

## **ALFASSIST: CAA e IA para reconhecimento de emoções de crianças não verbais**

Eduardo dos Santos Pereira <sup>1</sup>

Bianca Goulart dos Santos <sup>2</sup>

### **RESUMO**

Este projeto desenvolve o aplicativo Alfassist, destinado a tablets, que integra o Sistema de Comunicação por Troca de Figuras (PECS) com Inteligência Artificial (IA). O objetivo é facilitar a comunicação e interpretar as emoções de crianças não verbais, promovendo sua autonomia. O app utiliza um banco de dados criado no Roboflow, com cerca de mil imagens, permitindo que a IA reconheça expressões faciais e ofereça opções visuais para atender às necessidades diárias dos usuários. Além disso, a comunicação é tornada mais acessível por meio de pictogramas selecionados do site ARASAAC, combinados com técnicas de IA, como a plataforma Ultralytics YOLO, para o treinamento do modelo. A metodologia aplicada é qualitativa, com o desenvolvimento de um código funcional testado. A equipe é multidisciplinar, composta por um Engenheiro Eletrônico responsável pela programação em Python e uma psicopedagoga, que implementa intervenções e garante o funcionamento da aplicação. O referencial teórico abrange autores como Bondy e Frost (PECS), Resnick (IA na educação) e legislações sobre Tecnologia Assistiva e Inclusão. Resultados preliminares mostram que o aplicativo detecta emoções e sugere ações com base em movimentos da cabeça, com planos para aprimorar a seleção por meio do olhar. O sistema também apresenta potencial de personalização, adaptando-se às necessidades específicas de cada usuário. O projeto visa melhorar a qualidade de vida de crianças não verbais, promovendo uma comunicação mais adaptativa, com possibilidades de expansão para outras condições que afetam a fala.

**Palavras-chave:** Comunicação Aumentativa e Alternativa; Inteligência Artificial, Crianças não verbais; Inclusão, Tecnologia Assistiva.

### **INTRODUÇÃO**

Este estudo investigou como a Inteligência Artificial pode atuar como aliada da Educação Especial, explorando possibilidades para promover a comunicação de crianças com necessidades específicas. Por meio de um código funcional que utilizou a CAA (Comunicação Aumentativa e Alternativa), foi possível desenvolver a comunicação através de figuras PECS (Sistema de Comunicação por Troca de Figuras) relacionadas ao cotidiano de crianças não verbais. O recurso criado constitui uma ferramenta de tecnologia assistiva que apresenta possibilidade de personalização de acordo com as necessidades individuais de cada usuário.

---

<sup>1</sup> Mestrando em Ciência da Computação na linha de pesquisa Sistemas Embarcados da Universidade Federal do Rio Grande do Sul - RS, [@edu.nh.rs.gov.br](mailto:@edu.nh.rs.gov.br).

<sup>2</sup> Mestra em Educação na linha de pesquisa Tecnologias Digitais na Educação da Universidade Federal do Rio Grande do Sul - RS, [biancags@edu.nh.rs.gov.br](mailto:biancags@edu.nh.rs.gov.br).



O planejamento do projeto teve como base a utilização do site Roboflow<sup>3</sup> para a criação de um banco de dados contendo mais de 1000 imagens. Essas imagens foram cuidadosamente classificadas a partir de dois critérios principais: as emoções expressas e a posição do rosto, com foco específico na região dos olhos.

Para o treinamento do modelo, foi empregado o site Ultralytics YOLO<sup>4</sup>, plataforma que reuniu todas as imagens do banco de dados. Por meio deste processo, a rede neural foi ensinada a reconhecer as categorias estabelecidas, tornando-se capaz de identificar reações faciais do adulto cadastrado. Com base nesse reconhecimento, o sistema passou a oferecer opções comunicativas adequadas a cada situação identificada.

A seleção das opções comunicativas apresentadas pelo sistema pôde ser identificada por meio do movimento da cabeça, proporcionando uma forma alternativa de interação para as crianças não verbais. Este método de seleção permite que os usuários façam suas escolhas entre as figuras PECS oferecidas.

O estudo apresenta um produto funcional com prova de conceito disponibilizada neste artigo, demonstrando a viabilidade da aplicação da Inteligência Artificial como recurso de tecnologia assistiva na Educação Especial. Ainda assim, foi destacada a necessidade de futuros investimentos para o aprimoramento e refinamento da ferramenta desenvolvida, visando expandir suas funcionalidades e alcançar seu pleno potencial.

## METODOLOGIA

A metodologia aplicada neste estudo é de natureza qualitativa, estruturada em etapas sequenciais e complementares que viabilizaram o desenvolvimento da ferramenta Alfassist. O processo metodológico iniciou-se com a identificação da necessidade comunicativa apresentada por crianças não verbais, fase essencial para compreender as demandas específicas desse público e direcionar o desenvolvimento de uma solução tecnológica adequada.

Em seguida, procedeu-se à pesquisa de referencial teórico, etapa que fundamentou cientificamente o projeto por meio da revisão de estudos sobre CAA, sistemas PECS, tecnologia assistiva e aplicações de Inteligência Artificial na Educação

---

<sup>3</sup> <https://roboflow.com/>

<sup>4</sup> <https://universe.roboflow.com/ultralytics-yolo>



Especial. Esta revisão permitiu embasar as escolhas metodológicas e técnicas adotadas ao longo do desenvolvimento.

A terceira etapa consistiu na criação do banco de imagens, utilizando a plataforma Roboflow para organizar mais de 1000 imagens classificadas a partir de critérios específicos relacionados às emoções e à posição do rosto, com foco na região dos olhos. Paralelamente, realizou-se o planejamento das figuras PECS que comporiam as opções comunicativas do sistema, selecionando imagens relacionadas ao cotidiano das crianças e complementando o repertório visual com recursos disponibilizados pela plataforma ARASAAC (Portal Aragonês de Comunicação Aumentativa e Alternativa).

Subsequentemente, foi conduzido o treinamento do modelo de Inteligência Artificial por meio da plataforma Ultralytics YOLO, que processou o banco de dados imagético e ensinou a rede neural a reconhecer as categorias estabelecidas, capacitando o sistema a identificar reações faciais e oferecer opções comunicativas contextualizadas.

Por fim, foram realizados testes com o protótipo funcional, cujos resultados demonstraram a viabilidade da proposta e forneceram a prova de conceito apresentada neste artigo, evidenciando que o sistema é capaz de reconhecer expressões faciais e permitir que as crianças selecionem opções comunicativas por meio do movimento da cabeça.

## REFERENCIAL TEÓRICO

A partir da necessidade observada nas crianças não verbais, buscou-se identificar quais estratégias poderiam ser utilizadas e se adequar às demandas metodológicas de criação da Alfassist. Nesse contexto, os dispositivos de CAA, como pranchas de símbolos, leitores de voz, aplicativos para tablets e outros recursos que facilitam a comunicação por meios não convencionais, permitem que essas crianças desenvolvam habilidades comunicativas de forma independente, promovendo a autonomia, desenvolvimento emocional e a inclusão social (REINERT JUNIOR; COUTINHO, 2024).

Para a construção do sistema Alfassist, foram utilizadas como recurso imagético as figuras do PECS, embora não seguindo todos os passos do treinamento formal do método. O PECS foi desenvolvido como um sistema de treinamento comunicativo único, criado para contornar algumas limitações de outras estratégias utilizadas com crianças autistas. O sistema permite que crianças com autismo e deficiências



relacionadas desenvolvam um sistema de comunicação funcional, de rápida aquisição e auto-iniciado (BONDY; FROST, 1994). Complementando o banco de imagens, recorreu-se ao site ARASAAC<sup>5</sup>, plataforma que conta com a colaboração de criadores e programadores de diferentes tipos de software relacionados à comunicação aumentativa e acessibilidade cognitiva. Esses recursos disponibilizados pelo ARASAAC auxiliam as pessoas a melhorar sua interação e participação na sociedade.

A escolha por integrar tecnologia assistiva ao desenvolvimento da Alfassist fundamenta-se no papel que esses recursos desempenham na promoção da inclusão e da autonomia de pessoas com deficiência, especialmente no contexto educacional. A Lei Brasileira de Inclusão (Lei nº 13.146/2015) define tecnologia assistiva como produtos, equipamentos, dispositivos, recursos, metodologias, estratégias, práticas e serviços que objetivem promover a funcionalidade relacionada à atividade e à participação da pessoa com deficiência, visando garantir autonomia, independência, qualidade de vida e inclusão social.

No campo da Educação Especial, as ferramentas de tecnologia assistiva tornam-se essenciais para eliminar barreiras comunicativas que impedem a plena participação de estudantes não verbais ou com deficiências de comunicação no processo de aprendizagem.

A integração da Inteligência Artificial como ferramenta de apoio à CAA na Alfassist alinha-se à perspectiva defendida por Mitchel Resnick, professor do MIT (Instituto de Tecnologia de Massachusetts) e autor do livro "Lifelong Kindergarten: Cultivating Creativity through Projects, Passion, Peers, and Play" (2017). Resnick apresenta uma visão fundamental sobre como devemos integrar a Inteligência Artificial na educação, argumentando que, à medida que novas tecnologias surgem na sociedade, é necessário decidir conscientemente se e como integrá-las aos ambientes de aprendizado.

O autor enfatiza que essas escolhas são extremamente importantes, pois podem levar a resultados e implicações muito diferentes, dependendo de como a tecnologia é concebida e implementada. Nesse sentido, a Alfassist foi desenvolvida com o propósito de colocar a tecnologia a serviço das necessidades comunicativas reais das crianças, respeitando suas individualidades e promovendo sua autonomia expressiva.

---

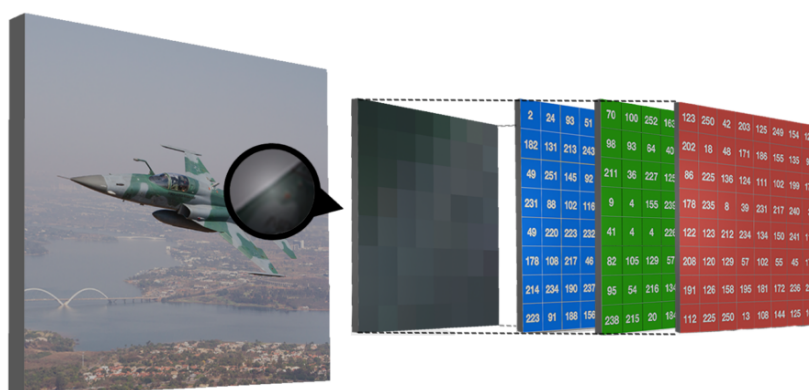
<sup>5</sup> <https://arasaac.org/world>



A Alfassist utiliza visão computacional, uma área da inteligência artificial dedicada a capacitar máquinas para interpretar e compreender informações visuais de forma automatizada.

O processo começa com a transformação de imagens em dados processáveis computacionalmente. Toda imagem digital é representada internamente como uma matriz numérica tridimensional, onde cada pixel armazena valores numéricos que codificam sua informação cromática. No caso de imagens coloridas no padrão RGB (Red, Green, Blue), cada pixel é composto por três canais de cor, com valores de intensidade que variam de 0 a 255 para cada componente - vermelho, verde e azul. A combinação desses três valores determina a cor final exibida em cada posição da imagem, como apresentado na figura a seguir.

**Figura 1-** Classificação das emoções



Fonte: <https://sigmoidal.ai/fundamentos-da-formacao-da-imagem-na-visao-computacional/>

Com essas matrizes numéricas em mãos, entram em ação as Redes Neurais Convolucionais (CNNs). Essas redes são especialmente projetadas para processar dados visuais e conseguem aprender automaticamente a identificar padrões nas imagens, desde características simples como bordas e texturas, até padrões complexos como formas de objetos específicos.

Para realizar a detecção de objetos em tempo real, a Alfassist utiliza o YOLOv8, uma das arquiteturas mais avançadas de CNN disponíveis atualmente. O nome "YOLO" vem de "You Only Look Once" (Você Olha Apenas Uma Vez), e essa nomenclatura descreve exatamente seu diferencial: ao contrário de métodos tradicionais que analisam a imagem em múltiplas etapas, o YOLO processa a imagem inteira de uma única vez, identificando e localizando múltiplos objetos simultaneamente. Essa abordagem não só



garante maior velocidade no processamento, como também mantém alta precisão na detecção, tornando possível a análise de vídeo em tempo real com latência mínima.

## RESULTADOS E DISCUSSÃO

A seguir será apresentado o passo a passo da criação da Alfassist para ilustrar o processo de desenvolvimento e a possibilidade de acesso ao repositório do GitHub.

### Etapa 1: Criação do Projeto no Roboflow

Foi realizada a criação do projeto na plataforma Roboflow, onde foram definidas e configuradas as classes de objetos que desejávamos detectar. Durante essa etapa, foi feito o planejamento das categorias de interesse, garantindo que cada classe representasse corretamente os elementos a serem reconhecidos pelo modelo de detecção, conforme a figura abaixo.

**Figura 2-** Classificação das emoções

**Classes & Tags**

Classes 5 Tags 0

[What is a class?](#) ☐ Lock Classes [+ Add](#) [Modify Classes](#)

Search classes...

Sort By Class Ascending

| COLOR                                                                               | CLASS NAME  | COUNT |
|-------------------------------------------------------------------------------------|-------------|-------|
|  | DOR         | 289   |
|  | FELIZ       | 285   |
|  | INDIFERENCA | 160   |
|  | ROSTO       | 1,117 |
|  | TRISTE      | 281   |

**Fonte:** Autores, 2025.

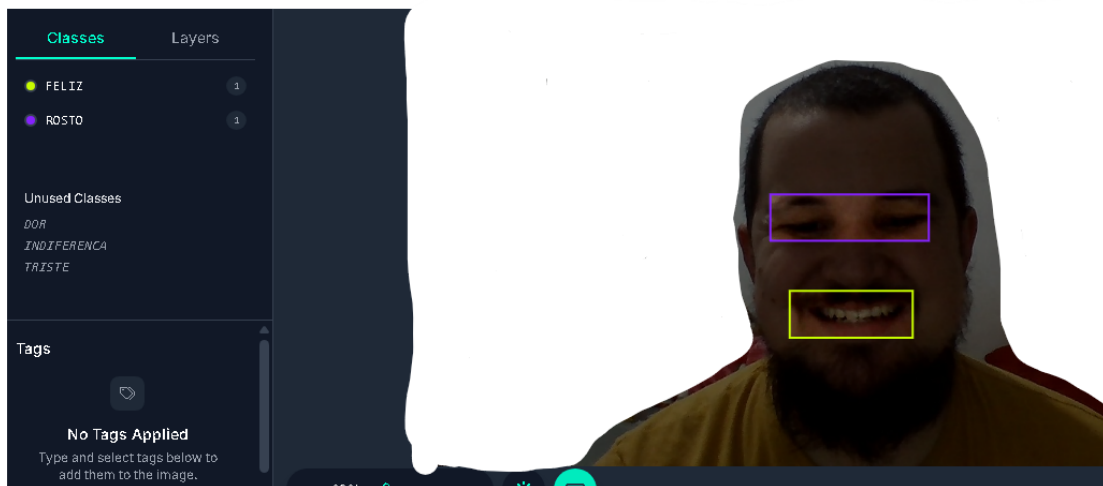
Após a criação das classes, realizamos o carregamento das imagens que comporiam nosso dataset, indicada na etapa 2.

### Etapa 2: Carregamento e Anotação das Imagens



Em seguida, procedemos com a anotação manual de cada imagem, identificando as classes correspondentes, como apresentado na figura 3.

**Figura 3-** Anotação das imagens



**Fonte:** Autores, 2025.

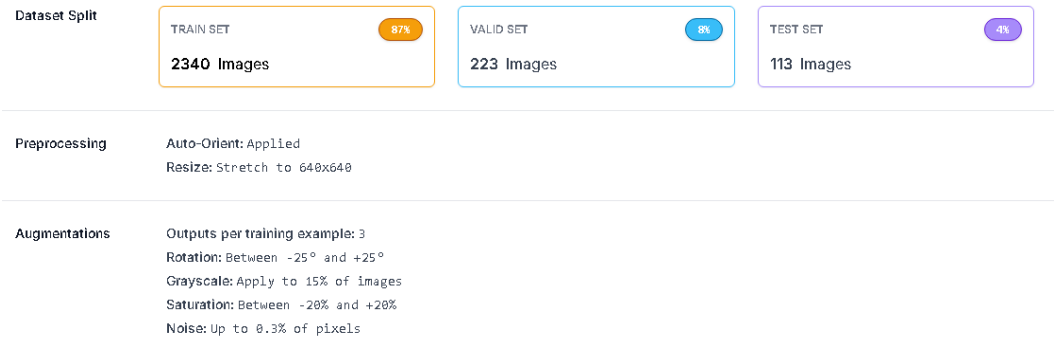
Em todas as imagens foi marcada a classe "rosto", enquanto em cada uma delas também foi indicada a emoção exibida pelo indivíduo, garantindo assim um conjunto de dados bem estruturado e coerente para o treinamento do modelo. Após a anotação de todas as imagens, a etapa seguinte consistiu na criação e preparação do dataset no Roboflow, exemplificado na etapa 3.

### **Etapas 3: Preparação e Processamento do Dataset**

Nessa fase, as imagens foram organizadas e divididas automaticamente em conjuntos de treinamento, validação e teste. Foram aplicadas diferentes configurações de pré-processamento e aumento de dados (data augmentation) para melhorar a diversidade e a qualidade do conjunto de treino. Segue abaixo a configuração utilizada e o tamanho final do dataset após o processamento.



Figura 4- Anotação das imagens



Fonte: Autores, 2025.

A configuração final ajustou automaticamente a orientação das imagens, redimensionou todas para 640x640 pixels, aplicou três variações por imagem com rotações entre -25° e +25°, converteu 15% das imagens para tons de cinza, variou a saturação entre -20% e +20% e adicionou ruído em até 0,3% dos pixels. Essas transformações tornaram o conjunto de dados mais robusto e adequado para o treinamento do modelo em diferentes condições visuais. Com o dataset finalizado no Roboflow, a etapa seguinte foi o treinamento do modelo utilizando a plataforma Ultralytics HUB.

Etapa 4: Treinamento do Modelo no Ultralytics HUB

Esse ambiente permite o carregamento direto do conjunto de dados e o treinamento do YOLOv8 em nuvem, dispensando configurações locais complexas. O uso de GPUs dedicadas é essencial para acelerar o processo, permitindo que milhares de imagens sejam processadas de forma paralela e reduzindo significativamente o tempo total de treinamento em comparação ao uso de CPUs. Durante essa fase, foram acompanhadas as métricas de desempenho e ajustadas as configurações para alcançar os melhores resultados.

Para dúvidas sobre o uso do Ultralytics HUB e o processo de treinamento, recomenda-se assistir ao vídeo "Roboflow Integration with Ultralytics HUB | Episode 27", disponível no canal Ultralytics (2024).

Etapa 5: Exportação e Integração do Modelo



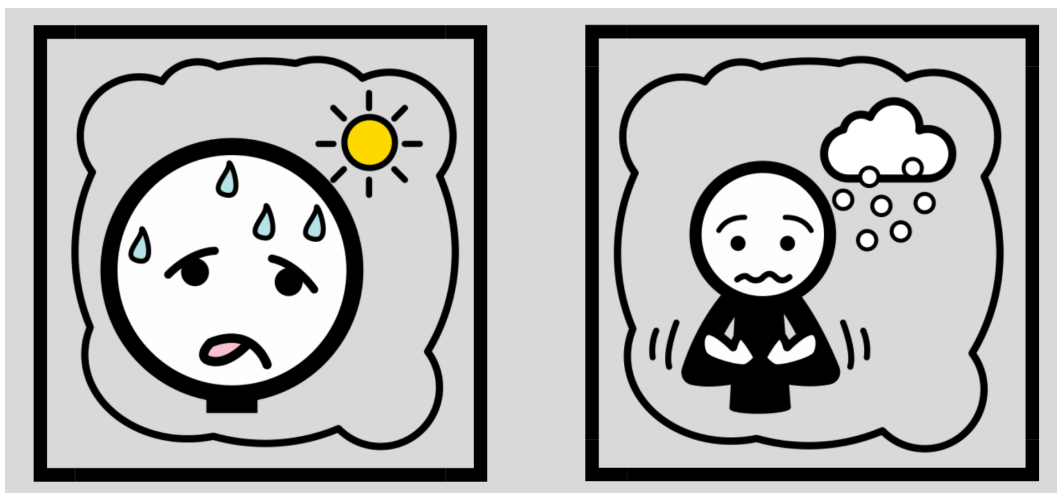
Após o término do treinamento, o modelo gerado é exportado em um arquivo com extensão .pt, que contém os pesos e parâmetros aprendidos durante o processo. Esse arquivo é então carregado em um script Python, permitindo a utilização do modelo dentro da lógica do programa. No código, é possível realizar a detecção e visualização das classes previstas a partir de uma imagem de entrada.

No caso deste projeto, a entrada é fornecida em tempo real por uma webcam, possibilitando que o sistema identifique rostos e emoções de forma dinâmica. Essa integração entre o modelo treinado e o código Python representa a etapa final do pipeline, em que o modelo passa a ser efetivamente aplicado e testado em situações práticas.

### **Etapa 6: Implementação das Figuras PECS**

A etapa seguinte consiste na aplicação das figuras PECS. As imagens 4, 5 e 6 apresentadas a seguir foram elaboradas no site Canva. Para ilustrá-las, foi produzido um vídeo que reúne todas as figuras selecionadas para este projeto, o qual pode ser totalmente personalizado de acordo com as necessidades de cada usuário.

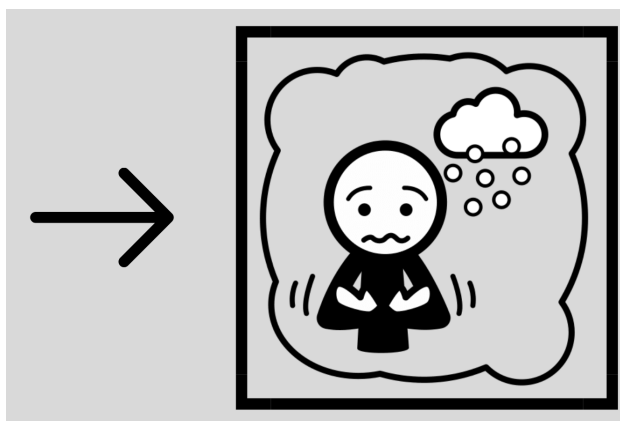
**Figura 5-** Opções de seleção



**Fonte:** Autores, 2025.

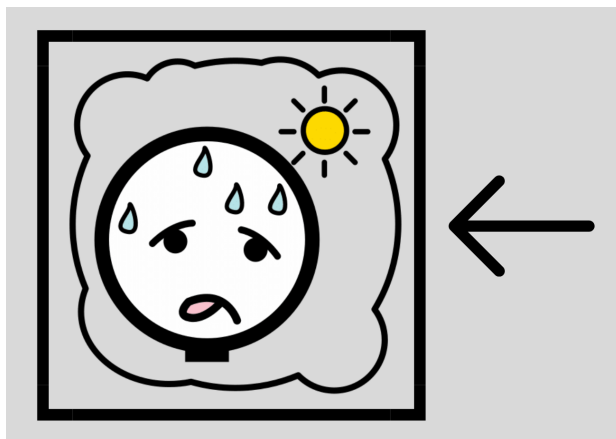
A imagem a seguir retrata a seleção para o lado esquerdo.



**Figura 6-** Sensação de frio

Fonte: Autores, 2025.

A imagem a seguir retrata a seleção para o lado direito.

**Figura 7-** Sensação de calor

Fonte: Autores, 2025.

As ações do sistema seguem as instruções da criança, cabendo ao cuidador atender às necessidades comunicadas, como sensações de frio ou calor. O aplicativo Alfassist funciona de forma dinâmica, atualizando os frames em tempo real conforme as classes detectadas pelo modelo, proporcionando respostas visuais interativas e adaptadas ao comportamento do usuário. O código completo está disponível no repositório do GitHub para acesso público, e recomenda-se o vídeo "YOLOv8: How to Train for Object Detection on a Custom Dataset" do canal Roboflow (2023) como material complementar para esclarecer dúvidas sobre o processo de treinamento em conjuntos de dados personalizados.



## CONSIDERAÇÕES FINAIS

Este estudo demonstrou que a integração entre Inteligência Artificial e Tecnologia Assistiva pode promover avanços significativos na comunicação de crianças não verbais, constituindo-se como um recurso para a Educação Especial e inclusiva. O desenvolvimento da Alfassist evidenciou a viabilidade técnica de sistemas baseados em visão computacional e YOLOv8 para o reconhecimento de expressões faciais e oferecimento de opções comunicativas contextualizadas por meio de figuras PECS, permitindo que as crianças expressem suas necessidades através do movimento da cabeça. A metodologia aplicada, estruturada desde a criação do banco de imagens no Roboflow até o treinamento do modelo no Ultralytics HUB, resultou em uma prova de conceito funcional e personalizável, disponibilizada publicamente no GitHub para beneficiar a comunidade. Embora os resultados obtidos sejam promissores e demonstrem o potencial da ferramenta como instrumento de autonomia comunicativa, reconhece-se a necessidade de investimentos futuros para o aprimoramento do sistema, incluindo testes com maior número de usuários, expansão do repertório de figuras PECS e refinamento dos algoritmos de detecção. Espera-se que este trabalho inspire novas pesquisas na área e contribua efetivamente para a construção de uma sociedade mais inclusiva, onde a tecnologia seja aliada na superação de barreiras comunicativas e na garantia do direito à expressão de todas as crianças.

## REFERÊNCIAS

BONDY, A. S.; FROST, L. A. The Picture Exchange Communication System. **Focus on Autistic Behavior**, v. 9, n. 3, p. 1-19, 1994.

BRASIL. Lei nº 13.146, de 6 de julho de 2015. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). **Diário Oficial da União**, Brasília, DF, 7 jul. 2015. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2015/lei/l13146.htm](https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm). Acesso em: 25 out. 2025.

DEVIOTEDULARDO. **Alfassist**. 2024. Disponível em: <https://github.com/DevIoTEduardo/Alfassist>. Acesso em: 25 out. 2025.

DEVIOTEDULARDO. **Demonstração ALFASSIST**. YouTube, 2024. Disponível em: [https://www.youtube.com/watch?v=0i7hIK\\_IDcQ](https://www.youtube.com/watch?v=0i7hIK_IDcQ). Acesso em: 25 out. 2025.



GOVERNMENT OF ARAGON. **ARASAAC**: Aragonese Portal of Augmentative and Alternative Communication. Zaragoza, 2008. Disponível em: <https://arasaac.org>. Acesso em: 25 ago. 2025.

KATTENBORN, T. et al. A review of convolutional neural networks in computer vision. **Artificial Intelligence Review**, v. 57, 2024. Disponível em: <https://link.springer.com/article/10.1007/s10462-024-10721-6>. Acesso em: 25 out. 2025.

REDMON, J. et al. Real-Time Flying Object Detection with YOLOv8. **arXiv preprint arXiv:2305.09972**, 2023. Disponível em: <https://arxiv.org/abs/2305.09972>. Acesso em: 25 out. 2025.

REINERT JUNIOR, A. J.; COUTINHO, D. J. G. O impacto dos dispositivos de Comunicação Alternativa e Aumentativa (CAA) na inclusão de alunos com deficiências de comunicação. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, São Paulo, v. 10, n. 11, p. 2020-2030, nov. 2024.

RESNICK, M. **Jardim de infância para a vida toda**: por uma aprendizagem criativa, mão na massa e relevante para todos. Tradução: Mariana Casetto Cruz, Livia Rulli Sobral. Revisão técnica: Carolina Rodeghiero, Leo Burd. Porto Alegre: Penso, 2020.

ROBOFLOW. **Roboflow Integration with Ultralytics HUB | Episode 27**. YouTube, 2024. Disponível em: <https://www.youtube.com/watch?v=Mp82DKBnQqw>. Acesso em: 25 out. 2025.

SIGMOIDAL. **Fundamentos da formação da imagem na visão computacional**. Disponível em: <https://sigmoidal.ai/fundamentos-da-formacao-da-imagem-na-visao-computacional/>. Acesso em: 25 out. 2025.

ULTRALYTICS. **Como treinar modelos YOLOv8 no Ultralytics HUB**. YouTube, 2024. Disponível em: <https://www.youtube.com/watch?v=MdF6x6Y3rFQ>. Acesso em: 25 out. 2025.

ULTRALYTICS. **YOLOv8**. Documentação oficial. Disponível em: <https://docs.ultralytics.com/pt/models/yolov8/>. Acesso em: 25 out. 2025.

VISÃO COMPUTACIONAL. **Identificação, detecção, reconhecimento e segmentação de imagem e objetos**. 2024. Disponível em: <https://visaocomputacional.com.br/identificacao-deteccao-reconhecimento-e-segmentacao-de-imagem-e-objetos/>. Acesso em: 25 out. 2025.

VISO.AI. **YOLOv8 guide**. 2024. Disponível em: <https://viso.ai/deep-learning/yolov8-guide/>. Acesso em: 25 out. 2025.

YOUTUBE. **YOLOv8: How to Train for Object Detection on a Custom Dataset**. 2024. Disponível em: <https://www.youtube.com/watch?v=wuZtUMEiKWY>. Acesso em: 25 out. 2025.

