

**CENTRO PAULA SOUZA
FACULDADE DE TECNOLOGIA DE FRANCA
“Dr. THOMAZ NOVELINO”**

TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

**MICHELE MAYRA DA SILVA
WANDERSON HONORIO ROBERTO DA SILVA**

**PROVADOR VIRTUAL MULTIMODAL COM KINECT PARA O VAREJO
DE MODA**

**FRANCA/SP
2025**

MICHELE MAYRA DA SILVA
WANDERSON HONORIO ROBERTO DA SILVA

**PROVADOR VIRTUAL MULTIMODAL COM KINECT PARA O VAREJO
DE MODA**

Trabalho de Graduação apresentado à Faculdade de Tecnologia de Franca - “Dr. Thomaz Novelino”, como parte dos requisitos obrigatórios para obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Orientador: Profa. Dra. Jaqueline Brigladori
Pugliesi

FRANCA/SP

2025

MICHELE MAYRA DA SILVA
WANDERSON HONORIO ROBERTO DA SILVA

PROVADOR VIRTUAL MULTIMODAL COM KINECT PARA O VAREJO
DE MODA

Trabalho de Graduação apresentado à Faculdade de Tecnologia de Franca - “Dr. Thomaz Novelino”, como parte dos requisitos obrigatórios para obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Orientador: Profa. Dra. Jaqueline Brigladori
Pugliesi

Trabalho avaliado e aprovado pela seguinte Banca Examinadora:

Orientador(a) : Profa. Dra. Jaqueline Brigladori Pugliesi
Nome : Orientador
Instituição : Faculdade de Tecnologia de Franca – “Dr. Thomaz Novelino”

Examinador(a) 1 :
Nome : Profa. Dra. Sílvia Regina Viel
Instituição : Faculdade de Tecnologia de Franca – “Dr. Thomaz Novelino”

Examinador(a) 2 :
Nome : Prof. Me. Claudio Eduardo Paiva
Instituição : Faculdade de Tecnologia de Franca – “Dr. Thomaz Novelino”

Franca, 09 de Dezembro de 2025

AGRADECIMENTO

Agradecemos, em primeiro lugar, a Deus, pela vida, pela força e pela oportunidade de concluir esta etapa tão importante da nossa formação.

Às nossas famílias, que nos sustentaram com amor, paciência e incentivo em todos os momentos de dúvida, cansaço e renúncia. Sem o apoio de vocês, este trabalho não teria sido possível.

À Profa. Jaqueline, pela orientação atenta, pela paciência ao longo de todo o projeto e por acreditar na relevância desta pesquisa, incentivando-nos a ir além do óbvio e do que já estava posto no meio acadêmico.

Ao Professor Carlos Eduardo de França Roland, por nos inspirar a enxergar a tecnologia de forma crítica e criativa, sempre nos provocando a buscar novas referências e a ampliar nossa visão de mundo e de projeto.

Ao Centro Paula Souza e à Fatec, pelo compromisso com a educação pública de qualidade e pela oportunidade de cursar uma graduação que alia teoria, prática e impacto social.

Por fim, agradecemos a todos os colegas, professores e profissionais que, de alguma forma, contribuíram com conversas, sugestões, críticas construtivas ou apoio moral ao longo desta jornada. A todos o nosso muito obrigado.

Dedicamos este Trabalho de Graduação a Deus, que sustentou cada passo desta jornada, e à nossa família, em especial aos nossos pais, que nos apoiaram com fé, afeto e coragem. À professora Jaqueline, pela paciência, orientação e apoio constante ao longo do desenvolvimento deste projeto. Ao professor Carlos Eduardo de França Roland, por sempre nos incentivar a buscar e enxergar além do que o meio acadêmico já ofertava.

“Nós moldamos nossos sistemas digitais, e eles, depois, moldam a forma como consumimos, trabalhamos e nos vemos.” — inspirado em Marshall McLuhan

RESUMO

O comércio eletrônico de moda no Brasil apresenta faturamento crescente, mas convive com taxas elevadas de devolução em razão da dificuldade de antecipar tamanho, caimento e adequação estética das peças no corpo do consumidor. Nesse contexto, este trabalho propõe um estudo conceitual sobre provedores virtuais de moda baseados em Realidade Aumentada, Inteligência Artificial e sensor Kinect, com interação multimodal por voz e gestos, tomando o projeto ARTificial como prova de conceito. A pesquisa combina revisão estruturada da literatura sobre provedores virtuais, sensores de profundidade, interfaces multimodais e recomendação de moda com a sistematização de requisitos funcionais e não funcionais para um provedor em ambiente de loja física integrado a aplicações móveis. A partir dessa base, são definidas uma arquitetura conceitual centrada no Kinect como hub de captura de corpo e voz e diretrizes de concepção de interface e experiência do usuário, materializadas em protótipos de alta fidelidade desenvolvidos em Figma. O estudo contribui ao explicitar um modelo de provedor virtual multimodal alinhado ao contexto brasileiro e organizado para orientar uma futura etapa de desenvolvimento técnico e avaliação experimental de usabilidade, desempenho de captura e impactos sobre a incerteza de compra.

Palavras-chave: Captura de movimento. Experiência do usuário. Interface multimodal. Kinect. Provedor virtual. Realidade aumentada.

ABSTRACT

The Brazilian online fashion market has shown consistent revenue growth but still faces high product return rates due to the difficulty of anticipating size, fit, and visual suitability of garments on the consumer's body. In this context, this study presents a conceptual investigation of fashion virtual fitting rooms based on Augmented Reality, Artificial Intelligence, and the Kinect sensor, with multimodal interaction through voice and gestures, using the ARTificial project as a proof of concept. The research combines a structured literature review on virtual fitting rooms, depth sensors, multimodal interfaces, and fashion recommendation models with the systematization of functional and non-functional requirements for an in-store fitting room integrated with mobile applications. Based on this foundation, a conceptual architecture is defined, positioning Kinect as a central hub for body and voice capture, alongside design guidelines for user interface and experience, which are materialized in high-fidelity prototypes developed in Figma. The study contributes by outlining a multimodal virtual fitting room model tailored to the Brazilian context and organized to guide a subsequent phase of technical development and experimental evaluation of usability, capture performance, and its impact on purchase uncertainty.

Keywords: Augmented Reality. Fashion. Kinect Sensor. Multimodal Interfaces. User Experience. Virtual Fitting Room.

LISTA DE FIGURAS

| | | |
|------------------|--|-----|
| Figura 1 | – Totem de provador virtual da Renner | 24 |
| Figura 2 | – Esquema de fusão multimodal de voz e gestos no Artificial | 26 |
| Figura 3 | – Componentes do sensor Kinect v2 | 28 |
| Figura 4 | – Componentes do Kinect v2 e o fluxo de dados entre seus elementos | 30 |
| Figura 5 | – Visão geral da arquitetura conceitual do sistema ARTificial | 35 |
| Figura 6 | – Visão Frontal Componentes | 36 |
| Figura 7 | – Visão Posterior Componentes | 36 |
| Figura 8 | – Business Model Canvas | 69 |
| Figura 9 | – BPMN | 75 |
| Figura 10 | – Estrutura geral do protótipo ARTificial em Figma | 90 |
| Figura 11 | – Tela de espera do espelho ARTificial | 91 |
| Figura 12 | – Tela de login do espelho ARTificial | 92 |
| Figura 13 | – Tela de tutorial de gestos do espelho Artificial | 92 |
| Figura 14 | – Telas de experimentação por categoria e produto | 93 |
| Figura 15 | – Telas de captura de look e de galeria Artificial | 94 |
| Figura 16 | – Telas finalização da experiência no espelho Artificial | 95 |
| Figura 17 | – Tela de perfil do usuário | 96 |
| Figura 18 | – Tela de login do aplicativo Cliente do sistema ARTificial | 97 |
| Figura 19 | – Tela de login do espelho via leitura de QR Code | 98 |
| Figura 20 | – Tela de carrinho e reserva | 99 |
| Figura 21 | – Tela de coleção de looks | 100 |
| Figura 22 | – Tela de perfil do usuário no aplicativo Cliente | 101 |
| Figura 23 | – Tela de login do Sistema lojista | 102 |
| Figura 24 | – Tela de cadastro de produtos | 103 |
| Figura 25 | – Tela de checkout e registro de venda | 104 |
| Figura 26 | – Tela de perfil do Lojista e painel de insights | 105 |

LISTA DE QUADROS

| | |
|---|----|
| Quadro 1 – Comparação entre tipos provedores | 21 |
| Quadro 2 – Termo de Abertura do Projeto (TAP) – Artificial | 61 |
| Quadro 3 – Matriz SWOT | 70 |
| Quadro 4 – Matriz 5W2H | 71 |
| Quadro 5 – Requisitos Funcionais do sistema | 77 |
| Quadro 6 – Requisitos Não Funcionais do sistema | 80 |
| Quadro 7 – Regras de Negócio do sistema | 82 |
| Quadro 8 – Casos de Uso | 84 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 – Síntese de trabalhos sobre provedores virtuais com Kinect e RA | 18 |
|--|----|

LISTA DE SIGLAS

- ABCOMM** – Associação Brasileira de Comércio Eletrônico
- ABIT** – Associação Brasileira da Indústria Têxtil e de Confecção
- ABNT** – Associação Brasileira de Normas Técnicas
- API** – Application Programming Interface (Interface de Programação de Aplicações)
- BPMN** – Business Process Model and Notation (Notação para Modelagem de Processos de Negócio)
- CRM** – Customer Relationship Management (Gestão de Relacionamento com o Cliente)
- DER** – Diagrama Entidade-Relacionamento
- EAP** – Estrutura Analítica do Projeto
- ERP** – Enterprise Resource Planning (Planejamento dos Recursos Empresariais)
- GAN** – Generative Adversarial Network (Rede Generativa Adversária)
- GPU** – Graphics Processing Unit (Unidade de Processamento Gráfico)
- IA** – Inteligência Artificial
- IHC** – Interação Humano-Computador
- IEMI** – Instituto de Estudos e Marketing Industrial
- PMBOK** – Project Management Body of Knowledge
- PMI** – Project Management Institute
- QR** – Quick Response (Código de Resposta Rápida)
- RA** – Realidade Aumentada
- SDK** – Software Development Kit (Kit de Desenvolvimento de Software)
- SEBRAE** – Serviço Brasileiro de Apoio às Micro e Pequenas Empresas
- TAP** – Termo de Abertura do Projeto
- UX** – User Experience (Experiência do Usuário)
- VTON** – Virtual Try-On (Experimentação Virtual de Roupas)

SUMÁRIO

| | |
|--|-----------|
| 1 INTRODUÇÃO | 13 |
| 2 TRABALHOS RELACIONADOS E PANORAMA TECNOLÓGICO | 16 |
| 2.1 Provedores virtuais e realidade aumentada na moda | 19 |
| 2.1.1 Sistemas baseados em Kinect e sensores de profundidade | 19 |
| 2.1.2 Abordagens de virtual try-on com redes neurais e GANs | 20 |
| 2.1.3 Sistemas focados em recomendação de tamanho | 22 |
| 2.2 Soluções brasileiras e estado da prática | 23 |
| 2.3 Interfaces multimodais: voz, gestos e visão computacional | 25 |
| 3 FUNDAMENTOS TECNOLÓGICOS | 28 |
| 3.1 Sensor Kinect, captura de movimento e captação de voz | 28 |
| 3.2 Redes neurais para virtual try-on e recomendação de tamanho | 30 |
| 3.3 Simulação e retexturização de tecidos no corpo do usuário | 31 |
| 3.4 Interfaces multimodais: fusão de voz, gesto e visão | 33 |
| 4 ARQUITETURA CONCEITUAL DA ARTIFICIAL | 35 |
| 4.1 Componentes físicos do sistema | 36 |
| 4.2 Camada de captura e interpretação | 37 |
| 4.3 Camada de interação e fluxo multimodal | 38 |
| 4.4 Camada de apresentação e protótipos em Figma | 40 |
| 5 METODOLOGIA | 42 |
| 5.1 Tipo de pesquisa e abordagem | 42 |
| 5.2 Procedimentos de revisão e mapeamento bibliográfico | 42 |
| 5.3 Planejamento do projeto e artefatos de gestão (TAP e EAP) | 43 |
| 5.4 Etapas de desenvolvimento conceitual e de prototipação | 44 |
| 5.5 Limitações metodológicas | 45 |
| 6 RESULTADOS ESPERADOS E DIRETRIZES PROPOSTAS | 47 |
| 6.1 Síntese conceitual esperada | 47 |
| 6.2 Diretrizes para uso do Kinect como hub de captura corpo + voz | 47 |
| 6.3 Diretrizes para o design de comandos de voz em provedores virtuais | 48 |
| 6.4 Diretrizes para combinação de voz e gestos em uma interface multimodal | 49 |
| 6.5 Implicações esperadas para UX, varejo e pesquisa futura | 50 |
| 7 DISCUSSÃO | 52 |
| 7.1 ARTificial frente a soluções baseadas em foto e aplicativos móveis | 52 |
| 7.2 ARTificial frente a sistemas de recomendação de tamanho | 52 |
| 7.3 ARTificial frente a experimentos de RA sem multimodalidade | 53 |
| 7.4 Limitações da proposta | 53 |
| 7.5 Implicações para pesquisas futuras | 54 |
| CONSIDERAÇÕES FINAIS | 56 |
| REFERÊNCIAS | 58 |

| | |
|--|----|
| Apêndice A – Termo de Abertura do Projeto (TAP) | 61 |
| Apêndice B – Estrutura Analítica do Projeto (EAP) | 63 |
| Apêndice C – Matrizes SWOT e Plano de Ação 5W2H do Projeto ARtificial | 68 |
| Apêndice D – Artefatos de Engenharia de Requisitos | 73 |
| Apêndice E – Protótipo de Interfaces no Figma – ARtificial | 89 |

1 INTRODUÇÃO

O setor de moda no Brasil se destaca como um dos pilares da economia criativa. Em 2022, a produção de vestuário alcançou R\$ 193,2 bilhões, o que representou 6,6% da indústria de transformação do país (ABIT/IEMI, 2023). No mesmo período, a digitalização do varejo impulsionou o comércio eletrônico, que movimentou R\$ 137,6 bilhões, respondendo por cerca de 20% das compras de moda no país (ABComm, 2023). Apesar desse crescimento, o e-commerce de moda convive com um problema estrutural: as altas taxas de devolução, estimadas entre 15% e 30% das compras, com maior incidência em categorias como vestuário e calçados (Sebrae, 2023). Além de pressionar margens e operações logísticas, esse cenário intensifica o impacto ambiental da logística reversa e do descarte de produtos.

A frustração do consumidor ao receber um produto que não corresponde às expectativas – em tamanho, caimento, proporção ou conforto visual – é um dos principais vetores dessas devoluções (Almeida, 2018). No ambiente online, a ausência de uma experiência tátil e da prova física desloca o risco da escolha de tamanho para o usuário, que decide com base em fotos, tabelas de medidas e, muitas vezes, tentativas sucessivas.

Nos últimos anos, o varejo passou a experimentar soluções baseadas em Realidade Aumentada (RA) e Inteligência Artificial (IA) para reduzir essa incerteza. A RA permite sobrepor elementos virtuais ao corpo do usuário ou ao ambiente físico, aproximando a experiência da prova de roupas (Lee; Xu, 2019). Em paralelo, redes neurais e modelos de aprendizado de máquina têm sido aplicados para estimar medidas corporais, recomendar tamanhos e personalizar ofertas a partir de imagens ou históricos de compra (Adikari et al., 2020; Mistry; Wu, 2022). Reportagens recentes sobre o mercado brasileiro apontam o avanço de provedores virtuais em redes como Renner, soluções de recomendação de tamanho como as da Sizebay e iniciativas de realidade aumentada no varejo de moda, mas em geral com foco em IA aplicada a fotos e recomendação de medidas, sem detalhamento técnico de captura de movimento em tempo real ou uso de sensores de profundidade como o Kinect.

Na literatura científica, estudos como os de Isikdogan e Kara (2012) e Elisson (2015) evidenciam a viabilidade de provedores virtuais baseados em sensores de profundidade – notadamente o Microsoft Kinect – para sobrepor modelos 3D de roupas ao corpo do usuário em tempo real, explorando o esqueleto estimado pelo

sensor e a retexturização de malhas. Outros trabalhos avançam em módulos específicos: captura de gestos para controle sem toque, estimativa de pose em 3D, recomendação de tamanhos a partir de medidas corporais e avaliação de realismo visual na aplicação de tecidos. No entanto, a maior parte dessas propostas é fragmentada: trata de reconhecimento de gestos, ou de try-on, ou de recomendação, mas raramente integra captura 3D, interação gestual e comando por voz em uma arquitetura multimodal única voltada ao varejo de moda.

No contexto brasileiro, essa lacuna é ainda mais evidente: embora existam soluções comerciais de “provador virtual” e recomendação de tamanho, há pouca produção acadêmica nacional documentando arquiteturas completas que combinem RA, IA, captura de movimento em tempo real e interfaces multimodais (voz e gestos) com base em sensores acessíveis como o Kinect. Esse vazio teórico-prático limita tanto a difusão de boas práticas de concepção quanto a possibilidade de replicação e avaliação sistemática dessas soluções.

Diante desse cenário, este trabalho utiliza o conceito de **ARtificial** como prova conceitual: um provador virtual de moda que opera em tempo real, apoiado em RA, IA e no sensor Kinect, articulando três frentes principais: (i) captura de movimento e pose corporal; (ii) aplicação de tecidos e peças virtuais sobre o corpo do usuário; e (iii) interação multimodal por gestos e comandos de voz, sem contato físico com a interface. Nesta primeira etapa, o foco está na concepção e prototipação em alta fidelidade (Figma) e na sistematização de diretrizes de UX e arquitetura de solução, sem implementação de software nem avaliação de desempenho em ambiente real.

Importante destacar que o ARtificial não corresponde a uma solução comercial já existente no mercado, mas a um projeto acadêmico concebido e desenvolvido pelos autores no âmbito do Trabalho de Graduação. Sua denominação, identidade visual, fluxos de interação e protótipos em alta fidelidade foram definidos especificamente para este estudo, com o objetivo de servir como prova de conceito (proof of concept) e como artefato de apoio à investigação sobre provadores virtuais de moda baseados em RA, IA e sensor Kinect. A grafia “ARtificial”, com as letras “AR” em caixa alta, é adotada intencionalmente ao longo do texto para evidenciar a centralidade da Realidade Aumentada na concepção do projeto.

A questão de pesquisa que orienta este estudo é quais requisitos, tecnologias e fatores críticos devem orientar a concepção e a prototipação de provadores virtuais

de moda com sensor Kinect, utilizando Realidade Aumentada e Inteligência Artificial em uma interface multimodal (voz e gestos), antes da sua implementação técnica?

A partir dessa questão, o trabalho assume as seguintes proposições orientadoras:

1. Sobreposição digital e caimento virtual: a sobreposição de elementos digitais ao corpo do usuário – com uso de sensores de profundidade como o Kinect – tende a reduzir a incerteza sobre tamanho e caimento e a aumentar a confiança na decisão de compra (Adikari et al., 2020; Lee; Xu, 2019).
2. Interfaces sem toque e multimodalidade: interfaces sem toque baseadas em gestos e comandos de voz podem ampliar acessibilidade, engajamento e fluidez da interação, especialmente em ambientes de loja física, reduzindo atrito na navegação por categorias e opções de look (Alzamzami et al., 2023).
3. Integração *omnichannel* e personalização por IA: a integração entre o provador virtual, aplicações mobile e sistemas de loja possibilita estratégias *omnichannel*, como salvar looks, gerar QR codes com peças e tamanhos, e sincronizar recomendações personalizadas com o histórico do cliente (Almeida, 2018).

O objetivo geral deste estudo é identificar e sistematizar diretrizes de concepção e experiência do usuário para provedores virtuais de moda com RA, IA e sensor Kinect, organizando-as em uma arquitetura conceitual e representando-as em protótipos de alta fidelidade desenvolvidos em Figma, tomando o ARTificial como prova de conceito.

Os objetivos específicos são: (a) sintetizar evidências sobre RA, IA, captura de movimento e interfaces multimodais aplicadas a provedores virtuais; (b) discutir sobre conceitos de realidade aumentada, inteligência artificial, captura de movimento e interfaces multimodais aplicadas a provedores virtuais; (c) mapear requisitos críticos de experiência do usuário, interação por gestos e comandos de voz em ambientes de loja física, identificando elementos essenciais para a concepção de interfaces sem toque; (d) organizar esses requisitos em um conjunto estruturado de diretrizes práticas de UX, interação multimodal e arquitetura conceitual para provedores virtuais baseados em sensores de profundidade; (e) representar essas diretrizes em protótipos de alta fidelidade (Figma), incluindo fluxos de uso, telas e pontos de integração *omnichannel*, tomando o ARTificial como prova de conceito.

2 TRABALHOS RELACIONADOS E PANORAMA TECNOLÓGICO

O desenvolvimento de provadores virtuais de moda com suporte a RA, IA e captura de movimento em tempo real tem sido tema de pesquisas acadêmicas e de soluções comerciais em diferentes países. De modo geral, a literatura internacional se organiza em três grandes frentes: (i) sistemas baseados em sensores de profundidade, com destaque para o Microsoft Kinect; (ii) abordagens de *virtual try-on* (VTON), incluindo tanto técnicas de RA quanto métodos baseados em visão computacional e redes neurais; e (iii) soluções de recomendação de tamanho, muitas vezes integradas a plataformas de e-commerce, mas nem sempre articuladas a interfaces imersivas ou multimodais.

No eixo dos provadores virtuais com sensores de profundidade, trabalhos como os de Isikdogan e Kara (2012) comprovam a viabilidade de utilizar o Kinect como base para provadores em tela ou espelho digital, alinhando malhas 3D de roupas ao esqueleto estimado pelo sensor em tempo real. Estudos subsequentes, como os de Araujo (2015), aprimoram o pipeline de *virtual dressing* ao combinar captura de profundidade, sobreposição de roupas e avaliação de desempenho visual, enfatizando a responsividade e o realismo da aplicação de tecidos sobre o corpo do usuário. Adikari et al. (2020) avançam na discussão ao investigar se um único sensor de profundidade é suficiente para sustentar a simulação 3D de roupas em corpo inteiro, explorando limites de alcance, ruído e estabilidade da captura.

A literatura de *virtual try-on* baseada em RA amplia esse cenário ao explorar diferentes configurações de hardware e técnicas de alinhamento de roupas virtuais. Hilsmann; Eisert (2013), no trabalho sobre *tracking and retexturing cloth* em tempo real, mostram como o Kinect pode ser usado para rastrear e retexturizar roupas sobre o corpo do usuário, aproximando o sistema de um provador virtual com forte ênfase na física do tecido e na plausibilidade visual. Joglekar; Gohokar (2018), em “*Virtual Cloth Try-On Using Augmented Reality*”, discutem a utilização de sensores de profundidade e RA em espelhos digitais, buscando equilibrar usabilidade, qualidade de sobreposição e realismo do caimento. Em paralelo, estudos como o de Erra, Scanniello e Colonnese (2018) investigam a efetividade de provadores com RA baseados apenas em câmera RGB, focando menos na precisão geométrica e mais na percepção de utilidade, confiança e intenção de compra em contextos de e-commerce.

As soluções que combinam provedores virtuais e recomendação de tamanho formam um terceiro eixo relevante. Kurniawati et al. (2020) propõem um sistema de recomendação de tamanho acoplado a uma simulação de *fitting* em tempo real, utilizando sensores de profundidade no estilo Kinect para estimar medidas corporais e sugerir tamanhos mais adequados. Esse tipo de abordagem aproxima o campo dos provedores virtuais das práticas hoje adotadas em plataformas comerciais, nas quais algoritmos de recomendação se apoiam em medidas, perfis de corpo e históricos de compra, ainda que nem sempre haja uma camada explícita de RA ou simulação 3D.

Do ponto de vista acadêmico, observa-se que grande parte desses estudos se concentra em componentes específicos da cadeia de valor do provedor virtual — como captura esquelética, retexturização de roupas, reconhecimento de gestos, recomendação de tamanho ou simulação física de tecidos — sem necessariamente integrá-los em uma arquitetura multimodal única orientada ao varejo de moda. O resultado é um panorama tecnicamente rico, mas fragmentado: há contribuições robustas em captura de movimento, em RA aplicada ao corpo do usuário e em modelos de recomendação, porém poucas propostas articulam, em um mesmo sistema, visão computacional, voz e gestos como canais complementares de interação.

No contexto de mercado, relatórios e cases de empresas de tecnologia e varejo indicam a difusão de provedores virtuais e soluções de recomendação de tamanho em *e-commerces* e lojas físicas, inclusive no Brasil. Plataformas comerciais exploram RA em espelhos digitais, provedores baseados em foto e algoritmos de ajuste de tamanho, mas frequentemente tratam o Kinect apenas como uma das possíveis fontes de dados ou substituem sensores de profundidade por câmeras RGB e modelos de IA. Além disso, o detalhamento técnico sobre arquitetura, fluxos de dados e desenho de interação multimodal é, em geral, limitado em documentos de mercado, o que reforça a necessidade de estudos acadêmicos que sistematizem requisitos e diretrizes de concepção.

A partir do mapeamento conduzido neste trabalho, foram selecionadas vinte e nove publicações centrais para a discussão sobre provedores virtuais, sensores de profundidade e RA aplicada à moda. Dentre elas, um subconjunto de sete artigos se destaca por abordar de forma direta o uso do Kinect ou de sensores de profundidade em provedores virtuais, com foco em simulação de roupas, RA e, em alguns casos, recomendação de tamanho. A Tabela 1 sintetiza esses estudos, destacando a

tecnologia principal, o tipo de aplicação, o modo de interação e o foco de cada proposta, e serve como base para a derivação dos requisitos conceituais que orientam a arquitetura do ARtificial.

Tabela 1 – Síntese de trabalhos sobre provadores virtuais com Kinect e RA

| Autor(es) / Ano | Tecnologia principal | Tipo de aplicação | Tipo de interação | Foco do estudo |
|---|--|---|-------------------------------|--|
| Isikdogan; Kara (2012) | Kinect (RGB + sensor de profundidade) | Provador virtual em tela / espelho digital | Gestos + menus na tela | Simulação em tempo real (geralmente mapeamento 2D/Warping). |
| Araujo (2015) | Kinect + RA | Provador virtual com sobreposição de roupas | Gestos + interface gráfica | Desenvolvimento de sistema de medidas livres de marcadores e sobreposição. |
| Adikari et al. (2020) | Sensor de profundidade único (Kinect) | Provador virtual de corpo inteiro com RA | Gestos + seleção em interface | Avaliar a aplicabilidade de um único sensor de profundidade para simulação 3D de roupas em tempo real. |
| Hilsmann; Eisert (2013) | Kinect / sensor de profundidade | Protótipo de provador virtual em ambiente controlado | Gestos / postura corporal | Rastreamento e retexturização de roupas em tempo real sobre o corpo do usuário. |
| Joglekar; Gohokar (2018) | Sensor de profundidade + RA | Provador virtual com espelho digital | Interface gráfica | Método de virtual try-on em RA, com foco em alinhamento de peças virtuais e realismo do tecido. |
| Erra; Scanniello; Colonnese (2018) | Câmera RGB + RA | “Provador” de RA para e-commerce | Toque / menus | Avaliar percepção de utilidade, confiança e intenção de compra em provador de RA baseado em câmera. |

| | | | | |
|------------------------------------|--|---|------------------------------------|--|
| <i>Kurniawati et al.</i> (2020) | Sensor de profundidade estilo Kinect | Provedor virtual com recomendação de tamanho | Menus + parâmetros corporais | Integração entre simulação em tempo real e recomendação de tamanho a partir de medidas corporais. |
|------------------------------------|--|---|------------------------------------|--|

Fonte: Elaborado pelos autores.

2.1 Provedores virtuais e realidade aumentada na moda

Esta seção apresenta um panorama conceitual e tecnológico dos provedores virtuais aplicados ao varejo de moda, com ênfase no uso da Realidade Aumentada como estratégia para simular a experimentação de vestuário no ambiente digital e físico. O objetivo é contextualizar as principais abordagens descritas na literatura, destacando diferenças entre soluções baseadas em imagens bidimensionais, sensores de profundidade e modelos de recomendação de tamanho, bem como suas implicações para a experiência do usuário e para a redução da incerteza na decisão de compra. A partir desse enquadramento geral, as subseções seguintes aprofundam as tecnologias, modos de interação e arquiteturas mais recorrentes nesse domínio.

2.1.1 Sistemas baseados em Kinect e sensores de profundidade

Os primeiros trabalhos que se aproximam da proposta do ARtificial utilizam sensores de profundidade, sendo o Kinect o protagonista em boa parte da literatura. Esses estudos exploram o esqueleto 3D gerado pelo sensor, a captura de pose em tempo real e a possibilidade de sobrepor roupas virtuais ao corpo do usuário. Em geral, eles se concentram em três linhas principais: (a) salas de prova virtuais em tela grande, com o usuário em pé, diante de uma TV ou espelho digital; (b) controle por gestos para navegação entre peças, troca de modelos e captura de fotos; e (c) avaliação de desempenho em termos de tempo de processamento, estabilidade do rastreamento e percepção de realismo do caimento. A Tabela 1 sintetiza alguns desses trabalhos que usam o Kinect ou sensores de profundidade em contextos de provedores virtuais.

Trabalhos de “*virtual dressing room*” com Kinect tendem a adotar uma pipeline semelhante: segmentação do corpo, estimativa de esqueleto, ajuste de uma malha ou silhueta de roupa e exibição do resultado em tempo real. Em alguns casos, há um

avatar intermediário (um “corpo virtual” que imita a pose do usuário), sobre o qual as roupas são aplicadas; em outros, a roupa é sobreposta diretamente à imagem RGB do usuário. Em todos, a presença do sensor de profundidade é determinante para estimar o volume do corpo e alinhar a peça em 3D, superando limitações das soluções que trabalham apenas com fotos bidimensionais.

Além da aplicação em moda, há uma base relevante de artigos que usam o Kinect para reconhecer gestos, postura, caminhar (*gait*), mãos e braços em diferentes domínios (jogos, navegação em museus, robótica, vigilância, automação residencial). Embora não sejam focados em vestuário, esses estudos consolidam técnicas de rastreamento esquelético, detecção de mãos e interpretação de gestos que podem ser adaptadas para o contexto de provadores virtuais — por exemplo, levantar a mão para iniciar a sessão, acenar para trocar o look ou apontar para um botão virtual na interface do espelho.

Esses trabalhos constituem a “espinha dorsal” tecnológica do ARtificial: confirmam que o Kinect é capaz de operar em tempo real, identificar o corpo do usuário com precisão suficiente para aplicações interativas e servir como hub de captura de movimento para experiências imersivas em tela grande. Ao mesmo tempo, evidenciam limitações que precisam ser consideradas nas diretrizes de concepção, como a sensibilidade a oclusões (braços cruzados, corpos múltiplos na cena) e a necessidade de distância e enquadramento adequados diante do sensor.

2.1.2 Abordagens de virtual *try-on* com redes neurais e GANs

Uma segunda vertente da literatura sobre provadores virtuais se apoia em modelos de *virtual try-on* baseados em redes neurais, muitas vezes utilizando arquiteturas de tipo *encoder-decoder* e redes generativas adversariais (GANs). Em vez de depender de sensores de profundidade, esses sistemas operam, em geral, apenas com imagens: uma foto da pessoa e uma foto da peça de roupa, gerando como saída uma imagem sintética da pessoa vestindo aquela peça.

Enquanto os sistemas baseados em Kinect privilegiam a interação em tempo real diante de um espelho digital, os modelos VTON concentram-se principalmente na qualidade visual da síntese: preservação da identidade facial, coerência anatômica, deformação adequada da roupa em diferentes poses e manutenção de sombras, texturas e detalhes de tecido. Em muitos casos, são empregados mecanismos de controle de pose e representações intermediárias (como mapas de segmentação,

malhas paramétricas ou espaços latentes estruturados), viabilizando a transferência de roupas entre corpos distintos ou poses variadas.

A comparação entre provedores virtuais baseados em Kinect e sistemas VTON baseados em imagem é sintetizada no Quadro 1, evidenciando diferenças quanto à fonte de dados, contexto de uso, tipo de interação, saídas geradas e aderência ao escopo do ARTificial.

Quadro 1 – Comparação entre tipos provedores

| Aspecto | Sistemas baseados em Kinect / Sensor de Profundidade | Sistemas de virtual try-on (VTON) baseados em imagem / GANs |
|---------------------------------|--|---|
| Fonte de dados principal | Fluxo de depth + RGB em tempo real (esqueleto 3D, pose, distância do usuário). | Imagens estáticas (foto da pessoa + foto da roupa, às vezes segmentadas). |
| Contexto típico de uso | Loja física, espelho digital / tela grande, usuário em pé diante do sensor. | E-commerce, app ou web: usuário envia foto ou vê um modelo “similar” vestindo a peça. |
| Tipo de interação | Interação em tempo real, com gestos (navegação sem toque) e, potencialmente, voz. | Interação assíncrona ou quase em tempo real, focada em clicar, escolher peças e visualizar combinações. |
| Saída principal (output) | Sobreposição de roupas 2D/3D ao corpo capturado ao vivo; avatar ou imagem do próprio usuário em movimento. | Imagem sintetizada da pessoa vestindo a peça, com forte ênfase em realismo visual e coerência anatômica. |
| Papel da profundidade | Central: profundidade permite estimar esqueleto, volume do corpo e alinhamento espacial da peça. | Geralmente ausente: o modelo “aprende” deformação e caimento a partir de grandes bases de imagens. |
| Pontos fortes | Sensação de “espelho” e presença corporal; experiência lúdica; possibilidade de gestos e comandos de voz; adequado para loja física. | Alto nível de realismo na imagem final; pode operar só com fotos; escalável para grandes catálogos em e-commerce. |
| Principais limitações | Sensível a oclusões, distância e enquadramento; depende de hardware específico (Kinect ou similar); instalação no ponto de venda. | Menor foco em interação ao vivo; pode exigir muitas imagens rotuladas; desafios de privacidade e de generalização para corpos diversos. |
| Aderência ao ARTificial | Base tecnológica direta: o ARTificial assume o Kinect como hub de captura corpo + gestos + voz em tempo real. | Inspira módulos futuros de “aplicação de tecidos” mais realistas, combinando pose capturada pelo Kinect com modelos neurais de VTON. |

Fonte: Elaborado pelos autores.

Essas abordagens de *virtual try-on* orientadas por redes neurais, portanto, deslocam o foco da experiência de “espelho ao vivo” para a qualidade da imagem gerada. Do ponto de vista deste estudo, elas são relevantes por, ao menos, dois motivos. Em primeiro lugar, indicam um caminho promissor para aperfeiçoar o realismo da “aplicação de tecidos” em uma fase posterior do ARTificial, combinando a captura de pose em tempo real via Kinect com modelos VTON capazes de produzir imagens mais plausíveis do usuário vestindo diferentes peças. Em segundo lugar, reforçam a necessidade de planejar, já na etapa de concepção, quais dados o sistema precisaria coletar (poses, fotos frontais, malhas de roupa, metadados de tecido e informações de caimento) para viabilizar, em um estudo subsequente, a implementação efetiva de um módulo VTON acoplado ao protótipo aqui proposto

2.1.3 Sistemas focados em recomendação de tamanho

Um terceiro grupo de trabalhos, tanto acadêmicos quanto comerciais, aborda o problema sob a ótica da recomendação de tamanho. Em vez de simular diretamente o caimento da peça, esses sistemas utilizam medidas corporais, tabelas de grade, histórico de compras ou fotos estáticas para sugerir o tamanho mais adequado. Nesse segmento, são comuns modelos que:

- estimam circunferências e proporções corporais a partir de fotos (ou de um breve questionário) e cruzam essas medidas com a grade de tamanhos de cada marca;
- constroem perfis de usuários “gêmeos” (*body twins*), recomendando tamanhos com base em pessoas com medidas semelhantes que tiveram experiência positiva com determinada peça;
- alimentam algoritmos de recomendação com dados de devolução, devolvendo menor probabilidade a combinações peça–tamanho com histórico de troca ou insatisfação.

Do ponto de vista do ARTificial, esses sistemas são importantes porque ajudam a separar duas camadas conceituais: (i) a camada de simulação visual (RA + Kinect + aplicação de tecidos), que melhora a percepção de caimento e estilo, e (ii) a camada de recomendação de tamanho, que pode operar “por baixo do capô”, sugerindo automaticamente um tamanho padrão ao carregar o look no espelho. A literatura mostra que uma boa solução de provador virtual tende a integrar as duas camadas: o

usuário vê o caimento da peça sugerida e, se desejar, pode testar visualmente outros tamanhos ou modelagens.

2.2 Soluções brasileiras e estado da prática

No contexto brasileiro, a produção acadêmica sobre provadores virtuais ainda é reduzida quando comparada ao cenário internacional, mas já existe um núcleo de trabalhos que sinaliza um campo em consolidação. Um dos marcos é a dissertação de Araujo (2015), dedicada ao desenvolvimento de um protótipo de compra de roupas on-line utilizando o Kinect como sensor de profundidade. Nessa pesquisa, o autor explora a captura do esqueleto do usuário, a segmentação do corpo e a sobreposição de modelos 3D de vestuário, estruturando um fluxo de processamento que vai da aquisição do sinal de profundidade à renderização das peças em tempo quase real. Embora o foco esteja em um único sensor e em um cenário controlado, o trabalho demonstra a viabilidade técnica de integrar Kinect, motores gráficos e lógica de provador virtual em uma arquitetura coerente, o que o aproxima diretamente dos objetivos deste estudo.

Em paralelo, há iniciativas que deslocam o problema da moda para outros domínios, mas utilizam conceitos muito próximos aos de provadores virtuais. O artigo sobre o “Espelho Virtual Interativo para Simulação de Maquiagem” Silva e Silva (2025), propõe um sistema em que o rosto do usuário é capturado por câmera e recebe, em tempo real, a aplicação virtual de cosméticos. A solução envolve detecção de regiões faciais, rastreamento de movimentos e renderização de texturas de maquiagem, compondo uma interface de espelho digital voltada ao varejo de beleza. Ainda que o foco não seja vestuário, o trabalho reforça a lógica de “provar sem tocar” e evidencia preocupações de UX semelhantes às do ARtificial: alinhamento preciso entre elementos virtuais e o corpo, tempo de resposta adequado e simplicidade na interação.

Outros estudos nacionais, como o trabalho de Queiroz; Kurafasi; Silva (2019) e Fernandes (2013), também dialogam com esse universo ao explorar temas como captura de movimento, realidade aumentada e interfaces interativas baseadas em visão computacional. Em geral, tratam de cenários específicos — jogos, treinamento, aplicações educacionais ou experimentos de interface —, mas consolidam o uso de câmeras RGB-D e técnicas de rastreamento em protótipos desenvolvidos em contextos brasileiros. Esses projetos reforçam que há capacidade técnica instalada

em grupos de pesquisa e cursos de tecnologia para trabalhar com sensores como o Kinect, mesmo que a ênfase ainda recaia em aplicações pontuais e em interações predominantemente visuais, sem integração com comandos de voz ou estratégias *omnichannel*.

Enquanto a produção acadêmica avança de forma gradual, o mercado brasileiro já opera experiências comerciais que se aproximam da ideia de provedores virtuais. Um exemplo emblemático é a ação da Renner em parceria com a Eletromidia, que instalou um provedor virtual em totens interativos fora das lojas físicas, em áreas de grande circulação. Nessa solução, o usuário interage com uma tela vertical de grande formato, navega por coleções por meio de toques e botões digitais e visualiza combinações de looks, em um ambiente fortemente orientado à experimentação de marca. A Figura 1 ilustra essa instalação, em que o totem funciona como vitrine interativa e extensão do e-commerce, aproximando a descoberta de produtos e entretenimento no espaço urbano.

Figura 1 – Totem de provedor virtual da Renner



Fonte: Divulgação Renner/Eletromidia, adaptado de Meio & Mensagem (2025).

Apesar do apelo midiático dessas iniciativas, é raro encontrar documentação técnica detalhada sobre sua arquitetura de software e hardware. As reportagens enfatizam indicadores de negócio — aumento de engajamento, impacto em reconhecimento de marca e potencial redução de devoluções —, mas pouco revelam

sobre como são capturados e tratados os dados do usuário, se há uso de sensores de profundidade ou apenas câmeras convencionais, que tipos de modelos de IA são empregados ou quais são as limitações de rastreamento e privacidade. Do ponto de vista científico, essas experiências aparecem mais como “caixas-pretas” inspiradoras do que como referências metodológicas plenamente descritas.

Nesse cenário, o ARTificial se posiciona como um esforço de síntese e explicitação. Ao dialogar com as pesquisas brasileiras que já exploram espelhos virtuais, maquiagem em RA e provadores com Kinect, o projeto assume essas contribuições como prova de viabilidade técnica e repertório de soluções de interface. Ao mesmo tempo, busca ir além da fragmentação observada tanto na literatura quanto nas soluções de mercado, propondo uma arquitetura conceitual que integra, de forma explícita, captura 3D em tempo real, interação por gestos e comandos de voz, prototipação de telas em alta fidelidade e integração com aplicações mobile para uso *omnichannel*. A principal lacuna que se pretende preencher é justamente a ausência, no contexto nacional, de descrições sistemáticas de provadores virtuais multimodais que combinem Kinect, RA e IA em um fluxo de uso completo, documentado em linguagem acadêmica e orientado por diretrizes de UX e engenharia de software.

2.3 Interfaces multimodais: voz, gestos e visão computacional

A terceira frente de referência do ARTificial está na literatura de interfaces multimodais, que discute a combinação coordenada de diferentes canais de entrada – fala, gesto, toque, visão, olhar – em um mesmo sistema interativo. Pesquisas de referência nessa área mostram que interfaces que integram voz e gestos podem reduzir a carga cognitiva, aumentar a expressividade do usuário e oferecer caminhos alternativos de interação em contextos em que o toque não é ideal (ambientes compartilhados, superfícies grandes, situações de higiene ou acessibilidade).

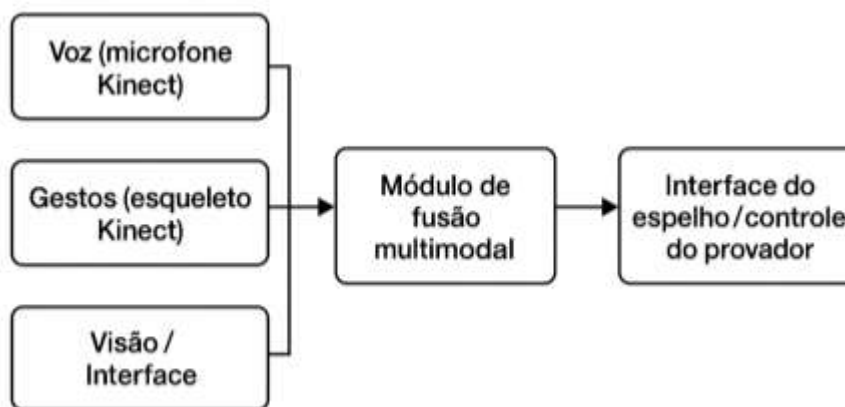
No contexto deste trabalho, a multimodalidade é entendida como um mecanismo de fusão entre três fluxos principais capturados pelo sistema:

- **Voz**, por meio do microfone array do Kinect;
- **Gestos e postura**, a partir do rastreamento esquelético do usuário;
- **Visão / Interface gráfica**, representada pelas telas do espelho interativo.

Esses fluxos convergem para um módulo de fusão multimodal, responsável por interpretar a intenção do usuário e acionar os comandos correspondentes na interface do provador. Essa arquitetura conceitual está sintetizada na Figura 2, em que se

destacam os caminhos de entrada (voz, gestos, visão), o bloco de fusão e a saída em forma de ações na interface do espelho.

Figura 2 – Esquema de fusão multimodal de voz e gestos no ARTificial



Fonte: Elaborado pelos autores.

Os modelos clássicos de multimodalidade descrevem diferentes estratégias de combinação, que inspiram diretamente o desenho do ARTificial:

- Fusão complementar – cada modalidade traz uma parte da informação, por exemplo, o usuário diz “*quero ver calças masculinas*” enquanto aponta para a área de “masculino” na tela.
- Fusão redundante – modalidades distintas carregam a mesma intenção, aumentando a robustez do reconhecimento, por exemplo, comando “*próximo look*” acompanhado de um gesto de deslizar para a direita.
- Fusão alternativa – o sistema aceita múltiplos caminhos para a mesma ação, permitindo que a pessoa escolha voz ou gesto, por exemplo, iniciar o provador apenas levantando a mão aberta ou apenas dizendo “*iniciar provador*”.

A literatura enfatiza ainda que interfaces multimodais operam sob incerteza temporal: voz e gesto não são eventos discretos, mas fluxos contínuos que precisam ser alinhados no tempo. Por isso, o esquema ilustrado na Figura 2 pressupõe três etapas:

- reconhecimento individual de cada modalidade: ASR para fala, *tracking* esquelético e de mãos para gestos;
- sincronização desses fluxos em janelas temporais;

- decisão sobre qual comando multimodal foi emitido, com base em regras de projeto ou modelos probabilísticos.

Aplicado ao contexto de provedores virtuais, esse arcabouço teórico orienta casos de uso como:

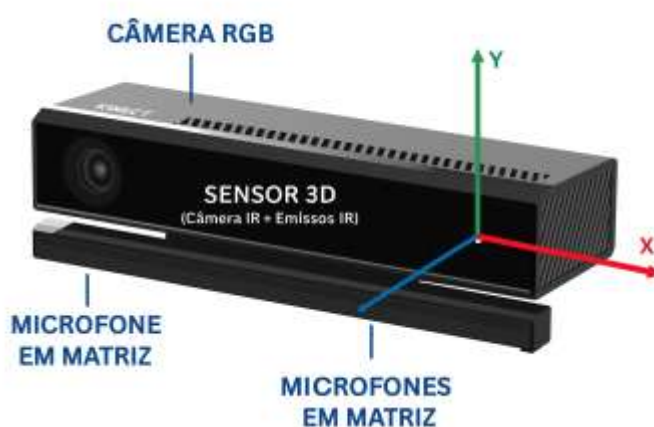
- o cliente em frente ao espelho dizendo “*iniciar provedor*” ou levantando a mão aberta (fusão alternativa ou redundante);
- comandos como “*ver tamanhos maiores*” acompanhados de um gesto de afastar as mãos, reforçando a intenção de aumento (fusão complementar);
- gestos simples (apontar, acenar, levantar a mão) associados a regiões da interface destacadas visualmente, facilitando a descoberta das possibilidades de interação.

Ao trazer a literatura de interfaces multimodais para o centro do referencial teórico e organizar essas relações no esquema da Figura 2, o ARtificial deixa de tratar voz e gestos como “recursos opcionais” e passa a enxergá-los como componentes estruturantes da arquitetura. Isso orienta tanto o desenho das telas em Figma – com pontos de ancoragem visuais para gestos e feedbacks explícitos para comandos de voz – quanto às decisões futuras de implementação, como o uso do microfone array do Kinect, a escolha de modelos de reconhecimento de fala em português e a definição do vocabulário e da gramática de comandos do sistema.

3 FUNDAMENTOS TECNOLÓGICOS

Esta seção consolida os principais fundamentos tecnológicos que estruturam o conceito do ARTificial, integrando captura corporal por profundidade, reconhecimento de voz, rastreamento gestual, modelos neurais aplicados à simulação de vestuário e princípios de interação multimodal. A proposta não é detalhar algoritmos específicos, mas organizar o estado da arte que fundamenta a arquitetura conceitual adotada neste estudo. O projeto parte da premissa de que a combinação entre sensoriamento 3D, modelos neurais e interfaces multimodais produz uma base técnica sólida para experiências de provador virtual em tempo real, particularmente quando apoiadas em hardware acessível e bem documentado. A Figura 3 ilustra essa estrutura tecnológica geral, evidenciando como os módulos de captura de voz, gestos e visão operam de forma combinada.

Figura 3 – Componentes do sensor Kinect v2



Fonte: Elaborado pelos autores.

3.1 Sensor Kinect, captura de movimento e captação de voz

O Microsoft Kinect, originalmente lançado como acessório para o Xbox e posteriormente disponibilizado em versões para PC, consolidou-se como um dos sensores de profundidade mais estudados em aplicações de interação natural, rastreamento corporal e visão computacional em tempo real. No contexto deste trabalho, utiliza-se especificamente o Kinect para Xbox One (v2), que combina câmera RGB, sensor de profundidade baseado em *time-of-flight* e um array de microfones,

reunindo em um único dispositivo os elementos necessários para capturar corpo, gestos e voz em ambiente de loja.

Do ponto de vista de hardware, o Kinect v2 fornece simultaneamente um fluxo de vídeo colorido em alta definição e um mapa de profundidade associado, permitindo estimar a distância de cada ponto da cena e derivar modelos esqueléticos tridimensionais. A partir desses dados, a SDK oficial disponibiliza a detecção automática de articulações — cabeça, tronco, ombros, cotovelos, quadris, joelhos e pés — atualizadas em tempo real. Trabalhos como Lee & Yoo (2014) mostram que a combinação RGB + profundidade é suficiente para reconstrução de vistas virtuais, correção de oclusões e síntese de novos pontos de vista, reforçando a robustez desse tipo de sensor para aplicações interativas. Estudos como Kotan & Öz (2015) mostram, adicionalmente, que o esqueleto 3D produzido pelo Kinect pode ser utilizado para posicionar roupas virtuais sobre avatares humanoides, ajustando dimensões e rotações com fidelidade adequada para protótipos de provedores virtuais.

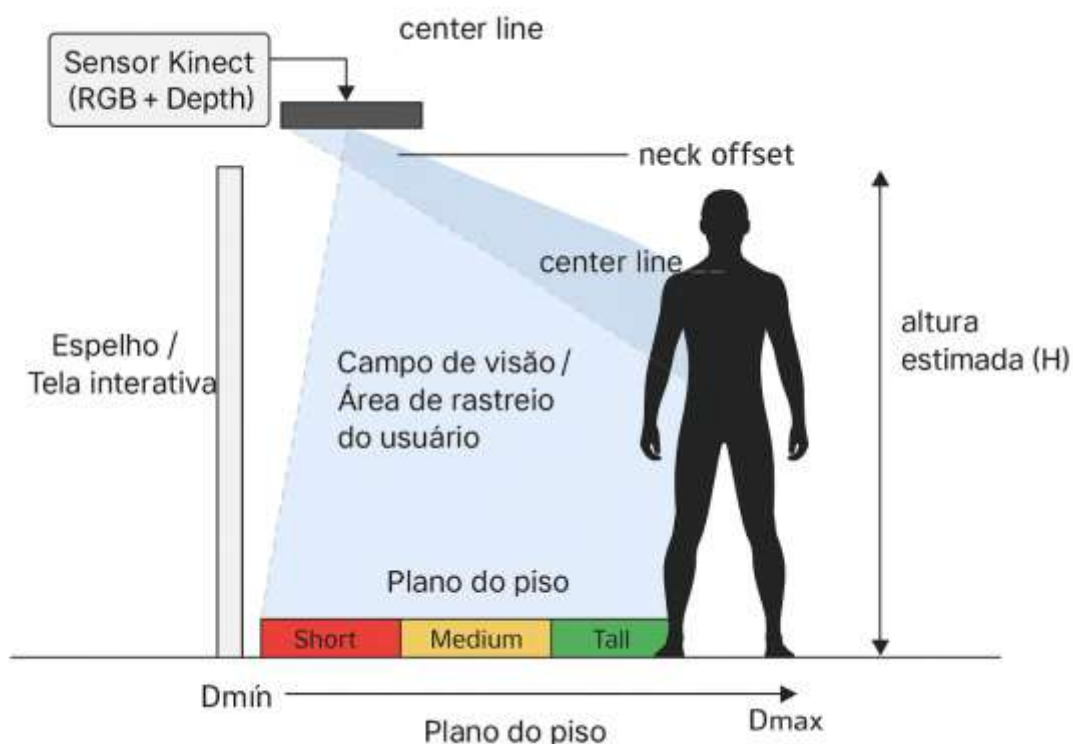
Além da visão, o Kinect incorpora um *array* de microfones projetado para reconhecimento de voz em ambientes domésticos, característica que se torna estratégica para interfaces sem toque em ambientes de varejo. O arranjo físico dos microfones, aliado a técnicas de *beamforming* e cancelamento de eco, permite atenuar ruídos e localizar a fonte de fala na direção do usuário. Isso viabiliza cenários de comando por voz sem headsets, nos quais o cliente simplesmente se posiciona diante do espelho e emite comandos naturais, como “mostrar tamanhos maiores” ou “ver opções masculinas”.

No ARTificial, o Kinect é adotado como hub único de captura de corpo e voz. A câmera RGB e o sensor de profundidade alimentam os módulos de detecção de pose, ajuste de silhueta e aplicação de roupas virtuais, enquanto o array de microfones fornece o sinal para reconhecimento de fala. Essa decisão de projeto simplifica a arquitetura física, reduz dependências externas e mantém alinhamento com a literatura que utiliza o Kinect como interface multimodal generalizada em jogos, robótica e sistemas interativos.

A Figura 4 apresenta, de forma esquemática, a arquitetura de fusão multimodal considerada no ARTificial, evidenciando como os fluxos de voz, gestos e visão convergem para um módulo de decisão que aciona a interface do espelho. No fluxo de gestos, o sistema consome os dados esqueléticos fornecidos pelo Kinect, incluindo juntas-chave como cabeça, pescoço (neck offset) e ombros, que servem de referência

tanto para alinhamento do avatar quanto para estimativas de altura e enquadramento do usuário no espelho virtual.

Figura 4 – Componentes do Kinect v2 e o fluxo de dados entre seus elementos



Fonte: Elaborado pelos autores.

3.2 Redes neurais para virtual *try-on* e recomendação de tamanho

O avanço recente dos provedores virtuais de moda está diretamente associado à evolução dos modelos de visão computacional baseados em redes neurais profundas. Mesmo em sistemas que não utilizam sensores de profundidade, é comum empregar arquiteturas convolucionais ou *encoder-decoder* para aprender representações de corpos e peças de vestuário a partir de imagens 2D.

Na literatura de virtual *try-on* (VTON), destacam-se modelos baseados em GANs e variações condicionais, como C-VTON e abordagens de garment transfer controlável. Esses modelos, em linhas gerais, seguem um pipeline conceitual semelhante: (i) extração de uma representação “agnóstica à roupa” do corpo-alvo (pose, silhueta, mapa de partes do corpo); (ii) codificação da peça de vestuário em um espaço latente que preserva textura e detalhes visuais; e (iii) síntese da imagem

composta, na qual a peça é “vestida” sobre o corpo, respeitando pose, orientação e deformações esperadas do tecido. Trabalhos mais recentes incorporam módulos de warping que aprendem campos de fluxo para deformar a imagem da roupa de acordo com a geometria do corpo, aumentando o realismo do caimento.

Em paralelo, há uma linha de pesquisa focada na recomendação de tamanho, em que redes neurais ou modelos probabilísticos utilizam medidas corporais, histórico de compras ou fotos do usuário para inferir o tamanho mais adequado em cada marca. Ferramentas comerciais como a Sizebay e soluções descritas em notícias de mercado exploram justamente essa dimensão, buscando reduzir devoluções ao ajustar a recomendação de tamanho ao biotipo do cliente.

No contexto deste artigo, as redes neurais são tratadas como componentes lógicos da arquitetura ARTificial, com dois papéis principais:

- Módulo de *try-on* visual: responsável, em uma fase futura do projeto, por gerar a imagem ou visualização da peça aplicada ao corpo do usuário, combinando dados de profundidade, pose e textura da roupa.
- Módulo de recomendação de tamanho: responsável por sugerir tamanhos com base nas medidas estimadas pelo Kinect (altura, proporções de membros) e em dados adicionais eventualmente fornecidos pelo usuário (peso, preferências de caimento).

É importante destacar que, nesta fase, esses módulos são tratados em nível conceitual: o estudo identifica quais tipos de modelos são mais adequados, quais dados seriam necessários e como esses componentes se encaixariam na arquitetura geral. A implementação concreta de um modelo de VTON ou recomendação de tamanho, bem como sua avaliação quantitativa, é explicitamente posicionada como agenda do segundo artigo, que tratará da dimensão técnica e experimental do ARTificial.

3.3 Simulação e retexturização de tecidos no corpo do usuário

Outra dimensão crítica para a experiência de provadores virtuais é a forma como o tecido é representado, deformado e aplicado ao corpo do usuário. Não basta apenas “colar” uma textura plana sobre a silhueta: a percepção de realismo depende de como o sistema lida com dobras, estiramento, queda do tecido e interação com a pose corporal.

Estudos de *tracking* e retexturização de roupas em tempo real utilizam mapas de profundidade para segmentar regiões do corpo, extrair malhas aproximadas e projetar texturas de tecido sobre essas superfícies, corrigindo o alinhamento a cada quadro (Hilsmann; Eisert, 2013; Kim; Kim, 2015). Em muitos casos, parte da roupa real é utilizada como “base geométrica”, e a textura é substituída, preservando a dinâmica física capturada pela câmera.

Outras abordagens trabalham com simulação física mais explícita, modelando o tecido como uma malha de partículas conectadas por molas (modelo massa–mola) ou por métodos de elementos finitos. Esses modelos permitem simular efeitos de gravidade, rigidez, atrito e colisão com o corpo, gerando resultados visualmente consistentes, mas com custo computacional elevado para aplicações em tempo real. Para o ARTificial, a literatura de simulação e retexturização tem duas funções principais:

- Fundamentar o nível de realismo que se pretende atingir: para uso em loja, é desejável um equilíbrio entre fidelidade visual e desempenho, de forma que a experiência seja fluida e responsiva. Isso indica a preferência por técnicas de retexturização e *warping* guiadas por profundidade, possivelmente combinadas com simplificações físicas, em vez de simulações completas de alta complexidade.
- Orientar requisitos de captura: a qualidade da simulação de tecido depende diretamente da fidelidade do mapa de profundidade, da estabilidade do rastreamento e da capacidade de lidar com oclusões (por exemplo, quando o usuário cruza os braços ou gira o corpo). Trabalhos que exploram preenchimento de buracos, filtragem temporal e reconstrução de *viewpoints* alternativos com Kinect ajudam a definir parâmetros mínimos de ambiente (iluminação, distância ao sensor, posicionamento da TV) para que o provador seja viável em contexto de varejo.

Na fase de prototipação em Figma, essa dimensão aparece principalmente como requisitos declarados (ex.: “o sistema deve preservar textura e queda visual do tecido”, “o sistema deve se adaptar a poses variadas sem distorções severas”), que serão posteriormente mapeados para escolhas de algoritmos na etapa de implementação técnica.

3.4 Interfaces multimodais: fusão de voz, gesto e visão

Interfaces multimodais são aquelas que combinam, de forma integrada, múltiplas modalidades de entrada (voz, gesto, toque, olhar, movimento corporal) e de saída (visual, auditiva, tátil). O *survey* clássico de interfaces multimodais discute como a combinação de modalidades pode aumentar robustez, naturalidade e acessibilidade das interações, desde que a fusão de sinais seja tratada de forma coerente no tempo e no significado (Oviatt; Cohen et al.).

Em termos de arquitetura, sistemas multimodais podem realizar a fusão em diferentes níveis:

- Fusão em nível de dados: sinais brutos, como áudio e coordenadas de esqueleto, são combinados diretamente, por exemplo, para estimar a direção da fala em relação à posição do corpo.
- Fusão em nível de características: extrai-se um conjunto de atributos (palavras reconhecidas, tipo de gesto, região da tela apontada) e, a partir dessa representação intermediária, decide-se o significado da interação.
- Fusão em nível de decisão: cada modalidade produz uma hipótese independente (ex.: “usuário pediu calças masculinas” pela voz; “usuário apontou para a categoria calças” pelo gesto), e essas decisões são combinadas por regras ou modelos probabilísticos.

No contexto de um provador virtual de moda, a multimodalidade oferece oportunidades específicas:

- Comandos de voz para navegação sem toque (“mostrar calças masculinas”, “filtrar por tamanho G”, “trocar cor para azul”), reduzindo fricção e aumentando inclusão de pessoas com limitações motoras.
- Gestos corporais para confirmação rápida e feedback de controle, como levantar a mão para iniciar, girar o corpo para ver o look em diferentes ângulos ou estender o braço em direção à TV para “selecionar” uma peça em destaque.
- Visão computacional para contextualizar a fala, por exemplo, interpretando “esse modelo em outro tamanho” com base na peça atualmente aplicada ao corpo do usuário.

Para que essa combinação funcione, o sistema precisa considerar o tempo como variável central: comandos de voz e gestos não acontecem isoladamente, mas em janelas de interação. Modelos de diálogo multimodal tratam justamente da sincronização entre modalidades, da resolução de ambiguidades (por exemplo,

quando o usuário fala algo impreciso, mas o gesto aponta para um item específico) e da definição de estratégias de feedback (respostas visuais na tela, sons de confirmação etc.).

No ARtificial, a literatura de interfaces multimodais é utilizada para justificar três decisões de concepção:

1. Uso do Kinect como ponto único de fusão sensorial, já que ele oferece simultaneamente visão de profundidade e captura de voz.
2. Prioridade a comandos curtos, contextuais e combináveis com gestos simples, evitando gramáticas de fala complexas que aumentariam erros de reconhecimento.
3. Representação explícita dos fluxos multimodais nos protótipos Figma, com anotações visuais indicando em que momentos a fala é esperada, quais gestos são compreendidos e como o sistema responde.

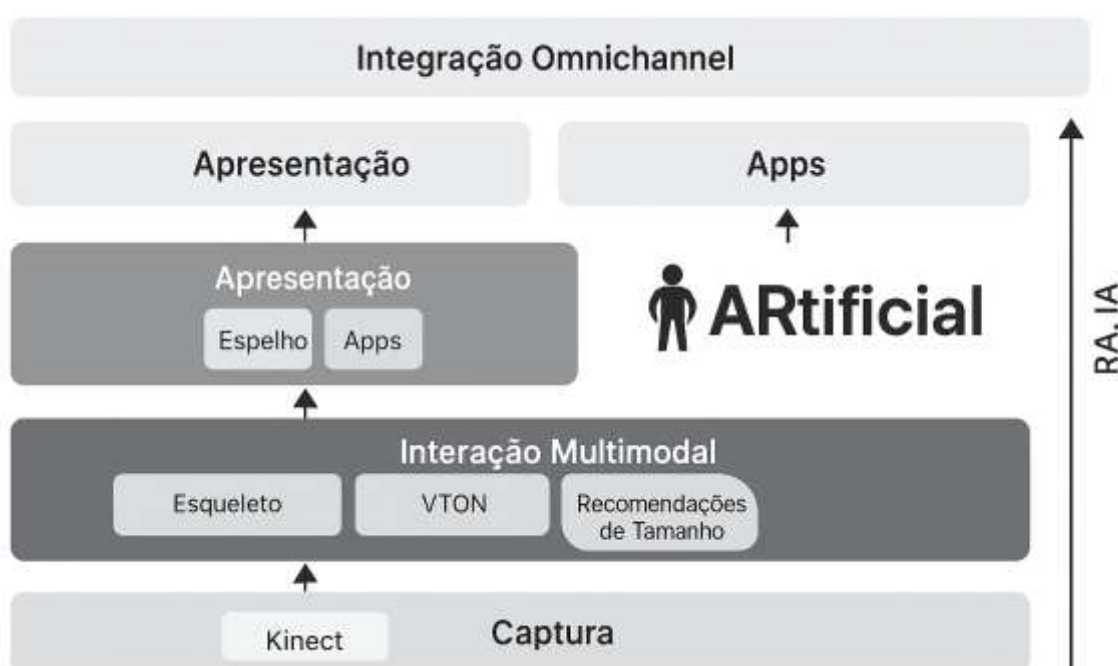
Esses fundamentos tecnológicos — Kinect, redes neurais para try-on e recomendação, simulação de tecidos e interfaces multimodais — formam a base sobre a qual a arquitetura conceitual do ARtificial é construída na seção seguinte.

4 ARQUITETURA CONCEITUAL DO ARTIFICIAL

A arquitetura conceitual do ARTificial organiza, em camadas, como os diferentes componentes físicos e lógicos se articulam para viabilizar um provedor virtual multimodal baseado em Kinect, Realidade Aumentada e Inteligência Artificial. Nesta etapa, trata-se de um desenho de referência: um modelo que descreve como o sistema deveria funcionar em termos de fluxos de dados, responsabilidades de cada módulo e pontos de integração com as interfaces projetadas em Figma, sem desenvolvimento de código ou implementação efetiva dos serviços.

Partindo do Kinect como hub de captura de corpo e voz, a arquitetura é estruturada em camadas que incluem: (i) os componentes físicos do sistema (sensor, espelho/tela, dispositivos de apoio); (ii) a camada de captura e interpretação de dados (profundidade, esqueleto, gestos, comandos de voz); (iii) a camada de interação multimodal, responsável por combinar voz, gesto e contexto visual; e (iv) a camada de apresentação e integração *omnichannel*, onde se situam a interface do espelho, os aplicativos móvel/lojista e os pontos de troca de informação com o varejo.

Figura 5 – Visão geral da arquitetura conceitual do ARTificial



Fonte: Elaborado pelos autores.

4.1 Componentes físicos do sistema

No nível físico, o ARTificial é concebido como um ecossistema enxuto, ancorado em hardware amplamente disponível e adequado a ambientes de varejo. A Figura 6 apresenta a disposição frontal dos componentes físicos do sistema ARTificial, enquanto a Figura 7 ilustra a vista posterior e as principais conexões entre eles.

Figura 6 - Visão Frontal Componentes

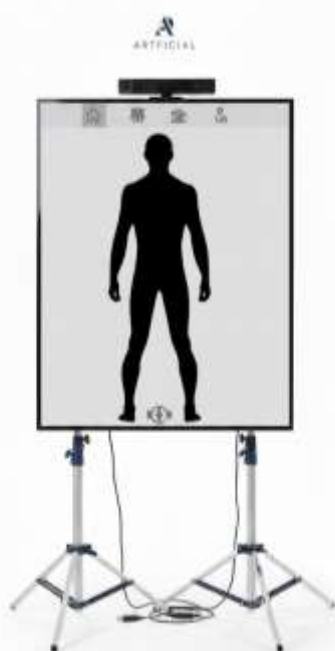


Figura 7 - Visão Posterior Componentes



Fonte: Elaborado pelos autores.

- TV/monitor Full HD (1920×1080): responsável por exibir a interface do espelho virtual em loja, permitindo que o usuário veja, em tempo real, a sobreposição de peças de roupa virtuais ao seu corpo, bem como menus, instruções e feedbacks de sistema.
- Sensor Microsoft Kinect para Xbox One (Kinect v2): atua como núcleo de captura, combinando câmera RGB, sensor de profundidade de tempo de voo e *array* de microfones integrado. É o Kinect que fornece a nuvem de pontos, o esqueleto tridimensional do usuário e o sinal de áudio necessário para comandos de voz.
- Unidade de processamento (PC ou mini-PC dedicado): equipada com GPU e capacidade de processamento suficiente para receber os fluxos do Kinect, executar os módulos de rastreamento, interpretação multimodal e renderização

gráfica, além de se comunicar com serviços em nuvem (por exemplo, APIs de recomendação ou banco de dados de produtos).

- **Conectividade em rede:** conexão à internet ou à rede interna da loja para sincronização de catálogos, atualização de looks, armazenamento de sessões e integração com aplicações móveis e sistemas do lojista.

Nessa configuração, o Kinect é explicitamente tratado como “hub sensorial” do sistema físico: ele concentra a captura de corpo e voz, reduz a necessidade de múltiplos periféricos e alinha a proposta do ARtificial à literatura que explora o Kinect como interface generalizada para interação natural em aplicações de varejo e entretenimento.

4.2 Camada de captura e interpretação

Acima da camada física, a camada de captura e interpretação é responsável por transformar sinais brutos (imagens, profundidade, áudio) em informações semânticas que podem ser usadas como comandos de interação. Conceitualmente, essa camada é composta por quatro módulos principais:

1. **Módulo de rastreamento corporal (*Skeleton Tracking*)**
 - Recebe os dados de profundidade e RGB do Kinect.
 - Estima o esqueleto do usuário (juntas, ângulos, posição no espaço) e identifica a região de interesse (tronco, braços, pernas) que servirá de base para a aplicação de roupas virtuais e para a detecção de gestos.
 - Fornece, em tempo real, parâmetros como altura estimada, distância da câmera e orientação corporal, que podem ser utilizados tanto para ajuste de escala de peças quanto para lógica de interface (por exemplo, exibir mensagens quando o usuário está fora da zona ideal de captura).
2. **Módulo de detecção e classificação de gestos**
 - Utiliza o esqueleto fornecido pelo módulo anterior para reconhecer gestos definidos no vocabulário de interação do ARtificial (por exemplo: levantar a mão para “iniciar”, virar o corpo para “ver costas”, mover o braço para “próximo look”).
 - Cada gesto reconhecido é traduzido em um evento simbólico (ex.: GESTO_INICIAR, GESTO_CONFIRMAR, GESTO_PROXIMO) encaminhado à camada de interação.

3. Módulo de captação e reconhecimento de Voz

- Usa o *array* de microfones do Kinect para capturar o áudio do usuário, com foco na área em frente ao espelho.
- Envia o fluxo de áudio para um serviço de reconhecimento de fala (local ou em nuvem), que converte a fala contínua em texto e, em seguida, em comandos de alto nível (por exemplo: “quero ver calças masculinas”, “mostrar promoções”, “salvar esse look”).
- O resultado é representado por eventos como `COMANDO_VOZ_FILTRO_CALCAS_MASCULINAS` ou `COMANDO_VOZ_SALVAR_LOOK`.

4. Módulo de pré-processamento e normalização de eventos

- Centraliza a chegada de eventos de voz e gesto, aplica regras de limpeza e descarte (por exemplo, eliminação de ruído ou gestos incompletos) e associa metadados como *timestamp*, confiança e contexto de sessão.
- O objetivo é entregar à camada seguinte um fluxo organizado de “intenções de usuário”, sem expô-la à complexidade de sinais de baixo nível.

Nessa camada, a preocupação principal não é implementar algoritmos específicos, mas explicitar quais dados serão capturados, como serão tratados e em que formato chegarão à camada de interação, preparando o terreno para uma futura implementação técnica em estudo subsequente.

4.3 Camada de interação e fluxo multimodal

A camada de interação e fluxo multimodal é responsável por orquestrar, em tempo real, os eventos provenientes das diferentes modalidades (voz e gesto) e traduzi-los em ações na interface e no motor de provador virtual. Ela funciona, conceitualmente, como um “cérebro de diálogo e navegação” do ARTificial.

Essa camada se organiza em três blocos:

1. Gerenciador de Contexto e Estados de Sessão

- Mantém o estado atual da interação (ex.: “tela inicial”, “modo provador”, “comparação de looks”) e o histórico recente de comandos.
- Garante que o mesmo comando possa ter significados diferentes dependendo do contexto. Por exemplo, um gesto de levantar a mão

pode iniciar a experiência na tela inicial, mas ser interpretado como “confirmar seleção” quando o usuário estiver na etapa de finalização do look.

2. Mecanismo de Fusão Multimodal

- Combina eventos de voz e gesto considerando janelas de tempo e regras de prioridade.
- Exemplos:
 - Se o usuário disser “quero ver calças masculinas” enquanto aponta com o braço para a área de “masculino” no menu, o sistema reforça a confiança no comando e aplica o filtro de categoria correspondente.
 - Se houver conflito (por exemplo, voz pede “calçados” e gesto indica “camisetas”), o sistema pode adotar estratégias de desambiguação, como priorizar a modalidade com maior confiança ou solicitar confirmação visual (“Você quis ver calças ou camisetas?”).
- Esse mecanismo se inspira nos modelos de fusão descritos na literatura de interfaces multimodais, mas é descrito aqui em nível de regra e fluxo, preservando o caráter conceitual do estudo.

3. Motor de Navegação e Ações de Sistema

- Recebe comandos multimodais já interpretados (por exemplo: “entrar no provador”, “trocar peça superior”, “salvar look e gerar QR code”) e aciona as transições de tela e consultas necessárias (catálogo de produtos, combinações de look, perfil do usuário).
- Opera como uma *state machine* que:
 - atualiza o layout do espelho;
 - registra eventos relevantes para métricas futuras (tempo de uso, categoria mais acessada, comandos de voz mais frequentes);
 - dispara interações com camadas de RA e IA responsáveis por selecionar peças e simular a aplicação de tecidos sobre o corpo.

A partir desse desenho, a interação no ARtificial é concebida como um fluxo contínuo em que o usuário pode, por exemplo, ativar o espelho com um gesto, dizer “quero ver calças masculinas”, navegar entre opções com movimentos do braço e, por

fim, salvar o look com um comando de voz, sem necessidade de toque em telas físicas.

4.4 Camada de apresentação e protótipos em Figma

A camada de apresentação corresponde às interfaces visuais do ecossistema ARtificial, que, neste estudo, são materializadas como protótipos de alta fidelidade em Figma. Essa camada foi derivada de forma sistemática da etapa de Engenharia de Requisitos, na qual foram elicitados e organizados requisitos funcionais, regras de negócio e casos de uso (Apêndice D), além dos diagramas estruturais de apoio, como o diagrama de classes e o diagrama entidade–relacionamento (Apêndice E).

É nessa camada que a arquitetura conceitual “ganha forma” para fins de comunicação, avaliação visual e discussão com *stakeholders*, permanecendo, entretanto, independente de qualquer implementação tecnológica específica. Ela se desdobra em três frentes principais:

1. Interface do espelho interativo em loja, projetada na resolução 1920×1080, apresenta:
 - área central ocupada pela imagem do usuário (feed do Kinect), sobre a qual serão aplicadas as peças virtuais;
 - barra superior com logotipo, estado de sessão e indicadores de modo ativo (por exemplo, o ícone de microfone animado quando o sistema está “ouvindo”);
 - menus laterais ou inferiores para categorias de produtos, ações rápidas e instruções de voz/gesto.

Nos protótipos, anotações visuais indicam em quais momentos a interação deve se dar por voz, por gesto ou pela combinação de ambas, antecipando a lógica multimodal que será implementada em fase posterior.

2. Aplicativo do usuário (app cliente): as telas conceituais do aplicativo móvel mostram como o cliente pode:
 - sincronizar-se com o espelho por meio de QR code;
 - visualizar e salvar os looks montados em loja;
 - consultar tamanhos, cores e disponibilidade das peças;
 - receber recomendações personalizadas com base em histórico e preferências.

Essa interface complementa a experiência física, reforçando a visão *omnichannel* e permitindo que a sessão iniciada no provador virtual continue no dispositivo móvel.

3. Portal / aplicativo do lojista: os protótipos destinados ao lojista contemplam painéis para:
 - gerenciamento de catálogo (peças, tamanhos, imagens e metadados necessários à simulação);
 - acompanhamento de uso do espelho (tempo médio de sessão, categorias mais exploradas, looks mais salvos);
 - visualização de reservas ou intenções de compra quando o usuário envia looks para o aplicativo.

Essas interfaces traduzem, em termos visuais, os requisitos de negócio e de operação identificados na etapa de elicitación e documentados nos artefatos de requisitos e nos diagramas estruturais.

Em todos os casos, os protótipos em Figma não são apenas “*mockups* estéticos”, mas representações visuais do contrato arquitetural definido nas camadas anteriores: cada botão, mensagem e ícone está associado a um comando, evento ou fluxo mapeado nas camadas de captura, interpretação e interação multimodal. A descrição detalhada de cada tela, incluindo funções, navegação e comandos de voz e gestos associados, encontra-se no Apêndice E.

Dessa forma, a arquitetura conceitual do ARTificial cumpre duas funções centrais neste estudo: (i) organizar, de maneira sistemática, como Kinect, RA, IA e interface multimodal se articulam em um provador virtual de moda; e (ii) servir como referência estruturante para a fase subsequente de desenvolvimento técnico e avaliação empírica, a ser explorada em um segundo artigo.

5 METODOLOGIA

5.1 Tipo de pesquisa e abordagem

Este estudo caracteriza-se como uma pesquisa aplicada, de natureza qualitativa e abordagem exploratória–descritiva, centrada na concepção de uma solução e na prototipação de alta fidelidade para um provador virtual multimodal de moda com Kinect.

Em termos de lógica de desenvolvimento, aproxima-se de uma abordagem de design science: o trabalho (i) estrutura o problema a partir de dados de mercado e literatura; (ii) revisa o estado da arte sobre RA, sensores de profundidade, IA e interfaces multimodais; (iii) deriva requisitos conceituais para a solução; e (iv) materializa tais requisitos em artefatos projetuais (arquitetura conceitual, fluxos, protótipos em Figma), sem ainda avançar para a implementação técnica. A fase de desenvolvimento do software e avaliação experimental com usuários é explicitamente planejada como Etapa 2 do projeto, a ser tratada em artigo subsequente.

5.2 Procedimentos de revisão e seleção bibliográfica

A revisão de literatura foi conduzida com foco em cinco eixos temáticos:

1. Realidade Aumentada (RA) aplicada a provedores virtuais de moda;
2. Sensores de profundidade e Microsoft Kinect (captura de movimento, reconstrução 3D, espelhos virtuais);
3. Interfaces multimodais (combinação de voz e gestos em sistemas interativos);
4. Inteligência Artificial aplicada à moda (virtual try-on, GANs, recomendação de tamanho);
5. Soluções brasileiras e estado da prática em varejo de moda (projetos acadêmicos e iniciativas comerciais).

Foram consultadas bases nacionais e internacionais, incluindo periódicos e anais indexados, bem como plataformas como ResearchGate, IEEE Xplore e portais brasileiros de pós-graduação, além de notícias de mercado e matérias especializadas sobre provedores virtuais e experiências omnichannel em redes de varejo.

O processo de seleção dos 29 artigos principais seguiu a lógica de um fluxo PRISMA simplificado, em quatro etapas:

- Identificação: busca inicial a partir de combinações de descritores em português e inglês (por exemplo, “*virtual fitting room*”, “*Kinect dressing room*”,

“*virtual try-on GAN*”, “*multimodal interface*”, “*augmented reality fashion*”, “provador virtual”, “espelho virtual”, “recomendação de tamanho em moda”), resultando em um conjunto amplo de publicações potenciais.

- **Triagem:** exclusão de duplicatas, estudos claramente fora do escopo (por exemplo, RA em contextos industriais sem relação com corpo humano/vestuário) e materiais sem texto completo disponível.
- **Elegibilidade:** leitura dos resumos e, quando necessário, de seções metodológicas para verificar aderência ao recorte do estudo (uso de RA em moda, provadores virtuais, Kinect/sensores de profundidade, VTON, interfaces voz–gesto). Nesta etapa, também foram incluídos trabalhos brasileiros específicos sobre espelhos virtuais e RA em moda, bem como projetos relacionados à captura de corpo com Kinect.
- **Inclusão:** consolidação de um conjunto final de 29 estudos, organizados em matrizes analíticas por eixo temático (RA/Kinect, VTON baseado em imagem, recomendação de tamanho, interfaces multimodais e soluções brasileiras).

O conjunto de artigos e dissertações brasileiras sobre RA, moda ou comportamento — incluindo, por exemplo, o espelho virtual interativo para maquiagem, a dissertação de Elisson sobre provador virtual com Kinect, e outros projetos nacionais recentes — foi analisado separadamente para compor o subtópico de soluções brasileiras na revisão, articulando produção acadêmica local e iniciativas de mercado (como o provador virtual da Renner com mídia OOH).

As sínteses individuais e a categorização dos artigos constam em matriz específica de apoio (não reproduzida integralmente no corpo do texto), utilizada para alimentar tanto o Capítulo 2 quanto a derivação de requisitos na Seção 5.3.

5.3 Planejamento do projeto e artefatos de gestão (TAP e EAP)

Para organizar o desenvolvimento do estudo, foram utilizados artefatos clássicos de gestão de projetos, adaptados ao contexto acadêmico.

- **Termo de Abertura do Projeto (TAP):**

Documento que define o objetivo geral e os objetivos específicos, o escopo desta primeira fase (concepção + prototipação em Figma, sem implementação), os principais stakeholders (orientadora, instituição, público-alvo do varejo de moda) e as entregas previstas (revisão de literatura, diretrizes, arquitetura conceitual, protótipos). O TAP ancora o recorte deste

artigo e explicita que a implementação técnica e a avaliação com usuários pertencem à fase seguinte do projeto.

- Estrutura Analítica do Projeto (EAP):

A EAP decompõe o projeto em duas grandes etapas:

1. Etapa 1 – Estudo conceitual e prototipação (objeto deste artigo);
 2. Etapa 2 – Desenvolvimento técnico e avaliação com usuários (planejada para estudo posterior).
- Dentro da Etapa 1, a EAP organiza o trabalho nas seguintes macrofases:
 3. Planejamento e levantamento de requisitos conceituais (revisão de literatura, análise de mercado, derivação de requisitos);
 4. Modelagem conceitual e arquitetura de solução (diagramas, fluxos, arquitetura conceitual);
 5. Prototipação de interfaces em Figma (espelho, app do usuário, portal do lojista);
 6. Documentação acadêmica (redação do artigo, organização de apêndices e anexos).

A versão resumida da EAP aparece no corpo do texto, enquanto a versão detalhada integra os apêndices de gestão do projeto (junto com a Matriz SWOT e o Plano de Ação 5W2H apresentados no Apêndice C).

5.4 Engenharia de requisitos e modelagem conceitual

A partir da revisão de literatura e da análise de mercado, foram definidos requisitos conceituais para o provador virtual multimodal ARTificial, por meio de um processo de engenharia de requisitos orientado ao domínio de provedores virtuais com Kinect.

Em vez de realizar um levantamento empírico extenso com usuários finais nesta etapa, optou-se por uma elicitação analítica, baseada em três fontes principais:

1. Resultados da revisão sistemática (Seção 2): identificação de funcionalidades recorrentes (*virtual try-on*, captura de movimento, recomendação de tamanho, integração com app), limitações técnicas e oportunidades de diferenciação (uso de voz e gestos em loja física).
2. Mapeamento de soluções existentes (científicas e comerciais): análise comparativa de provedores virtuais Kinect-based, VTON baseados em imagem, recomendadores de tamanho e iniciativas brasileiras de RA em moda.

3. Artefatos anteriores do projeto ARtificial (planos, canvas e materiais de trabalho): consolidados e atualizados para refletir o recorte atual do estudo.

Com base nisso, foram organizados:

- Requisitos funcionais conceituais, especificando o que o sistema deve ser capaz de fazer (por exemplo, rastrear corpo via Kinect, sobrepor roupas em tempo real, permitir navegação por comandos de voz e gestos, integrar-se ao aplicativo do usuário e ao painel do lojista).
- Requisitos não funcionais conceituais, relacionados a qualidades como desempenho (tempo de resposta aceitável para uso em loja), usabilidade, acessibilidade, privacidade de dados (LGPD) e compatibilidade de hardware.
- Regras de negócio conceituais, que expressam restrições e políticas de uso do provedor em contexto de varejo (por exemplo, limites de tempo de sessão, requisitos de consentimento para uso de imagem, controle de catálogo pelo lojista).

Esses requisitos foram descritos textualmente e organizados em quadros e tabelas (requisitos funcionais, requisitos não funcionais e regras de negócio), que se encontram consolidados no Apêndice D. Complementarmente, foram elaborados:

- Diagramas de processos em BPMN, representando a jornada do usuário diante do espelho, desde a aproximação, autenticação/sincronização via QR code, interação por voz/gestos, até o registro de looks favoritos;
- Casos de uso conceituais, descrevendo cenários típicos (ativar espelho virtual, aplicar roupas, salvar look, sincronizar com app, consultar looks salvos).

5.5 Prototipação de interfaces em Figma

A etapa final desta primeira fase metodológica consistiu na prototipação de interfaces do ecossistema ARtificial em Figma, tomando os requisitos conceituais como guia.

Foram criados protótipos de alta fidelidade para três núcleos:

1. Espelho interativo em loja (resolução 1920×1080), exibindo:
 - área central de vídeo (feed do Kinect) com espaço reservado para aplicação de peças virtuais;
 - barra superior com logotipo, estado de sessão e indicador de escuta de voz (ícone de microfone);

- áreas de menu e feedback multimodal (instruções de gestos e comandos de voz).
2. Aplicativo do Usuário (App Cliente), incluindo telas conceituais para:
- leitura de QR code para sincronização com o espelho;
 - visualização e salvamento de looks experimentados em loja;
 - consulta a tamanhos, cores e disponibilidade;
 - continuidade da jornada em ambiente mobile (omnichannel).
3. Portal / Aplicativo do Lojista, com telas para:
- cadastro e manutenção de catálogo (peças, tamanhos, imagens, metadados);
 - acompanhamento de uso do espelho (indicadores de sessão, categorias mais usadas, looks mais salvos);
 - visão de reservas e intenção de compra derivadas das sessões no espelho.

Em todos os casos, as decisões de layout e fluxo foram diretamente derivadas da arquitetura conceitual apresentada no Capítulo 4, e das diretrizes de interação multimodal descritas na Seção 3.4. Os protótipos incluem anotações visuais sobre:

- em quais momentos a interação deve ocorrer por voz, gesto ou combinação de ambas;
- quais feedbacks visuais são exibidos para cada comando;
- como o espelho se integra ao app do usuário e ao painel do lojista.

A descrição detalhada de cada tela e fluxo, bem como capturas e links de navegação dos protótipos, encontra-se organizada no Apêndice E.

6 RESULTADOS ESPERADOS E DIRETRIZES PROPOSTAS

Esta pesquisa, por ter caráter conceitual e de prototipação em alta fidelidade, não resulta em um sistema funcional, mas na sistematização de diretrizes de concepção, arquitetura e experiência do usuário para provedores virtuais multimodais com sensor Kinect. Esses resultados funcionam como um “mapa de projeto” para a fase de implementação técnica e para o segundo artigo derivado deste estudo.

6.1 Síntese conceitual esperada

Como primeiro resultado, espera-se a consolidação de uma arquitetura conceitual de referência para provedores virtuais de moda baseados em:

- captura de movimento e pose corporal via sensor de profundidade (Kinect);
- aplicação virtual de peças de vestuário sobre o corpo do usuário, com suporte de RA e, em etapa futura, modelos de virtual *try-on* (VTON) baseados em IA;
- interação multimodal com comandos de voz e gestos, reduzindo a dependência de toque físico em telas;
- integração com aplicações móveis e sistemas do lojista, compondo uma jornada *omnichannel*.

Essa arquitetura é representada tanto em diagramas de processos e artefatos de requisitos (BPMN, requisitos, regras de negócio e casos de uso, reunidos no Apêndice D – Artefatos de Engenharia de Requisitos) quanto nas telas prototipadas em Figma (Apêndice E). Dessa forma, engenharia, UX e negócio conseguem dialogar sobre a mesma visão sistêmica antes de qualquer linha de código.

6.2 Diretrizes para uso do Kinect como hub de captura corpo + voz

Com base na literatura sobre sensores de profundidade e nos estudos que empregam o Kinect em aplicações de rastreamento corporal, reconhecimento de gestos e interfaces naturais, o trabalho propõe um conjunto de diretrizes para o uso do dispositivo como hub principal de captura no contexto de provedores virtuais de moda:

- Centralidade do Kinect na arquitetura física: o sensor deve ser posicionado de forma fixa, em altura e distância adequadas, de modo a capturar todo o corpo do usuário em postura ereta, minimizando oclusões e garantindo estabilidade do esqueleto estimado.

- Aproveitamento combinado de RGB + profundidade: a concepção do sistema deve considerar, desde o início, que o Kinect fornece simultaneamente imagem de cor, mapa de profundidade e dados de skeleton tracking, permitindo alinhar detecção de corpo, ajuste de vestuário e avaliação de caimento virtual em um único fluxo de dados.
- Uso do array de microfones para reconhecimento de voz: o próprio Kinect deve ser priorizado como fonte de áudio para comandos de voz, explorando seu array de microfones com cancelamento de ruído e orientação para captação frontal, reduzindo a necessidade de microfones externos na fase inicial do projeto.
- Delimitação da “zona de interação”: recomenda-se definir, no software e na comunicação visual do espelho, uma área de uso ideal (distância mínima e máxima do sensor), a fim de garantir qualidade na captura de corpo e voz e evitar experiências frustrantes em ambientes de loja.
- Sensibilidade a condições de iluminação e ruído: a arquitetura conceitual deve prever a necessidade de calibração e tratamento de ruído (visual e sonoro), sobretudo em shoppings e lojas movimentadas, influenciando desde o layout físico até regras de fallback da interface (por exemplo, voltar a menus mais simples em caso de perda de rastreamento).

Essas diretrizes ancoram o Kinect não apenas como um acessório de captura, mas como núcleo técnico da experiência, ao redor do qual a interface e os fluxos de interação são desenhados.

6.3 Diretrizes para o design de comandos de voz em provadores virtuais

Outra frente central de resultados esperados diz respeito ao desenho de comandos de voz adequados ao contexto de moda, à dinâmica de loja física e às limitações de reconhecimento automático de fala. O estudo propõe, entre outras, as seguintes orientações:

- Vocabulário restrito e orientado a tarefas: privilegiar comandos simples e diretamente associados a ações do sistema, como “quero ver calças masculinas”, “mostrar tamanhos maiores”, “trocar a cor da blusa”, “salvar este look”, evitando frases longas e ambiguidades.
- Estruturação por intenções principais: agrupar comandos em categorias claras – navegação (entrar/sair de modos), filtragem (categoria, gênero, tamanho,

cor), interação com looks (salvar, comparar, excluir) e integração omnichannel (gerar QR Code, enviar para app) – facilitando tanto o desenho da interface quanto o desenvolvimento futuro do módulo de compreensão de linguagem.

- Feedback imediato e multimodal: cada comando de voz deve gerar uma resposta clara na interface (animação, highlight de itens filtrados, mensagem textual curta) e, quando possível, um retorno sonoro, reforçando para o usuário que o sistema “entendeu” a intenção, mesmo em cenários com pequena latência.
- Frases de ativação e cancelamento: recomenda-se definir expressões consistentes para ativar e desativar o modo de escuta (“espelho, ouvir comandos”, “parar comandos de voz”), reduzindo a chance de acionamentos acidentais em ambientes barulhentos.
- Apoio visual na interface: o microfone animado, já presente no conceito ARTificial, deve comunicar estados como “escutando”, “processando” e “não entendi”, permitindo ao usuário ajustar o comando sem precisar tocar na tela.

Mesmo sem implementar o módulo de reconhecimento de fala nesta fase, o protótipo em Figma já antecipa esses comportamentos, o que é essencial para que a etapa de desenvolvimento não trate a voz como um “adendo” tardio, mas como parte estrutural do design de interação.

6.4 Diretrizes para combinação de voz e gestos em uma interface multimodal

A literatura de interfaces multimodais indica que a combinação de modalidades – voz, gesto, visão – gera interações mais naturais, desde que a fusão seja tratada de forma coerente. Tomando essa base teórica, o estudo consolida diretrizes específicas para o contexto de provedores virtuais:

- Complementaridade de modalidades: comandos de voz devem ser usados para expressar a intenção semântica (“quero ver vestidos longos”), enquanto gestos assumem papel de manipulação espacial (girar o corpo virtual, avançar ou retroceder no carrossel de looks, confirmar seleção levantando uma mão).
- Gestos canônicos e facilmente aprendidos: o conjunto de gestos deve ser reduzido, intuitivo e inspirado em ações cotidianas (levantar a mão para “chamar” o sistema, acenar para trocar de peça, aproximar as mãos para zoom), evitando coreografias complexas que exigiriam treinamento intensivo.

- Fusão temporal voz–gesto: a arquitetura conceitual deve considerar que certos comandos podem combinar fala e gesto quase simultaneamente, por exemplo, o usuário dizer “trocar a cor desta camisa” enquanto aponta para o tronco. Isso exige que o sistema, na fase futura, trate voz e skeleton tracking em janelas de tempo sincronizadas.
- Redução de toque físico em superfícies: a combinação de voz e gestos é desenhada, desde o protótipo, com o objetivo explícito de minimizar o contato com telas em ambiente de loja, o que tem implicações de usabilidade (fluidez) e de higiene.
- Estratégias de *fallback*: deve existir sempre um caminho alternativo – por exemplo, menus navegáveis por gesto ou botões virtuais – para situações em que a voz não é compreendida ou o gesto não é detectado, preservando a continuidade da experiência.

Essas diretrizes são materializadas nas telas do espelho ARtificial por meio de ícones, microinterações e anotações de fluxo, ilustrando concretamente como o usuário combinaria fala e movimento para explorar o catálogo e experimentar roupas.

6.5 Implicações esperadas para UX, varejo e pesquisa futura

Por fim, espera-se que o conjunto de diretrizes propostas gere impactos em três frentes:

- Na experiência do usuário (UX): ao oferecer um modelo de interação mais próximo da prova física de roupas, reduzindo a incerteza de tamanho e caimento e tornando a navegação por categorias e looks mais fluida e engajante.
- No varejo de moda: ao indicar caminhos concretos para diminuir taxas de devolução, enriquecer a jornada phygital e criar novos pontos de contato entre loja física, provador virtual e aplicativos móveis, com potencial de geração de dados mais ricos sobre comportamento de experimentação.
- Na pesquisa acadêmica: ao preencher uma lacuna nacional na documentação de arquiteturas multimodais baseadas em Kinect para o varejo de moda, oferecendo um modelo replicável para estudos posteriores que possam implementar e avaliar, empiricamente, os efeitos dessas diretrizes sobre usabilidade, aceitação do usuário e indicadores de negócio.

Dessa forma, os resultados esperados e as diretrizes aqui sistematizadas funcionam como ponte entre a literatura técnica e a prática de design de sistemas, preparando o terreno para que o segundo estudo – focado na implementação tecnológica do ARtificial – possa partir de uma base conceitual sólida, alinhada tanto às demandas do mercado quanto às discussões contemporâneas em Interação Humano–Computador e interfaces multimodais.

7 DISCUSSÃO

A discussão articula o modelo proposto com o estado da prática em provedores virtuais e com as limitações inerentes à abordagem multimodal baseada em Kinect. O foco é posicionar o ARTificial não apenas como um protótipo de interface, mas como uma resposta estruturada a lacunas identificadas na literatura e no mercado brasileiro.

7.1. ARTificial frente a soluções baseadas em foto e aplicativos móveis

Grande parte das soluções de provedor virtual atualmente em uso comercial no Brasil trabalha a partir de fotos estáticas ou avatares parametrizados, acessados via aplicativo ou navegador. Essas abordagens oferecem vantagens importantes: baixo custo de hardware, escalabilidade e integração direta ao e-commerce. No entanto, elas tendem a operar em:

- cenários não imersivos (tela pequena do celular, ausência de corpo em escala real);
- interação predominantemente **tátil**, baseada em toques e deslizes;
- ausência de captura de movimento em tempo real.

O ARTificial se diferencia justamente por propor um provedor em escala humana, projetado para loja física, com captura de corpo em tempo real e interação multimodal. Isso não substitui as soluções baseadas em foto, mas as complementa: enquanto o app cobre a jornada remota, o provedor com Kinect foca na experiência phygital em ponto de venda, reforçando a convergência entre canais.

7.2. ARTificial frente a sistemas de recomendação de tamanho

Ferramentas como as utilizadas pela Sizebay e por outras retail techs focam na recomendação de tamanho com base em tabelas, medidas do corpo e estatísticas de devolução. Elas atacam diretamente o problema da numeração, mas não necessariamente oferecem uma visualização rica de caimento ou uma interação imersiva.

O modelo proposto neste estudo incorpora a lógica de recomendação como um módulo de IA dentro de uma arquitetura mais ampla: o provedor não só indica o tamanho ideal, como também apresenta o look com aplicação virtual de tecidos e peças ao corpo, permitindo que o usuário perceba proporção, comprimento e combinação entre itens.

Na discussão, essa diferença é relevante:

- sistemas puramente algorítmicos → atacam o “número certo”;
- ARTificial → combina “número” com experiência visual e interativa, o que pode gerar confiança adicional e valor de marca.

7.3. ARTificial frente a experimentos de RA sem multimodalidade

A literatura internacional apresenta vários estudos de provadores em RA que utilizam Kinect ou câmeras de profundidade para simular roupas em tempo real, mas em muitos casos a interação é limitada a gestos ou menus visuais. Voz e multimodalidade aparecem de forma pontual, geralmente em outros domínios (navegação indoor, robótica, interfaces para pessoas com deficiência visual).

O ARTificial avança ao trazer para o contexto do varejo de moda uma arquitetura explicitamente multimodal, na qual voz e gesto são concebidos desde o início como partes da interface, e não como acréscimos tardios. Essa integração permite explorar cenários de uso mais ricos, como:

- o cliente verbalizar categorias e filtros (“mostrar looks sociais”, “trocar para jeans escuro”);
- o uso de gestos como confirmação rápida, navegação entre looks e seleção de peças;
- a possibilidade futura de adaptar a experiência a pessoas com restrição de mobilidade ou de visão, por meio de feedbacks sonoros e comandos por voz mais extensos.

7.4. Limitações da proposta

Apesar do potencial identificado, a concepção do ARTificial traz limitações que precisam ser reconhecidas:

- Dependência de hardware específico

A arquitetura parte do Kinect como hardware central. Embora seja um sensor relativamente acessível, sua descontinuação comercial e a necessidade de adaptadores em ambientes atuais podem dificultar a adoção em larga escala. A proposta, portanto, deve ser vista como referência conceitual que pode futuramente ser reimplementada com sensores equivalentes.

- Sensibilidade ao ambiente físico

O desempenho de rastreamento e reconhecimento de voz depende de condições de iluminação, espaço livre em frente ao espelho e nível de ruído da loja. Ambientes muito apertados ou extremamente ruidosos podem prejudicar a experiência. Isso sugere que, em um segundo estudo, testes de campo em diferentes tipos de ponto de venda serão essenciais.

- Ausência de avaliação empírica nesta fase

Este primeiro artigo concentra-se na concepção e prototipação em Figma, sem implementação nem testes com usuários. As diretrizes propostas são fundamentadas na literatura e em boas práticas de UX, mas ainda não foram validadas empiricamente no contexto brasileiro. Essa limitação é assumida explicitamente e transformada em agenda para o estudo subsequente.

- Complexidade de integração com ecossistemas legados

A integração do provador com ERPs, CRMs, sistemas de estoque e plataformas de e-commerce de cada varejista configura um desafio técnico e organizacional relevante, em razão da heterogeneidade dos sistemas legados, da ausência de padrões de interoperabilidade e das restrições de segurança e governança de dados. A arquitetura conceitual prevê APIs e sincronização em nuvem, mas a viabilidade concreta dependerá de esforços técnicos e organizacionais que extrapolam o escopo desta pesquisa.

7.5. Implicações para pesquisas futuras

As discussões apresentadas reforçam que o ARTificial deve ser entendido como um modelo de referência, e não como uma solução fechada. Entre as principais implicações para trabalhos futuros, destacam-se:

- necessidade de desenvolver um protótipo funcional com Kinect ou sensores equivalentes, para testar na prática as diretrizes de multimodalidade;
- realização de estudos de usabilidade e experiência do usuário, comparando a interação multimodal com interfaces tradicionais (toque/aplicativo);
- investigação de impactos em indicadores operacionais, como taxa de devolução, ticket médio e tempo de atendimento em loja;
- análise de aceitação tecnológica por parte de consumidores e vendedores, considerando aspectos culturais, de privacidade e de exposição corporal em provadores digitais.

Com isso, a discussão posiciona o ARtificial como um passo intermediário: um elo entre o estado fragmentado da literatura (gestos, RA, recomendação) e a necessidade de arquiteturas completas e replicáveis para o varejo de moda brasileiro.

CONSIDERAÇÕES FINAIS

Este estudo partiu de um problema concreto do varejo de moda brasileiro – a elevada taxa de devoluções no e-commerce e a dificuldade de experimentar produtos à distância – para propor uma resposta conceitual ancorada em tecnologias emergentes. A partir de um mapeamento inicial de aproximadamente 190 registros bibliográficos e da análise em profundidade de 29 estudos centrais sobre provedores virtuais, Realidade Aumentada, Inteligência Artificial, captura de movimento e interfaces multimodais, foi possível estruturar o ARtificial como um provedor virtual de moda que toma o Kinect como hardware central de captura corpo+voz e organiza a experiência do usuário em torno de uma interface sem toque, baseada em gestos e comandos de voz.

Do ponto de vista teórico, o trabalho contribui ao sistematizar, em um único modelo, elementos que na literatura aparecem frequentemente de forma fragmentada: o uso de sensores de profundidade como o Kinect para rastreamento corporal em tempo real; a aplicação de roupas virtuais e simulação de tecidos sobre o corpo do usuário; a personalização de recomendações por IA; e a fusão de modalidades (voz, gesto, visão) em uma arquitetura multimodal voltada especificamente ao contexto do varejo de moda. Ao propor diretrizes de concepção e uma arquitetura conceitual para esse ecossistema, o estudo dialoga com a literatura de Interação Humano-Computador, RA aplicada ao varejo e interfaces multimodais, preenchendo uma lacuna especialmente sensível no contexto brasileiro, em que predominam soluções comerciais pouco documentadas em termos de desenho de interação e decisões tecnológicas.

No plano prático, a pesquisa organiza esse conjunto de decisões em artefatos que podem ser diretamente apropriados por equipes de produto e UX: (i) uma arquitetura conceitual que explicita camadas de captura, interpretação e apresentação; (ii) um conjunto de requisitos funcionais e não funcionais para provedores virtuais multimodais com Kinect; e (iii) protótipos de alta fidelidade em Figma que materializam fluxos-chave do ARtificial, incluindo navegação por voz (“quero ver calças masculinas”), seleção de peças por gestos e integração omnichannel via QR code e aplicativos móveis. Ainda que se trate de uma fase estritamente conceitual, esses artefatos funcionam como um “guia de projeto” para iniciativas futuras que desejem avançar da ideia à implementação.

Como todo estudo de primeira fase, entretanto, este trabalho apresenta limitações que precisam ser explicitadas. Em primeiro lugar, a pesquisa não envolve implementação de software nem avaliação empírica com usuários, o que impede, neste momento, medir o impacto efetivo do provador virtual em métricas como redução de devoluções, tempo de atendimento em loja ou aumento de conversão. Em segundo lugar, as diretrizes propostas assumem um ambiente de loja relativamente controlado (espaço físico, iluminação, distância do usuário ao sensor), o que pode exigir adaptações em contextos mais complexos ou heterogêneos. Por fim, embora a literatura sobre VTON, redes neurais e simulação de tecidos tenha sido incorporada como base conceitual, o estudo não implementa nem compara modelos específicos, o que também se configura como uma limitação do escopo.

Essas limitações, por outro lado, abrem uma agenda clara de trabalhos futuros, que se organiza em três frentes principais. A primeira diz respeito à implementação técnica do ARTificial: desenvolvimento do *backend* (por exemplo, em Django/Python), integração com o Kinect para captura de esqueleto e voz, uso de bibliotecas de RA e motores de simulação de tecido, e conexão com aplicativos móveis de cliente e lojista. A segunda frente envolve a avaliação empírica com usuários em ambiente controlado de loja, explorando métodos de teste de usabilidade, análise de jornada e experimentos quase-experimentais que relacionem o uso do provador virtual a indicadores de negócio (devoluções, tíquete médio, conversão). A terceira frente, mais prospectiva, aponta para a comparação entre diferentes arranjos tecnológicos – por exemplo, variações de modelos de rede neural para try-on e recomendação de tamanho, uso de motores de reconhecimento de fala distintos, ou mesmo comparação entre Kinect e outros sensores de profundidade – de modo a refinar as diretrizes propostas e avançar em direção a um framework de referência para provedores virtuais multimodais no varejo de moda.

Em síntese, este trabalho não pretende oferecer uma solução fechada, mas sim um marco conceitual para orientar a próxima etapa da pesquisa: sair do protótipo navegável em Figma para um protótipo funcional capaz de ser testado, medido e iterado. A expectativa é que, ao articular RA, IA, sensores de profundidade e interfaces multimodais em uma arquitetura coerente, o ARTificial contribua para qualificar o debate acadêmico sobre provedores virtuais e, ao mesmo tempo, sirva como referência prática para projetos que buscam reinventar a experiência de compra de moda em um cenário cada vez mais phygital.

REFERÊNCIAS

ABCOMM. Associação Brasileira de Comércio Eletrônico. Relatório ABComm News – Números do E-commerce Brasileiro 2022. São Paulo, 2023.

ABIT; IEMI. Associação Brasileira da Indústria Têxtil e de Confecção; Instituto de Estudos e Marketing Industrial. Perfil do Setor Têxtil e de Confecção no Brasil 2023. São Paulo, 2023.

ABNT. ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 6023: Informação e documentação – Referências – Elaboração. Rio de Janeiro, 2018. 2. versão corrigida, publicada em 24 set. 2020.

ADIKARI, Sasadara B. et al. Applicability of a single depth sensor in real-time 3D clothes simulation: augmented reality virtual dressing room using Kinect sensor. *Inventi Impact: Human-Computer Interaction*, [s.l.], v. 2020, n. 1, p. 1-8, 2020.

ALMEIDA, Juliana. Comportamento do consumidor e devoluções no e-commerce de moda. São Paulo: [s.n.], 2018.

ALZAMZAMI, Ohoud et al. Intelligent gloves: an IT intervention for deaf-mute people. *Journal of Intelligent Systems*, v. 32, n. 1, 2023.

AUDACES. Audaces adquire a Sizebay, startup de provedores virtuais para e-commerce de moda. *ND Mais*, Florianópolis, jul. 2024. Disponível em: <https://ndmais.com.br>. Acesso em: 15 nov. 2025.

BLUMER. Penalty – provedor virtual interativo. *Blumer – Cases*, [s.l.], 2023. Disponível em: <https://blumer.com.br>. Acesso em: 15 nov. 2025.

CAMPOS, Vicente Falconi. *TQC – Controle da Qualidade Total (no estilo japonês)*. 8. ed. Nova Lima: Falconi Editora, 2004.

CLIENTESA. Provedor virtual com IA analisa formato do corpo e recomenda tamanho ideal. *ClienteSA*, São Paulo, jun. 2024. Disponível em: <https://www.clientesa.com.br>. Acesso em: 15 nov. 2025.

DUMAS, Bruno; LALANNE, Denis; OVIATT, Sharon. Multimodal interfaces: a survey of principles, models and frameworks. In: LALANNE, Denis; NIELSEN, Jakob (org.). *Human Machine Interaction*. Berlin: Springer, 2009. p. 3-26. (Lecture Notes in Computer Science, v. 5440).

DUMAS, Bruno; OVIATT, Sharon. Multimodal interfaces: principles and models. In: *Human Machine Interaction*. Berlin: Springer, 2010.

ELETROMÍDIA; RENNER. Renner e Eletromidia lançam provedores virtuais com IA generativa em shoppings. *Meio & Mensagem*, São Paulo, ago. 2025. Disponível em: <https://www.meioemensagem.com.br>. Acesso em: 15 nov. 2025.

ELISON, R. D. et al. Virtual dressing using augmented reality. ITM Web of Conferences, v. 36, art. 01003, 2021.

ERRA, Ugo; SCANNIELLO, Giuseppe; COLONNESE, Valerio. Exploring the effectiveness of an augmented reality dressing room for online shopping. Multimedia Tools and Applications, v. 77, p. 25077–25107, 2018.

HILSMANN, Anna; EISERT, Peter. Tracking and retexturing cloth for real-time virtual try-on. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.I.]: IEEE, 2013.

ISIKDOGAN, Fatih; KARA, L. Burak. Real-time 3D garment simulation and fitting using Kinect. In: Proceedings of the ASME 2012 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference. [S.I.]: ASME, 2012.

JOGLEKAR, Prajakta; GOHOKAR, Vinaya. Virtual cloth try-on using augmented reality – marker based approach. Applied Sciences, v. 8, n. 9, 2018.

KURNIAWATI, Arik; KUSUMANINGSIH, Ari; ALIFFIO, Yanuar. Clothing size recommender on real-time fitting simulation using skeleton tracking and rigging. Jurnal Teknologi dan Sistem Komputer, v. 8, n. 2, p. 127-132, 2020.

LARMAN, Craig. Utilizando UML e padrões: uma introdução à análise e ao projeto orientados a objetos e ao desenvolvimento iterativo. 3. ed. Porto Alegre: Bookman, 2004.

LEE, Hanna; XU, Yingjiao. Augmented reality in fashion retail: a review of consumer adoption and usage. International Journal of Fashion Design, Technology and Education, v. 13, n. 1, p. 1-9, 2019.

MCLUHAN, Marshall. Os meios de comunicação como extensões do homem. São Paulo: Cultrix, 1974.

MISTRY, Purvi; WU, Peng. AI-powered fashion retail: recommendation systems, personalization and virtual try-on. Journal of Retailing and Consumer Services, [s.l.], 2022.

PMI. PROJECT MANAGEMENT INSTITUTE. Um Guia do Conhecimento em Gerenciamento de Projetos (Guia PMBOK). 7. ed. Newtown Square, PA: Project Management Institute, 2021.

PRESSMAN, Roger S.; MAXIM, Bruce R. Engenharia de Software: uma abordagem profissional. 8. ed. Porto Alegre: AMGH, 2016.

SEBRAE. SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS. Boas práticas em beleza online: dicas e orientações sobre o e-commerce de produtos de beleza. Brasília, 2013. Disponível em: <https://bibliotecas.sebrae.com.br>. Acesso em: 9 mar. 2024.

SINGH BHARDWAJ, Shreya et al. A comprehensive review on virtual trial rooms: integration of AR, AI, and 3D technologies for enhanced e-commerce experience. *Webology*, v. 18, n. 4, art. 189, 2021.

SOMMERVILLE, Ian. *Engenharia de Software*. 9. ed. São Paulo: Pearson, 2011.

UMODE. AR e provadores virtuais na moda: como a realidade aumentada está sendo adotada por marcas brasileiras. *U Mode Blog*, ago. 2025. Disponível em: <https://umode.com.br>. Acesso em: 15 nov. 2025.

VIEIRA FILHO, Geraldo. *Ferramentas de gestão empresarial*. São Paulo: Atlas, 2017.

Apêndice A – Termo de Abertura do Projeto (TAP)

O Apêndice A apresenta o Termo de Abertura do Projeto (TAP) do ARTificial, sintetizado no Quadro 2. Esse artefato organiza, de forma resumida, o título, a justificativa, os objetivos, o escopo, as entregas, a equipe, as premissas e as restrições que orientam o desenvolvimento deste estudo, tomando o ARTificial como prova de conceito para provedores virtuais multimodais no varejo de moda.

Logo abaixo você coloca, por exemplo:

Quadro 2 – Termo de Abertura do Projeto (TAP) – ARTificial

| Campo | Descrição |
|---------------------------------|--|
| Título do Projeto | Diretrizes para provedores virtuais no varejo de moda com RA e IA: estudo de concepção e prototipação (Figma), tendo o ARTificial como prova de conceito. |
| Gerente do Projeto | Profa. Dra. Jaqueline Brigladori Pugliese — Orientadora do Trabalho de Conclusão. |
| Sponsor | Não se aplica (projeto acadêmico). |
| Justificativa do Projeto | A experiência de compra em moda, especialmente no e-commerce, é impactada pela impossibilidade de experimentação física, o que contribui para elevadas taxas de devolução (em torno de 15% a 30%). Provedores virtuais baseados em Realidade Aumentada (RA) e Inteligência Artificial (IA) podem reduzir incertezas de compra e aperfeiçoar a tomada de decisão do consumidor, ao aproximar, de forma virtual, a experiência do provedor físico. |
| Objetivo do Projeto | Investigar e propor diretrizes conceituais para provedores virtuais no varejo de moda utilizando RA e IA, desenvolvendo um protótipo conceitual em Figma (ARTificial) como prova de conceito para etapas posteriores de implementação técnica e validação empírica. |

| | |
|------------------------------------|---|
| Escopo do Projeto | O estudo contempla: (i) revisão bibliográfica temática e análise de soluções existentes; (ii) definição de requisitos conceituais para provadores virtuais multimodais com sensor Kinect; e (iii) concepção de um protótipo navegável em Figma representando o provador virtual ARTificial. |
| Entregas do Projeto | <p>1- Protótipo conceitual navegável em Figma (fluxos do espelho, app do usuário e painel do lojista);</p> <p>2- Diretrizes de referência para provadores virtuais multimodais com RA, IA e Kinect;</p> <p>3- Análise comparativa de evidências bibliográficas e soluções de mercado;</p> <p>4- Documento técnico-científico do TCC consolidando fundamentação teórica, metodologia e resultados esperados.</p> |
| Membros da Equipe e Funções | <p>Wanderson Honorio Roberto da Silva – pesquisa, concepção e prototipação do sistema ARTificial;</p> <p>Michele Mayra da Silva – apoio na prototipação visual e na organização dos artefatos de design.</p> |
| Premissas | Acesso à orientação acadêmica; disponibilidade de ferramentas de prototipação (Figma); acesso à literatura científica e a casos de mercado sobre provadores virtuais, RA, IA e interfaces multimodais. |
| Restrições | Prazo acadêmico institucional para conclusão do TCC; ausência de recursos financeiros para desenvolvimento de um sistema em tempo real; escopo limitado à fase de concepção e prototipação, sem entrega de produto comercial ou implementação funcional do ARTificial. |

Fonte: Elaborado pelos autores.

Apêndice B – Estrutura Analítica do Projeto (EAP)

A Estrutura Analítica do Projeto (EAP) do ARTificial organiza o estudo em fases e sub-entregas conceituais, permitindo visualizar com clareza o que foi efetivamente desenvolvido neste primeiro artigo (estudo conceitual e prototipação em Figma) e o que permanece planejado para a fase de implementação técnica e avaliação experimental, a ser tratada em um segundo trabalho.

De forma sintética, o projeto se desdobra em duas grandes etapas:

- **Etapa 1 – Estudo conceitual e prototipação** (objeto deste artigo);
- **Etapa 2 – Desenvolvimento técnico e avaliação com usuários** (planejada para artigo subsequente).

A seguir, detalha-se a EAP com foco na **Etapa 1**, sinalizando, quando pertinente, os pontos que se conectam de forma prospectiva à **Etapa 2**.

1. Planejamento e levantamento de requisitos conceituais

Esta fase estabelece a base teórica, contextual e estratégica do estudo. A partir dela são definidos o problema, a relevância prática, o recorte tecnológico (Kinect, RA, IA, voz e gestos) e os requisitos que orientam a prototipação do provador virtual multimodal.

1.1 Pesquisa bibliográfica temática

A revisão de literatura foi estruturada em eixos temáticos, com o objetivo de mapear o estado da arte de tecnologias e abordagens diretamente relacionadas ao escopo do ARTificial.

1.1.1 Realidade Aumentada (RA) aplicada a provedores virtuais

Levantamento de conceitos, arquiteturas e aplicações de RA que utilizam sobreposição de elementos digitais ao corpo do usuário, com foco em provedores virtuais e experiências imersivas no varejo de moda.

1.1.2 Sensores de profundidade e Microsoft Kinect

Análise de trabalhos que utilizam o Kinect como sensor de profundidade e captura de movimento para rastreamento corporal, reconhecimento de gestos e geração de viewpoints virtuais. Esse eixo fundamenta a adoção do Kinect como dispositivo principal de entrada (voz, gesto e esqueleto).

1.1.3 Interfaces multimodais (voz e gestos)

Revisão de pesquisas sobre interfaces sem toque, combinando comandos de voz e gestos corporais, especialmente em contextos de navegação em ambientes

virtuais, museus, robótica e sistemas embarcados. Esse eixo sustenta a ideia de um provador virtual acessível, higienizado e fluido, comandado por fala e movimentos.

1.1.4 Inteligência Artificial e recomendação de moda

Análise de modelos de IA e redes neurais aplicados ao varejo de moda, incluindo sistemas de recomendação, previsão de tamanho, personalização de vitrine e análise de preferências do usuário. Nesse eixo são identificadas as funções de IA que podem dialogar com os dados coletados pelo Kinect.

1.1.5 Provedores virtuais existentes (estado da arte e da prática)

Mapeamento de soluções científicas e comerciais (nacionais e internacionais) de provedores virtuais, incluindo casos baseados apenas em IA/visão computacional, soluções de recomendação de tamanho e experiências em varejo físico. A partir desses exemplos são identificadas lacunas, especialmente a ausência de soluções brasileiras que integrem, de forma robusta, captura 3D via Kinect, RA, IA e interface multimodal (voz e gestos).

1.2 Análise de mercado

A análise de mercado reforça a relevância prática do problema e ancora a pesquisa na realidade do varejo brasileiro.

1.2.1 E-commerce de moda e taxa de devolução

Levantamento de dados sobre faturamento do setor, participação do e-commerce e índices de devolução (especialmente em vestuário e calçados), evidenciando impacto econômico e ambiental e justificando a pertinência de uma solução mais precisa de experimentação virtual.

1.2.2 Tendências omnichannel e experiências phygital

Identificação de evidências sobre a convergência entre canais físicos e digitais, destacando o papel de provedores virtuais, espelhos inteligentes e integrações com aplicativos móveis como elementos centrais de estratégias omnichannel no varejo de moda.

1.3 Derivação de requisitos conceituais

Com base na literatura e no contexto de mercado, são organizados requisitos que orientam a prototipação em Figma e a futura implementação técnica.

1.3.1 Requisitos funcionais conceituais

Mapeamento das principais funções esperadas no provador virtual multimodal, tais como:

- rastreamento corporal via Kinect;

- sobreposição de peças de roupa em tempo real (RA);
- navegação por comandos de voz e gestos;
- recomendação de tamanhos e peças com apoio de IA;
- integração com aplicativo do cliente e painel do lojista.

1.3.2 Requisitos não funcionais conceituais

Identificação de atributos desejáveis de qualidade, incluindo:

- realismo visual e estabilidade da sobreposição;
- responsividade e baixa latência;
- legibilidade e clareza de feedbacks na interface;
- fluidez da interação multimodal (voz + gestos) em ambiente de loja.

1.4 Documentação inicial do estudo

Formalização da identidade acadêmica do projeto, contemplando:

- problema de pesquisa;
- questão central de pesquisa;
- objetivo geral e objetivos específicos;
- delimitação da Etapa 1 (concepção + prototipação em Figma).

1.5 Modelagem conceitual do sistema

Construção de modelos visuais e textuais que organizam o fluxo da solução em nível conceitual.

1.5.1 Diagramas de processos (por exemplo, BPMN)

Elaboração de diagramas representando a jornada do usuário no provador virtual, desde a aproximação do espelho até a seleção de peças, interação por gestos/voz e registro de preferências.

1.5.2 Casos de uso e cenários de interação

Descrição de casos de uso conceituais para perfis distintos (cliente, lojista, sistema), com ênfase em cenários de uso multimodal, como comandos do tipo “quero ver calças masculinas”, “trocar de cor” ou “finalizar provador”.

2. Design e prototipação em Figma

Nesta fase, o estudo avança da modelagem abstrata para a materialização visual da experiência, **sem implementação de código**. O foco está em comunicar claramente a proposta de interação e o comportamento esperado da interface.

2.1 Criação das interfaces centrais do ecossistema ARTificial

Foram projetadas as principais telas e fluxos dos três núcleos do ecossistema:

- espelho interativo em loja;

- aplicativo do cliente;
- painel/portal do lojista.

2.1.1 Wireframes de baixa fidelidade

Construção inicial de wireframes com foco em hierarquia da informação, distribuição de elementos, zonas de captura pelo Kinect e áreas de feedback visual dos comandos de voz/gesto.

2.1.2 Protótipos de alta fidelidade (Figma)

Evolução dos wireframes para protótipos visualmente refinados, em Figma, incluindo:

- layout em resolução 1920×1080 para o espelho;
- simulação de estados de interface para diferentes comandos de voz e gestos;
- fluxos de navegação entre telas do espelho, app do usuário e ambiente do lojista.

2.2 Estudos de arquitetura conceitual da solução

Embora não haja desenvolvimento técnico nesta etapa, é proposta uma arquitetura lógica que demonstre a viabilidade futura da implementação.

2.2.1 Componentes principais da arquitetura

Documentação conceitual dos módulos:

- captura de corpo, profundidade, voz e gestos com Kinect;
- módulo de processamento (RA e IA);
- camada de APIs para comunicação entre espelho, backend e aplicativos;
- banco de dados de produtos, perfis e interações.

2.2.2 Fluxos de dados e integração entre módulos

Descrição de como os dados capturados pelo Kinect poderiam alimentar o módulo de IA, atualizar o espelho em tempo real e sincronizar informações com app e painel do lojista, reforçando a visão omnichannel.

3. Desenvolvimento técnico do sistema (previsto para a Etapa 2)

Esta fase **não é executada neste estudo**, mas está planejada como continuidade natural do projeto e base para um segundo artigo científico, de natureza mais aplicada e experimental.

3.1 Implementação do módulo Kinect

Desenvolvimento da camada de captura de movimento e voz utilizando o Kinect, incluindo:

- calibração de esqueleto;

- mapeamento de gestos relevantes ao provador;
- integração com o canal de reconhecimento de fala.

3.2 Interface multimodal (voz e gestos)

Construção da lógica de comandos de voz e gestos, definição de gramática de comandos, feedbacks visuais e tratamento de erros (como ruído de ambiente ou gestos ambíguos).

3.3 Módulo de IA e recomendação

Implementação de modelos de IA para:

- recomendação de tamanhos;
- sugestão de looks;
- análise de padrões de uso do provador.

3.4 Integrações e camadas de serviço

Integrações com cadastros de produtos, sistemas de estoque, CRM e, eventualmente, meios de pagamento ou geração de QR codes para checkout em outros canais.

4. Testes e validação (previstos para a Etapa 2)

Planejamento de estudos de usabilidade, avaliação de experiência do usuário, testes de desempenho do rastreamento por Kinect, taxa de acerto da IA e impacto percebido nas incertezas de compra e intenção de compra.

5. Documentação e apresentação

Na Etapa 1, a documentação concentra-se em:

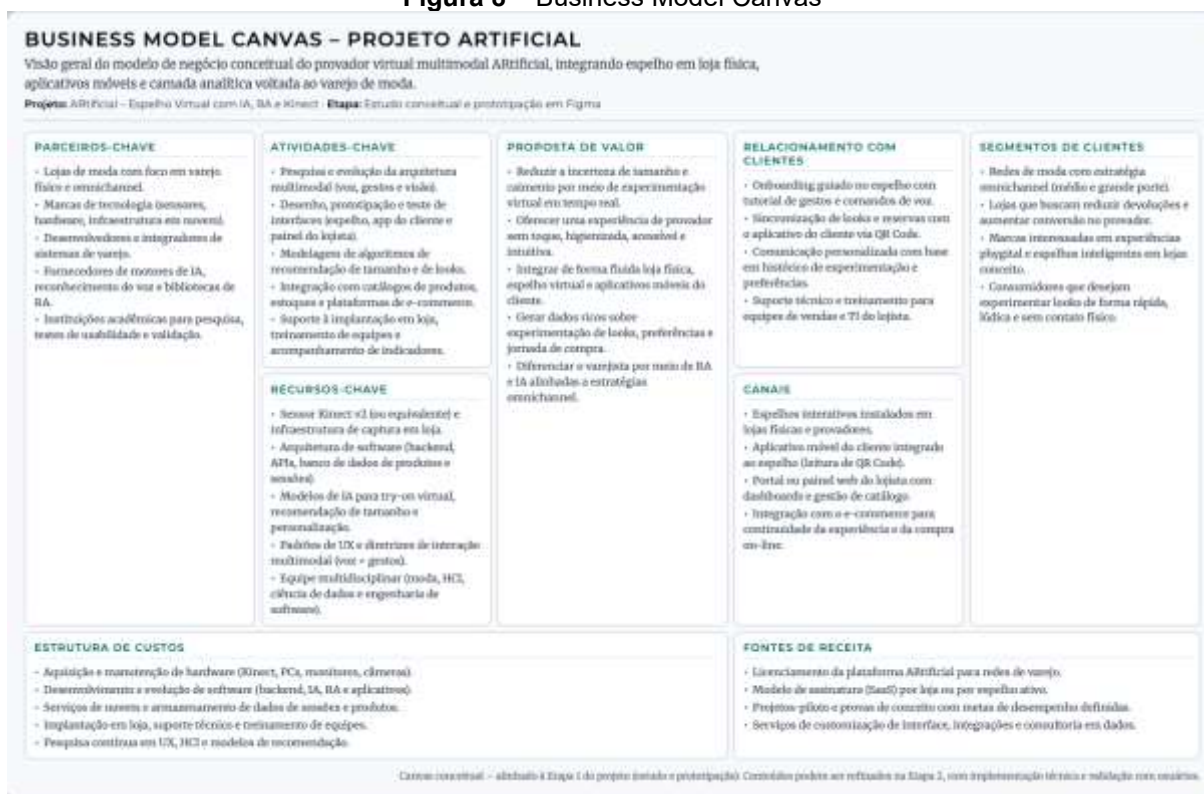
- revisão de literatura;
- sistematização da questão de pesquisa;
- organização da EAP e da metodologia;
- protótipos navegáveis em Figma;
- anexos conceituais (diagramas, fluxos, matrizes de requisitos).

Apêndice C – Matrizes SWOT e Plano de Ação 5W2H do Projeto ARtificial

O Apêndice C reúne os principais artefatos gerenciais elaborados para o projeto ARtificial – o provador virtual multimodal com sensor Kinect –, com o objetivo de explicitar sua lógica de negócio, suas forças e fragilidades estratégicas e o plano de ação que orienta o desenvolvimento do protótipo. Esse apêndice é composto por três instrumentos complementares: o Business Model Canvas (BMC) específico do ARtificial (Figura 8), a matriz SWOT do projeto (Quadro 3) e o plano de ação 5W2H (Quadro 4).

O Business Model Canvas sintetiza, em um único quadro visual, os principais blocos do modelo de negócio do ARtificial: parceiros-chave (varejistas de moda, fornecedores de hardware e software, instituições acadêmicas, entre outros), atividades-chave (pesquisa e desenvolvimento, implantação em loja, análise de dados de experimentação), recursos-chave (Kinect, plataformas de RA e IA, infraestrutura de dados), proposta de valor (redução de devoluções, experiência phygital mais imersiva e geração de insights sobre comportamento de prova), segmentos de clientes (lojas de moda, marcas diretas ao consumidor, shopping centers) e fluxos de receitas e custos associados. A Figura 8 apresenta o Canvas preenchido, permitindo visualizar como o protótipo conceitual se desdobra em um modelo de negócio potencialmente viável para o varejo de moda.

Figura 8 – Business Model Canvas



Fonte: Elaborado pelos autores.

A matriz SWOT (Quadro 3) aprofunda essa análise ao organizar, de forma estruturada, as forças (como o uso de tecnologias já validadas em provedores virtuais, o foco em um problema real de negócio – devoluções de 15–30% – e a abordagem omnichannel), as fraquezas (dependência de hardware específico como o Kinect, complexidade técnica da simulação de tecidos em tempo real, desafios de integração com sistemas legados), as oportunidades (crescimento do e-commerce de moda, interesse do varejo em experiências imersivas com RA e IA, potencial de parcerias B2B) e as ameaças (concorrência de soluções comerciais já consolidadas, preocupação com privacidade de dados e eventual resistência de usuários a tecnologias de RA). Essa matriz ajuda a posicionar o Artificial no cenário competitivo e orienta decisões futuras de escopo e priorização.

Quadro 3 - Matriz SWOT

| Forças | Fraquezas |
|---|--|
| Uso de tecnologia já validada em pesquisas acadêmicas com Kinect e provedores virtuais, reduzindo incertezas sobre a viabilidade técnica. | Dependência de hardware específico (Kinect ou equivalente), com risco de descontinuidade ou dificuldade de reposição. |
| Foco em um problema real do varejo de moda: altas taxas de devolução (15–30%) por incerteza de tamanho e caimento. | Complexidade técnica para simular tecidos e caimento em tempo real com boa performance em ambiente de loja. |
| Proposta de experiência multimodal sem toque (voz + gestos + visão), mais higienizada e acessível em comparação a telas físicas. | Necessidade de bases de dados específicas de peças, tecidos e corpos para treinar modelos de IA de try-on e recomendação. |
| Integração conceitual entre espelho em loja, app do cliente e painel do lojista, permitindo visão analítica das interações. | Exigência de infraestrutura mínima em loja (espaço, energia, rede, posicionamento de câmera) que pode limitar adoção em pequenos varejistas. |
| Oportunidades | Ameaças |
| Crescimento contínuo do e-commerce de moda no Brasil e no mundo, ampliando a demanda por soluções de experimentação virtual. | Concorrência de empresas e startups que utilizam outras abordagens de provedor virtual (aplicativos mobile, IA somente por foto etc.). |
| Tendência de estratégias omnichannel e experiências “phygital”, integrando loja física, espelhos inteligentes e aplicativos móveis. | Evolução tecnológica rápida, que pode tornar determinadas soluções de hardware ou frameworks obsoletos em curto prazo. |
| Maior sensibilidade do varejo à temática de sustentabilidade, com interesse em reduzir logística reversa e desperdício. | Riscos relacionados à privacidade e proteção de dados de imagem e métricas corporais dos usuários, exigindo forte governança. |
| Possibilidade de atuação B2B com grandes marcas e redes, oferecendo plataforma white label ou solução customizada. | Possibilidade de atuação B2B com grandes marcas e redes, oferecendo plataforma white label ou solução customizada. |

Fonte: Elaborado pelos autores.

Por fim, o plano de ação 5W2H (Quadro 4) organiza, em formato tabular, as principais ações previstas para o desenvolvimento do ARtificial, respondendo às sete

questões clássicas da ferramenta: What (o que será feito), Why (por que), Where (onde), When (quando), Who (quem), How (como) e How much (quanto). Conforme discutido na seção de fundamentação teórica, o 5W2H é uma técnica de planejamento amplamente utilizada para estruturar atividades de forma clara e objetiva (Campos, 2004; Vieira Filho, 2017). No contexto deste trabalho, o Quadro 4 detalha tanto as ações realizadas na Etapa 1 – como levantamento bibliográfico, concepção da arquitetura conceitual, elaboração dos protótipos em Figma e documentação acadêmica – quanto ações projetadas para a Etapa 2, relacionadas à implementação técnica, testes com usuários e ajustes iterativos do sistema.

Quadro 4 - Matriz 5W2H

| Elemento | Pergunta | Definição Aplicada ao Projeto |
|--------------|----------|---|
| What | O quê? | Desenvolvimento de um estudo conceitual e de um protótipo navegável em Figma para o provador virtual multimodal ARTificial , sintetizando diretrizes para provedores virtuais de moda com Realidade Aumentada (RA), Inteligência Artificial (IA) e sensor Kinect como prova de conceito. |
| Why | Por quê? | Mitigar a incerteza de tamanho e caimento nas compras de moda, associada a taxas de devolução entre 15% e 30%, explorando o potencial de RA e IA no varejo omnichannel e produzindo um referencial acadêmico que oriente futuras implementações técnicas e estudos experimentais. |
| Where | Onde? | Projeto desenvolvido no contexto acadêmico do TCC , aplicado conceitualmente ao varejo de moda omnichannel , contemplando cenários de uso em e-commerce e em lojas físicas com espelhos inteligentes integrados a aplicativos móveis. |
| When | Quando? | Durante o período de elaboração do TCC , desde a aprovação do projeto até a entrega e defesa, incluindo as etapas de revisão bibliográfica, modelagem conceitual, prototipação em Figma e sistematização dos resultados. |
| Who | Quem? | Wanderson Honorio Roberto da Silva (autor – pesquisa, modelagem e prototipação) e Michele Mayra da Silva (apoio na prototipação visual), sob orientação da Profa. Dra. Jaqueline Brigladori Pugliesi . |

| | | |
|-----------------|---------|---|
| How | Como? | Por meio de revisão de literatura , análise de mercado , definição de requisitos conceituais, modelagem de processos e casos de uso , desenho de arquitetura conceitual e desenvolvimento de protótipos de alta fidelidade no Figma para o espelho, aplicativo do cliente e painel do lojista. |
| How much | Quanto? | Utilização prioritária de recursos não financeiros : tempo de estudo, orientação acadêmica e ferramentas gratuitas/educacionais (como Figma e bases de artigos). Não há investimento em desenvolvimento em tempo real ou infraestrutura adicional além do ambiente acadêmico disponível. |

Fonte: Elaborado pelos autores.

Em conjunto, o Canvas, a matriz SWOT e o plano 5W2H apresentados neste apêndice funcionam como um enlace entre a dimensão técnica e a dimensão gerencial do projeto ARTificial. Eles evidenciam que o protótipo conceitual não é apenas um exercício de interface e tecnologia, mas está ancorado em um modelo de negócio plausível, em uma leitura crítica de riscos e oportunidades e em um roteiro de ação estruturado para a transição da fase conceitual para a fase de implementação e avaliação empírica.

Apêndice D – Artefatos de Engenharia de Requisitos

Este apêndice reúne os principais artefatos de engenharia de requisitos elaborados para o projeto ARtificial, estruturando, em nível técnico, os elementos que sustentam a concepção do espelho virtual multimodal com Kinect, RA e IA. Esses artefatos detalham o comportamento esperado do sistema, os fluxos de interação e as restrições de negócio, funcionando como uma formalização das decisões já discutidas ao longo do trabalho.

Inicialmente, apresenta-se o diagrama de processos em notação BPMN, que representa a jornada do usuário no provador virtual, desde a aproximação do espelho e ativação por gesto até a seleção de peças, experimentação virtual, comparação de looks e geração de QR Code para reserva. Esse modelo de processo permite visualizar, de forma integrada, a interação entre usuário, espelho, aplicativo móvel e painel do lojista, evidenciando pontos de decisão, eventos de início e fim e possíveis exceções no fluxo operacional.

Em seguida, são descritos os requisitos funcionais e não funcionais do sistema, organizados em quadros próprios. Os requisitos funcionais contemplam, por exemplo, cadastro e autenticação de usuários, cadastro e exibição de produtos, ativação do espelho virtual, aplicação de roupas virtuais, recomendações inteligentes, comandos de voz, reconhecimento de gestos e geração de QR Code. Já os requisitos não funcionais especificam propriedades como tempo de resposta, usabilidade, acessibilidade, disponibilidade, segurança, privacidade (LGPD), compatibilidade com sensores (Kinect e câmeras RGB), manutenibilidade, portabilidade e escalabilidade. Em consonância com Sommerville (2011) e Pressman e Maxim (2016), esses artefatos reforçam que tanto os requisitos funcionais quanto os não funcionais são críticos para a aceitação e o sucesso do sistema, pois definem não só “o que” o ARtificial deve fazer, mas também “como” essa experiência deve ser entregue em termos de qualidade.

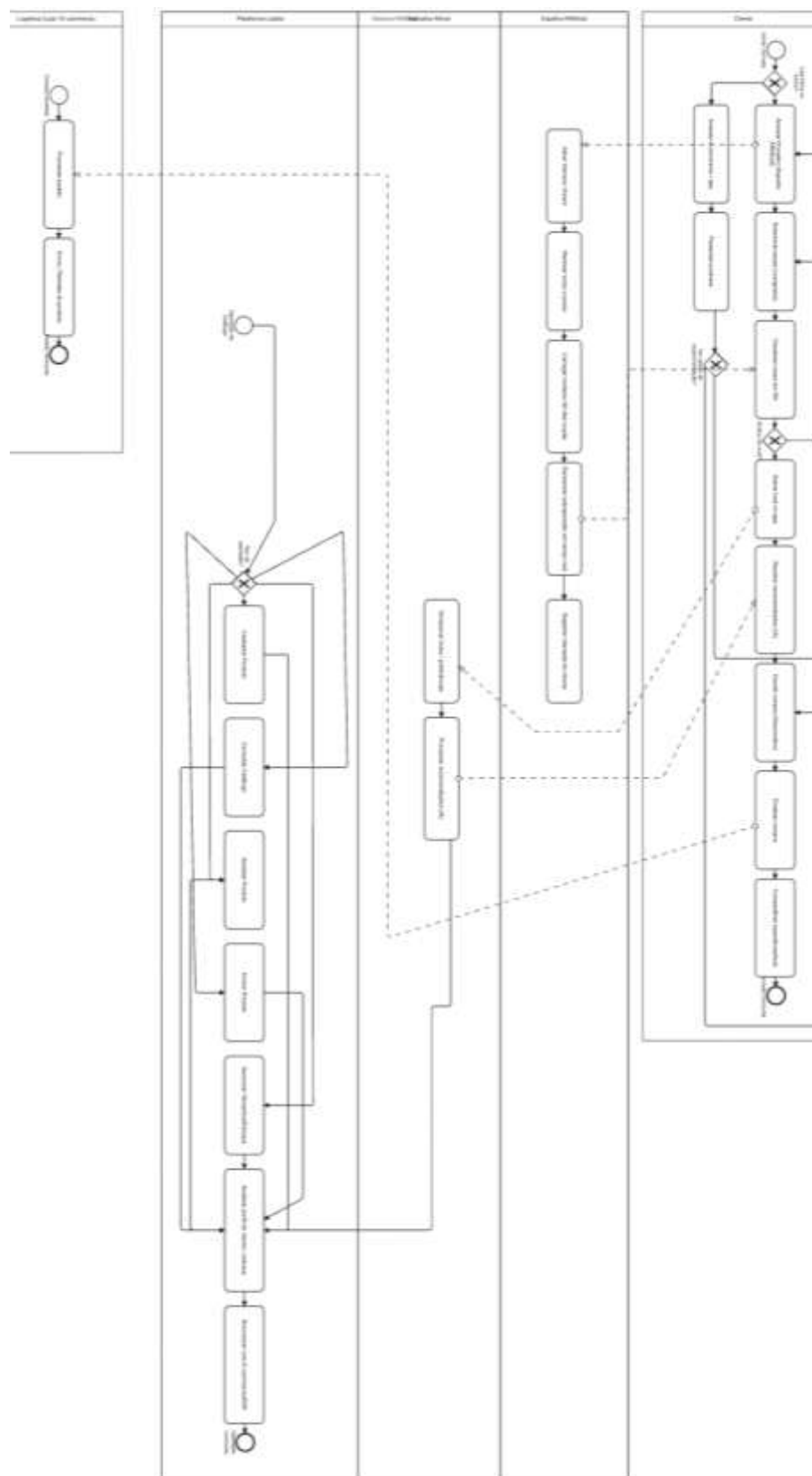
Na sequência, o apêndice apresenta o conjunto de regras de negócio que orientam o funcionamento do espelho virtual com IA. Essas regras traduzem políticas e restrições do domínio, como critérios para cadastro de usuários, formatos e limites para upload de roupas virtuais, condições para captura de imagem, requisitos de consentimento e privacidade, controle de acesso ao catálogo, limites de sessão e parâmetros para o motor de recomendação. Como destacam Sommerville (2011),

Pressman e Maxim (2016) e Larman (2004), tais regras não são funcionalidades em si, mas restringem e qualificam os requisitos, garantindo consistência, segurança e alinhamento às normas da organização e à LGPD.

Por fim, são apresentados os casos de uso selecionados para o sistema, descrevendo, em linguagem orientada ao usuário, os principais cenários de interação: cadastrar usuário, realizar login, selecionar roupas no catálogo, ativar o espelho virtual, capturar imagem, aplicar roupas virtualmente, visualizar o look em tempo real, alterar/remover itens do look, salvar looks favoritos, consultar histórico, compartilhar looks e encerrar sessão. Cada caso de uso explicita atores, pré-condições, fluxos principais e alternativos, além das pós-condições esperadas. Esses artefatos, conforme defendem Sommerville (2011) e Pressman e Maxim (2016), facilitam a comunicação entre equipe técnica, orientadora e usuários potenciais, servindo de base para planejamento, desenvolvimento futuro e testes.

Dessa forma, os artefatos de engenharia de requisitos apresentados neste apêndice materializam, em nível técnico, os procedimentos descritos na metodologia, constituindo a ponte entre o estudo conceitual do ARTificial e as etapas futuras de implementação e avaliação do sistema.

Figura 9 – BPMN



Fonte: Elaborado pelos autores.

Na sequência, são organizados os requisitos funcionais (Quadro 5) e não funcionais (Quadro 6) do sistema. Os requisitos funcionais descrevem o que o sistema deve fazer, abrangendo desde o cadastro e autenticação de usuários até a ativação do espelho virtual via Kinect, exibição de catálogo, aplicação de roupas virtuais, geração de recomendações por IA, comandos de voz, reconhecimento de gestos e integração com QR Codes. Já os requisitos não funcionais tratam de características de qualidade, como desempenho, usabilidade, acessibilidade, segurança, privacidade (LGPD), disponibilidade e compatibilidade com o Kinect e outros sensores, refletindo as preocupações técnicas necessárias para a operação em ambiente de loja.

Em complemento, o Quadro 7 reúne as regras de negócio que condicionam o comportamento do sistema, traduzindo em termos formais as políticas e restrições de uso do ARtificial. Essas regras abordam, por exemplo, critérios de cadastro, formatos e tamanhos de arquivos aceitos, limites de sessão, controle de acesso ao catálogo, consentimento para uso de imagem, tratamento de tentativas de login inválidas e parâmetros para o módulo de recomendação, garantindo aderência às diretrizes do domínio (varejo de moda, RA, IA e LGPD).

Por fim, são apresentados casos de uso exemplificativos (Quadro 8), descrevendo, em linguagem próxima ao usuário, os principais cenários de interação com o sistema, tais como cadastrar usuário, realizar login, selecionar roupas, ativar o espelho virtual, aplicar peças virtualmente, salvar looks e encerrar sessão. Esses casos de uso sintetizam, em nível narrativo, os requisitos funcionais previamente elencados, facilitando a comunicação entre as partes interessadas e servindo como referência para futuras atividades de projeto, desenvolvimento e teste.

Dessa forma, os artefatos de engenharia de requisitos apresentados neste apêndice materializam, em nível técnico, os procedimentos descritos na metodologia, constituindo a ponte entre o estudo conceitual do ARtificial e as etapas futuras de implementação e avaliação do sistema.

O levantamento de requisitos foi realizado com o objetivo de identificar as necessidades do usuário e garantir que o sistema atenda de forma eficaz às expectativas e demandas levantadas.

Para a obtenção dos requisitos, foram empregadas as seguintes técnicas de elicitación:

- Entrevistas: realizadas com os futuros usuários do sistema, a fim de compreender seus objetivos, dificuldades e necessidades específicas.

- Questionários: aplicados para coletar informações de maneira mais ampla e rápida, possibilitando analisar padrões de respostas.
- Observação/Visitas in loco: visitas ao ambiente onde o sistema será utilizado para identificar problemas reais e oportunidades de melhoria.

Importância da Elicitação

A elicitação de requisitos é uma das etapas mais importantes do processo de desenvolvimento de software, pois permite compreender as necessidades reais do cliente e do usuário final, reduzindo riscos de falhas no sistema. Segundo Sommerville (2011), “a elicitação adequada de requisitos é fundamental para o sucesso do projeto, visto que requisitos mal definidos são uma das principais causas de fracasso em sistemas de software”. De acordo com Pressman (2016), “a engenharia de requisitos é essencial para estabelecer a base sólida de um projeto, permitindo uma comunicação clara entre desenvolvedores e clientes”.

O processo de levantamento foi conduzido em três etapas principais:

1. Entrevistas estruturadas com os principais stakeholders, cujas perguntas e respostas completas encontram-se no Apêndice A.
2. Aplicação de questionários direcionados aos usuários finais, coletando percepções sobre funcionalidades desejadas.
3. Visitas presenciais ao ambiente de uso, possibilitando observar rotinas e identificar requisitos que não foram explicitamente mencionados.

As especificações foram organizadas em requisitos funcionais, apresentados na forma textual e em formato de tabela para melhor visualização. Também foram utilizadas histórias de usuários, como exemplo:

Quadro 5 – Requisitos Funcionais do sistema

| | | |
|---|---|---|
| RF001 -Cadastro de Usuários | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deve permitir o cadastro de novos usuários, armazenando dados pessoais, credenciais e informações de contato. | | |
| RF002 - Autenticar Usuários | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |

| | | |
|--|---|---|
| Descrição: O sistema deverá permitir o login por meio de autenticação segura com e-mail e senha.. | | |
| RF003- Cadastro de Produtos | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá permitir cadastrar, editar e excluir produtos com informações completas, como nome, preço, categoria e imagens. | | |
| RF004- Exibir Catálogo de Produtos | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá apresentar ao usuário o catálogo completo de produtos disponíveis para experimentação e compra. | | |
| RF005- Ativar Espelho Virtual | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá ativar a câmera/sensor e exibir a imagem do usuário em tempo real. | | |
| RF006- Redimensionar Roupa Automaticamente | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá ajustar automaticamente o tamanho da roupa virtual conforme as medidas detectadas. | | |
| RF007- Gerar Recomendações Inteligentes | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá recomendar produtos com base no histórico e preferências do usuário utilizando algoritmos de IA. | | |
| RF008- Gerenciar Carrinho de Compras | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |

| | | |
|--|---|---|
| Descrição: O sistema deverá permitir adicionar e remover itens do carrinho de compras antes da finalização. | | |
| RF009 -Processar Pagamento Online | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá oferecer métodos de pagamento como cartão, boleto e PIX. | | |
| RF010 -Gerar QR Code de Acesso | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá gerar QR Codes para facilitar o acesso ao espelho virtual em dispositivos | | |
| RF011 -Exibir Histórico de Compras | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input checked="" type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deve gerar QR Code para acesso rápido ao espelho virtual em dispositivos móveis. | | |
| RF012 - Gerar Relatórios de Vendas | Categoria: <input checked="" type="checkbox"/> Oculto <input type="checkbox"/> Evidente | Prioridade: <input type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input checked="" type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá gerar relatórios contendo métricas como produtos mais vistos e mais vendidos. | | |
| RF013 -Permitir Comandos de Voz | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá permitir interação por comandos de voz para navegação e seleção de itens. | | |
| RF014 -Reconhecer Gestos pelo Kinect | Categoria: <input type="checkbox"/> Oculto <input checked="" type="checkbox"/> Evidente | Prioridade: <input checked="" type="checkbox"/> Altíssima <input type="checkbox"/> Alta <input type="checkbox"/> Média <input type="checkbox"/> Baixa |
| Descrição: O sistema deverá detectar gestos do usuário para navegar no catálogo e seleciona | | |

Fonte: Elaborado pelos autores.

Os Requisitos Não Funcionais (RNFs) descrevem restrições e características de qualidade que o sistema deve atender, mas que não estão diretamente ligadas às funcionalidades. Eles definem critérios como desempenho, usabilidade, segurança, confiabilidade e portabilidade.

Segundo Sommerville (2011), “os requisitos não funcionais definem propriedades e restrições do sistema, como tempo de resposta, capacidade de armazenamento e requisitos de segurança”. Já Pressman e Maxim (2016) destacam que “estes requisitos são críticos para a aceitação do sistema pelos usuários, pois determinam a experiência de uso e a eficiência operacional”.

Portanto, os RNFs são essenciais para assegurar que o Espelho Virtual com Inteligência Artificial atenda padrões de qualidade, garantindo sua viabilidade técnica e aceitação pelos usuários.

No Quadro 6 estão descritos os Requisitos Não Funcionais levantados para o sistema proposto. Cada requisito está devidamente identificado com código (RNF001, RNF002, etc.), acompanhado de sua descrição, tipo e classificação quanto à obrigatoriedade e permanência.

Quadro 6 – Requisitos Não Funcionais do sistema

| | | | | |
|--|---|----------------------------------|----------------------------------|-----------------------------------|
| RNF001- Tempo de resposta | O sistema deve processar a captura da imagem e renderizar a roupa em até 2 segundos. | Desempenho | () Desejável (X) Obrigatório | (X) Permanente () Transitório |
| RNF002- Carregament o Rápido | O catálogo de roupas deve carregar em no máximo 3 segundos | Desempenho | () Desejável (X) Obrigatório | (X) Permanente () Transitório |
| RNF003- Usabilidade | A interface deve ser intuitiva, com ícones e mensagens de orientação para o usuário | Usabilidade | () Desejável (X) Obrigatório | (X) Permanente () Transitório |
| RNF004- Acessibilidade e | O sistema deve oferecer instruções visuais e textuais para diferentes perfis de usuários. | Acessibilidade | () Desejável (X) Obrigatório | (X) Permanente () Transitório |
| RNF005- Disponibilidad e | O sistema deve garantir no mínimo 95% de disponibilidade. | Disponibilidade / Confiabilidade | () Desejável (X) Obrigatório | (X) Permanente () Transitório |
| RNF006- Confiabilidad e | O sistema deve validar imagens e evitar falhas em capturas não reconhecidas. | Confiabilidade | () Desejável (X) Obrigatório | (X) Permanente () Transitório |
| RNF007- Segurança de Dados | As imagens do usuário não devem ser armazenadas sem consentimento. | Segurança | () Desejável (X) Obrigatório | (X) Permanente () Transitório |

| | | | | |
|---|---|----------------------------|---|--|
| RNF008- Privacidade (LGPD) | O sistema deve estar em conformidade com a LGPD, garantindo a proteção dos dados pessoais. | Legal Privacidade / | <input type="checkbox"/> Desejável <input checked="" type="checkbox"/> Obrigatório | <input checked="" type="checkbox"/> Permanente <input type="checkbox"/> Transitório |
| RNF009- Compatibilidade de | O sistema deve ser compatível com sensores Kinect e câmeras RGB convencionais. | Compatibilidade / Hardware | <input type="checkbox"/> Desejável <input checked="" type="checkbox"/> Obrigatório | <input checked="" type="checkbox"/> Permanente <input type="checkbox"/> Transitório |
| RNF010- Manutenibilidade | O sistema deve ser modular e ter código documentado para facilitar ajustes e melhorias. | Manutenibilidade | <input checked="" type="checkbox"/> Desejável <input type="checkbox"/> Obrigatório | <input checked="" type="checkbox"/> Permanente <input type="checkbox"/> Transitório |
| RNF011- Portabilidade | O sistema deve poder ser implantado tanto em ambiente local quanto em servidores na nuvem. | Portabilidade | <input checked="" type="checkbox"/> Desejável <input type="checkbox"/> Obrigatório | <input checked="" type="checkbox"/> Permanente <input type="checkbox"/> Transitório |
| RNF012- Escalabilidade e | O sistema deve suportar aumento de usuários simultâneos sem perda de desempenho. | Escalabilidade | <input checked="" type="checkbox"/> Desejável <input type="checkbox"/> Obrigatório | <input checked="" type="checkbox"/> Permanente <input type="checkbox"/> Transitório |
| RNF013- Resolução Mínima da Interface | Resolução Mínima da Interface | Escalabilidade | <input type="checkbox"/> Desejável <input checked="" type="checkbox"/> Obrigatório | <input checked="" type="checkbox"/> Permanente <input type="checkbox"/> Transitório |
| RNF014- Latência de Interação | O tempo de resposta entre a entrada do usuário (gesto, voz ou toque) e a reação do sistema não deve exceder 100 milissegundos (ms). | Escalabilidade | <input type="checkbox"/> Desejável <input checked="" type="checkbox"/> Obrigatório | <input checked="" type="checkbox"/> Permanente <input type="checkbox"/> Transitório |

Fonte: Elaborado pelos autores.

As Regras de Negócio são políticas, diretrizes ou restrições que refletem o funcionamento da organização e influenciam diretamente o comportamento do sistema. Segundo Sommerville (2011), “às regras de negócio são restrições derivadas das políticas da empresa ou do domínio de aplicação que devem ser seguidas pelo sistema”.

Já Pressman e Maxim (2016) afirmam que “as regras de negócio têm papel essencial na modelagem de sistemas, pois estabelecem critérios que controlam a forma como os dados podem ser inseridos, processados e utilizados”.

Portanto, no contexto do Espelho Virtual com Inteligência Artificial, as regras de negócio garantem que o sistema opere de acordo com os objetivos definidos, respeitando limitações de segurança, uso adequado e padronização de informações.

As regras de negócio são fundamentais porque:

- Garantem consistência nos processos.
- Definem os limites de atuação do sistema.
- Ajudam a manter segurança, padronização e integridade das informações.
- Reduzem falhas humanas, já que o sistema impede entradas ou operações inválidas.

Conforme destaca Larman (2004), “as regras de negócio não são requisitos funcionais por si mesmas, mas influenciam e restringem os requisitos funcionais e não funcionais”.

No Quadro 7 a seguir, são apresentadas as regras de negócio que devem ser consideradas no desenvolvimento do sistema.

Quadro 7 – Regras de Negócio do sistema.

| |
|--|
| RN001 - Cadastro de Usuários |
| Descrição: Só serão permitidos cadastros de usuários com e-mail válido e senha de no mínimo 8 caracteres. |
| RN002 - Upload de Roupas |
| Descrição: Apenas serão aceitos uploads de roupas nos formatos de imagem PNG ou JPEG, com tamanho máximo de 5MB. |
| RN003 - Captura de Imagem |
| Descrição: Serão permitidas capturas de imagem apenas se o sensor detectar a presença completa do usuário em frente à câmera. |
| RN004 – Privacidade de Dados |
| Descrição: As imagens capturadas não poderão ser armazenadas ou compartilhadas sem o consentimento expresso do usuário. |
| RN005 – Acesso ao Catálogo |
| Descrição: O acesso ao catálogo de roupas virtuais será permitido somente a usuários autenticados no sistema. |
| RN006 – Compatibilidade de Dispositivos |
| Descrição: O sistema somente poderá ser utilizado em dispositivos compatíveis (Kinect ou câmeras RGB aprovadas). |
| RN007 – Atualização do Catálogo |
| Descrição: Apenas administradores autorizados poderão cadastrar, remover ou atualizar roupas no catálogo virtual. |
| RN008 – Restrições de Tamanho das Roupas |
| Descrição: As roupas virtuais devem ser cadastradas com medidas padrão para garantir ajuste proporcional ao corpo do usuário. |

| |
|---|
| RN009 – Sessão de Uso |
| Descrição: Cada sessão de uso terá limite máximo de 30 minutos por usuário, para otimizar os recursos do sistema. |
| RN010 – Consentimento para Uso de Imagem |
| Descrição: O sistema só poderá ser utilizado após o usuário aceitar os termos de uso e a política de privacidade. |
| RN011 – Registro de Tentativas Inválidas |
| Descrição: Em caso de 3 tentativas inválidas de login, o sistema deve bloquear o acesso temporariamente por 15 minutos. |
| RN012 – Exclusão de Conta |
| Descrição: O usuário poderá solicitar a exclusão de sua conta e dados a qualquer momento, conforme previsto na LGPD. |
| RN013 – Política de Recomendações |
| Descrição: O sistema de Recomendações Inteligentes (RF007) deve utilizar uma combinação de filtragem colaborativa e filtragem baseada em conteúdo. As recomendações devem ser atualizadas a cada 10 minutos ou a cada nova interação significativa do usuário. |
| RN014 – Limite de Produtos no Carrinho |
| Descrição: O número máximo de itens (diferentes produtos) que um usuário pode adicionar ao Carrinho de Compras (RF008) é limitado a 20 itens. |
| RN015 – Validade de Sessão de Autenticação |
| Descrição: A sessão de autenticação do usuário (login) no aplicativo móvel ou no espelho virtual tem uma validade máxima de 24 horas a partir do último acesso. |

Fonte: Elaborado pelos autores.

Um Caso de Uso (Use Case) descreve uma sequência de ações que um sistema realiza em interação com atores externos (usuários ou outros sistemas), para alcançar um objetivo específico. Ele é utilizado para capturar os requisitos funcionais do sistema de forma compreensível por analistas, desenvolvedores e usuários.

Segundo Sommerville (2011), “os casos de uso são uma técnica de modelagem que descreve o comportamento do sistema do ponto de vista do usuário, facilitando a comunicação e o entendimento dos requisitos”. Já Pressman e Maxim (2016) destacam que “casos de uso ajudam a identificar e organizar funcionalidades, cenários alternativos e fluxos de exceção, sendo essenciais para a documentação de requisitos”.

A importância dos casos de uso está em:

- Garantir clareza sobre como o sistema deve se comportar em diferentes situações;
- Facilitar o planejamento, desenvolvimento e testes;

- Servir como base para diagramas, documentação de requisitos e validação com usuários;
- Permitir identificação de cenários principais e alternativos, melhorando a robustez do sistema (LARMAN, 2004).

No Quadro 8, apresenta-se o índice dos casos de uso do sistema Espelho Virtual com IA, indicando cada caso, ID, ator principal e objetivo.

Quadro 8 – Casos de Uso

| Caso de Uso – Cadastrar Usuário | |
|--|--|
| ID | UC 001 |
| Descrição | Descrição: Este caso de uso tem por objetivo cadastrar novos usuários no sistema, garantindo que apenas dados válidos sejam aceitos. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Nenhuma |
| Cenário Principal | <p>O usuário seleciona “Cadastro de Usuário”.</p> <p>O sistema carrega o formulário.</p> <p>O usuário preenche nome, e-mail e senha.</p> <p>O sistema valida os dados.</p> <p>O sistema gera um código único.</p> <p>O usuário confirma os dados.</p> <p>O sistema registra o usuário no banco de dados.</p> <p>O sistema exibe mensagem de confirmação.</p> <p>Pós-condição: Usuário cadastrado e apto a realizar login.</p> |
| Pós-condição | Usuário cadastrado e apto a realizar login. |
| Cenário Alternativo | <p>4a – O usuário informa dados inválidos.</p> <p>4a.1 – O sistema exibe mensagem de erro e solicita correção.</p> |

| Caso de Uso – Realizar Login | |
|-------------------------------------|--|
| ID | UC 002 |
| Descrição | Este caso de uso tem por objetivo permitir que o usuário acesse o sistema utilizando suas credenciais. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Usuário já cadastrado. |

| | |
|----------------------------|--|
| Cenário Principal | 1.O usuário seleciona “Login”. 2.O sistema carrega o formulário de autenticação. 3.O usuário informa e-mail e senha. 4.O sistema valida as credenciais. 5.O sistema concede acesso ao usuário. |
| Pós-condição | Sessão iniciada. |
| Cenário Alternativo | 4a – O usuário informa credenciais incorretas. 4a.1 – O sistema exibe mensagem de erro. |

| | |
|--|--|
| Caso de Uso – Selecionar Roupas/Itens | |
| ID | UC 003 |
| Descrição | Este caso de uso tem por objetivo permitir que o usuário navegue e selecione roupas no catálogo. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Usuário logado. |
| Cenário Principal | <ol style="list-style-type: none"> 1. O usuário acessa o catálogo de roupas. 2. O sistema exibe as opções disponíveis. 3. O usuário seleciona um item. 4. O sistema adiciona o item à lista de experimentação. |
| Pós-condição | Item disponível para visualização no espelho virtual. |
| Cenário Alternativo | |

| | |
|---|--|
| Caso de Uso – Ativar Espelho Virtual | |
| ID | UC 004 |
| Descrição | Este caso de uso tem por objetivo ativar a câmera Kinect e iniciar a captura da imagem do usuário. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Usuário logado e Kinect conectado. |
| Cenário Principal | O usuário seleciona “Ativar Espelho Virtual”. O sistema inicia a captura de imagem. O sistema calibra o posicionamento. O sistema exibe a imagem em tempo real. |
| Pós-condição | Espelho Virtual ativado. |
| Cenário Alternativo | |

| |
|---|
| Caso de Uso – Capturar Imagem do Usuário |
|---|

| | |
|----------------------------|---|
| ID | UC 005 |
| Descrição | Este caso de uso tem por objetivo capturar a imagem do usuário para sobreposição das roupas. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Espelho Virtual ativado. |
| Cenário Principal | O sistema identifica a silhueta do usuário. O usuário ajusta sua posição. O sistema registra a captura inicial. |
| Pós-condição | Imagem capturada com sucesso. |
| Cenário Alternativo | |

Caso de Uso – Aplicar Roupas Virtualmente

| | |
|----------------------------|---|
| ID | UC 006 |
| Descrição | Este caso de uso tem por objetivo permitir a aplicação de roupas virtuais sobre a imagem do usuário. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Roupa selecionada e imagem capturada. |
| Cenário Principal | O usuário seleciona uma roupa do catálogo. O sistema projeta a roupa sobre a imagem do usuário. O sistema ajusta o tamanho e proporção automaticamente. |
| Pós-condição | Roupa aplicada virtualmente. |
| Cenário Alternativo | |

Caso de Uso – Visualizar Look em Tempo Real

| | |
|----------------------------|--|
| ID | UC 007 |
| Descrição | Este caso de uso tem por objetivo permitir ao usuário visualizar o look em tempo real. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Roupa aplicada. |
| Cenário Principal | O sistema exibe a sobreposição da roupa em tempo real. O usuário se movimenta para observar os ajustes. |
| Pós-condição | Look visualizado em tempo real. |
| Cenário Alternativo | |

| Caso de Uso – Alterar/Remover Item do Look | |
|---|---|
| ID | UC 008 |
| Descrição | Este caso de uso tem por objetivo permitir que o usuário altere ou remova roupas aplicadas. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Look montado. |
| Cenário Principal | O usuário seleciona a roupa que deseja alterar ou remover. O sistema atualiza a visualização conforme a escolha. |
| Pós-condição | Look atualizado. |
| Cenário Alternativo | |

| Caso de Uso – Salvar Look Favorito | |
|---|---|
| ID | UC 009 |
| Descrição | Este caso de uso tem por objetivo salvar um look escolhido pelo usuário em seu perfil. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Look montado. |
| Cenário Principal | O usuário seleciona a opção “Salvar Look”. O sistema registra o look no perfil do usuário. O sistema confirma a operação. |
| Pós-condição | Look salvo no perfil. |
| Cenário Alternativo | |

| Caso de Uso – Consultar Último Histórico | |
|---|---|
| ID | UC 010 |
| Descrição | Este caso de uso tem por objetivo exibir o último look ou ação realizada pelo usuário. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Existe histórico registrado. |
| Cenário Principal | O usuário seleciona “Último Histórico”. O sistema carrega os dados da última ação. O sistema exibe as informações para o usuário. |
| Pós-condição | Último histórico exibido. |

| | |
|----------------------------|--|
| Cenário Alternativo | 1a – Não há histórico disponível. 1a.1 – O sistema informa “Nenhum histórico encontrado”. |
|----------------------------|--|

| Caso de Uso – Compartilhar Look | |
|--|--|
| ID | UC 011 |
| Descrição | Este caso de uso tem por objetivo permitir que o usuário compartilhe um look salvo em redes sociais ou e-mail. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Look salvo. |
| Cenário Principal | O usuário seleciona “Compartilhar”. O sistema exibe as opções de compartilhamento. O usuário escolhe o canal (e-mail, rede social). O sistema envia o look. |
| Pós-condição | Look compartilhado. |
| Cenário Alternativo | |

| Caso de Uso – Encerrar Sessão | |
|--------------------------------------|---|
| ID | UC 012 |
| Descrição | Este caso de uso tem por objetivo encerrar com segurança a sessão do usuário. |
| Ator Primário | Usuário do sistema |
| Pré-condição | Sessão iniciada. |
| Cenário Principal | O usuário seleciona “Sair”. O sistema encerra a sessão. O sistema retorna à tela inicial. |
| Pós-condição | Sessão finalizada com segurança. |
| Cenário Alternativo | |

Fonte: Elaborado pelos autores.

Apêndice E – Protótipo de Interfaces no Figma – ARtificial

O Apêndice E apresenta o protótipo de alta fidelidade do sistema ARtificial, desenvolvido na ferramenta Figma, contemplando o fluxo completo de interação no espelho virtual (telas de espera, login, tutorial, experimentação de categorias e produtos, captura de fotos, comparação de looks, finalização e perfil do usuário).

O objetivo deste apêndice é registrar visualmente a proposta de interface e apoiar futuras etapas de implementação, permitindo que equipes de desenvolvimento, UX e negócio compreendam, de forma integrada, a jornada do usuário no provador virtual multimodal.

Para facilitar a visualização dinâmica das transições entre telas e dos microcomportamentos de interface, o protótipo navegável pode ser acessado no endereço a seguir:

<https://www.figma.com/design/b39F7PLXhNBgvTYvRiHQ8/ARtificial?node-id=0-1&t=PInydAhlHdu2wAS0-1>.

A Figura 10 sintetiza a estrutura geral do protótipo desenvolvido em Figma para o ecossistema ARtificial. O arquivo está organizado em três blocos principais de telas: (i) Espelho – Virtual Fitting, que reúne o fluxo completo do provador físico, incluindo tela de espera, login por QR code ou uso sem login, tutorial de gestos, telas de experimentação por categoria e produto, captura de fotos, comparação de looks, finalização com QR code e tela de perfil do usuário no espelho; (ii) App – Lojista, com telas de login, página inicial com navegação inferior, cadastro de produtos, área de checkout/reservas e perfil do lojista; e (iii) App – Cliente, contemplando login, home, busca/navegação, carrinho ou área de reservas, perfil do usuário e telas de coleção/galeria de looks salvos. Em conjunto, esses três conjuntos de wireframes representam, de forma integrada, os fluxos essenciais de interação entre espelho, aplicativo do cliente e aplicativo do lojista.

Figura 10 – Estrutura geral do protótipo ARTificial em Figma.



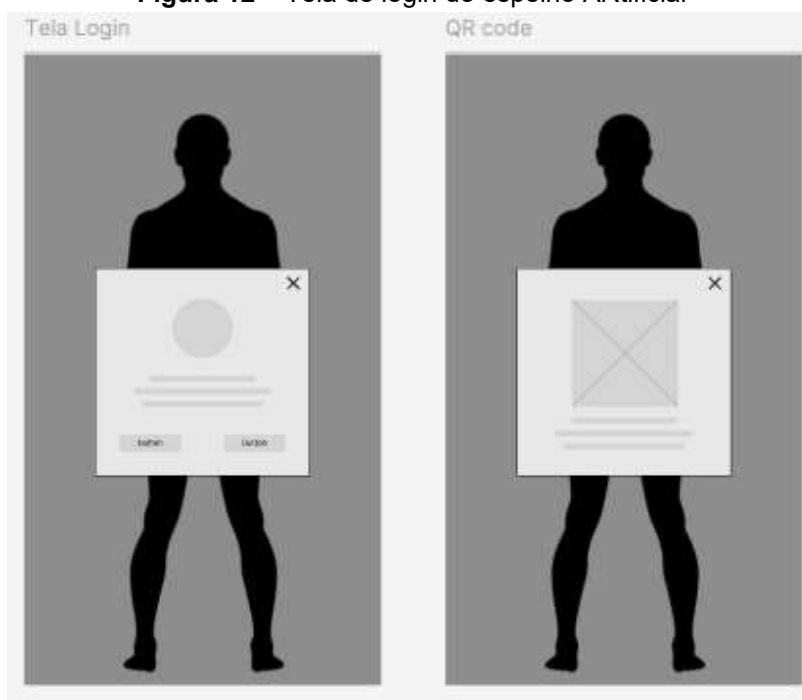
Fonte: Elaborado pelos autores.

A Figura 11 representa a tela de espera do espelho ARTificial. Essa tela é acionada automaticamente após cerca de três minutos de inatividade e permanece visível para quem circula em frente ao equipamento, aplicando virtualmente peças de destaque conforme o sexo estimado das pessoas que passam e exibindo chamadas de categorias em promoção (por exemplo, “acessórios com 20% off”). No centro da interface é apresentada uma indicação visual do gesto de levantar a mão espalmada para “chamar” o espelho, comunicando ao usuário que esse movimento inicia a sessão interativa e mantendo a coerência com o vocabulário gestual definido para o sistema.

Figura 11 - Tela de espera do espelho ARtificial

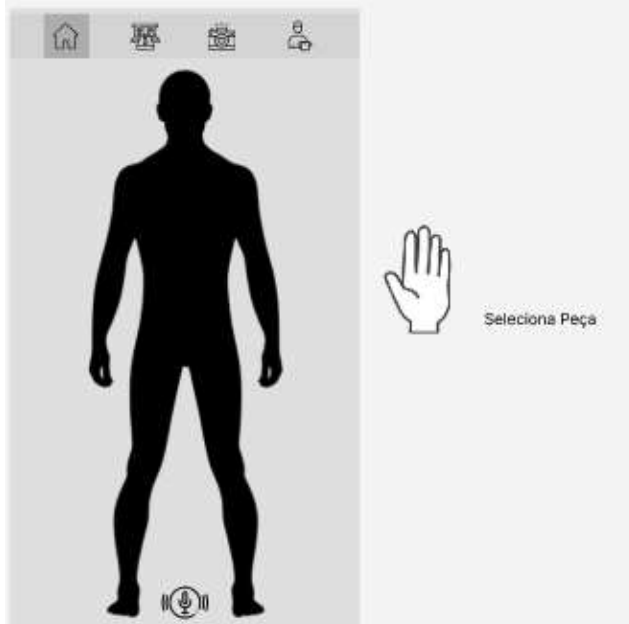
Fonte: Elaborado pelos autores.

A Figura 12 ilustra a tela de login apresentada após o reconhecimento do gesto de ativação (levantamento de mão). Sobre a imagem de fundo do usuário, o sistema exibe uma janela central com duas possibilidades de acesso: (i) autenticação via QR code, que deve ser lido pelo aplicativo móvel do ARtificial para carregar preferências e histórico do usuário; e (ii) a opção “usar sem login”, na qual o espelho é liberado apenas com funcionalidades básicas de experimentação, sem recuperação de dados personalizados. Essa tela marca a transição entre o estado passivo de espera e o início efetivo da sessão de provedor virtual.

Figura 12 – Tela de login do espelho ARTificial

Fonte: Elaborado pelos autores.

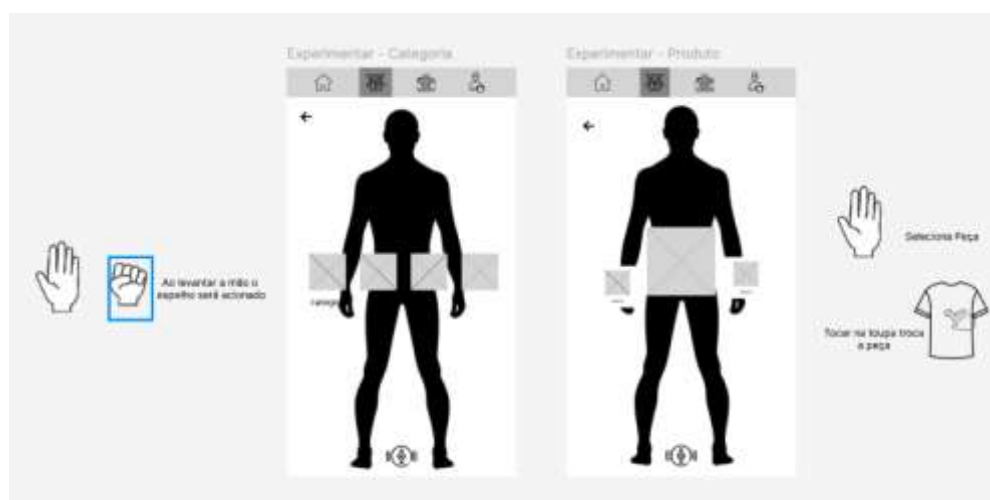
A Figura 13 apresenta a tela de tutorial do espelho ARTificial, exibida tanto no primeiro acesso autenticado quanto quando o usuário opta por utilizar o sistema sem login. Nessa tela são ilustrados os gestos básicos de interação — como o movimento de mão para seleção de peças — orientando o usuário sobre como comandar o provador virtual por meio da interface multimodal.

Figura 13 – Tela de tutorial de gestos do espelho ARTificial

Fonte: Elaborado pelos autores.

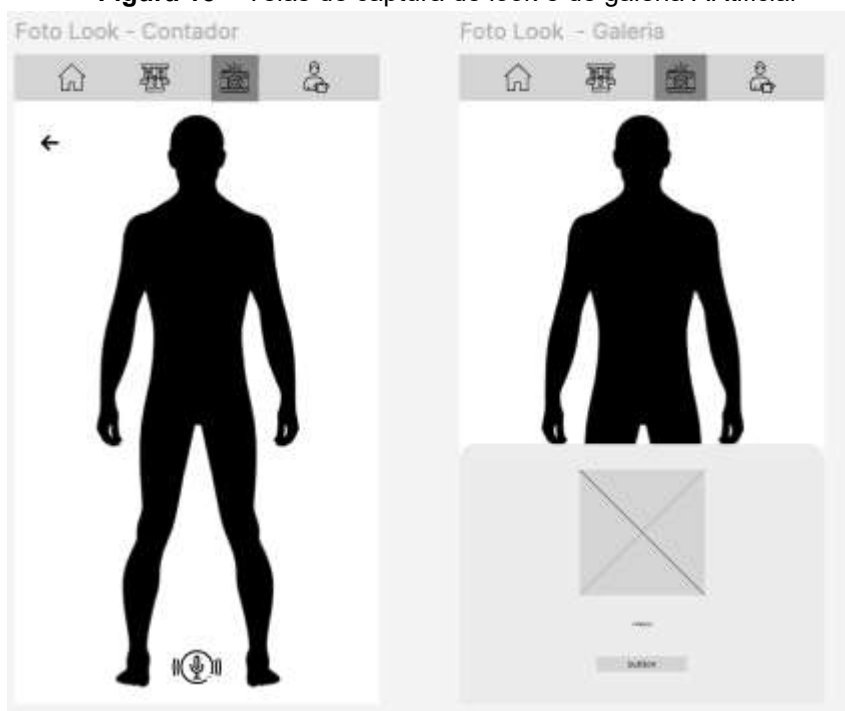
A Figura 14 ilustra o fluxo de experimentação de peças no espelho virtual, contemplando tanto a seleção de categorias quanto a navegação entre produtos. Na etapa de categorias, as peças são organizadas em um carrossel na altura do tronco do usuário, permitindo percorrer as opções por meio do gesto de “pegar e arrastar” com a mão fechada, enquanto o gesto de mão espalmada é utilizado para selecionar a categoria desejada. Em seguida, na etapa de produto, o usuário pode trocar a peça exibida aproximando a mão do item desejado, conforme exemplificado na figura, bem como emitir comandos de voz, como “mostrar saias”, para filtrar o catálogo apresentado no espelho.

Figura 14 – Telas de experimentação por categoria e produto



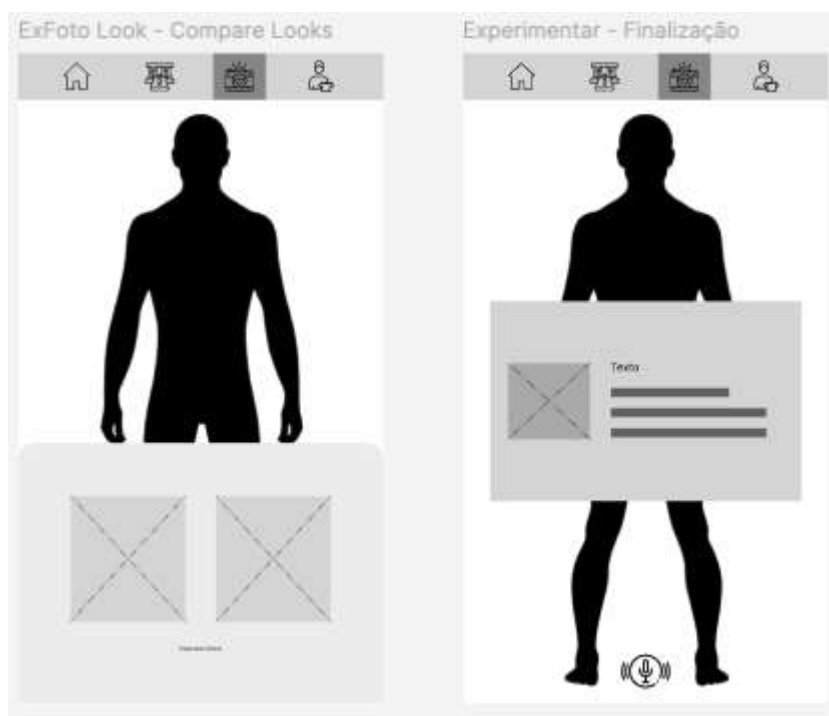
Fonte: Elaborado pelos autores.

A Figura 15 apresenta as telas de captura e de galeria de looks do espelho virtual. Ao acionar o ícone de câmera ou emitir o comando de voz “tirar foto”, o sistema registra o look que está sendo experimentado e o envia para a galeria do usuário. Em seguida, é exibida uma tela em que o usuário pode apenas salvar o registro ou encaminhá-lo para a funcionalidade de comparação de looks, que o redireciona para a tela de composição de looks.

Figura 15 – Telas de captura de look e de galeria ARTificial

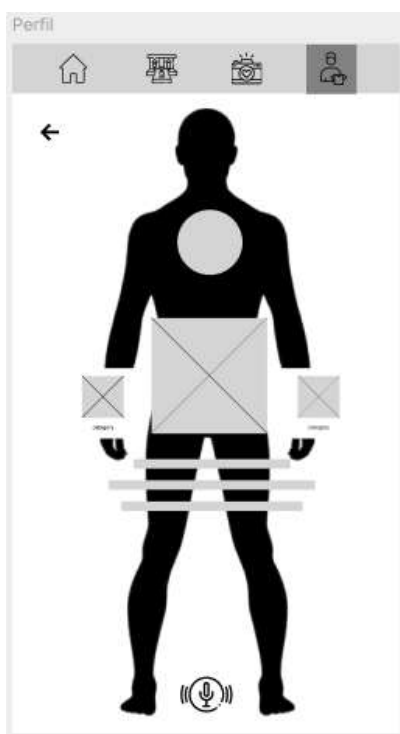
Fonte: Elaborado pelos autores.

A Figura 16 apresenta duas etapas finais do fluxo de uso do espelho ARTificial: a tela de comparação de looks e a tela de finalização da experimentação. À esquerda, a tela de comparação exibe, lado a lado, as fotografias de dois looks previamente capturados, permitindo que o usuário visualize simultaneamente as combinações e escolha aquela que melhor atende às suas preferências estéticas. À direita, a tela de finalização apresenta um quadro com a descrição dos itens que compõem o look selecionado e um código QR gerado automaticamente, que pode ser escaneado pelo dispositivo móvel do usuário para reservar as peças na loja, concluindo a jornada de experimentação no provador virtual.

Figura 16 – Telas finalização da experiência no espelho ARTificial

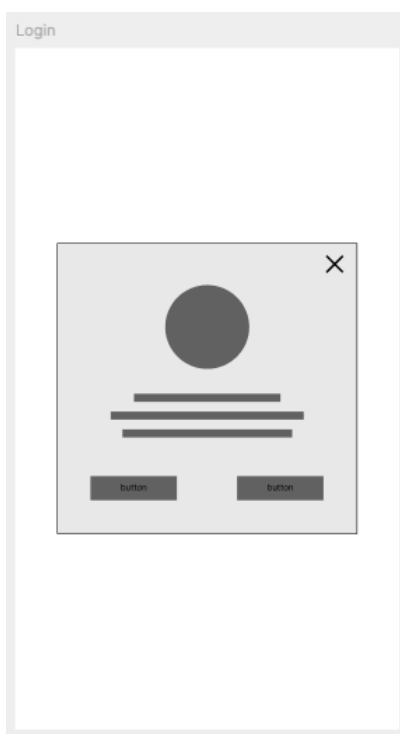
Fonte: Elaborado pelos autores.

A Figura 17 apresenta a tela de perfil do usuário no espelho ARTificial. Nessa interface são exibidos a fotografia do usuário, uma minibio com informações pessoais relevantes e os principais dados corporalmente captados pelo Kinect, como medidas, gênero e estilo de preferência configurado. Além disso, mantém-se, na parte inferior da tela, o ícone do microfone, indicando que os comandos de voz permanecem disponíveis mesmo no contexto de visualização do perfil, em consonância com a proposta de interação multimodal do sistema.

Figura 17 – Tela de perfil do usuário

Fonte: Elaborado pelos autores.

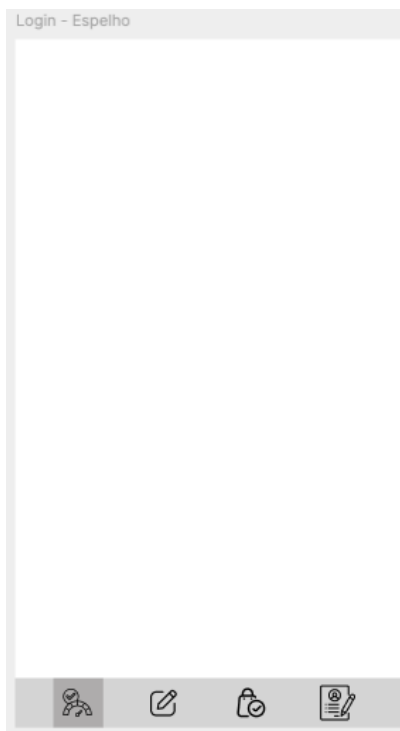
A Figura 18 representa a tela de login do aplicativo Cliente do sistema ARtificial. Nessa interface, o usuário pode realizar o acesso por meio do cadastro com endereço de e-mail ou optar por métodos de autenticação social, como Google, Apple ou Facebook. A proposta dessa tela é reduzir barreiras de entrada e agilizar o processo de identificação do usuário, especialmente em contextos de uso associados ao espelho virtual no ambiente de varejo. O aplicativo é projetado para funcionamento em smartphones e tablets, garantindo compatibilidade com diferentes tamanhos de tela e reforçando a estratégia omnichannel do sistema, ao permitir a integração fluida entre o dispositivo móvel do usuário e o provedor virtual.

Figura 18 – Tela de login do aplicativo Cliente do sistema ARTificial

Fonte: Elaborado pelos autores.

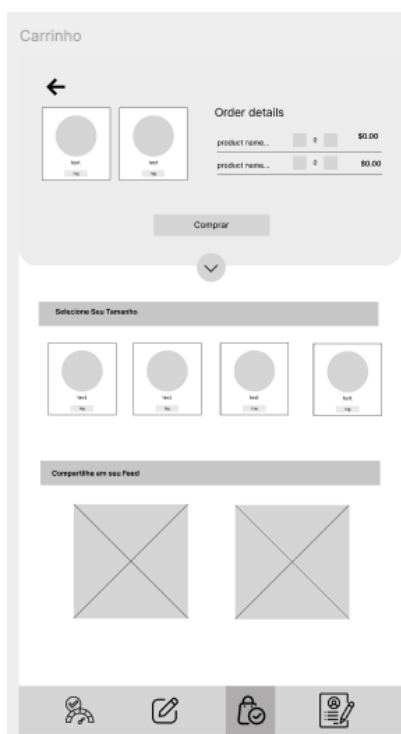
A Figura 19 apresenta a tela de login do espelho ARTificial realizada por meio do aplicativo Cliente. Nessa interface, o aplicativo aciona a câmera do dispositivo móvel do usuário para a leitura do QR Code exibido na tela do espelho virtual, permitindo a autenticação e a autorização da sessão. Esse mecanismo viabiliza a integração segura entre o dispositivo pessoal do usuário e o espelho em ambiente físico de varejo, dispensando a digitação de credenciais no próprio espelho. A abordagem reforça a estratégia omnichannel do sistema e contribui para a preservação da privacidade, ao concentrar o processo de autenticação no smartphone do usuário.

Figura 19 –Tela de login do espelho via leitura de QR Code



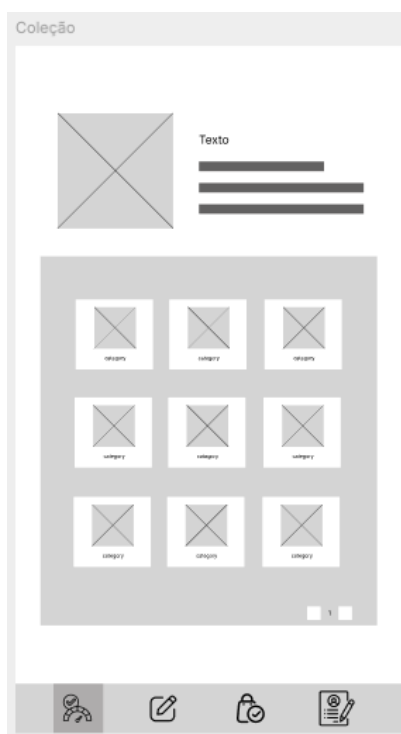
Fonte: Elaborado pelos autores.

A Figura 20 apresenta a tela de carrinho do aplicativo Cliente do sistema ARtificial. Nessa interface, são exibidas as peças previamente experimentadas pelo usuário por meio do espelho virtual, permitindo a visualização dos itens selecionados, a conferência de detalhes do pedido e a escolha do tamanho desejado. A partir dessa tela, o usuário pode realizar a reserva das peças diretamente pelo aplicativo, integrando a experiência digital de experimentação à etapa de decisão e intenção de compra. Adicionalmente, a interface oferece recursos de compartilhamento das combinações em seu feed, ampliando o engajamento e a dimensão social da experiência, em consonância com a proposta omnichannel do sistema.

Figura 20 – Tela de carrinho e reserva

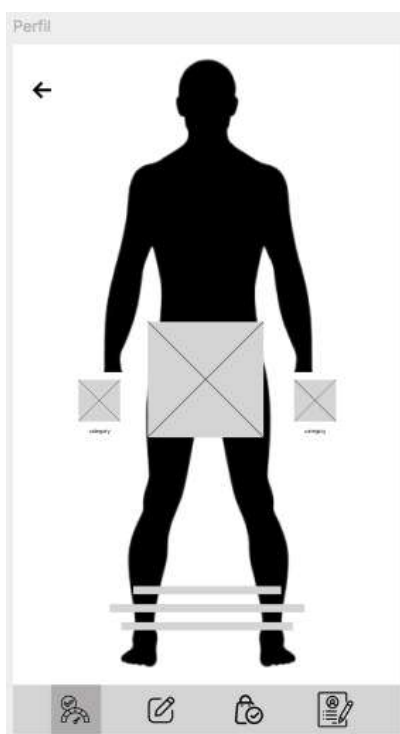
Fonte: Elaborado pelos autores.

A Figura 21 apresenta a tela de coleção de looks do aplicativo Cliente do sistema ARtificial. Nessa interface, o usuário visualiza uma galeria com todas as imagens capturadas durante as experimentações realizadas no espelho virtual, organizadas em formato de grade para facilitar a navegação. A tela permite a filtragem das imagens por estação, possibilitando ao usuário revisar seus looks de acordo com o contexto sazonal das peças experimentadas. Essa funcionalidade favorece a comparação visual entre diferentes combinações, apoia o processo de decisão e reforça a continuidade da experiência entre o ambiente físico do provador e o aplicativo móvel, alinhando-se à proposta omnichannel do sistema.

Figura 21 –Tela de coleção de looks

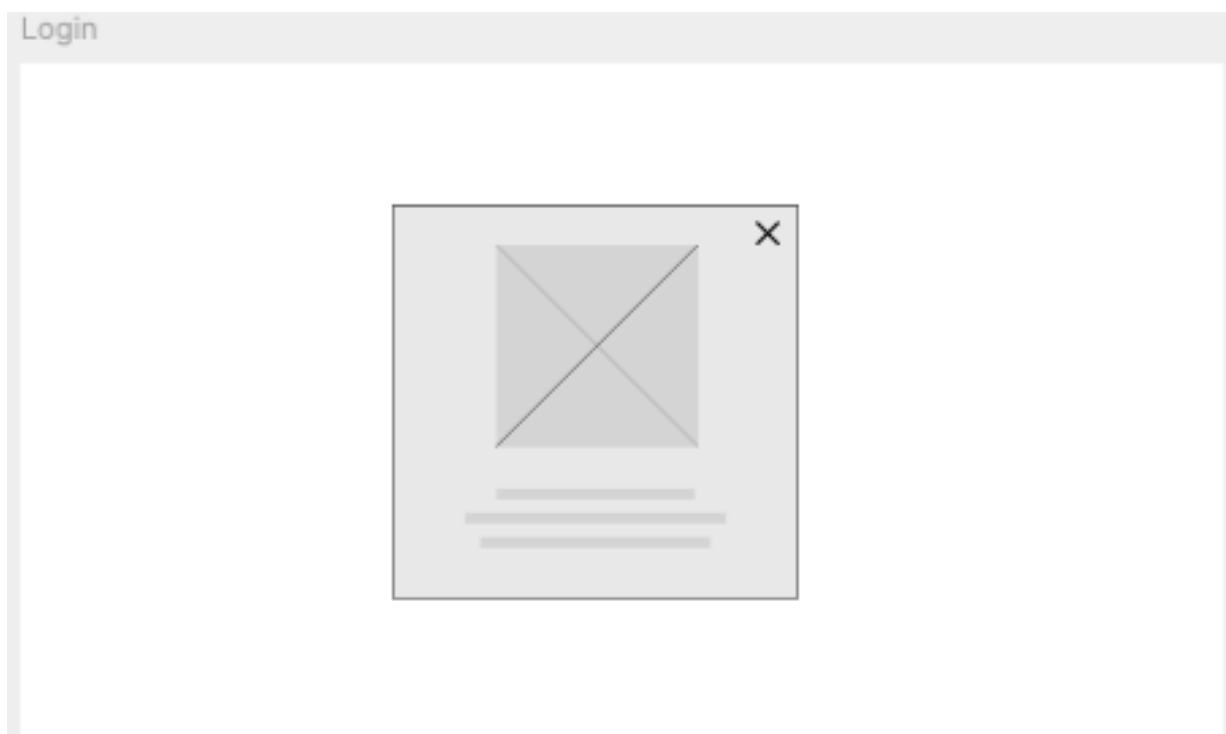
Fonte: Elaborado pelos autores.

A Figura 22 apresenta a tela de perfil do usuário no aplicativo Cliente do sistema ARtificial. Nessa interface, o usuário pode visualizar os dados corporais captados pelo espelho virtual, como medidas e informações associadas ao tamanho recomendado, além de consultar o histórico de acessos e interações realizadas no sistema. A tela também permite a navegação pelas peças favoritas e pelas peças reservadas, funcionando como um ponto central de gerenciamento da experiência do usuário. Essa organização reforça a continuidade entre a experimentação no espelho e o acompanhamento das escolhas no ambiente móvel, consolidando a proposta omnichannel do sistema.

Figura 22 –Tela de perfil do usuário no aplicativo Cliente

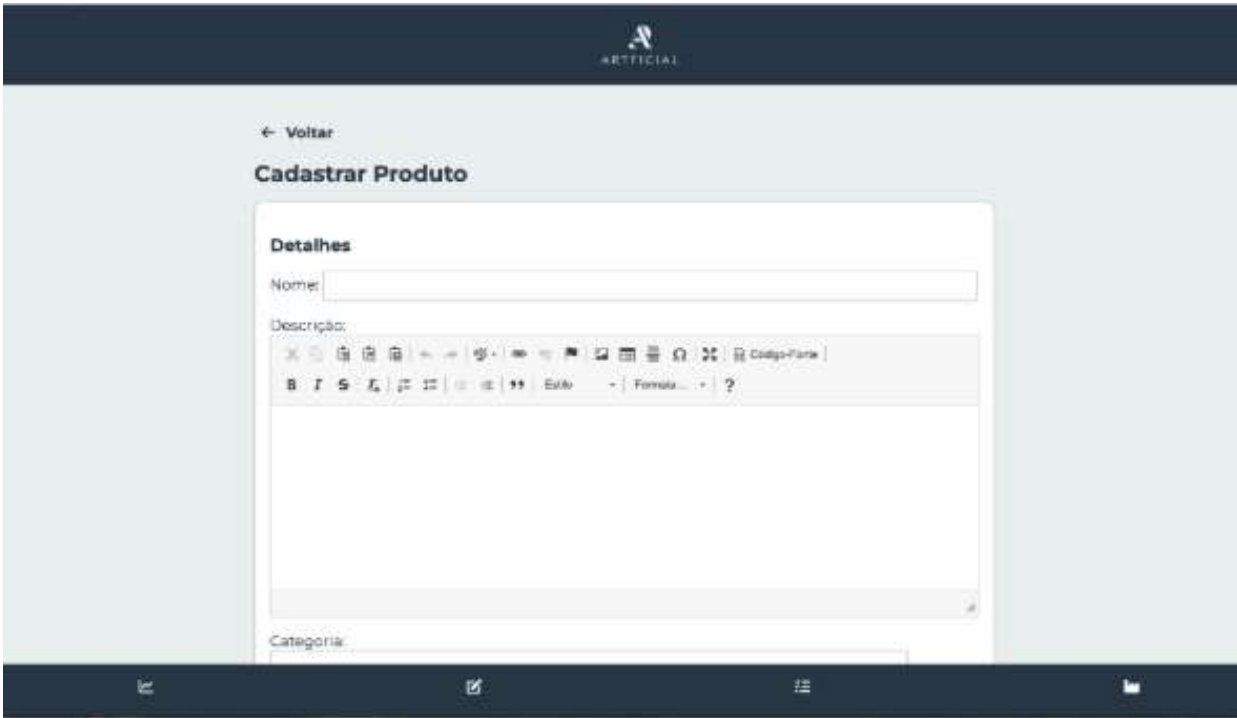
Fonte: Elaborado pelos autores.

A Figura 23 apresenta a tela de login do aplicativo Lojista do sistema ARtificial, uma plataforma dedicada à gestão operacional do provedor virtual no contexto do varejo de moda. Nessa interface, o acesso é restrito a usuários autorizados, permitindo a autenticação para utilização das funcionalidades de cadastro de produtos, controle de reservas realizadas pelos clientes e gerenciamento de caixa. A separação entre os aplicativos Cliente e Lojista reforça a arquitetura modular do sistema, garantindo maior segurança, organização das informações e adequação às diferentes necessidades dos atores envolvidos na operação do provedor virtual.

Figura 23 – Tela de login do Sistema lojista

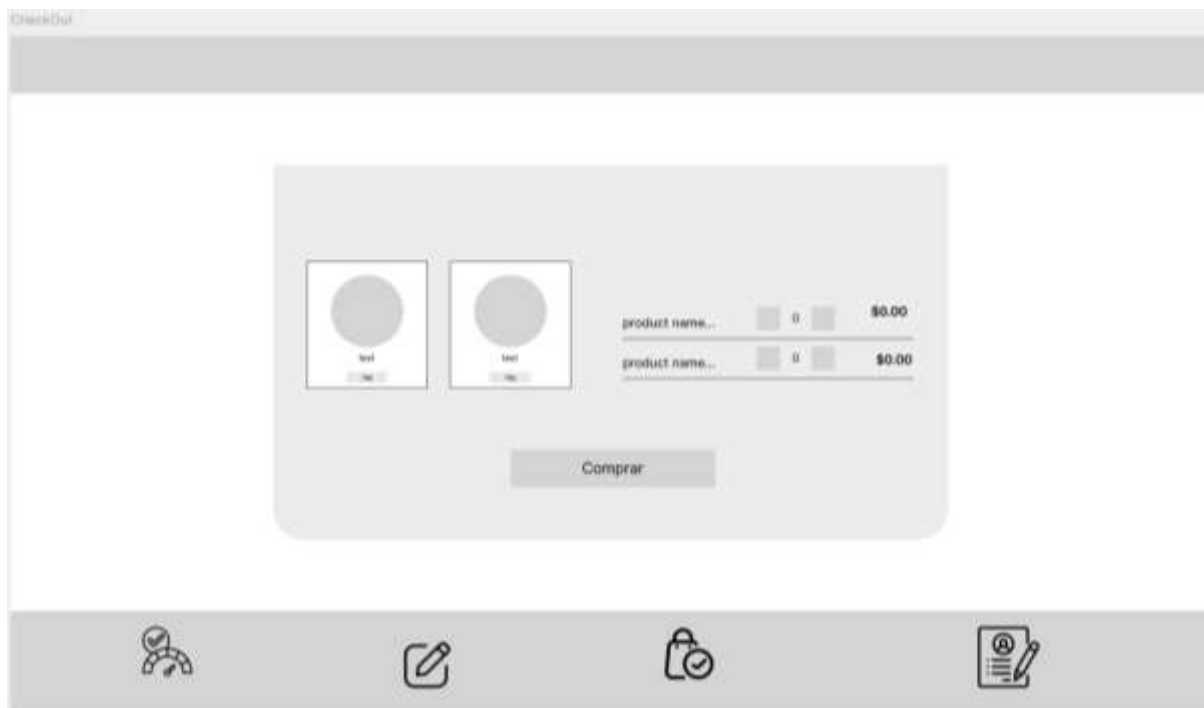
Fonte: Elaborado pelos autores.

A Figura 24 apresenta a tela de cadastro de produtos do aplicativo Lojista do sistema ARTificial. Nessa interface, o lojista pode inserir e gerenciar as informações necessárias para disponibilização das peças no provedor virtual, como nome do produto, descrição detalhada, categoria, além de demais atributos relevantes para a experimentação digital. O editor de texto integrado permite a formatação das descrições, facilitando a padronização das informações exibidas ao usuário final. Essa tela constitui um elemento central da operação do sistema, pois viabiliza a integração entre o catálogo físico da loja, o provedor virtual e o aplicativo Cliente, assegurando consistência informacional e eficiência no processo de reserva de produtos.

Figura 24 –Tela de cadastro de produtos

Fonte: Elaborado pelos autores.

A Figura 25 apresenta a tela de checkout do aplicativo Lojista do sistema ARTificial, utilizada para o registro da venda dos produtos na loja física. Nessa interface, são exibidos os itens selecionados pelo cliente, com a identificação das peças, controle de quantidade e visualização do valor correspondente. A partir dessa tela, o lojista pode finalizar a transação, integrando as reservas realizadas previamente pelo aplicativo Cliente ou pelo espelho virtual ao processo de venda presencial. Essa funcionalidade contribui para a centralização das informações de consumo, assegura maior controle operacional e reforça a integração entre o ambiente digital do provador virtual e o fluxo tradicional de caixa no varejo.

Figura 25 – Tela de checkout e registro de venda

Fonte: Elaborado pelos autores.

A Figura 26 apresenta a tela de perfil do Lojista no sistema ARtificial, integrando informações institucionais da loja e um painel de insights operacionais. À esquerda da interface, é exibido o perfil público da loja, contendo imagem de destaque, descrição institucional e acesso às páginas e categorias de produtos em evidência. À direita, é apresentado um dashboard de insights, no qual são consolidados dados relacionados às experimentações realizadas no espelho virtual, reservas efetuadas pelos clientes e compras finalizadas na loja física. Essa visualização analítica permite ao lojista acompanhar o desempenho das peças, identificar padrões de interesse do consumidor e apoiar a tomada de decisão estratégica, reforçando o caráter orientado a dados da proposta do sistema.

Figura 26 –Tela de perfil do Lojista e painel de insights

Fonte: Elaborado pelos autores.