/Ultralytics/YOLOv8>. Acesso em: 5 dez. 2024.

Ultralytics YOLOv8. Ultralytics. Disponível em: <https://github.com/ultralytics/ultralytics>. Acesso em: 5 dez. 2024.