

**Recomendação personalizada de produtos a partir de características demográficas do consumidor usando YOLO**

**Renyclay Quadros de Souza**

Trabalho de Conclusão de Curso  
MBA em Inteligência Artificial e Big Data

# UNIVERSIDADE DE SÃO PAULO

## Instituto de Ciências Matemáticas e de Computação

---

Recomendação personalizada de produtos a partir  
de características demográficas do consumidor  
usando YOLO

*Renyclay Quadros de Souza*

---



Renyclay Quadros de Souza

## Recomendação personalizada de produtos a partir de características demográficas do consumidor usando YOLO

Trabalho de conclusão de curso apresentado ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial.

Orientador: Prof. Dr. Francisco A. Rodrigues.

USP - São Carlos

2025

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi  
e Seção Técnica de Informática, ICMC/USP,  
com os dados inseridos pelo(a) autor(a)

Qlr QUADROS DE SOUZA, RENYCLAY  
Recomendação personalizada de produtos a partir  
de características demográficas do consumidor usando  
YOLO / RENYCLAY QUADROS DE SOUZA; orientador  
FRANCISCO APARECIDO RODRIGUES. -- São Carlos, 2025.  
66 p.

Trabalho de conclusão de curso (MBA em  
Inteligência Artificial e Big Data) -- Instituto de  
Ciências Matemáticas e de Computação, Universidade  
de São Paulo, 2025.

1. visão computacional. 2. YOLO. 3. inteligência  
artificial de borda. 4. sistemas de recomendação. I.  
APARECIDO RODRIGUES, FRANCISCO, orient. II. Título.

## ERRATA

Errata			
Folha	Linha	Onde se lê	Leia-se



## DEDICATÓRIA

*A minha esposa pela compreensão,  
carinho e apoio incansável. Aos  
meus amados filhos.*



## AGRADECIMENTOS

Ao Dr. Francisco A. Rodrigues, pela sua paciência e disponibilidade.

A Dra. Solange O. Rezende, pela compreensão das dificuldades encontradas para a finalização do projeto

## EPÍGRAFE

A primeira regra de qualquer tecnologia utilizada nos negócios é que a automação aplicada a uma operação eficiente aumentará a eficiência. A segunda é que a automação aplicada a uma operação ineficiente aumentará a ineficiência.

Bill Gates (1995)

## RESUMO

SOUZA, R. Q. **Título: Recomendação personalizada de produtos a partir de características demográficas do consumidor usando YOLO.** 2025. 66 f. Monografia (MBA em Ciências de Dados) – Centro de Ciências Matemáticas Aplicadas à Indústria, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2025.

O presente trabalho propõe o desenvolvimento de uma solução embarcada de recomendação personalizada de produtos no varejo físico, baseada na análise demográfica automatizada de clientes utilizando técnicas de visão computacional. O objetivo principal é criar um sistema capaz de identificar características como gênero e faixa etária dos consumidores em tempo real, visando aumentar a taxa de conversão e melhorar a experiência de compra através de recomendações contextualizadas. A metodologia empregou o algoritmo YOLO11n para detecção de pessoas e o framework DeepFace com backend RetinaFace para inferência de atributos demográficos, implementados em um dispositivo Raspberry Pi 5. O sistema foi validado em três etapas complementares em uma farmácia na cidade de Santo Antônio de Jesus, Bahia. A solução opera integralmente de forma local, sem armazenamento ou transmissão de dados biométricos, em conformidade com a Lei Geral de Proteção de Dados (LGPD). Com os resultados obtidos conclui-se que a integração entre algoritmos de aprendizado profundo e dispositivos de borda viabiliza aplicações práticas de personalização em ambientes físicos de varejo, conciliando eficiência técnica, baixo custo de implementação e respeito à privacidade dos usuários, abrindo perspectivas para o desenvolvimento de sistemas inteligentes centrados no consumidor.

Palavras-chave: visão computacional, YOLO, inteligência artificial de borda, sistemas de recomendação.



## ABSTRACT

SOUZA, R. Q. **Personalized product recommendation based on consumer demographic characteristics using YOLO**. 2025. 66 f. Monografia (MBA em Ciências de Dados) – Centro de Ciências Matemáticas Aplicadas à Indústria, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2025.

This work proposes the development of an embedded solution for personalized product recommendation in physical retail, based on automated demographic analysis of customers using computer vision techniques. The main objective is to create a system capable of identifying characteristics such as gender and age group of consumers in real time, aiming to increase conversion rates and improve the shopping experience through contextualized recommendations. The methodology employed the YOLO11n algorithm for person detection and the DeepFace framework with RetinaFace backend for demographic attribute inference, implemented on a Raspberry Pi 5 device. The system was validated in three complementary stages at a pharmacy in Santo Antônio de Jesus, Bahia. The solution operates entirely locally, without storing or transmitting biometric data, in compliance with the General Data Protection Law (LGPD). Based on the results obtained, it is concluded that the integration between deep learning algorithms and edge devices enables practical personalization applications in physical retail environments, combining technical efficiency, low implementation cost, and respect for user privacy, opening perspectives for the development of consumer-centered intelligent systems.

Keywords: computer vision, YOLO, edge artificial intelligence, recommendation systems.



## LISTA DE ILUSTRAÇÕES

Figura 1 – Agrupamento de áreas de conhecimento.....	29
Figura 2 – Funcionamento do YOLO.....	33
Figura 3 – Estabelecimento comercial selecionado.....	37
Figura 4 – Local de Visualização de Recomendações e Visão de Captura de Imagens.....	38
Figura 5 – Pipeline utilizada na solução.....	38
Figura 6 – Produtos Raspberry utilizados no projeto .....	42
Figura 7 – Unificação de tarefas do RetinaFace .....	44
Figura 8 – Estrutura do RetinaFace .....	45
Figura 9 – Processo criativo de propaganda de produtos .....	47
Figura 10 – Métricas de Desempenho Técnico .....	49
Figura 11 – Avaliação Média dos Voluntários (escala 1–5).....	51
Figura 12 – Distribuição por faixa etária e gênero .....	52
Figura 13 – Distribuição total por gênero .....	53



## LISTA DE TABELAS

Tabela 1 – Critérios para seleção dem empresa colaboradora com projeto de pesquisa.....	35
Tabela 2 – Especificação Técnica de Produtos utilizados no projeto.....	39
Tabela 3 – Avaliação Média dos Voluntários (escala 1–5).....	49



## LISTA DE ABREVIATURAS E SIGLAS

CNN	–	<i>Convolutional Neural Networks</i>
FPS	–	<i>Frame por segundo</i>
IA	–	Inteligência Artificial
LGPD	–	Lei Geral de Proteção de Dados
LLM	–	<i>Large Language Models</i>
mAP	–	<i>Mean Average Precision</i>
ROI	–	<i>Region of Interest</i>
SSD	–	<i>Single Shot MultiBox Detector</i>
YOLO	–	<i>You Only Look Once</i>



## LISTA DE SÍMBOLOS

°C Graus Celsius

" Polegadas



## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	26
<b>2 REVISÃO BIBLIOGRÁFICA</b> .....	28
<b>2.1 Inteligência Artificial e Big Data no Varejo</b> .....	28
<b>2.2 Inteligência Artificial e Aprendizado Profundo</b> .....	29
2.2.1 Redes Neurais Convolucionais (CNNs).....	30
<b>2.3 Visão Computacional e Detecção de Objetos</b> .....	31
2.3.1 Evolução da Família YOLO.....	31
<b>2.4 Reconhecimento Facial e Análise Demográfica</b> .....	33
<b>2.5 Inteligência Artificial de Borda</b> .....	34
<b>2.6 Aspectos Éticos, Viés e Legislação (LGPD)</b> .....	35
<b>3 METODOLOGIA</b> .....	36
<b>3.1 Ambiente Experimental</b> .....	36
3.1.1 Escolha do Local.....	37
<b>3.2 Estrutura Geral do Projeto</b> .....	38
<b>3.3 Infraestrutura utilizada na solução</b> .....	39
3.3.1 Definição de Equipamento de Edge IA.....	39
3.3.2 Especificação de Produtos Utilizados.....	40
<b>3.4 Ambiente de Desenvolvimento</b> .....	42
3.4.1 YOLO versão 11.....	43
3.4.2 Análise Facial e Inferência demográfica com DeepFace.....	44
<b>3.5 Modelo de Recomendação</b> .....	45
<b>3.6 Considerações Éticas e de Privacidade</b> .....	47
<b>4 EXPERIMENTOS E ANÁLISE DE RESULTADOS</b> .....	48
<b>4.1 Validação Técnica de Detecção e Contagem de Pessoas</b> .....	48
<b>4.2 Avaliação com Voluntários e Validação das Inferências</b> .....	49
<b>4.3 Operação Contínua e Análise de Comportamento em Campo</b> .....	51
<b>4.4 Discussão Geral dos Resultados</b> .....	53
<b>5 CONCLUSÕES</b> .....	54
<b>5.1 Contribuições do Trabalho</b> .....	54
<b>5.2 Limitações Identificadas</b> .....	55
<b>5.3 Trabalhos Futuros</b> .....	55

<b>5.4 Considerações Finais.....</b>	<b>56</b>
<b>REFERÊNCIAS.....</b>	<b>57</b>
<b>APÊNDICE A – Modelo de Aviso ao Público.....</b>	<b>64</b>
<b>APÊNDICE B – Modelo de Termo de Consentimento.....</b>	<b>65</b>

# 1 INTRODUÇÃO

Em um cenário de negócios cada vez mais competitivo, muitas empresas enfrentam dificuldades em oferecer produtos que realmente atendam às preferências e necessidades de seus clientes em tempo real (CHRISTOPHER, 2000). No varejo físico, oferecer experiências de compra personalizadas ainda é um desafio, principalmente devido à limitação de dados coletados em tempo real e à dificuldade de compreender o perfil do cliente durante a jornada de compra. Como muitas vezes as decisões de compra são influenciadas por fatores como idade, gênero e outros aspectos demográficos (RYBINTSEV, 2017), a falta de um sistema inteligente para capturar e analisar esses dados de forma automatizada prejudica a criação de estratégias eficazes de *marketing* e vendas, o que pode resultar em recomendações irrelevantes, impactando negativamente a experiência do cliente e as taxas de conversão.

A visão computacional tem avançado significativamente com o uso de redes neurais profundas (GOODFELLOW; BENGIO; COURVILL, 2016). Entre as abordagens mais utilizadas, o algoritmo *You Only Look Once* - YOLO (ULTRALYTICS, 2025) vem ganhando reconhecimento por sua capacidade de realizar detecção de objetos em tempo real com precisão e eficiência (RYBINTSEV, 2017) (REDMON; FARHADI, 2017, 2018). A arquitetura YOLO realiza a detecção e a classificação de múltiplos objetos em uma única passagem pela rede neural, tornando-a particularmente adequada para aplicações embarcadas e sistemas que exigem respostas imediatas (TERVEN, CÓRDOVA-ESPARZA, 2023).

Paralelamente, o conceito de computação de borda (*edge computing*) vem se consolidando como uma alternativa viável ao processamento em nuvem. Nesse paradigma, o processamento de dados é executado localmente, no próprio dispositivo embarcado, reduzindo a latência, preservando a privacidade dos usuários e diminuindo a dependência de conectividade constante com servidores externos. Essa abordagem é especialmente relevante em aplicações de varejo e atendimento ao público, onde há necessidade de resposta imediata e de preservação de dados sensíveis (SHI et al, 2016).

Assim, o objetivo da pesquisa é desenvolver uma solução automatizada e embarcada que, utilizando tecnologia de reconhecimento e análise de dados demográfica baseada no YOLO, permita sugerir produtos personalizados com base no gênero e idade dos clientes, visando aumentar a taxa de conversão e melhorar a experiência de compra.

A implementação dessa solução também pode fornecer *insights* valiosos sobre o comportamento dos clientes, possibilitando ajustes contínuos nas estratégias de *marketing*,

melhoramento da percepção da marca e fidelizando o público com abordagem mais centrada no consumidor. Ou seja, espera-se não só a recomendação produtos com maior probabilidade de serem adquiridos, mas também uma oportunidade da empresa ajustar sua comunicação para um formato mais assertivo, gerando um vínculo emocional e aumentando a satisfação do cliente.

Além da dimensão tecnológica, este trabalho considera aspectos éticos e legais envolvidos no uso de dados visuais. A coleta e o processamento de imagens são realizados de forma anônima, sem armazenamento de dados pessoais identificáveis, em conformidade com os princípios da Lei Geral de Proteção de Dados - LGPD (BRASIL, 2018). Está sendo considerada uma abordagem que busque equilibrar o potencial inovador da solução com a responsabilidade social e a proteção da privacidade do usuário, princípios fundamentais no uso ético da IA (JOBIN; IENCA; VAYENA, 2019).

Durante o desenvolvimento e a redação deste Trabalho de Conclusão de Curso, foram utilizadas ferramentas de Inteligência Artificial baseadas em modelos de linguagem de grande porte (*Large Language Models* – LLM's), especificamente ChatGPT (OPENAI, 2025) e Gemini (GOOGLE, G., 2025), com a finalidade exclusiva de apoiar atividades de revisão linguística, aprimoramento textual, sugestões de reformulação e busca bibliográfica. O uso dessas ferramentas foi conduzido de forma ética e responsável, sem substituição das etapas autorais de análise, interpretação e redação. Todo o conteúdo conceitual, metodológico e conclusivo apresentado é de autoria integral do pesquisador, que assume total responsabilidade pelo texto final.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 Inteligência Artificial e *Big Data* no Varejo

A Inteligência Artificial (IA) compreende o campo da ciência da computação dedicado ao estudo e à construção de sistemas capazes de perceber seu ambiente, processar informações e tomar ações que antes eram exclusivas da inteligência humana (RUSSELL; NORVIG, 2020). Segundo Stonebraker (2012), *Big Data* é um conceito que engloba as vertentes: grande volume de dado, seja para a realização de análise com escopo grandes ou pequenos; grande velocidade para análise e processamento de intenso fluxo de entrada de dados; e grande variedade de dados com as mais diversas fontes.

A personalização é um dos fatores mais relevantes na fidelização de clientes e no aumento das taxas de conversão (JANNACH, 2021). Em especial, recomendações baseadas em características demográficas tendem a aumentar a percepção de relevância das ofertas, gerando maior engajamento e satisfação. Contudo, tais abordagens exigem modelos robustos, capazes de realizar inferências rápidas e confiáveis a partir de dados visuais, sem comprometer a privacidade dos indivíduos. A capacidade de interpretar imagens, estimar idade, gênero e até estados emocionais, permite o desenvolvimento de soluções inteligentes voltadas à experiência do consumidor e à melhoria de processos de recomendação em tempo real (MOLLAHOSSEINI; HASANI; MAHOOR, 2019).

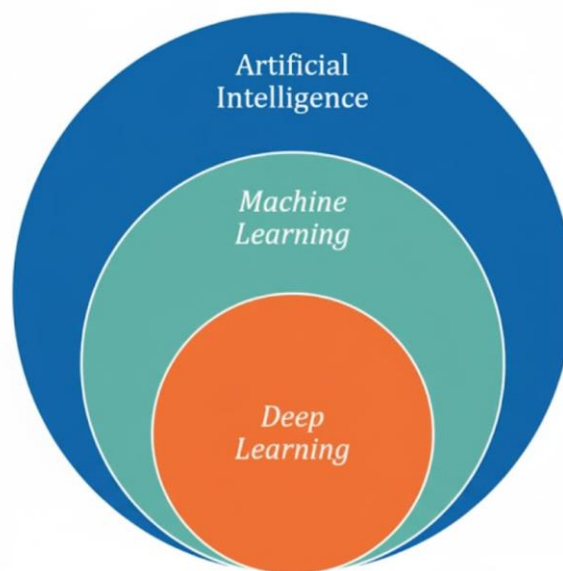
Sistemas de recomendação baseados em IA estão sendo empregados em plataformas digitais, mas sua aplicação em ambientes físicos ainda enfrenta barreiras, especialmente pela limitação de dados (AGGARWAL, 2016). O uso de câmeras, sensores e novas abordagens de aprendizado de máquina tem aberto novas possibilidades para compreender o comportamento do consumidor em tempo real.

No contexto do varejo, a integração entre IA e *Big Data* tem viabilizado a criação de estratégias centradas no cliente, baseadas em dados comportamentais, históricos de compra e interações em tempo real (DAVENPORT; RONANKI, 2018) que tem impulsionado a personalização de experiências, o aumento da eficiência operacional e a otimização de campanhas de *marketing*.

## 2.2 Inteligência Artificial e Aprendizado Profundo

Dentro da área de IA, o Aprendizado de Máquina (*Machine Learning*) figura como uma subárea responsável pelo desenvolvimento de algoritmos que permitem que sistemas aprendam padrões e tomem decisões com base em dados, dispensando regras fixas e explicitamente programadas (MITCHELL, 1997). O Aprendizado Profundo (*Deep Learning*), por sua vez, constitui um subconjunto do Aprendizado de Máquina e é caracterizado pelo uso de redes neurais artificiais compostas por múltiplas camadas, capazes de extrair representações hierárquicas e complexas a partir de dados brutos (LECUN; BENGIO; HINTON, 2015). A relação entre as áreas de conhecimento mencionadas podem ser observadas na Figura 1 abaixo.

Figura 1 –Agrupamento de áreas de conhecimento



Fonte: Adaptada de Goodfellow, Bengio e Courvill (2016)

Um dos pilares fundamentais do Aprendizado Profundo está no algoritmo de retropropagação do erro (*backpropagation*), que ajusta iterativamente os pesos da rede para reduzir o erro entre as previsões e os valores reais (ZHANG; BENGIO, 2021). Diversos tipos de algoritmos tem sido utilizados ao longo do tempo e com os mais diversos formatos de conexão entre camadas das redes utilizadas. Desde soluções com camadas totalmente conectadas, a alternativas com conexões locais, as proposta tem como principal objetivo o de fazer com que os modelos aprendam sobre representações e identifiquem correspondência nos dados.

Embora ofereça avanços de desempenho em desafios como reconhecimento de imagens, fala e linguagem natural, o aprendizado profundo também impõe desafios, como a necessidade de grandes volumes de dados rotulados, alto custo computacional e dificuldades de interpretabilidade (GOODFELLOW; BENGIO; COURVILLE, 2016). Diante disso, o uso de modelos de Aprendizado Profundo tem sido impulsionada por três fatores principais: o aumento do poder computacional, a disponibilidade de grandes volumes de dados e avanços em algoritmos de otimização.

### 2.2.1 Redes Neurais Convolucionais (CNNs)

As Redes Neurais Convolucionais (*Convolutional Neural Networks* – CNNs) constituem um marco significativo do avanço da Aprendizagem Profunda aplicada à visão computacional. Inspiradas na organização do córtex visual humano, as CNNs são compostas por camadas convolucionais capazes de extrair automaticamente padrões espaciais e hierárquicos de uma imagem (LECUN; BENGIO; HINTON, 2015).

O uso de filtros convolucionais permite capturar relações locais entre pixels e formar representações invariantes a deslocamentos e ruídos, eliminando grande parte da necessidade de engenharia manual de características — uma limitação recorrente em métodos clássicos de processamento de imagem.

A popularização das CNNs ocorreu a partir da arquitetura AlexNet, que venceu o ImageNet Large Scale Visual Recognition Challenge (ILSVRC) em 2012, demonstrando desempenho significativamente superior a todas as abordagens anteriores (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). Desde então, redes mais profundas e eficientes como VGGNet (SIMONYAN; ZISSERMAN, 2015) e ResNet (HE et al., 2016) revolucionaram o campo, tornando-se referência em aplicações de reconhecimento e detecção de objetos.

Mais recentemente, técnicas como *transfer learning*, quantização e compressão de modelos têm viabilizado a execução de CNNs em dispositivos embarcados e ambientes de borda (*edge computing*), reduzindo a latência e o consumo de energia. Esse avanço é fundamental para aplicações em ambientes físicos de varejo, nos quais o processamento local é essencial para oferecer respostas rápidas, personalizadas e em conformidade com princípios de privacidade e segurança de dados.

## 2.3 Visão Computacional e Detecção de Objetos

A visão computacional é um ramo da Inteligência Artificial dedicado a capacitar máquinas a interpretar e compreender o conteúdo de imagens e vídeos, buscando emular a percepção visual humana. Seu objetivo é extrair informações significativas de dados visuais, permitindo que sistemas automatizados realizem tarefas como detecção, segmentação, reconhecimento.

A detecção de objetos consiste em identificar e localizar instâncias de diferentes classes em uma imagem. Essa tarefa exige a combinação de precisão e eficiência computacional, uma vez que o modelo deve delimitar com exatidão as regiões dos objetos por meio de *bounding boxes* e, simultaneamente, classificá-los

Historicamente, a detecção de objetos evoluiu de abordagens clássicas baseadas em engenharia manual de atributos, como *Haar Cascades* (VIOLA; JONES, 2001) e *Histogram of Oriented Gradients* (DALAL; TRIGGS, 2005), para arquiteturas de Aprendizado Profundo, capazes de aprender automaticamente as representações discriminativas necessárias para a tarefa (LECUN; BENGIO; HINTON, 2015).

Entre os modelos modernos, destacam-se o Faster R-CNN (REN et al., 2015), o Single Shot MultiBox Detector (SSD) (LIU et al., 2016) e a família YOLO (REDMON et al., 2016). Esses métodos diferenciam-se pelo compromisso entre acurácia e velocidade de inferência. Enquanto o Faster R-CNN apresenta elevado desempenho de detecção, ele é mais custoso em termos computacionais. Já o YOLO foi projetado desde sua primeira versão para processamento em tempo real, dividindo a imagem em regiões e realizando simultaneamente a detecção e a classificação de objetos em uma única passagem pela rede (REDMON; FARHADI, 2017).

Essa característica confere ao YOLO um diferencial importante frente a outras abordagens: sua arquitetura unificada e otimizada para velocidade e eficiência, o que o torna particularmente adequado para aplicações embarcadas e de resposta imediata, como monitoramento em tempo real, veículos autônomos e sistemas inteligentes no varejo (TERVEN; CÓRDOVA-ESPARZA, 2023).

### 2.3.1 Evolução da Família YOLO

A arquitetura YOLO foi introduzida por Redmon et al. (2016) com o propósito de realizar detecção de objetos em tempo real a partir de uma única passagem da imagem pela rede — uma abordagem inovadora frente aos métodos de múltiplas etapas. Desde então, a família

YOLO evoluiu continuamente, incorporando avanços estruturais, novos *backbones* e estratégias de treinamento mais robustas.

O YOLOv1 (REDMON et al., 2016) foi pioneiro ao propor a unificação da detecção em um único modelo convolucional. Entretanto, apresentava limitações em objetos pequenos e sobrepostos. O YOLOv2 (REDMON; FARHADI, 2017) introduziu o uso de *anchor boxes*, *batch normalization* e redimensionamento dinâmico, aprimorando a precisão e a estabilidade do treinamento. Já o YOLOv3 (REDMON; FARHADI, 2018) adicionou camadas residuais e detecção em múltiplas escalas, o que o tornou mais eficiente em cenários complexos.

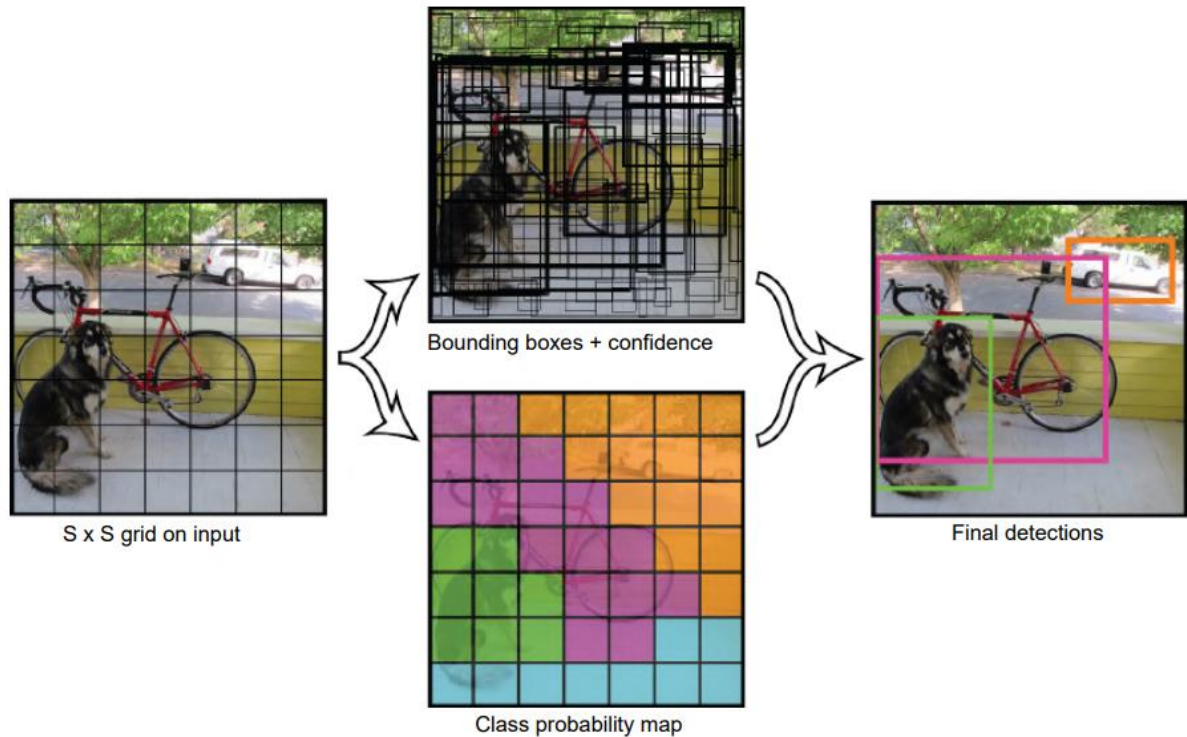
Com o YOLOv4 (BOCHKOVSKIY; WANG; LIAO, 2020), a arquitetura incorporou inovações como CSPDarknet53, *mosaic data augmentation* e *self-adversarial training*, alcançando desempenho de ponta sem depender de hardware especializado. O YOLOv5 (ULTRALYTICS, 2025) migrou completamente para o framework PyTorch, facilitando o uso, o treinamento e a exportação de modelos para diferentes formatos como ONNX, TensorRT e CoreML.

Em seguida, o YOLOv6 (LI et al., 2022), criado pela Meituan, aprimorou a eficiência em edge devices, enquanto o YOLOv7 (WANG et al., 2022) introduziu o conceito de *trainable bag of freebies*, melhorando o desempenho sem aumentar o custo de inferência. O YOLOv8 (ULTRALYTICS, 2025) trouxe um design modular e suporte nativo a segmentação e rastreamento.

Mais recentemente, o YOLOv11 (ULTRALYTICS, 2025) consolidou uma nova geração de detectores, com backbone C2f-Darknet, neck baseado em BiFPN e estratégias avançadas de dynamic label assignment, atingindo melhor mAP (*mean average precision*) e menor tempo de inferência. O menor tipo de modelo do YOLOv11n (nano), apresenta cerca de 4,3 milhões de parâmetros, operando com excelente eficiência em dispositivos embarcados como o Raspberry Pi 5, mantendo precisão próxima aos modelos maiores (TERVEN; CÓRDOVA-ESPARZA, 2023).

Em síntese, a evolução da família YOLO reflete o avanço contínuo das técnicas de aprendizado profundo aplicado à visão computacional, conciliando velocidade, precisão e portabilidade — atributos fundamentais para aplicações práticas de reconhecimento e recomendação em tempo real. Seu funcionamento é ilustrado na Figura 2.

Figura 2 – Funcionamento do YOLO



Fonte: Redmon et al. (2016)

## 2.4 Reconhecimento Facial e Análise Demográfica

O reconhecimento facial tem se consolidado como uma das aplicações mais expressivas da visão computacional, permitindo que sistemas automatizados detectem, localizem e analisem rostos humanos em imagens e vídeos (TAIGMAN et al., 2014). Embora, historicamente, o foco dessa tecnologia tenha sido a identificação de indivíduos, abordagens recentes passaram a explorar sua utilização em contextos não identificatórios, voltados à análise demográfica, ou seja, à inferência de atributos como idade e gênero sem associação direta a uma identidade específica (ROTHER; TIMOFTE; VAN GOOL, 2015).

Do ponto de vista técnico, o reconhecimento facial é uma etapa necessária para a análise demográfica, pois permite delimitar e padronizar a região da face antes da inferência. A razão fundamental da necessidade da delimitação da face é que as características que permitem inferir gênero e idade estão concentradas principalmente no rosto como, por exemplo: estrutura óssea, textura da pele e proporções faciais. A detecção precisa da face possibilita que o sistema alinhe e normalize a imagem, reduzindo distorções causadas por ângulos, iluminação ou expressões.

Essa padronização é essencial para que os modelos de inferência demográfica, que trabalham com variações sutis nos traços faciais, atinjam resultados consistentes.

Nos últimos anos, o avanço do aprendizado profundo tem impulsionado fortemente o estado da arte nessa área. Modelos baseados em redes neurais convolucionais (CNNs) extraem representações profundas (*embeddings*) das faces, permitindo reconhecer ou estimar atributos com elevado grau de precisão (DENG et al., 2019). Bibliotecas de código aberto como DeepFace (TAIGMAN et al., 2014) e InsightFace (GUO, 2021) tornaram-se amplamente utilizadas por combinar robustez, facilidade de integração e suporte a estimativas de gênero e idade. Ambas são baseadas em extração de *embeddings* faciais que podem ser aplicados a tarefas de análise demográfica.

Apesar dos avanços, o reconhecimento e a análise facial ainda enfrentam desafios significativos. Entre eles, destacam-se o viés demográfico nos conjuntos de dados, que pode levar a estimativas menos precisas para determinados grupos (BUOLAMWINI; GEBRU, 2018), e as variações ambientais — como iluminação, oclusões e expressões — que afetam diretamente o desempenho dos modelos. Para aplicações comerciais, como a proposta deste trabalho, a busca por equilíbrio entre precisão, ética e privacidade é essencial: o objetivo não é identificar indivíduos, mas entender perfis de público de forma agregada e respeitosa.

## 2.5 Inteligência Artificial de Borda

A evolução da Inteligência Artificial tem caminhado lado a lado com o avanço da infraestrutura computacional. Nos últimos anos, o aumento da demanda por processamento em tempo real tem impulsionado um novo paradigma tecnológico: a Computação de Borda (*Edge Computing*). Quando esse paradigma se funde com tecnologias e modelos de Inteligência Artificial, temos a Inteligência Artificial de Borda (*Edge AI*), isto é, é abordagem que consiste na execução modelos de Aprendizado de Máquina diretamente em dispositivos locais sem depender exclusivamente da nuvem para processar e gerar inferências (SHI et al., 2016).

Assim, ao contrário do modelo tradicional baseado em *cloud computing*, no qual grandes volumes de dados são enviados a servidores remotos para análise, a *Edge AI* realiza o processamento junto à origem dos dados. Essa mudança reduz significativamente a latência, o consumo de banda de rede e os riscos associados ao tráfego de informações sensíveis, além de garantir maior autonomia em ambientes com conectividade limitada. Em cenários como o varejo físico, onde decisões precisam ser tomadas em frações de segundo, a resposta imediata oferecida pelo processamento local torna-se um diferencial competitivo.

Além da velocidade, a *Edge AI* traz benefícios em privacidade e conformidade legal. Como o processamento ocorre no próprio dispositivo, dados não precisam ser transmitidos nem armazenados externamente, o que reduz riscos e facilita a adequação à LGPD (BRASIL, 2018).

O avanço da computação de borda tem sido impulsionado por melhorias tanto em *hardware* quanto em modelos de IA otimizados. Dispositivos como o Raspberry Pi 5, o NVIDIA Jetson Nano, o Google Coral e placas com aceleradores de inferência Neural Processing Units (NPU) possibilitam a execução de redes neurais profundas com desempenho cada vez mais eficiente (CHENG et al., 2020). Paralelamente, *frameworks* como TensorFlow Lite, ONNX Runtime e PyTorch Mobile permitem converter e otimizar modelos para ambientes embarcados, aplicando técnicas de quantização, *pruning* e compilação específica de *hardware*, sem perdas significativas de precisão. Esses avanços tornaram viável o uso de algoritmos complexos de visão computacional, como os modelos da família YOLO, em plataformas de baixo custo.

A convergência entre IA embarcada, privacidade de dados e personalização em tempo real possibilita uma oportunidade estratégica nas aplicações modernas de Inteligência Artificial. No varejo, essa integração amplia as possibilidades de interação com o cliente, combinando eficiência operacional e respeito à privacidade.

## 2.6. Aspectos Éticos, Viés e Legislação (LGPD)

O uso de tecnologias de visão computacional levanta desafios éticos e legais relacionados à privacidade, consentimento e segurança dos dados. Diversas diretrizes internacionais, como as da *European Commission's Ethics Guidelines for Trustworthy AI* (2019) enfatizam que sistemas de IA devem ser transparentes, explicáveis e alinhados aos direitos fundamentais dos indivíduos (JOBIN; IENCA; VAYENA, 2019).

No Brasil, a LGPD (BRASIL, 2018) — estabelece diretrizes rigorosas sobre o tratamento de dados pessoais, incluindo imagens que possam identificar indivíduos. De acordo com essa lei, a coleta de dados biométricos requer consentimento explícito, finalidade específica e mecanismos de anonimização sempre que possível. É de suma importância que modelos de inferência que operem de forma anônima, sem armazenamento permanente dos dados originais e de maneira a reduzir riscos de violação de privacidade.

### 3 METODOLOGIA

#### 3.1 Ambiente Experimental

Para a avaliação das hipóteses propostas no presente trabalho, considerou-se importante utilizar a solução em um estabelecimento físico, com variação demográfica de consumidores ao longo dos dias e que tivesse um espaço físico nem tão grande, que tornasse prejudicasse a personalização da recomendação, nem tão pequeno que prejudicasse a correta segmentação dos clientes. Considerando esses e outras questões, foram criados alguns critérios de seleção dos estabelecimento parceiro, conforme Tabela 1 abaixo.

Tabela 1 – Critérios para seleção de empresa colaboradora com projeto de pesquisa

<b>Critério</b>	<b>Requisito</b>
Requisito técnico obrigatório	<ul style="list-style-type: none"> <li>▪ TV/monitor com entrada HDMI disponível e liberada para uso durante o período do projeto.</li> <li>▪ Ponto de energia (tomada próxima) para alimentar o equipamento.</li> </ul>
Condições de aceitação	<ul style="list-style-type: none"> <li>▪ Aceita instalação de câmera e equipamento associado (equipe de instalação associada ao projeto).</li> <li>▪ Autoriza a transmissão local (HDMI) para exibição de recomendações na TV/monitor.</li> <li>▪ Disponibilidade para permitir testes em horário comercial (08 até 18 horas).</li> <li>▪ Aceita a divulgação de sinalização informando “ambiente monitorado / coleta de imagem para pesquisa” e o termo de consentimento.</li> </ul>
Requisitos legais/éticos	<ul style="list-style-type: none"> <li>▪ Empresa concorda em cumprir obrigações de transparência previstas na LGPD (BRASIL, 2018).</li> </ul>
Valor agregado (para priorizar)	<ul style="list-style-type: none"> <li>▪ Tráfego de clientes diário acima de 50 pessoas por dia.</li> <li>▪ Tamanho físico do local deve comportar clientes de maneira confortável.</li> <li>▪ Diversidade de público (faixa etária/gênero) se isso ajudar validação.</li> <li>▪ Localização conveniente para equipe (custos/logística).</li> </ul>

Fonte: Elaborado pelo autor (2025).

### 3.1.1 Escolha do Local

O autor optou por escolher uma cidade aleatória dentro da região conhecida como Recôncavo Baiano. Essa escolha ocorreu como uma homenagem ao que essa região representa e representou para a história do Brasil (TAVARES, 2005). Dentre as opções de cidades possíveis, a cidade de Santo Antônio de Jesus foi a escolhida. É uma cidade que representa a maior economia da região, com um PIB que aumentou de R\$ 379 milhões para R\$ 2,2 bilhões entre 2002 e 2019, e que possui economia fortemente baseada em serviços (BAHIA, 2025).

Dentre os estabelecimentos que formam o Centro Comercial da cidade de Santo Antônio de Jesus, a empresa que melhor se adequou aos critérios considerados se chama Farmácia do Povo, localizada na Praça Padre Mateus, 137 – Centro. Na figura 3 são apresentadas imagens do estabelecimento selecionado.

Figura 3 – Estabelecimento comercial selecionado



Fonte: Elaborado pelo autor (2025).

A seguir na figura 4, é ilustrado o local considerado adequado para a instalação da solução. A primeira parte da imagem apresenta o monitor em que serão apresentadas as recomendações de produto, na segunda, o ponto escolhido para captura de imagens, com visão frontal de quem entra no estabelecimento.

Figura 4 – Local de Visualização de Recomendações e Visão de Captura de Imagens



Fonte: Elaborado pelo autor (2025).

### 3.2 Estrutura Geral do Projeto

O pipeline considerado no projeto é composto por cinco módulos, conforme representado na figura 5:

Figura 5 – Pipeline utilizada na solução



Fonte: Elaborado pelo autor (2025).

A responsabilidade de cada módulo é:

- 1) Captura de imagem: câmera realiza a aquisição contínua dos quadros. O processo de captura é executado em janelas configuráveis e para o projeto foi deividido o intervalo de janelas de 1 hora de duração;
- 2) Detecção de pessoas: o modelo YOLO reconhece pessoas em tempo real, retornando as coordenadas para cada pessoa detectada no *frame*. YOLO foi configurado de maneira a somente identificar a classe de pessoas e a manter um mesmo identificador entre os

*frames* da gravação. Como isso, espera-se reduzir o ruído de contar a mesma pessoa por diversas vezes numa mesma sessão de gravação;

- 3) Extração facial e análise demográfica: somente as região de interesse (*Region of Interest* - ROI) contendo pessoa detectadas são processada. Desse estágio são esperadas inferências de gênero e idade das pessoas detectadas no estágio anterior;
- 4) Recomendação de conteúdo: informações de gênero e idade, obtidos da etapa anterior, são utilizados com chaves de mapeamento de conteúdos previamente estabelecidos. Vídeos e imagens foram selecionados e agrupados por gênero e grupos de idade. Após escolha de um conteúdo, aguarda-se um tempo configurável para nova recomendação. Para os objetivos da pesquisa, optou-se por apresentação de conteúdos de 30 segundos entre vídeos e imagens estáticas;
- 5) Registro de métricas demográficas: os dados inferidos são armazenados e agrupados por faixa de horário de captura. Esses dados somente consideram um identificador anonimizado que identifica a pessoa no vídeo e suas inferências.

Todo o processamento e armazenamento ocorre localmente, sem envio de dados à nuvem e evitando riscos de vazamento de informações sensíveis e de maneira inadvertida.

### **3.3 Infraestrutura utilizada na solução**

#### **3.3.1 Definição de Equipamento de *Edge IA***

A definição do equipamento de Inteligência Artificial de Borda (Edge AI) utilizado neste projeto foi orientada por critérios de desempenho, custo, escalabilidade e adequação ao cenário de uso proposto: a recomendação de produtos em tempo real no varejo físico, com base na análise demográfica de clientes. Dentre as alternativas existentes, como o NVIDIA Jetson Nano, o Google Coral e placas com *Neural Processing Units* (NPU) dedicadas, optou-se pela utilização do Raspberry Pi 5 em conjunto com a Raspberry Pi Global Shutter Camera, por apresentar um equilíbrio balanceado entre capacidade de processamento, acessibilidade e compatibilidade com o ecossistema de visão computacional requerido.

O Raspberry Pi 5, lançado em 2023, representa um avanço significativo em relação às versões anteriores da plataforma, incorporando uma CPU ARM Cortex-A76 de quatro núcleos a 2,4 GHz, GPU VideoCore VII, suporte a PCIe 2.0, além de melhorias substanciais em desempenho gráfico e eficiência energética (RASPBERRY, 2023). Essa configuração nos

permitem executar modelos de aprendizado profundo mantendo baixo consumo de energia e operação silenciosa, características essenciais em ambientes de varejo.

Outro ponto decisivo foi o forte suporte da comunidade *open-source* e a ampla compatibilidade com frameworks de IA embarcada, como TensorFlow Lite, PyTorch Mobile e OpenCV, que facilitam o desenvolvimento e a manutenção do sistema (CHENG et al., 2020). Essa abertura tecnológica favorece a reprodutibilidade dos experimentos e a integração futura com novas arquiteturas de aprendizado profundo, garantindo a longevidade da solução.

A escolha da Raspberry Pi Global Shutter Camera complementa o sistema ao oferecer captura de imagens sem distorções por movimento, fator crucial para aplicações em ambientes dinâmicos. Diferentemente das câmeras com obturador de rolagem, que podem causar artefatos em imagens com movimento rápido, o obturador escolhido captura todos os pixels simultaneamente, assegurando maior precisão na detecção de faces e objetos (RASPBERRY PI, 2023). Essa característica contribui diretamente para a qualidade das inferências de idade e gênero, etapas fundamentais para as recomendações de produtos neste projeto.

Em comparação com dispositivos mais robustos, como o Jetson Nano, o Raspberry Pi 5 oferece desempenho competitivo a um custo inferior, sendo mais acessível para implantação em múltiplas unidades de varejo.

Dessa forma, a escolha do Raspberry Pi 5 e da Global Shutter Camera oferece a infraestrutura necessária para garantir inferência rápida, precisa e segura, requisitos importantes para o sucesso de soluções inteligentes embarcadas no varejo contemporâneo.

### 3.3.2 Especificação de Produtos Utilizados

O sistema foi implantado em produtos Raspberry Pi (RASPBERRY PI, 2025). Para a apresentação das recomendações selecionadas, foi considerado como um requisito obrigatório que a empresa selecionada para o estudo de caso possua equipamento de tv/monitor que possa compartilhar com o estudo. Todos os equipamentos utilizados para a execução do projeto estão listados na tabela 2 abaixo e ilustrados na figura 6.

Tabela 2 – Especificação Técnica de Produtos utilizados no projeto

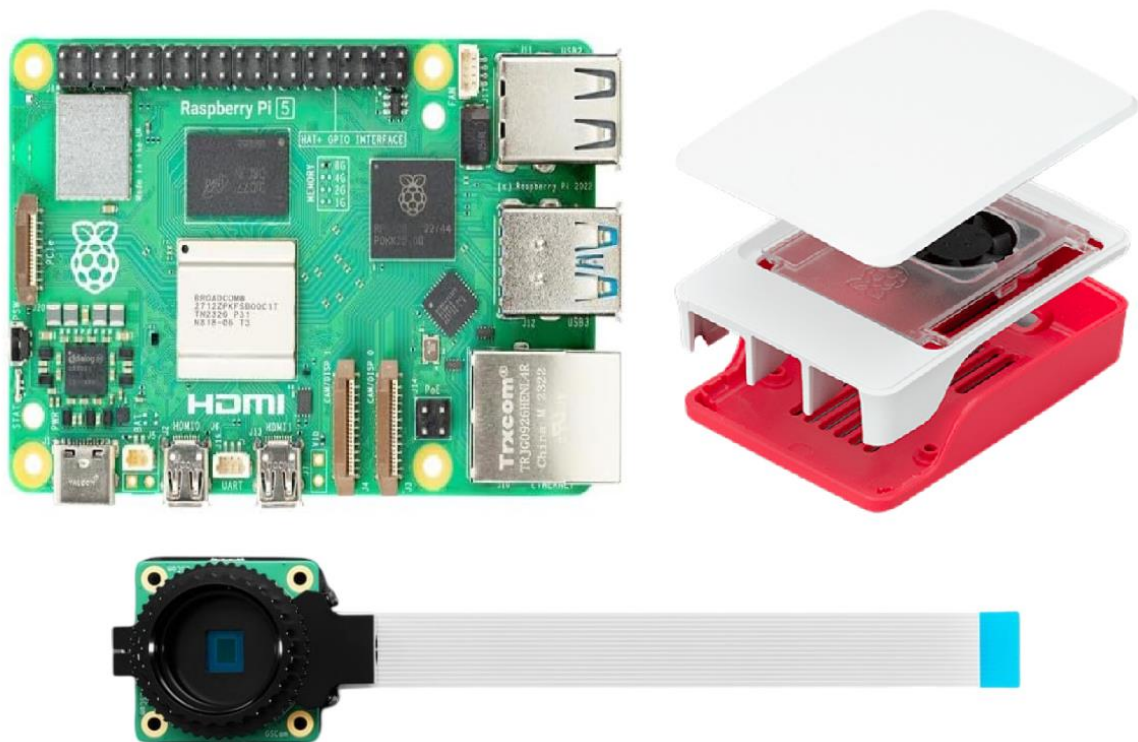
<b>Produto</b>	<b>Especificação Técnica</b>
Raspberry Pi 5	<ul style="list-style-type: none"> <li>▪ Processador (CPU): Broadcom BCM2712, quad-core ARM Cortex-A76 de 64 bits a 2,4 GHz, com extensões de criptografia, cache L2 de 512 KB por núcleo e cache L3 compartilhado de 2 MB.</li> </ul>

	<ul style="list-style-type: none"> <li>▪ GPU: VideoCore VII, com suporte a OpenGL ES 3.1 e Vulkan 1.3.</li> <li>▪ Saída de vídeo: Dupla saída HDMI® 4Kp60 (60 fps) com suporte a HDR.</li> <li>▪ Decodificação de vídeo: Suporte a decodificador HEVC 4Kp60.</li> <li>▪ Memória RAM: LPDDR4X-4267 SDRAM de 8 GB.</li> <li>▪ Conectividade sem fio: Wi-Fi® dual-band 802.11ac.</li> <li>▪ Bluetooth: Versão 5.0 com suporte a Bluetooth Low Energy (BLE).</li> <li>▪ Armazenamento: Slot para cartão microSD, compatível com o modo de alta velocidade SDR104.</li> <li>▪ Portas USB: <ul style="list-style-type: none"> <li>○ 2 × USB 3.0 (suportam operação simultânea a 5 Gbps)</li> <li>○ 2 × USB 2.0</li> <li>○ Rede cabeada: Ethernet Gigabit, com suporte a PoE+ (requer placa PoE+ HAT vendida separadamente).</li> </ul> </li> <li>▪ Interfaces de câmera e display: 2 × transceptores MIPI de 4 vias (para câmeras e telas).</li> <li>▪ Expansão: Interface PCIe 2.0 x1 para periféricos de alta velocidade (requer adaptador M.2 HAT ou outro adaptador compatível).</li> <li>▪ Alimentação: Entrada de 5 V/5 A via USB-C, com suporte a Power Delivery.</li> <li>▪ Conector GPIO: Padrão Raspberry Pi de 40 pinos.</li> <li>▪ Relógio em tempo real (RTC): Alimentado por bateria externa.</li> </ul>
Raspberry Pi Case for Raspberry Pi 5	<ul style="list-style-type: none"> <li>▪ Tensão de entrada: 5 V DC fornecidos através do conector de quatro pinos para ventoinha no Raspberry Pi 5.</li> <li>▪ Controle de velocidade da ventoinha: Controle por modulação por largura de pulso (PWM) com tacômetro para monitoramento da rotação.</li> <li>▪ Fluxo de ar máximo: 2,79 CFM (pés cúbicos por minuto).</li> <li>▪ Velocidade máxima da ventoinha: 8000 RPM ± 15%.</li> </ul>
Raspberry Pi Global Shutter Camera	<ul style="list-style-type: none"> <li>▪ Formato físico (Form factor): 38 × 38 × 19,8 mm (29,5 mm com adaptador e tampa contra poeira)</li> <li>▪ Peso: 34 g (41 g com adaptador e tampa contra poeira)</li> <li>▪ Sensor: Sony IMX296LQR-C</li> <li>▪ Resolução: 1,58 megapixels (colorida)</li> <li>▪ Tamanho do sensor: Diagonal de 6,3 mm</li> <li>▪ Tamanho do pixel: 3,45 μm × 3,45 μm</li> <li>▪ Saída de imagem: RAW10 (formato bruto de 10 bits)</li> <li>▪ Distância focal posterior (back focus): Ajustável (12,5 – 22,4 mm)</li> </ul>

<p>Monitor LG 22ma33n (disponibilizado pela empresa selecionada para pesquisa)</p>	<ul style="list-style-type: none"> <li>▪ Tipo: TV Monitor IPS LED</li> <li>▪ Tamanho da tela (Polegadas): 22"</li> <li>▪ Tempo de Resposta: 10ms</li> <li>▪ Sinal de Video: Digital</li> <li>▪ HDMI: Sim</li> <li>▪ Resolução Máxima: 1366 x 768</li> </ul>
--	---

Fontes: Raspberry Pi (2025) e LG (2025).

Figura 6 – Produtos Raspberry utilizados no projeto



Fonte: Raspberry (2025).

### 3.4 Ambiente de Desenvolvimento

O ambiente de desenvolvimento foi baseado em Python 3.11 com bibliotecas: ultralytics (para execução do modelo yolo), opencv-python (captura de vídeo), deepface e retinaface (análise facial e inferência de atributos demográficos), numpy e matplotlib (manipulação e visualização de dados em cenários de teste).

### 3.4.1 YOLO versão 11

Para a detecção de pessoas, a escolha da versão do algoritmo YOLO considerou que o desempenho do sistema dependeria diretamente do equilíbrio entre precisão, velocidade de inferência e viabilidade em *hardware* embarcado. Após a análise comparativa das versões optou-se pela adoção da versão YOLO11, mais especificamente o modelo YOLO11n (*nano*), desenvolvido pela Ultralytics (2025), por apresentar desempenho em dispositivos de baixo consumo energético sem comprometer a acurácia nas tarefas de detecção facial e análise demográfica.

Desde sua criação, o YOLO tem sido reconhecido por sua capacidade de realizar detecção de objetos em tempo real com eficiência superior a abordagens mais complexas, como o Faster R-CNN ou o SSD (REDMON et al., 2016; LIU et al., 2016). Com o avanço das versões, a arquitetura passou a incorporar técnicas cada vez mais sofisticadas de aprendizado profundo, como camadas residuais e ajuste dinâmico de rótulos, por exemplo; que aumentaram progressivamente a precisão e reduziram o custo computacional (BOCHKOVSKIY; WANG; LIAO, 2020; WANG et al., 2022; ULTRALYTICS, 2025).

Para este projeto, que visa à análise demográfica e recomendação de produtos em tempo real no varejo, o YOLO11n (*nano*) mostrou-se adequado por conter 4,3 milhões de parâmetros, com capacidade de operar a 2–3 quadros por segundo (FPS) em um Raspberry Pi 5 com 8GB de RAM, mantendo precisão comparável às versões maiores (TERVEN; CÓRDOVA-ESPARZA, 2023). Esse desempenho é suficiente para aplicações interativas em ambientes físicos, em que a latência inferior a um segundo é aceitável e a prioridade recai sobre respostas imediatas e confiáveis.

Como limitações, é importante reconhecer que modelos *nano* tendem a apresentar redução de precisão em objetos muito pequenos e menor capacidade de generalização em cenários altamente variados. Contudo, esses impactos são aceitáveis no contexto deste projeto, onde a detecção se restringe a faces humanas em ambiente controlado e iluminado.

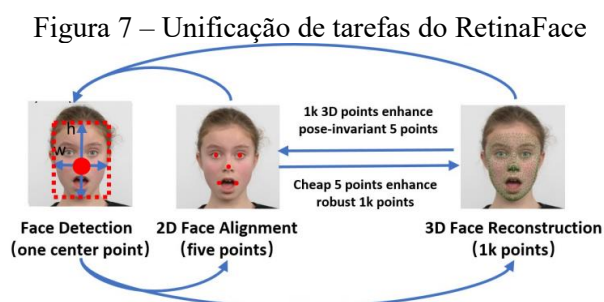
Portanto, a seleção dessa versão foi fundamentada em uma análise técnica e contextual: o YOLO11n combina a robustez e maturidade de uma arquitetura de ponta com a leveza e eficiência necessárias para execução local em dispositivos de Edge AI, atendendo plenamente aos requisitos de baixo custo, tempo real e respeito à privacidade estabelecidos para a solução proposta.

### 3.4.2 Análise Facial e Inferência demográfica com DeepFace

Após a etapa de detecção de pessoas realizada pelo modelo YOLO11n, o sistema desenvolvido necessita identificar e analisar regiões de interesse (faces) para realizar a inferência de gênero e faixa etária dos clientes, de forma anônima e em tempo real. Para essa etapa, optou-se pela utilização do *framework* DeepFace, empregando o RetinaFace como modelo base de detecção facial.

O DeepFace, proposto originalmente por Taigman et al. (2014), é uma das arquiteturas mais consolidadas para reconhecimento e análise facial. Sua principal vantagem está em oferecer uma estrutura modular e unificada, que integra diferentes soluções e permite o uso de detectores e extratores de características avançados. Essa flexibilidade é essencial para o desenvolvimento de um pipeline adaptável, capaz de alternar entre modelos conforme as restrições computacionais do dispositivo embarcado. Outro fator determinante para a escolha do DeepFace foi sua compatibilidade com ambientes de baixo consumo energético, como o Raspberry Pi 5, e o suporte a inferência otimizada em CPU. Algumas das possibilidades que temos acesso com o uso do DeepFace são: verificação de face (identificação de correspondências entre duas imagens), reconhecimento de face (identificar uma face a partir de uma base de dados previamente criada) e análise de atributos faciais (extração de atributos como idade, gênero e etnia).

A detecção é uma importante tarefa do reconhecimento de face. Tradicionalmente, a detecção facial era uma série de etapas sequenciais: primeiro, detectar a face; depois, estimar a pose, alinhar os *landmarks* (pontos de referência) ou reconstruir a face em 3D. Desenvolvido por Deng et al. (2019), o RetinaFace unifica essas três tarefas cruciais de localização facial sob um alvo comum: a regressão de pontos no plano da imagem. Essa abordagem entende que, embora essas tarefas visem diferentes níveis de detalhe de localização, elas compartilham o mesmo objetivo fundamental de prever pontos precisos na imagem, conforme figura 7.



Fonte: Deng et al. (2019).



pela localização, alinhamento e inferência das variáveis de gênero e idade. Essas informações são utilizadas de forma não identificatória, isto é, sem retenção de dados pessoais ou biométricos, assegurando conformidade com os princípios de anonimização definidos pela LGPD (BRASIL, 2018).

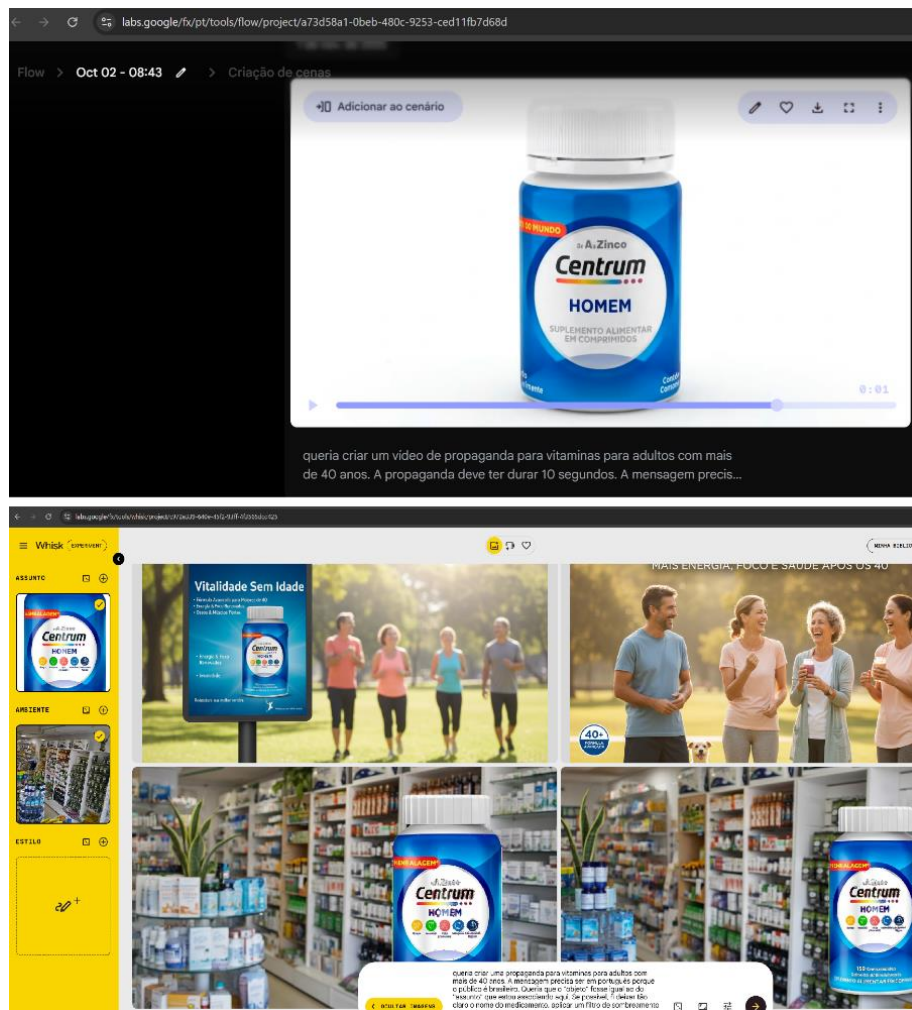
Com base nas inferências retornadas, o sistema agrupa as estimativas de idade em faixas etárias predefinidas, permitindo associar cada grupo a um conjunto específico de conteúdos audiovisuais (imagens ou vídeos promocionais) a serem exibidos no monitor localizado no ponto de venda. Essa abordagem busca simular um mecanismo de recomendação contextual imediata, ajustando o tipo de produto e o estilo de comunicação às características inferidas do público diante da tela.

A literaturas sobre estimativa de idade demonstram que, embora os modelos modernos baseados em aprendizado profundo, como DeepFace, ArcFace e InsightFace, apresentem desempenho satisfatório, há diferenças significativas na precisão conforme a faixa etária analisada. Trabalhos recentes, como os de Ahn et al. (2024) e Turner et al. (2024), mostram que as maiores taxas de erro ocorrem entre adolescentes e jovens adultos, em virtude da alta variabilidade facial nessa fase e do desbalanceamento de dados nos datasets públicos utilizados para treinamento. Da mesma forma, os estudos de Rothe, Timofte e Van Gool (2018), responsáveis pelo *dataset* IMDB-WIKI, e de Levi e Hassner (2015) com o *dataset* Adience, apontam que a acurácia tende a ser mais estável em faixas adultas (25–45 anos), mas declina em extremos (crianças pequenas e idosos). Essas constatações reforçam a necessidade de agrupar idades em intervalos discretos, em vez de tentar representar estimativas contínuas, o que aumenta a confiabilidade das inferências e a consistência das recomendações. Considerando tais evidências, as estimativas de idade neste projeto foram mapeadas para seis grupos etários: crianças (0–9 anos), adolescentes (10–17 anos), jovens adultos (18–24 anos), adultos (25–39 anos), adultos maduros (40–59 anos) e idosos (60 anos ou mais).

Esses agrupamentos seguem práticas comuns em estudos de percepção de idade e segmentação de mercado (AHN et al., 2024) e também refletem divisões usadas por sistemas comerciais de publicidade personalizada (GUO et al., 2021).

Para cada grupo, foram definidos conjuntos de mídia correspondentes a produtos potencialmente mais relevantes, considerando a experiência da responsável pela farmácia. Assim, jovens adultos receberam suplementos e adultos maduro produtos de saúde e bem-estar. Algumas dos conteúdos utilizados foram disponibilizados pela própria responsável da farmácia, outros fora construídos com ferramentas de apoio com o uso da IA como Flow (GOOGLE, F., 2025), para vídeos e Whisk (GOOGLE, W., 2025) para imagens, conforme figura 9.

Figura 9 – Processo criativo de propaganda de produtos



Fonte: Elaborado pelo autor (2025).

Com o intuito de reduzir erros e evitar recomendações incorretas, o sistema adota limiares de confiança para a inferência de idade e gênero. Quando a confiança do modelo é inferior a 55% o sistema exibe conteúdos genéricos, não personalizados, evitando inferências imprecisas.

### 3.6 Considerações Éticas e de Privacidade

Desde a sua concepção, a solução foi projetada de forma a não manter qualquer dado sensível das pessoas detectadas. Toda a inferência é feita localmente, e apenas os resultados agregados (idade e gênero estimados) são armazenados em arquivos de métricas.

Além disso, o modelo registra apenas inferência de características demográficas, o que minimiza riscos de violação da privacidade e mantém o projeto em conformidade com a LGPD (BRASIL, 2018).

Após o aceite da responsável pelo estabelecimento na colaboração com a presente pesquisa, foi elaborado um aviso, para ser direcionado ao público em geral, e um termo de consentimento. Ambos os modelos utilizados estão registrados em Anexos A e B.

## 4 EXPERIMENTOS E ANÁLISE DE RESULTADOS

Os testes foram realizados em etapas complementares ao longo de quatro dias de observação e com o objetivo de validar o desempenho técnico, a percepção dos usuários e o comportamento operacional contínuo da solução proposta. Ao total foram 3 etapas que estão explicadas nas seções a seguir.

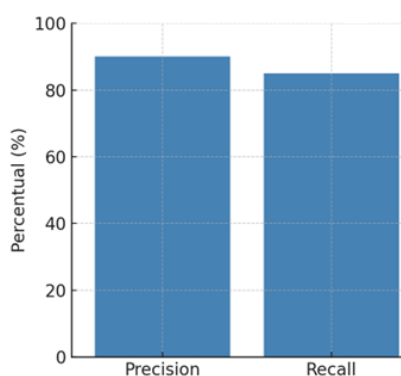
### 4.1 Validação Técnica de Detecção e Contagem de Pessoas

A primeira fase teve como foco avaliar a capacidade de detecção e contagem de indivíduos pelo modelo YOLO11n executado no dispositivo Raspberry Pi 5 e acoplada a uma Raspberry Pi Global Shutter Camera.

Essa etapa buscou validar se o sistema conseguiria identificar pessoas em tempo real, com precisão suficiente para as etapas subsequentes de reconhecimento facial e recomendação de produtos. Durante o teste, o equipamento foi posicionado de forma a abranger a área de entrada do estabelecimento. Foram observadas manualmente cerca de 200 passagens de clientes, cujos eventos foram comparados com as detecções automáticas realizadas pelo modelo, por meio de observação direta, sem qualquer gravação de vídeo.

Os resultados indicaram elevada consistência entre as contagens reais e as inferências automáticas. De *precision* foi obtido um resultado de 0,90, ou seja, 90% das detecções corresponderam a pessoas reais. Sobre *recall*, 85% das pessoas que passaram foram efetivamente detectadas. A figura 10 apresenta um resumo das principais métricas de desempenho técnico observadas ao longo dos testes.

Figura 10 – Métricas de Desempenho Técnico



Fonte: Elaborado pelo autor (2025).

A variação observada ocorreu principalmente em situações de oclusão parcial e iluminação lateral intensa, que reduziram a visibilidade de partes do corpo.

Apesar disso, o desempenho foi considerado suficiente para o contexto de recomendação em tempo real, em que pequenas imprecisões não comprometem a experiência do usuário (REDMON; FARHADI, 2018; WANG et al., 2023).

Em termos operacionais, o sistema manteve taxa média de 2,3 quadros por segundo (FPS), temperatura estável de 67 °C e nenhuma falha de processamento, demonstrando boa eficiência energética e estabilidade em ambiente real.

#### 4.2 Avaliação com Voluntários e Validação das Inferências

A segunda fase consistiu em testes controlados com 50 pessoas, entre voluntários e consumidores diretos, conduzidos para avaliar a precisão perceptiva das inferências de gênero e faixa etária, utilizando o pipeline com DeepFace + RetinaFace.

Nessa etapa, cada participante foi informado sobre o processo e convidado a confirmar se os resultados exibidos pelo sistema correspondiam à sua. A percepção dos voluntários foi coletada por meio de questionário de escala de 1 a 5, contemplando critérios de adequação, clareza visual, relevância do conteúdo e percepção de privacidade. A Tabela 3 apresenta as médias obtidas.

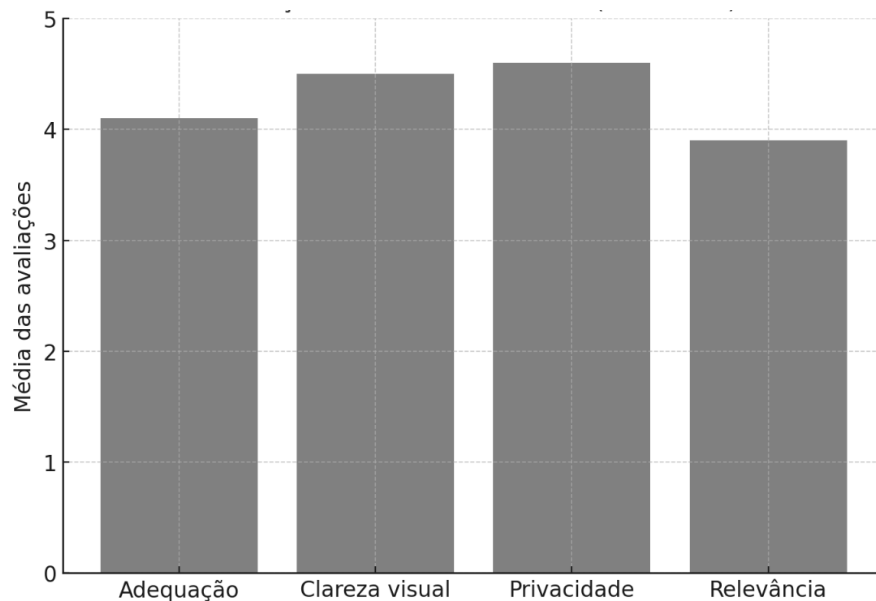
Tabela 3 – Avaliação Média dos Voluntários (escala 1–5)

<b>Critério avaliado</b>	<b>Média (1–5)</b>	<b>Desvio-padrão</b>
<b>Adequação da recomendação</b>	4,1	0,7
<b>Clareza visual da apresentação</b>	4,5	0,5
<b>Sensação de privacidade preservada</b>	4,6	0,6
<b>Relevância percebida do conteúdo</b>	3,9	0,8

Fonte: Elaborado pelo autor (2025).

Os resultados indicam aceitação positiva da solução, com destaque para a clareza visual e a preservação de privacidade, ambas acima de 4,5 pontos. A figura 11 apresenta as avaliações médias dos voluntários em cada critério analisado.

Figura 11 – Avaliação Média dos Voluntários (escala 1–5)



Fonte: Elaborado pelo autor (2025).

A adequação percebida das recomendações (4,1) mostra que os conteúdos apresentados foram considerados coerentes com o perfil inferido, o que confirma o potencial de estratégias de personalização baseadas em análise demográfica visual.

Cerca de 68% dos voluntários declararam que os produtos exibidos “pareciam adequados” ao seu perfil, e 63% afirmaram que considerariam adquirir o item apresentado, sugerindo viabilidade comercial para recomendações contextuais em ambientes físicos de varejo.

Os dados coletados permitiram estimar métricas indiretas de acurácia, sem necessidade de registro de imagem. As métricas foram acurácia de gênero de 88%, erro médio absoluto (MAE) na estimativa de idade de 5,9 anos, e distribuição de erros maior em adolescentes ( $\pm 8$  anos) e menor em adultos ( $\pm 4$  anos).

A análise qualitativa evidenciou boa aceitação da experiência interativa. 68% dos participantes consideraram que os produtos exibidos eram adequados à sua faixa etária. Além disso, 63% afirmaram que poderiam se interessar por adquirir o produto recomendado.

Esses indicadores apontam que as recomendações baseadas em inferência demográfica foram percebidas como coerentes e úteis, sem gerar desconforto quanto à privacidade.

### 4.3 Operação Contínua e Análise de Comportamento em Campo

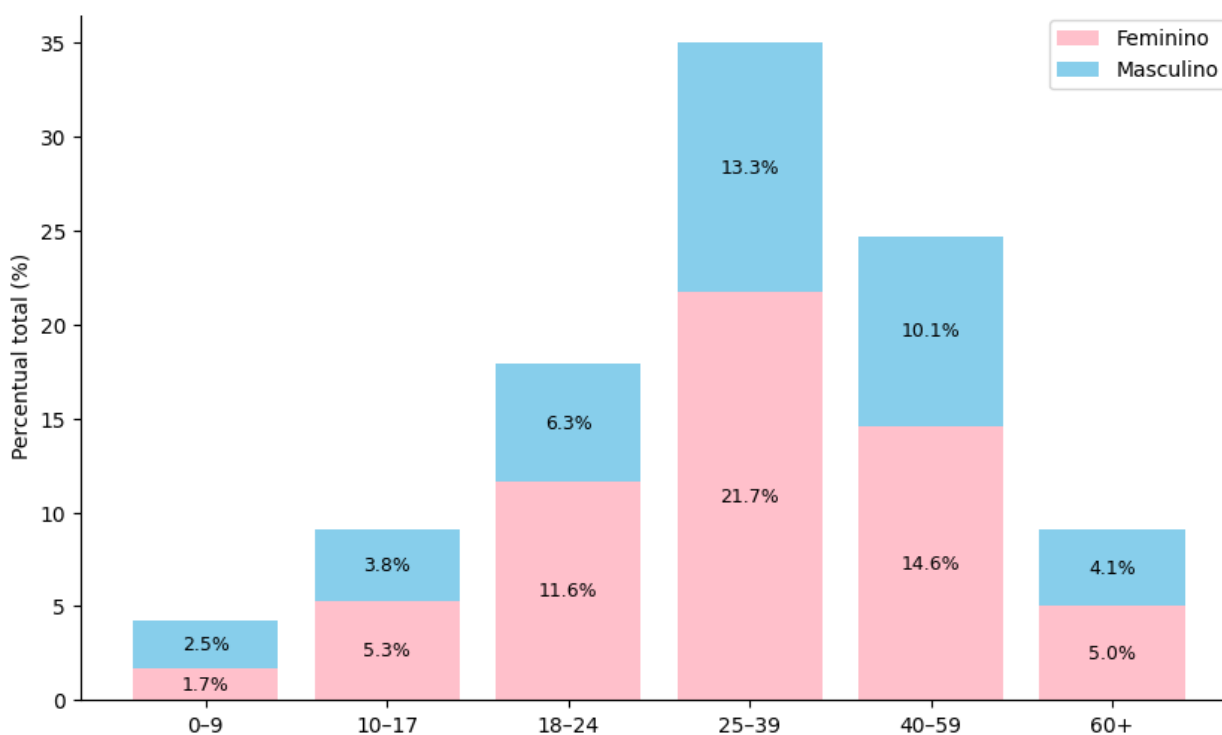
A terceira fase correspondeu à operação contínua do sistema por dois dias, durante o horário comercial (7h às 17h). Nesse período, foram registradas 420 inferências válidas, processadas localmente, sem armazenamento de imagens ou dados biométricos.

Os logs coletados indicaram taxa de baixa confiança de 7,8%, que ativou o mecanismo de *fallback* — substituindo automaticamente recomendações com incerteza elevada por conteúdos genéricos.

A temperatura média da CPU permaneceu abaixo de 70 °C e não foram observadas reinicializações, demonstrando robustez térmica e estabilidade de software.

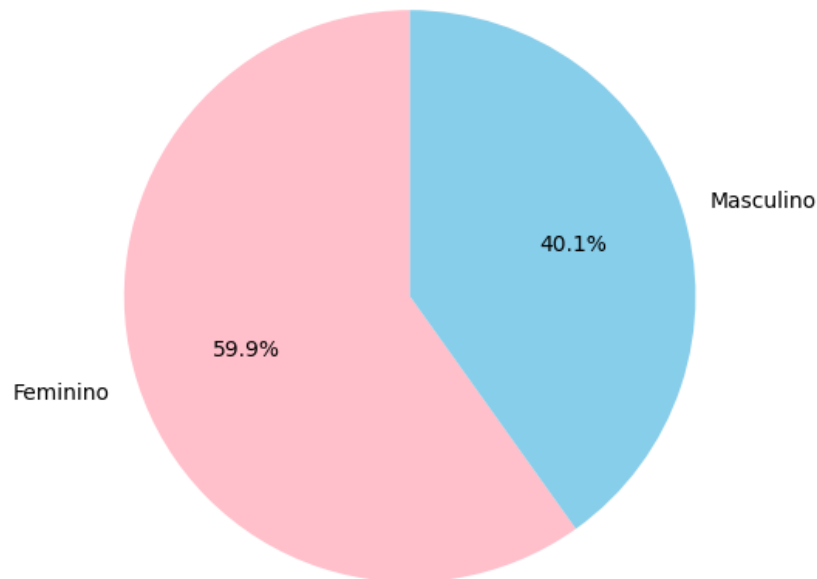
A distribuição etária inferida ao longo da operação é apresentada na Figura 8, com predominância das faixas de 25 a 39 anos (35%) e 40 a 59 anos (24,7%), o que reflete o perfil real de clientes informado pelo responsável do estabelecimento. As figuras 12 e 13 complementam essa análise, ilustrando graficamente a proporção de clientes detectados em cada grupo etário.

Figura 12 – Distribuição por faixa etária e gênero



Fonte: Elaborado pelo autor (2025).

Figura 13 – Distribuição total por gênero



Fonte: Elaborado pelo autor (2025).

As faixas extremas (0–9 e 60+) representaram menos de 10% do total, reforçando a coerência do modelo com o contexto de aplicação.

#### 4.4 Discussão Geral dos Resultados

De forma integrada, os resultados das três fases demonstram que a solução foi tecnicamente viável, eticamente responsável e operacionalmente estável.

A detecção e contagem mostraram desempenho comparável a modelos mais robustos, enquanto a inferência demográfica atingiu níveis satisfatórios de precisão para aplicações comerciais em tempo real.

A estratégia de validação indireta, combinando observação presencial, autorrelato e coleta de métricas agregadas, mostrou-se adequada para contornar restrições legais e éticas, e com o objetivo de manter o rigor científico do experimento.

Do ponto de vista de negócio, as inferências coletadas permitiram identificar padrões de comportamento e horários de maior fluxo, além de oferecer subsídios para futuras análises de impacto sobre vendas.

Não foi possível inferir diretamente o efeito das recomendações no faturamento da empresa. O que foi possível comprovar nesse aspecto, segundo relato da responsável pela

farmácia é que “a empresa teve um leve aumento de clientes que souberam da pesquisa e quiseram participar”.

Assim, a pesquisa confirma o potencial da Inteligência Artificial de Borda no varejo como instrumento de personalização e engajamento, capaz de unir tecnologia acessível, respeito à privacidade e aplicabilidade prática.

## 5 CONCLUSÃO

O presente trabalho apresentou o desenvolvimento e a validação de um sistema embarcado de detecção de pessoas, análise demográfica em tempo real e recomendação de produtos, utilizando o modelo YOLO11n para detecção e o framework DeepFace para inferência de idade e gênero.

O sistema foi implementado integralmente em um Raspberry Pi 5 com 8 GB de RAM e uma câmera acoplada, demonstrando a viabilidade de execução de modelos de visão computacional avançados em dispositivos de baixo custo.

A metodologia adotada contemplou desde a revisão das principais técnicas de detecção de objetos e reconhecimento facial até a implementação e análise da solução em ambiente real.

Foram exploradas diferentes possibilidades arquiteturais e utilizada uma estrutura modular que permite o avanço da pesquisa em trabalhos futuros.

Os resultados alcançados confirmam o potencial da solução para aplicações em interação inteligente em pontos de venda e sistemas de recomendação baseados em perfil demográfico.

### 5.1 Contribuições do Trabalho

O desenvolvimento deste trabalho resultou em contribuições, tanto do ponto de vista prático quanto científico. Em primeiro lugar, foi possível demonstrar de forma concreta a viabilidade da execução de modelos de visão computacional baseados em redes neurais profundas em dispositivos embarcados de baixo custo. A implementação do sistema utilizando o YOLO11n para detecção de pessoas e o DeepFace para análise facial mostrou que, mesmo em um ambiente restrito em recursos computacionais, é possível alcançar resultados compatíveis com aplicações de tempo real. Essa prova de conceito reforça o potencial da Inteligência Artificial de Borda (*Edge AI*) como alternativa para sistemas que exigem processamento local e resposta imediata.

Além da contribuição prática, o trabalho também apresenta uma contribuição analítica relevante ao comparar diferentes abordagens e definindo critérios os critérios de escolha das soluções adotadas, discutindo seus impactos no desempenho, precisão e viabilidade em plataformas embarcadas.

Por fim, o projeto reforça o compromisso com aspectos éticos e legais no uso de tecnologias de reconhecimento facial, implementando uma arquitetura local, que processa todas

as informações no próprio dispositivo e não realiza armazenamento ou transmissão de dados pessoais. Essa decisão está em consonância com os princípios da LGPD (BRASIL, 2018), evidenciando a preocupação do estudo com o desenvolvimento responsável e seguro de soluções baseadas em IA.

## 5.2 Limitações Identificadas

Embora os resultados obtidos tenham sido satisfatórios, algumas limitações foram observadas durante o processo de desenvolvimento e experimentação do sistema. O DeepFace, embora ofereça grande flexibilidade e precisão satisfatória, depende de modelos de *embedding* facial genéricos (como o VGG-Face e o Facenet), que podem apresentar viés em populações sub-representadas nos datasets de treinamento originais. Essa limitação reforça a necessidade de explorar, em trabalhos futuros, bases de dados mais diversificadas e técnicas de mitigação de viés para melhorar a generalização do sistema.

Além disso, observou-se que o YOLO11n, apesar de ser uma das versões mais otimizadas da família, ainda não possui suporte completo para aceleração via NPU no Raspberry Pi 5, o que restringe seu desempenho máximo potencial. Essas limitações, contudo, não comprometem a viabilidade geral da proposta, mas indicam caminhos claros para aprimoramentos futuros.

## 5.3 Trabalhos Futuros

A continuidade deste trabalho pode seguir diferentes direções, tanto em aspectos técnicos quanto científicos. Uma possibilidade é a otimização do modelo YOLO11n por meio de sua conversão para formatos compatíveis com aceleradores dedicados, que poderiam reduzir a latência e elevar a taxa de quadros para valores próximos de 10 fps, conforme estudos recentes sobre otimização de modelos em dispositivos ARM (Choudhary et al., 2022). Essa melhoria tornaria o sistema mais fluido e ampliaria sua aplicabilidade em contextos de interação direta com o usuário.

Outra linha de evolução interessante consiste na substituição do *backend* facial do DeepFace e comparar com outras alternativas disponíveis. Essa alteração poderia elevar substancialmente a precisão da detecção de rostos e a estabilidade da inferência sob condições adversas de iluminação e movimento, sem modificar a estrutura geral do pipeline implementado.

Adicionalmente, uma melhoria interessante no processo de detecção de pessoas seria ajustar o treinamento do modelo para não considerar os funcionários do estabelecimento na detecção.

O trabalho também poderia evoluir em direção à integração com sistemas de recomendação fortemente personalizados, de modo que as inferências demográficas realizadas em tempo real possam ser utilizadas para oferecer produtos, informações ou serviços de forma ainda mais contextualizada. Essa integração traria um valor comercial adicional à solução e reforçaria seu potencial de aplicação em ambientes de varejo e atendimento automatizado.

Por fim, é essencial que futuros desenvolvimentos considerem também aspectos de equidade algorítmica e transparência, promovendo a análise do desempenho do sistema em diferentes grupos demográficos e mitigando potenciais vieses de classificação. Trabalhos como o de Mehrabi et al. (2021) destacam a importância de incorporar métricas de justiça e auditoria nos ciclos de desenvolvimento de sistemas de IA, garantindo que soluções baseadas em reconhecimento facial respeitem princípios éticos e de responsabilidade social.

#### **5.4 Considerações Finais**

A realização deste trabalho permitiu constatar a maturidade e aplicabilidade das técnicas de visão computacional baseadas em redes neurais profundas mesmo em dispositivos com restrições de hardware.

A combinação entre o YOLO11n e o DeepFace mostrou-se tecnicamente consistente, eficiente e escalável, oferecendo uma base sólida para projetos de IA embarcada.

Em termos práticos, o sistema cumpre seu propósito de detectar pessoas e estimar atributos demográficos em tempo real, com confiabilidade e conformidade legal.

Em termos acadêmicos, a pesquisa contribui para o fortalecimento da área de Edge AI aplicada à visão computacional, fornecendo um estudo de caso que une fundamentação teórica, implementação real e análise crítica de resultados.

Dessa forma, conclui-se que o presente trabalho alcançou seus objetivos propostos, comprovando a viabilidade da aplicação de modelos de detecção e análise facial em dispositivos embarcados, e abrindo caminho para soluções mais sofisticadas e socialmente responsáveis no campo da Inteligência Artificial.

## REFERÊNCIAS

AGGARWAL, Charu C. **Recommender Systems: The Textbook**. Springer, 2016.

AHN, I. et al. **Perceived Age Estimation from Facial Image and Its Impact on Demographic Prediction**. Scientific Reports, v. 14, 2024. Disponível em <https://www.nature.com/articles/s41598-024-78695-7>

BAHIA. **Diagnóstico Socioeconômico do município de Santo Antônio de Jesus de 2000 a 2021**. Disponível em <https://www.ba.gov.br/justica/sites/site-sjdh/files/2024-11/DIAGN%C3%93STICO%20SOCIOECON%C3%94MICO%20DO%20MUNIC%C3%8DPIO%20DE%20SANTO%20ANT%C3%94NIO%20DE%20JESUS%20DE%202000%20A%202021%20%281%29.pdf>

BOCHKOVSKIY, A.; WANG, C. Y.; LIAO, H. Y. M. **YOLOv4: Optimal Speed and Accuracy of Object Detection**. 2020. Disponível em: [https://openaccess.thecvf.com/content/CVPR2021/html/Wang\\_Scaled-YOLOv4\\_Scaling\\_Cross\\_Stage\\_Partial\\_Network\\_CVPR\\_2021\\_paper.html](https://openaccess.thecvf.com/content/CVPR2021/html/Wang_Scaled-YOLOv4_Scaling_Cross_Stage_Partial_Network_CVPR_2021_paper.html).

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Diário Oficial da União, 2018.

BUOLAMWINI, J.; GEBRU, T. **Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification**. Proceedings of Machine Learning Research (PMLR), v. 81, p. 1–15, 2018. Disponível em <https://proceedings.mlr.press/v81/buolamwini18a.html>

CHENG, Y. et al. **Model Compression and Acceleration for Deep Neural Networks: The Principles, Progress, and Challenges**. IEEE Signal Processing Magazine, v. 37, n. 1, p. 126–136, 2020. Disponível em <https://ieeexplore.ieee.org/document/8253600>

CHOUDHARY, T. et al. **A Comprehensive Survey on Model Compression and Acceleration**. ACM Computing Surveys, v. 54, n. 10, p. 1-36, 2022. Disponível em <https://dl.acm.org/doi/10.1007/s10462-020-09816-7>

CHRISTOPHER, M. (2000). **The Agile Supply Chain: Competing in Volatile Markets**. *Industrial Marketing Management*, 29(1), 37-44. Disponível em <https://www.sciencedirect.com/science/article/abs/pii/S0019850199001108?via%3Dihub>

DAVENPORT, T., & RONANKI, R. (2018, January 9). **Artificial Intelligence for the Real World**. *Harvard Business Review (HBR)*. Disponível em <https://hbr.org/2018/01/artificial-intelligence-for-the-real-world>.

DALAL, N.; TRIGGS, B. **Histograms of Oriented Gradients for Human Detection**. *IEEE Computer Vision and Pattern Recognition (CVPR)*, p. 886–893, 2005. Disponível em <https://ieeexplore.ieee.org/document/1467360>

DENG, J. et al. **RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild**. *IEEE CVPR*, 2020. Disponível em [https://openaccess.thecvf.com/content\\_CVPR\\_2020/papers/Deng\\_RetinaFace\\_Single-Shot\\_Multi-Level\\_Face\\_Localisation\\_in\\_the\\_Wild\\_CVPR\\_2020\\_paper.pdf](https://openaccess.thecvf.com/content_CVPR_2020/papers/Deng_RetinaFace_Single-Shot_Multi-Level_Face_Localisation_in_the_Wild_CVPR_2020_paper.pdf)

EUROPEAN COMMISSION. **Ethics guidelines for trustworthy AI**. Luxembourg: Publications Office of the European Union, 2019. Disponível em: <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. MIT Press, 2016. Disponível em: <https://www.deeplearningbook.org/>

GOOGLE. Flow [ferramenta de inteligência artificial]. Disponível em: <https://labs.google/fx/pt/tools/flow/>

GOOGLE. Gemini [ferramenta de inteligência artificial]. Disponível em: <https://gemini.google.com/>

GOOGLE. Whisk [ferramenta de inteligência artificial]. Disponível em: <https://labs.google/fx/pt/tools/whisk>

GUO, J. et al. **InsightFace: An Open Source 2D and 3D Deep Face Analysis Toolbox**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021. Disponível em <https://www.insightface.ai/>

HAN, S. et al. **Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding**. arXiv preprint, arXiv:1510.00149, 2016.

HE, K. et al. **Deep Residual Learning for Image Recognition**. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 770–778, 2016. Disponível em <https://ieeexplore.ieee.org/document/7780459>

JANNACH, D. et al. **A survey on conversational recommender systems**. ACM Computing Surveys, v. 54, n. 5, p. 1–36, 2021. Disponível em <https://dl.acm.org/doi/10.1145/3453154>

JOBIN, A., IENCA, M., VAYENA, E. **The Global Landscape of AI Ethics Guidelines**. Nature Machine Intelligence, 1(9), 389–399. 2019 Disponível em <https://www.nature.com/articles/s42256-019-0088-2>

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. **ImageNet Classification with Deep Convolutional Neural Networks**. Communications of the ACM, v. 60, n. 6, p. 84–90, 2012. Disponível em <https://dl.acm.org/doi/10.1145/3065386>

LECUN, Y.; BENGIO, Y.; HINTON, G. **Deep Learning**. Nature, v. 521, p. 436–444, 2015. Disponível em <https://www.nature.com/articles/nature14539>

LEVI, G.; HASSNER, T. **Age and Gender Classification Using Convolutional Neural Networks**. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), p. 34–42, 2015. Disponível em <https://ieeexplore.ieee.org/document/7301352>

LG. Disponível em: <https://www.lg.com/br/monitores/monitores-de-tv/22ma33n/>

LI, C. et al. **YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications**. Disponível em <https://arxiv.org/pdf/2209.02976>

LIU, W. et al. **SSD: Single Shot MultiBox Detector**. ECCV, 2016. Disponível em [https://link.springer.com/chapter/10.1007/978-3-319-46448-0\\_2](https://link.springer.com/chapter/10.1007/978-3-319-46448-0_2)

MEHRABI, N. et al. **A Survey on Bias and Fairness in Machine Learning**. ACM Computing Surveys, v. 54, n. 6, 2021. Disponível em <https://dl.acm.org/doi/10.1145/3457607>

MITCHELL, T. **Machine Learning**. New York: McGraw-Hill, 1997. Disponível em <https://www.cs.cmu.edu/~tom/files/MachineLearningTomMitchell.pdf>

MOLLAHOSSEINI, A. ; HASANI, B.; MAHOOR, M. H. **AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild**. IEEE Transactions on Affective Computing, vol. 10, no. 1, pp. 18-31, 1 Jan.-March 2019. Disponível em <https://ieeexplore.ieee.org/document/8013713>

OPENAI. ChatGPT (GPT-5) [ferramenta de inteligência artificial]. São Francisco, CA: OpenAI, 2025. Disponível em: <https://chat.openai.com/>

PATEL, Jitendra; BANSAL, Alok. **Effect of demographic variables on e-marketing strategies: A review**. International Journal of Academic Research and Development, v. 3, n. 1, p. 311-321, 2018.

PINE, B. Joseph; GILMORE, James H. **The Experience Economy: Work Is Theatre & Every Business a Stage**. Harvard Business Review Press, 1999.

RAUT, Rakesh D. et al. "Big Data and Artificial Intelligence Applications in Retail Supply Chain Management: A Review and Research Agenda." Annals of Operations Research, 2021.

RASPBERRY PI. Disponível em: <https://www.raspberrypi.com/products/>

REDMON, J. et al. **You Only Look Once: Unified, Real-Time Object Detection**. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. Disponível em <https://ieeexplore.ieee.org/document/7780460>

REDMON, J.; FARHADI, A.. **YOLO9000: better, faster, stronger**. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. p. 7263-7271. Disponível em <https://arxiv.org/pdf/1612.08242>

REDMON, J.; FARHADI, A. **YOLOv3: An Incremental Improvement**. Disponível em <https://arxiv.org/pdf/1804.02767>

REN, S. et al. **Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks**. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), v. 39, n. 6, p. 1137–1149, 2015. Disponível em <https://ieeexplore.ieee.org/document/7485869>

ROTHER, R.; TIMOFTE, R.; VAN GOOL, L. **Deep Expectation of Real and Apparent Age from a Single Image Without Facial Landmarks**. International Journal of Computer Vision, v. 126, n. 2–4, p. 144–157, 2018. Disponível em <https://link.springer.com/article/10.1007/s11263-016-0940-3>

RUSSELL, S. ; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 4. ed. Hoboken, NJ: Prentice Hall, 2020

RYBINTSEV, A. **Age Estimation from a Face Image in a Selected Gender-Race Group based on Ranked Local Binary Patterns**. Complex & Intelligent Systems, v. 3, p. 93-104, 2017. DOI: 10.1007/s40747-017-0035-y. Disponível em: <https://link.springer.com/article/10.1007/s40747-017-0035-y>.

SHI, W. et al. **Edge Computing: Vision and Challenges**. IEEE Internet of Things Journal, v. 3, n. 5, p. 637–646, 2016. Disponível em <https://ieeexplore.ieee.org/document/7488250>

SIMONYAN, K.; ZISSERMAN, A. **Very Deep Convolutional Networks for Large-Scale Image Recognition**. 2015. Disponível em <https://arxiv.org/pdf/1409.1556>

STONEBRAKER, Michael. **What Does ‘Big Data’ Mean?** In: BLOG@CACM: Data and Information. Communications of the ACM, [s. 1.], 21 set. 2012. Disponível em: <https://cacm.acm.org/blogcacm/what-does-big-data-mean/>

TAIGMAN, Y. et al. **DeepFace: Closing the Gap to Human-Level Performance in Face Verification**. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 1701–1708, 2014. Disponível em <https://ieeexplore.ieee.org/document/6909616>

TAVARES, Luís Henrique Dias. **A Independência do Brasil na Bahia**. Salvador: EDUFBA, 2005.

TERVEN, J.; CÓRDOVA-ESPARZA, D. **A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas**. Machine Learning and Knowledge Extraction. 2023. Disponível em <https://arxiv.org/pdf/2304.00501>

TURNER, C. et al. **Deep Learning Predicted Perceived Age Is a Reliable Measure of Biological Age**. PeerJ, 2024. Disponível em <https://pmc.ncbi.nlm.nih.gov/articles/PMC11587682/>

ULTRALYTICS. **YOLO Documentation**. Disponível em: <https://docs.ultralytics.com/>

VIOLA, P.; JONES, M. **Rapid Object Detection Using a Boosted Cascade of Simple Features**. IEEE Computer Vision and Pattern Recognition (CVPR), p. 511–518, 2001. Disponível em <https://ieeexplore.ieee.org/document/990517>

WANG, C.-Y. et al. **YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors**. 2022. Disponível em <https://ieeexplore.ieee.org/document/10204762>

ZHANG, Y.; BENGIO, Y. **Dive into Deep Learning**. 2021. Disponível em: <https://d2l.ai>



## **Apêndice A – Modelo de Aviso ao Público**

Neste espaço são capturadas imagens para fins de estudo sobre exibição de conteúdo. As imagens serão tratadas pela equipe do projeto: [nome do responsável / instituição]. Finalidade: avaliar recomendações de produto por faixa etária e gênero inferidos e recomendar produtos. Para dúvidas ou para exercer direitos previstos na LGPD, entre em contato: [e-mail / telefone].

## **Apêndice B – Modelo de Termo de Consentimento**

A Empresa [nome] autoriza a instalação temporária do equipamento de filmagem e exibição conectado à TV (HDMI) no endereço [endereço].

Finalidade: pesquisa acadêmica [título do TCC].

Responsável pelo tratamento: [seu nome / instituição / e-mail].

Tipos de dados: imagens capturadas e metadados para inferência demográfica (idade/gênero).

Retenção: dados anonimizados serão mantidos localmente no estabelecimento.

Segurança: dados armazenados em ambiente seguro e acessíveis apenas à equipe autorizada.

Condições: não serão instaladas qualquer equipamentos em locais de privacidade; sinalização visível; cumprir LGPD; empresa permanecerá informada de qualquer incidente.

Assinatura representante da Empresa / data.