

Classificador de lixo reciclável*

Aline Y. Higa- 10402138 ¹, Gustavo G. Munhoz- 10409258 ², Karine Y. Lim Choi- 10403237 ¹, Paula A. Oliveira- 10403270 ³

¹ Faculdade de Computação e Informática (FCI)

² Universidade Presbiteriana Mackenzie – São Paulo, SP – Brasil

³ Programa de pós-graduação em Computação Aplicada – Faculdade de Computação e Informática (FCI) – Universidade Presbiteriana Mackenzie – São Paulo, SP – Brasil

{10402138, 10409258, 10409258, 10403270}@mackenzista.com.br

Abstract. *The project aims to automate the identification and separation of recyclable materials, contributing to sustainability. The system should be capable of recognizing different types of waste—such as plastic, paper, glass, organic, metal, and non-recyclables—through images captured by cameras, for example, via a mobile app. The solution can be applied in recycling centers, smart bins, and industrial processes, optimizing time and reducing operational costs. The project's innovation lies in its ability to identify multiple different objects in a single capture and classify them according to the aforementioned categories. The project seeks to encourage recycling, reduce the environmental impact of improper waste disposal, and promote a circular economy. Furthermore, the project can be expanded to recognize non-recyclable materials, increasing the efficiency of solid waste management.*

Resumo. *O projeto tem como objetivo automatizar a identificação e separação de materiais recicláveis, contribuindo para a sustentabilidade. O sistema deve ser capaz de reconhecer diferentes tipos de resíduos, como plástico, papel, vidro, orgânico, metal e não-recicláveis, por meio de imagens capturadas por câmeras, por um app mobile, por exemplo. A solução pode ser aplicada em centros de reciclagem, lixeiras inteligentes e processos industriais, otimizando o tempo e reduzindo custos operacionais. A inovação do projeto consiste em identificar, em uma única captura, vários objetos diferentes e classificá-los de acordo com as opções supracitadas. O projeto busca incentivar a reciclagem, diminuir o impacto ambiental do descarte inadequado e promover uma economia circular. Além disso, o projeto pode ser expandido para reconhecer materiais não recicláveis, aumentando a eficiência do gerenciamento de resíduos sólidos.*

*Adaptação para OpenOffice.org 1.1 feita por Roland Teodorowitsch (roland@ulbra.tche.br) em 29 mar. 2005.

1. Introdução

A geração de resíduos sólidos urbanos aumentou drasticamente e se tornou uma questão ambiental significativa no século 21. À medida que a população e o consumo aumentam, soluções devem ser desenvolvidas para auxiliar na correta disposição do lixo e na reutilização de matérias-primas recicláveis. Nesse sentido, a tecnologia é uma grande aliada, ao oferecer soluções automatizadas e eficientes para a gestão de resíduos (FACELI et al., 2021).

Neste trabalho, nosso objetivo é criar um sistema automatizado que utiliza imagens de câmeras, inclusive aquelas em dispositivos móveis, para analisá-las e classificá-las, a fim de identificar e separar recicláveis. Este sistema é projetado para reconhecer diferentes classes de resíduos — como plásticos, papel, vidro, orgânicos, metais e resíduos não recicláveis — aplicando técnicas de inteligência artificial, especialmente redes neurais convolucionais (CNNs), sendo comumente usadas para tarefas de classificação de imagens (GOODFELLOW; BENGIO; COURVILLE, 2016).

Por outro lado, um dos aspectos mais inovadores do projeto é sua concepção inteligente da viabilidade de reconhecer múltiplas coisas em uma captura, mesmo sem haver um processo de classificação correspondente, para permitir ainda maior eficiência para o objetivo prático do empreendimento. Entre as possíveis alternativas de emprego, destacam-se unidades de reciclagem, lixeiras inteligentes e processos industriais, onde a automação é vista como uma impressão de custo operacional e ganho de tempo na triagem (RUSSEL; NORVIG, 2022).

O conjunto de dados TACO (Trash Annotations in Context) é usado para treinar e aprender o modelo de classificação, disponível em tacodataset.org. Esses dados são compostos por fotografias de resíduos em seus ambientes do mundo real com uma profundidade de anotações e segmentações que possibilita a aplicação de métodos de aprendizado supervisionado em cenários mais realistas (TA-CO, 2025).

A eficácia do projeto não somente auxilia no aprimoramento dos sistemas de gestão de resíduos sólidos, mas também sublinha a pertinência da tecnologia em galvanizar a reciclagem e reduzir os custos ambientais associados a práticas inadequadas de disposição, fomentando uma economia circular. A base teórica e técnica vem de autores renomados em inteligência artificial e aprendizado de máquina (Goodfellow et al., 2016; Russel e Norvig, 2010, 2022; Chollet, 2021, entre outros) e o uso de bancos de dados públicos TACO (TA-CO, 2025), fundamentais para a validação dos modelos desenvolvidos.

2. Descrição do Problema

O aumento da população e o crescimento do consumo levam a um aumento considerável da geração de resíduos sólidos urbanos ao redor do mundo. De acordo com Faceli et al. (2021), a má gestão dos resíduos leva à degradação ambiental e à poluição do solo e dos corpos hídricos e impõe um ônus à saúde pública. Ainda existe uma abundância de materiais recicláveis descartados incorretamente na maioria das cidades — porque as políticas e programas de reciclagem não estão ativos ou porque os sistemas de triagem apropriados não estão em vigor ou onde as pessoas simplesmente não estão cientes. A segregação manual de lixo, ainda existente na maioria das estações de reciclagem, é um processo lento e caro que expõe os trabalhadores a riscos sanitários (TAN; STEINBACH; KUMAR, 2018).

Além disso, a classificação incorreta de resíduos afeta toda a cadeia de reutilização, tornando a reciclagem ineficiente. Soluções baseadas em tecnologias de inteligência artificial são exploradas para enfrentar esse desafio, particularmente incorporando modelos de visão computacional e aprendizado profundo para classificação de imagens (GOODFELLOW; BENGIO; COURVILLE, 2016). Desses métodos, as CNNs são as mais reconhecidas por sua capacidade de aprender padrões visuais complexos e têm sido amplamente utilizadas para realizar tarefas como reconhecimento de objetos e segmentação de imagens (GÉRON, 2019).

No entanto, a eficácia desses modelos está intrinsecamente ligada à qualidade e diversidade dos conjuntos de dados nos quais foram treinados. Portanto, o conjunto de dados selecionado para o trabalho foi o TACO (TA-CO, 2025), sendo um passo crucial na construção de sistemas resilientes para a identificação automatizada de resíduos. Concedido, a maioria das soluções disponíveis de ponta no mercado ou na literatura parece estar focada em detectar somente um objeto por imagem, o que as torna de uso limitado em ambientes do mundo real onde os resíduos heterogêneos estão misturados. Assim, cria-se uma lacuna prática: não há solução acessível e eficaz que permita identificar e classificar uma pluralidade de materiais recicláveis presentes em uma única captura de imagem, em tempo real e com precisão satisfatória (CHOLLET, 2021).

Dessa forma, precisamos de um sistema capaz de classificar os tipos de resíduos conforme as imagens capturadas e oferecer uma solução para classificar automaticamente os resíduos, diminuindo os custos operacionais enquanto aumenta a eficiência da reciclagem (AGGARWAL, 2021). Esta é parte da solução, em outros cenários, como lixeiras inteligentes, centros de reciclagem urbanos e indústrias, facilitando a sustentabilidade e o estabelecimento da economia circular (HRUSSELL; NORVIG, 2022).

3. Aspectos Éticos e Responsabilidades

Desde a concepção dos sistemas baseados em IA até a execução, há não somente desafios técnicos, mas também questões éticas levantadas neste trabalho que devem ser consideradas ao longo do caminho. Ao decidir ou automatizar processos que influenciam direta ou indiretamente a sociedade, a IA precisa estar vinculada por princípios como justiça, explicação, responsabilidade e segurança (RUSSEL; NORVIG, 2010).

As implicações sociais do uso de IA para este projeto específico — automação da triagem de resíduos sólidos — são vastas. A substituição da separação manual por sistemas inteligentes pode influenciar as comunidades que dependem economicamente da coleta seletiva — como catadores de materiais recicláveis. Portanto, o desenvolvimento e a aplicação da tecnologia devem considerar medidas de inclusão social, como treinar esses trabalhadores para operar e monitorar sistemas inteligentes (FACELI et al., 2021).

Outro tópico crítico é a transparência algorítmica, ou seja, a capacidade de entender e explicar as decisões tomadas por modelos de IA. Sistemas de aprendizado profundo, particularmente redes neurais convolucionais (CNNs) usadas neste projeto, são "caixas pretas" (GOODFELLOW; BENGIO; COURVILLE, 2016) — os resultados fornecidos por esses sistemas podem ser difíceis de interpretar. Para reduzir esse problema, é necessário trabalhar com classes de explicabilidade (IA explicável), tornando os resultados da classificação compreensíveis para usuários não especialistas (RASCHKA; MIRJALILI, 2017).

Por outro lado, questões éticas também surgem da representação e qualidade dos dados com os quais este sistema foi treinado. Outra armadilha é com dados assimétricos — como ter imagens com muito mais frequência de certo tipo de resíduo ou ambientes — e se o sistema é aplicado a um contexto de renderização, isso pode levar a resultados discriminatórios ou imprecisos (VANDERPLAS, 2017). Para isso, o emprego de conjuntos de dados públicos como o TACO (TA-CO, 2025) deve ser complementado por uma avaliação cuidadosa de sua cobertura, equilíbrio e variação.

Além disso, os desenvolvedores de sistemas de IA têm responsabilidades legais e éticas quando se trata da confiabilidade e segurança de seus sistemas. Para sistemas que tomam decisões que impactam o meio ambiente e a cidade, como sistemas de triagem de resíduos, tais falhas podem afetar o sucesso de toda a cadeia de reciclagem. Nesse sentido, práticas de validação, testes rigorosos e auditoria contínua devem ser adotadas, garantindo que o sistema opere em padrões aceitáveis de erro e desempenho (AGGARWAL, 2021). A implementação é um aspecto: é crucial alinhar o desenvolvimento de tecnologia com os Objetivos de Desenvolvimento Sustentável (ODS) da ONU para que os avanços técnicos se traduzam não somente em ganhos técnicos, mas também em benefícios sociais, ambientais e econômicos. Estas são as áreas de três subcomponentes importantes para uma IA responsável e ética (LUGER, 2015): redução de impactos ambientais, economia circular e justiça social.

4. Dataset

O conjunto de dados utilizado neste projeto é o TACO (*Trash Annotations in Context*),

cuja finalidade é disponibilizar imagens de diferentes ambientes — como ruas, praias e parques — onde o lixo se apresenta em seu contexto real (TA-CO, 2025). Cada fotografia conta com anotações em formato JSON, trazendo informações sobre as regiões onde os materiais aparecem (via *bounding boxes* ou máscaras de segmentação) e as categorias dos itens (por exemplo, “plástico”, “metal”, “papel”, “vidro”, entre outras). Para compreender a fundo o TACO, foi realizada uma análise exploratória no ambiente Python, registrada em um arquivo Jupyter Notebook. Inicialmente, efetuou-se o carregamento dos arquivos JSON e sua conversão para *dataframes* com auxílio de bibliotecas como *pandas*. Em seguida, geraram-se gráficos por meio de *matplotlib* com o objetivo de examinar a frequência das categorias. Essa análise apontou para um desbalanceamento dos dados, sobretudo pelo grande volume de itens plásticos em comparação a outras classes, como metal ou vidro (Chollet, 2021; Goodfellow; Bengio; Courville, 2016).

Além disso, observou-se que algumas imagens apresentavam baixa qualidade — iluminação precária, foco impreciso ou ruídos — o que poderia impactar negativamente o processo de aprendizado. Diante disso, foram feitas duas intervenções principais. A primeira consistiu em agrupar subclasses muito específicas (por exemplo, “garrafa plástica” e “embalagem plástica”) em categorias mais abrangentes, assegurando que a classificação final esteja alinhada às práticas comuns de separação de resíduos (Gerón, 2019).

5. Metodologia e Resultados Esperados

Para o desenvolvimento do sistema de classificação automática de resíduos sólidos recicláveis, adotou-se uma abordagem baseada em aprendizado supervisionado com o uso de redes neurais convolucionais (CNNs), técnica consolidada na literatura para tarefas de visão computacional, como reconhecimento e classificação de imagens (GOODFELLOW; BENGIO; COURVILLE, 2016; GÉRON, 2019; CHOLLET, 2021). O processo metodológico dividiu-se em etapas que englobam desde o tratamento dos dados até o treinamento e validação do modelo de inteligência artificial.

O ponto de partida foi a utilização do conjunto de dados TACO (Trash Annotations in Context), uma base pública que reúne imagens de resíduos em ambientes do mundo real, como ruas, parques e praias, contendo anotações detalhadas em formato JSON. Essas anotações incluem informações sobre a localização dos objetos (*bounding boxes*) e suas respectivas categorias, como “garrafa plástica” ou “lata de refrigerante” (TA-CO, 2025). Para manipular essas informações, utilizou-se um ambiente Python com o auxílio das bibliotecas *pandas*, *json*, *matplotlib* e *seaborn*, possibilitando uma análise exploratória inicial dos dados.

Durante a exploração, identificou-se um desbalanceamento significativo entre as classes, com predominância de resíduos plásticos em relação a categorias como metal, vidro ou papel. Além disso, observou-se a existência de subclasses muito específicas que poderiam dificultar o processo de generalização do modelo. Diante disso, realizou-se um processo de reclassificação e agregação das categorias, agrupando termos como “garrafa plástica”, “copos descartáveis” e “sacos plásticos” sob uma única classe abrangente: “Plástico”. Esse mapeamento semântico permitiu a criação de uma nova estrutura hierárquica de classes, mais adequada às práticas comuns de separação de resíduos urbanos (FACELI et al., 2021; TAN; STEINBACH; KUMAR, 2018).

Com base na metodologia descrita, espera-se que o modelo desenvolvido seja capaz de

realizar a classificação de resíduos sólidos com elevado grau de precisão, mesmo em imagens complexas e em ambientes ruidosos. A expectativa é que o sistema alcance uma acurácia global superior a 85% e apresente desempenho satisfatório também nas classes minoritárias, como vidro e metal, graças ao uso de técnicas de balanceamento e aumento de dados (AGGARWAL, 2021; CHOLLET, 2021).

Além disso, o classificador deve demonstrar capacidade de identificar múltiplos tipos de resíduos presentes em uma única captura de imagem — fator essencial para seu uso prático em contextos reais, como centros de reciclagem urbanos, lixeiras inteligentes e indústrias de processamento de materiais. A meta é que o sistema funcione em tempo quase real, viabilizando sua integração com dispositivos móveis e sensores embarcados.

Do ponto de vista prático, espera-se que a adoção dessa solução contribua para a redução de custos operacionais com a triagem manual, estimada em cerca de 20%, além de aumentar a eficiência geral da cadeia de reciclagem. O sistema também poderá ser utilizado como ferramenta educativa, promovendo a conscientização ambiental por meio de sua incorporação em aplicativos interativos.

No aspecto social, prevê-se que a tecnologia, se aplicada de forma responsável, contribua para a transição para uma economia circular, conforme preconizado nos Objetivos de Desenvolvimento Sustentável (ODS) da ONU. Para isso, recomenda-se que sua implementação leve em conta aspectos éticos, como a inclusão de trabalhadores da reciclagem por meio de capacitação para operar os novos sistemas, e a mitigação de vieses relacionados à qualidade e diversidade do conjunto de dados utilizado (LUGER, 2015; VANDERPLAS, 2017; RUSSEL; NORVIG, 2022).

6. Referencias bibliograficas

As referências abaixo foram utilizadas diretamente no corpo do texto do projeto para embasar conceitos, justificar escolhas metodológicas e fundamentar teoricamente as soluções propostas:

AGGARWAL, Charu C. *Artificial Intelligence: A Textbook*. New York: Springer, 2021.

dwsadwCHOLLET, François. *Deep Learning with Python*, 2ª ed. Shelter Island: Manning, 2021.

FACELI, Katti et al. *Inteligência artificial: uma abordagem de aprendizado de máquina*, 2ª ed. Rio de Janeiro: LTC, 2021.

GÉRON, Aurélien. *Hands-On Machine Learning with Scikit-Learn and TensorFlow*, 2ª ed. Sebastopol: O'Reilly, 2019.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. *Deep Learning*. Cambridge: MIT Press, 2016.

HRUSSELL, Stuart; NORVIG, Peter. *Inteligência Artificial: Uma Abordagem Moderna*, 4ª ed. GEN LTC, 2022.

LUGER, George F. *Inteligência Artificial*, 6ª ed. São Paulo: Pearson Education do Brasil, 2015.

RASCHKA, Sebastian; MIRJALILI, Vahid. *Python Machine Learning*, 3ª ed. Birmingham: Packt, 2017.

RUSSEL, Stuart; NORVIG, Peter. *Artificial Intelligence: A Modern Approach*, 3ª ed. Upper Saddle River: Pearson, 2010.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. *Introduction to Data Mining*, 2ª ed. Upper Saddle River: Pearson, 2018.

TA-CO. *TACO Dataset – Trash Annotations in Context*. [S.l.]: TACO Dataset, [2020]. Disponível em: <https://tacodataset.org/dataset>. Acesso em: 5 abr. 2025.

VANDERPLAS, Jake. *Python Data Science Handbook*. Sebastopol: O'Reilly, 2017.

7. Bibliografia

A bibliografia a seguir constitui o conjunto de obras utilizadas como base teórica para o desenvolvimento do projeto. Ela inclui tanto as referências citadas diretamente no texto quanto outras fontes de apoio técnico e conceitual:

AGGARWAL, Charu C. *Artificial Intelligence: A Textbook*.

CHOLLET, François. *Deep Learning with Python*.

FACELI, Katti et al. *Inteligência artificial: uma abordagem de aprendizado de máquina*.

GÉRON, Aurélien. *Hands-On Machine Learning with Scikit-Learn and TensorFlow*.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. *Deep Learning*.

HRUSSELL, Stuart; NORVIG, Peter. *Inteligência Artificial: Uma Abordagem Moderna*.

LUGER, George F. *Inteligência Artificial*.

RASCHKA, Sebastian; MIRJALILI, Vahid. *Python Machine Learning*.

RUSSEL, Stuart; NORVIG, Peter. *Artificial Intelligence: A Modern Approach*.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. *Introduction to Data Mining*.

VANDERPLAS, Jake. *Python Data Science Handbook*

8. Referncias

TA-CO. *TACO Dataset – Trash Annotations in Context*. [S.l.]: TACO Dataset, [2020]. Disponível em: <https://tacodataset.org/dataset>. Acesso em: 5 abr. 2025.

AGGARWAL, Charu C. *Artificial Intelligence: A Textbook*. New York: Springer: 2021.

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

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

GOODFELLOW, Ian; BENGIO, Yoshua, COURVILLE, Aaron. *Deep Learning*. Cambridge: MIT Press, 2016.

RASCHKA, Sebastian; MIRJALILI, Vahid. Python Machine Learning. 3 ed. Birmingham: Packt, 2017.

RUSSEL, Stuart; NORVIG, Peter. Artificial Intelligence: A Modern Approach. 3 ed. Upper Saddle River: Pearson, 2010.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. Introduction to Data Mining. 2 ed. Upper Saddle River: Pearson, 2018.

VANDERPLAS, Jake. Python Data Science Handbook. Sebastopol: O'Reilly, 2017.

FACELI, Katti et al. Inteligência artificial: uma abordagem de aprendizado de máquina. 2ª Ed. Rio de Janeiro: LTC Livros Técnicos e Científicos, 2021.

LUGER, George F. Inteligência Artificial - 6ª ed. São Paulo: Pearson Education do Brasil, 2015.

HRUSSELL, Stuart; NORVIG, Peter. Inteligência artificial: Uma Abordagem Moderna - 4ª. Ed. GEN LTC, 2022.