

Detecção e Classificação Automática de Rótulos Frontais de Produtos Alimentícios em Ênfase nos Selos da ANVISA^{1*}

Bruno Roveri Custodio(10401752)¹, Vinicius Serpa Pansan(10254346)²

¹ Faculdade de Computação e Informática – Universidade Presbiteriana Mackenzie Higienópolis

{10401752,10254346}@mackenzista.br

Resumo. Este projeto propõe o uso de técnicas de Visão Computacional para auxiliar na acessibilidade de pessoas com deficiência visual por meio da detecção de rótulos frontais em embalagens de alimentos. O foco será identificar selos obrigatórios da ANVISA que indicam “Alto em” açúcar adicionado, gordura saturada e/ou sódio. Para isso, será utilizado um modelo de detecção baseado em YOLOv8 nano, treinado com imagens de prateleiras reais, closes de embalagens em diferentes condições e datasets complementares. A solução tem como objetivo oferecer maior autonomia no cotidiano dessas pessoas, fornecendo informações nutricionais essenciais de forma rápida e acessível.

1. Introdução

A deficiência visual é uma condição que afeta milhões de pessoas no Brasil e compromete diretamente a autonomia em atividades cotidianas, como a escolha de alimentos. Desde 2020, a Agência Nacional de Vigilância Sanitária regulamentou a rotulagem nutricional frontal obrigatória em embalagens, destacando selos de advertência quando há excesso de açúcar adicionado, gordura saturada e sódio (Brasil, 2020a; Brasil, 2020b). Esses símbolos visam fornecer informações claras e acessíveis ao consumidor, mas ainda representam uma barreira para pessoas com deficiência visual. Nesse contexto, a Inteligência Artificial se apresenta como uma ferramenta eficaz para ampliar a acessibilidade, permitindo a detecção e identificação automática desses selos em produtos alimentícios. Entre as soluções mais promissoras, destacam-se os modelos da família YOLO, reconhecidos pela eficiência em tarefas de detecção de objetos em tempo real e otimizados para execução em dispositivos móveis (Ultralytics, 2023). Assim, este projeto tem como objetivo aplicar técnicas de Machine Learning e Visão Computacional para desenvolver um sistema capaz de identificar rótulos frontais obrigatórios, promovendo maior autonomia e inclusão social.

¹

2. Fundamentação Teórica

A rotulagem nutricional frontal foi regulamentada no Brasil pela Resolução RDC nº 429/2020 e pela Instrução Normativa nº 75/2020, que determinam a obrigatoriedade de símbolos informativos na face principal de embalagens quando há excesso de açúcar adicionado, gordura saturada e sódio (Brasil, 2020a; Brasil, 2020b). Essa medida visa auxiliar consumidores em suas escolhas alimentares, mas ainda apresenta barreiras de acessibilidade para pessoas com deficiência visual. Nesse contexto, soluções baseadas em Inteligência Artificial e Visão Computacional tornam-se relevantes para promover inclusão social e autonomia.

Entre os algoritmos de detecção de objetos, a família YOLO (You Only Look Once) se destaca pela eficiência em aplicações em tempo real. Sua versão mais recente, YOLOv8, disponibilizada pela Ultralytics, apresenta modelos otimizados que permitem execução em dispositivos móveis, mantendo alto desempenho e precisão (Ultralytics, 2023). Para a implementação deste projeto, será utilizada a linguagem Python, devido à sua ampla aceitação na comunidade científica, vasta biblioteca de suporte para Machine Learning e Visão Computacional, e documentação consolidada que facilita o desenvolvimento de soluções robustas e escaláveis (Python Software Foundation, 2024). Dessa forma, a proposta fundamenta-se na combinação entre a obrigatoriedade da rotulagem nutricional, a eficiência de modelos modernos de detecção e a flexibilidade do Python como linguagem de desenvolvimento.

3. Descrição do problema

O problema central abordado neste projeto reside na barreira de acessibilidade imposta a pessoas com deficiência visual em atividades rotineiras, como a escolha de produtos alimentícios em supermercados. Com a implementação da Resolução da Diretoria Colegiada nº 429 pela ANVISA em 2020, os alimentos passaram a exibir selos frontais de advertência (triângulos indicando "Alto em Açúcar Adicionado", "Alto em Gordura Saturada" e/ou "Alto em Sódio"). No entanto, esses símbolos visuais, embora essenciais para a saúde pública, não são traduzidos em informações acessíveis para quem não pode enxergá-los. A ausência de uma tecnologia que converta essa informação visual em auditiva ou tátil representa um desafio significativo à autonomia e à segurança alimentar de milhões de brasileiros. O projeto busca solucionar essa lacuna, desenvolvendo um sistema de Visão Computacional que detecte e classifique esses rótulos com precisão, permitindo que a informação seja lida em voz alta ou transmitida de forma alternativa.

4. Aspectos éticos do uso da IA e sua responsabilidade no desenvolvimento da solução

O desenvolvimento de um sistema de detecção de selos frontais de advertência envolve fatores éticos importantes, que começam desde a coleta e organização do dataset. O uso de imagens foi conduzido de forma consciente, evitando dados pessoais sensíveis e garantindo que as fotografias não expusessem indivíduos. Como o código evidência, todas as imagens foram tratadas exclusivamente como objetos de estudo técnico, sem qualquer propósito comercial ou de identificação pessoal. O processo de anotação e

validação das bounding boxes, reforçado pelas funções de visualização implementadas no notebook, foi projetado para reduzir vieses indesejados. Um modelo treinado com anotações inconsistentes pode se tornar injusto ou impreciso, especialmente para usuários dependentes de informações acessíveis. Por isso, a verificação da fidelidade das labels mostrada pela renderização gráfica das caixas é também um componente ético, pois evita que a IA produzida forneça respostas enganosas que possam afetar decisões de alimentação.

Outro aspecto ético importante é a transparência. O código deixa claro quais bibliotecas foram utilizadas, como os dados foram estruturados, como as labels foram processadas e como a distribuição das classes afeta o desempenho do modelo. Isso impede que o sistema funcione como uma "caixa-preta", fortalecendo a responsabilidade no desenvolvimento da solução. Além disso, o projeto reforça o compromisso com a inclusão: o dataset e o modelo foram desenhados para transformar uma informação visual muitas vezes inacessível em um recurso assistivo, promovendo autonomia para pessoas com deficiência visual. Esse compromisso exige que a solução mantenha baixa taxa de erros críticos, evitando tanto falsos positivos quanto falsos negativos, que poderiam comprometer a confiança do usuário no sistema.

5. DataSet

Para o desenvolvimento do projeto, será utilizada uma base de dados composta por imagens reais de prateleiras de supermercados e closes de embalagens de alimentos, capturadas em diferentes condições de iluminação, ângulos e distâncias. Ao todo, o conjunto de dados é formado por 50 imagens, distribuídas da seguinte forma: 33 imagens na classe AD_GS, em que a embalagem apresenta simultaneamente o selo de “alto em açúcar adicionado” e o selo de “alto em gordura saturada”, 8 imagens na classe GS, contendo apenas o selo de “alto em gordura saturada”, 2 imagens na classe GS_SD, em que aparecem juntos o selo de “alto em gordura saturada” e o selo de “alto em sódio”, 4 imagens na classe SD, contendo apenas o selo de “alto em sódio”, e 3 imagens na classe AD, contendo apenas o selo de “alto em açúcar adicionado”. Em alguns casos, mais de um selo aparece na mesma embalagem, o que justifica a existência das classes combinadas. O objetivo é garantir diversidade de cenários e maior robustez no treinamento do modelo. As imagens serão organizadas conforme a regulamentação da ANVISA (Brasil, 2020a; Brasil, 2020b). Sempre que necessário, os dados serão anonimizados, assegurando que não haja exposição de informações pessoais ou sensíveis. A preparação do conjunto de dados inclui a anotação manual das regiões de interesse por meio de ferramentas como LabelImg ou Roboflow, seguida de etapas de pré-processamento em Python. Esse pré-processamento envolverá redimensionamento das imagens, normalização dos pixels e técnicas de aumento de dados, como rotação, variação de brilho e recorte, com o objetivo de aumentar a capacidade de generalização do modelo. Adicionalmente, será realizada uma análise exploratória utilizando bibliotecas como pandas, matplotlib e OpenCV, o que garantirá uma base sólida para o treinamento do modelo YOLOv8 (Ultralytics, 2023; Python Software Foundation, 2024).

6. Metodologia

A metodologia do projeto segue o ciclo padrão de desenvolvimento de um modelo de Visão Computacional, estruturado em etapas de preparação de dados, treinamento e avaliação. Inicialmente, o conjunto de imagens foi coletado e anotado manualmente em ferramentas específicas, gerando arquivos no formato compatível com o YOLOv8 e organizando o dataset em pastas e arquivos de configuração YAML. Em seguida, foi realizada uma análise exploratória em Python, utilizando bibliotecas como pandas, matplotlib e OpenCV, para contabilizar a quantidade de imagens por classe, inspecionar a distribuição dos selos nas combinações de classes e analisar estatísticas dos retângulos delimitadores, como posição, área e proporção. Na etapa de preparação, aplicou-se aumento de dados por meio da própria configuração do YOLOv8, utilizando, por exemplo, espelhamento horizontal com probabilidade de 0,5, variação de matriz de até 0,015, variação de saturação de até 0,7, variação de brilho de até 0,4, translação de até 10 por cento, escala de até 50 por cento e técnica de random erasing com intensidade 0,4, além do uso de mosaico nas primeiras épocas de treino, com o objetivo de aumentar a diversidade visual e reduzir o risco de overfitting.

O modelo escolhido foi o YOLOv8 nano, inicializado com pesos pré treinados e ajustado ao problema específico dos selos frontais por meio de fine tuning, com parâmetros definidos de tamanho de imagem, tamanho de lote e número de épocas. Após o treinamento, o desempenho é avaliado no conjunto de validação utilizando métricas padrão em detecção de objetos, em especial o mean Average Precision em diferentes limiares de interseção sobre união (mAP em 0,5 e mAP em 0,5 a 0,95), além de precisão e recall globais e por classe, complementados por matriz de confusão e gráficos de evolução das métricas por época, que permitem analisar de forma sistemática a qualidade do modelo e orientar ajustes nos hiperparâmetros e nos dados.

7. Resultados

A análise realizada a partir do código permitiu avaliar detalhadamente a qualidade e a consistência do dataset utilizado para o reconhecimento dos selos frontais da ANVISA. O processamento dos arquivos de anotação no formato YOLOv8 mostrou que as três classes definidas alto-acucar, alto-gordurasaturada e alto-sodio estão presentes nas imagens e distribuídas entre diferentes combinações de selos, refletindo situações reais de embalagens com múltiplas advertências nutricionais. A varredura automática dos arquivos de label identificou corretamente os objetos e confirmou que as anotações seguem o padrão YOLO, com coordenadas normalizadas referentes ao centro, largura e altura de cada bounding box. A visualização gráfica das caixas em imagens dos diretórios train, valid e test mostrou que os selos foram anotados com precisão, mantendo alinhamento e proporção adequados às regiões verdadeiras dos triângulos de advertência.

A distribuição das classes no dataset revelou maior concentração nas categorias com múltiplos selos, especialmente em embalagens que apresentavam simultaneamente “alto

em açúcar” e “alto em gordura saturada”. Essa característica do conjunto de dados foi observada pelo agrupamento das anotações e pela contagem automática de classes por imagem, evidenciando um padrão realista do ambiente de supermercado, no qual embalagens com mais de um selo são frequentes. Também foi possível visualizar, por meio dos gráficos gerados, a presença de variação significativa nos tamanhos dos bounding boxes, resultado das diferentes distâncias de captura e ângulos das imagens. Essa diversidade visual é benéfica para o aprendizado do modelo, pois amplia a capacidade de generalização da futura detecção de selos em contextos reais.

Além disso, a verificação visual das amostras anotadas confirmou a integridade dos diretórios images e labels, e assegurou que os mapeamentos entre imagens e anotações estavam corretos. Isso garante que cada imagem possui sua respectiva descrição textual de classes e coordenadas, condição essencial para o desempenho consistente do modelo. Os resultados obtidos comprovam que o dataset está completo, bem estruturado, coerente com a taxonomia proposta e validado em termos de precisão espacial dos selos anotados. Essa organização adequada dos dados constitui a base necessária para um treinamento eficaz de modelos YOLO posteriormente e assegura que o sistema final poderá operar com maior acurácia na identificação dos selos frontais obrigatórios.

8. Conclusão

A análise detalhada do dataset e das anotações realizadas neste projeto permitiu estabelecer uma base consistente para o desenvolvimento de um sistema de detecção automática de selos frontais da ANVISA. A inspeção programática das labels, acompanhada da visualização das bounding boxes, confirmou que as anotações seguem corretamente o formato YOLOv8 e representam com precisão os selos presentes nas imagens, garantindo integridade estrutural e alinhamento semântico entre imagens e labels. A distribuição observada entre as classes com predomínio de embalagens contendo múltiplos selos, especialmente as combinações envolvendo açúcar e gordura saturada reflete fielmente o padrão encontrado em produtos alimentícios reais. Além disso, a variedade de tamanhos, proporções e posições dos selos nas imagens reforça a diversidade necessária para que modelos de Visão Computacional possam generalizar adequadamente.

A verificação da qualidade das anotações, somada à organização do dataset em diretórios de treino, validação e teste, assegura que o conjunto de dados está preparado para sustentar o treinamento de um modelo de detecção robusto. Essa preparação rigorosa reduz erros de aprendizado, melhora a estabilidade do modelo e contribui para o desenvolvimento de uma solução capaz de auxiliar pessoas com deficiência visual na identificação de informações nutricionais essenciais. Em síntese, o trabalho realizado demonstra que a construção de um sistema de IA confiável começa pela preparação minuciosa dos dados. A estruturação, validação e análise crítica do dataset fornecem a base necessária para que modelos como o YOLOv8 alcancem desempenho adequado e cumpram seu papel social na promoção da acessibilidade e autonomia do usuário final.

9. Referencia Bibliografica

BRASIL. Agência Nacional de Vigilância Sanitária. Resolução da Diretoria Colegiada nº 429, de 8 de outubro de 2020. Dispõe sobre a rotulagem nutricional dos alimentos embalados. Brasília, DF: Ministério da Saúde, 2020a. Disponível em: <https://www.in.gov.br/en/web/dou/-/resolucao-de-diretoria-colegiada-rdc-n-429-de-8-de-outubro-de-2020-282070599>.

BRASIL. Agência Nacional de Vigilância Sanitária. Rotulagem nutricional: escolha do símbolo informativo frontal e critérios para declaração de “alto em” açúcar adicionado, gordura saturada ou sódio. Instrução Normativa nº 75, de 8 de outubro de 2020. Brasília, DF: Ministério da Saúde, 2020b. Disponível em: <https://www.gov.br/anvisa/pt-br/assuntos/alimentos/rotulagem/rotulagem-nutricional>.

PYTHON SOFTWARE FOUNDATION. Python 3 Documentation. 2024. Disponível em: <https://docs.python.org/3/>.

ULTRALYTICS. Explore Ultralytics YOLOv8: documentação Ultralytics YOLOv8. 2023. Disponível em: <https://docs.ultralytics.com/pt/models/yolov8/>.