

1 *Introdução*

A inteligência artificial (IA) tem se consolidado como uma das tecnologias mais impactantes da contemporaneidade, transformando significativamente as formas como as pessoas interagem com máquinas, processam informações e compreendem o mundo ao seu redor. Dentro da ampla área da ciência da computação, a visão computacional desponta como um campo estratégico, especialmente quando combinada a sensores de profundidade capazes de capturar não apenas imagens, mas também a estrutura tridimensional dos ambientes. Essa integração viabiliza a construção de representações espaciais ricas em informação geométrica, o que é essencial para aplicações como robótica autônoma, monitoramento ambiental, sistemas inteligentes de navegação, modelagem arquitetônica, realidade aumentada e interação em ambientes virtuais.

O uso de sensores de profundidade, como o Microsoft Kinect, o Intel RealSense e sensores LiDAR, representa um avanço crucial nesse cenário. Tais dispositivos oferecem dados complementares à visão tradicional baseada em câmeras RGB, fornecendo a distância entre o sensor e os objetos da cena. Com isso, tornam-se possíveis abordagens mais robustas de segmentação, reconstrução de superfícies, detecção de obstáculos e mapeamento tridimensional. No entanto, esses avanços técnicos devem ser acompanhados por uma preocupação crescente com a acessibilidade, a explicabilidade e a usabilidade dos sistemas desenvolvidos, de forma a garantir que essas soluções não apenas funcionem bem, mas também sejam compreendidas, confiáveis e utilizáveis por um público amplo e diverso.

Segundo Sichman (2021)(1), observa-se um novo ciclo de entusiasmo com as tecnologias de IA, impulsionado por três pilares: a redução dos custos computacionais, a enorme disponibilidade de dados e os avanços significativos em algoritmos, sobretudo no campo do aprendizado de máquina. A combinação desses fatores tem levado a um florescimento de aplicações baseadas em modelos como redes neurais profundas, que têm demonstrado alto desempenho em tarefas de reconhecimento, classificação e predição. No entanto, muitos desses modelos funcionam como caixas-pretas, oferecendo poucas pistas sobre os critérios

e justificativas de suas decisões. Essa falta de explicabilidade torna-se ainda mais crítica em sistemas que atuam no espaço físico e interagem com humanos, como robôs móveis ou sistemas de segurança autônomos.

Nesse ponto, a interação humano-computador (IHC) surge como uma área fundamental. De acordo com Carvalho (2003)(2), a dificuldade de operação das máquinas digitais, mesmo por usuários com bom nível educacional ou cultural, é um dos fatores mais relevantes na exclusão digital. A simples disponibilidade de equipamentos não garante o acesso, pois a interface entre o humano e a máquina muitas vezes é um obstáculo. Para ampliar o acesso às tecnologias digitais, é necessário que os sistemas sejam projetados com foco na experiência do usuário, adotando princípios como o desenho universal, que busca tornar as interfaces acessíveis a todos, independentemente de idade, formação ou condição física. Nesse sentido, a usabilidade, a acessibilidade e a transparência dos sistemas são tão importantes quanto seu desempenho técnico.

Considerando esse panorama, foi realizada uma revisão sistemática da literatura sobre visão computacional com sensores de profundidade, que revelou um número crescente de trabalhos voltados à reconstrução 3D, localização e mapeamento simultâneo (SLAM), detecção semântica de objetos e fusão de dados multimodais. Contudo, mesmo diante desses avanços, a análise da literatura permitiu identificar lacunas relevantes. Observou-se, por exemplo, a escassez de sistemas que integrem as capacidades de percepção 3D com interfaces acessíveis e intuitivas. Muitos estudos priorizam a precisão dos algoritmos, mas negligenciam a interação com o usuário final. Além disso, poucos trabalhos abordam a explicabilidade dos modelos, o que compromete a confiabilidade e a capacidade de auditoria dos sistemas em contextos críticos, como navegação autônoma, saúde, educação e segurança pública.

Outro ponto crítico é que muitos dos sistemas analisados na literatura dependem de arquiteturas de aprendizado profundo altamente especializadas e opacas, sem mecanismos que permitam ao usuário compreender por que determinada decisão foi tomada. Essa falta de transparência é especialmente problemática em ambientes onde as decisões computacionais têm implicações éticas, legais ou sociais. Como argumenta Virginia Dignum (2019)(3), é fundamental que o desenvolvimento de sistemas inteligentes seja guiado por princípios de responsabilidade, prestação de contas e transparência. Isso implica projetar sistemas que não apenas funcionem, mas que também sejam capazes de justificar suas ações, adaptando-se a contextos humanos e respeitando valores sociais.

Nesse contexto, esta pesquisa propõe-se a investigar soluções baseadas em visão com-

putacional com sensores de profundidade para o mapeamento e o processamento de ambientes tridimensionais, tendo como diferencial o foco na acessibilidade, explicabilidade e na integração centrada no usuário. Busca-se, assim, alinhar o desenvolvimento tecnológico à reflexão crítica sobre os impactos sociais e éticos da IA, incorporando os aprendizados da IHC, da ciência da informação e dos estudos sobre agentes autônomos. A proposta é contribuir com sistemas que não apenas percebam o ambiente em três dimensões, mas que também se comuniquem de forma clara com os usuários humanos, oferecendo interfaces compreensíveis, interpretações justificadas e interações seguras.

Diante desse cenário, define-se o seguinte problema de pesquisa: como desenvolver soluções baseadas em visão computacional com sensores de profundidade que permitam o mapeamento e o processamento de ambientes tridimensionais de forma acessível, explicável e centrada no usuário? Esta questão orienta o desenvolvimento do projeto, que buscará integrar os conhecimentos das áreas envolvidas para propor abordagens que sejam, ao mesmo tempo, tecnicamente eficazes e socialmente responsáveis.