

SEGMENTAÇÃO DE OBJETOS COM NUVEM DE PONTOS PARAMANIPULADORES DE ROBÔS DESERVIÇO AUTÔNOMOS

Cauan Vinicius Espinha de Sousa¹, Plínio Thomaz Aquino Junior¹
¹Departamento de Ciência da Computação, Centro Universitário FEI
unifcasousa@fei.edu.br e plinio.aquino@fei.edu.br

Resumo: A robótica tem se destacado como uma tecnologia promissora em diversos campos, como a produção industrial e a assistência médica. Uma tarefa essencial dos robôs é a manipulação de objetos em ambientes desconhecidos, viabilizada pela segmentação de objetos por nuvem de pontos, que reconstrói a geometria 3D dos objetos e os identifica através da detecção de bordas. Esta pesquisa explora essa técnica para que robôs autônomos possam reconhecer e manipular objetos de forma eficiente. Simulações e análises experimentais indicam que o método proposto, utilizando o MoveIt, é eficaz na segmentação e manipulação precisa de objetos.

1. Introdução

A segmentação de objetos utilizando nuvens de pontos [2] para manipuladores robóticos autônomos é um desafio significativo na robótica. Este processo envolve identificar regiões de interesse em uma imagem tridimensional, capturada por câmeras RGB-D, para que o robô possa manipular os objetos de maneira eficaz. A segmentação pode ser feita combinando técnicas de análise de cores, texturas e formas [3, 5], sendo fundamental em várias aplicações robóticas [7], como manipulação em fábricas e serviços de assistência.

Para desenvolver essa habilidade, o artigo sugere a criação de um *dataset* sintético em um ambiente simulado com o Omniverse, permitindo o treinamento de uma rede neural convolucional (CNN) que possa segmentar objetos em ambientes novos. Este treinamento requer ajustes cuidadosos nos parâmetros da rede e uma análise detalhada do comportamento do modelo e do *dataset*. No entanto, as limitações do hardware utilizado podem afetar a precisão da simulação, comprometendo os resultados.

2. Metodologia

O projeto teve início com a criação de um conjunto de dados sintético utilizando o Omniverse Replicator, uma ferramenta que permite a simulação de ambientes tridimensionais com alta fidelidade às condições do mundo real, incluindo detalhes precisos como texturas, iluminação, a física dos objetos e os metadados necessários. Para garantir a diversidade e a qualidade dos dados gerados, como demonstrado na Figura 3 – Fluxograma usando replicator, foi desenvolvido um código especial que randomiza aspectos importantes do ambiente, como a disposição dos objetos, a iluminação, e até a posição e o ângulo das câmeras de captura. Essa abordagem é essencial para produzir um *dataset* robusto [1], adequado para o treinamento de modelos de

inteligência artificial, especificamente redes neurais convolucionais (CNNs) [4].



Figura 1 - Saída da segmentação do replicator

Após a geração e anotação dos dados, como é mostrado na Figura 1 - Saída da segmentação do replicator e na Figura 2 – Saída da marcação 2D do replicator, a CNN foi treinada para detectar e segmentar objetos em tempo real com alta precisão. A segmentação utiliza técnicas de *deep learning* [6], que permitem a extração de características complexas diretamente da nuvem de pontos e das bordas dos objetos [8]. Com a segmentação realizada, o próximo passo foi o planejamento de movimento, onde foi calculada a melhor trajetória para o robô manipular os objetos de forma segura e eficiente, levando em consideração as características físicas do manipulador.



Figura 2 - Saída da marcação 2D do replicator

A metodologia foi rigorosamente avaliada por meio de análises experimentais e simulações computacionais, que consideraram diferentes tipos de objetos, ambientes e condições de iluminação. Esses testes revelaram áreas para aprimoramento, permitindo ajustes contínuos para melhorar a precisão da segmentação e aumentar a taxa de sucesso na manipulação dos objetos. Além disso, toda a pesquisa foi documentada de forma detalhada, com o

código compartilhado e controlado por versão no Github, garantindo a replicabilidade do estudo.

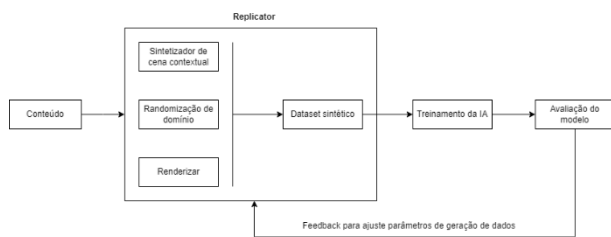


Figura 3-Fluxograma usando o replicator.

Finalmente, os resultados foram testados na plataforma robótica HERA da Equipe RoboFEI@Home, utilizando os equipamentos disponíveis nos laboratórios. Essa etapa final teve como objetivo validar a eficácia da metodologia em cenários práticos, demonstrando seu potencial para aplicações em diversas áreas, incluindo a aplicação em ambientes domésticos e saúde. O projeto, ao integrar esses elementos, contribui significativamente para o avanço da robótica de serviço autônoma, otimizando processos e aumentando a eficiência dos sistemas robóticos em ambientes complexos e variados.

3. Resultados

Após a criação inicial de um *dataset* sintético utilizando o Omniverse Replicator, o projeto chegou a um estágio estratégico de avaliação e aprimoramento. Embora os dados gerados já fossem de alta qualidade, foi identificada a necessidade de refinar ainda mais o *dataset* para garantir uma representação mais fiel do mundo real, especialmente em termos de diversidade e complexidade das condições ambientais. Para abordar essas limitações, aprimorou-se o código de randomização, que é essencial para gerar uma variedade maior de cenários, abrangendo variações em iluminação, texturas, e posições das câmeras, garantindo assim um treinamento mais robusto dos modelos de inteligência artificial.

Adicionalmente, houve um refinamento das texturas e propriedades físicas dos objetos, como refletância e materialidade, para aumentar a autenticidade visual e a precisão das simulações. Esses aprimoramentos foram continuamente testados para avaliar seu impacto no desempenho da rede neural convolucional (CNN) treinada com os dados gerados, com o objetivo de aumentar a precisão e a capacidade de generalização dos algoritmos de segmentação e detecção [9].

Apesar desses esforços, a precisão do modelo não alcançou o desempenho esperado, em parte devido à necessidade de ajustes contínuos nos parâmetros da rede neural, como taxa de aprendizado, número de camadas ocultas, e funções de ativação. A combinação de um *dataset* ainda em refinamento e a necessidade de ajustes constantes na rede neural mostrou que o processo iterativo de aprimoramento é essencial. Este ciclo contínuo de refinamento do *dataset* e ajuste dos parâmetros da rede neural é fundamental para superar as limitações atuais e alcançar um modelo eficaz para aplicação em cenários reais. A pesquisa avança com foco nesse processo, buscando otimizar a precisão e a

eficiência da CNN para aplicações práticas na robótica autônoma.

4. Conclusões

As conclusões do projeto mostram tanto o potencial quanto as limitações da metodologia proposta. A abordagem de segmentação em ambientes simulados demonstrou ser teoricamente robusta, sugerindo um caminho promissor para a identificação e manipulação autônoma de objetos por robôs. No entanto, o desempenho inferior em ambientes reais destaca a necessidade de refinamento, devido à maior complexidade desses cenários. As limitações de hardware emergiram como um grande obstáculo, afetando a capacidade de replicar a precisão e os detalhes obtidos nas simulações. Isso ressalta a importância de investir em hardware mais avançado, como GPUs compatíveis com RTX, e de explorar novos algoritmos de aprendizado de máquina para melhorar a robustez da segmentação em ambientes reais. A integração de sensores adicionais e a diversificação dos dados sensoriais também são recomendadas para aprimorar a precisão e a eficiência dos robôs em diferentes contextos.

5. Referências

- [1] ANTONIOU, Antreas; STORKEY, Amos; EDWARDS, Harrison. Data Augmentation Generative Adversarial Networks. arXiv preprint arXiv:1711.04340, 2017.
- [2] CHENG, Mingmei et al. SSPC-Net: Semi-supervised Semantic 3D Point Cloud Segmentation Network. ArXiv, abs/2104.07861, 2021.
- [3] CORKE, Peter. Robotics, Vision and Control: Fundamental Algorithms in MATLAB. [S.l.]: Springer Science & Business Media, 2011.
- [4] CRAIG, John J. Introduction to Robotics: Mechanics and Control. [S.l.]: Pearson Prentice Hall, 2005.
- [5] FAN, Linxi "Jim" et al. Eureka: A Research Breakthrough in Robot Learning with Generative AI and Omniverse. NVIDIA Blog, 2023. Disponível em: <https://blogs.nvidia.com/blog/2023/10/03/eureka-robot-learning/>.
- [6] GONZALEZ, Rafael C.; WOODS, Richard E. Digital Image Processing. [S.l.]: Pearson Education, 2008.
- [7] GOODFELLOW, Ian et al. Generative Adversarial Nets. Advances in Neural Information Processing Systems, p. 2672–2680, 2014.
- [8] GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. Deep Learning. [S.l.]: MIT Press, 2016.
- [9] HEYDEN, Anders et al. 3D Computer Vision: Efficient Methods and Applications. [S.l.]: Springer Science & Business Media, 2009.
- [10] KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. [S.l.: s.n.], 2012. p. 1097–1105.

¹ Aluno de IC do Centro Universitário FEI. Projeto com vigência de 06/2023 a 05/2024.