

CAPS-BQ: Uma Abordagem Compacta Baseada em Atenção para Segmentação de Partes em Objetos 3D

Calleo B. Barroso^{†*}, Hector L. M. Moreira^{††}, Francisco Hércules dos S. Silva^{‡*},
José Daniel de A. Santos^{†*} e Pedro P. Rebouças Filho^{§*}

^{*}Laboratório de Processamento de Imagens, Sinais e Computação Aplicada (LAPISCO)

[†]Programa de Pós-Graduação em Ciências da Computação (PPGCC)
Instituto Federal do Ceará (IFCE), Fortaleza, Ceará, Brasil

[‡]Instituto Federal do Ceará (IFCE)

Autor correspondente: pedrosarf@ifce.edu.br

Abstract—Technological progress has led to the development of high-precision sensors for three-dimensional scanning, such as LiDAR. Part segmentation of point clouds generated by these sensors is a fundamental task in data processing and analysis, yielding valuable insights for decision-making processes. This study introduces a lightweight, attention-based neural network architecture for point cloud part segmentation that employs a radius-based search algorithm for point selection. By leveraging a widely recognized benchmark dataset, we developed a model that is lightweight, robust, and computationally efficient. The proposed model achieves a competitive mIoU of 84.26% and shows an inference speed approximately 45.56% faster than its leading competitor on a test set of 2874 point clouds, highlighting its computational efficiency. We conducted a comparative analysis against the Local and Global APES models, the PartField network, and the consolidated PointNet, confirming the proposed architecture's efficacy in delivering high performance with an optimized inference time.

Resumo—O progresso tecnológico impulsiona o surgimento de sensores de alta precisão para representação de escaneamento tridimensional, como o LiDAR. A Segmentação de instâncias de nuvens de pontos, originárias de sensores desta natureza, são tarefas fundamentais no processamento e análise desses dados, trazendo informações valiosas e úteis para tomadas de decisão. Esse estudo propõe uma abordagem de rede neural leve, baseada em atenção, para segmentação de partes de nuvens de pontos, usando como critério seletor de pontos um algoritmo de busca baseada em raio. Utilizando um conjunto de dados amplamente difundido na literatura, essa pesquisa propõe o desenvolvimento de um modelo leve, robusto e computacionalmente eficiente. O modelo proposto alcançou um resultado competitivo de 84,26% de mIoU, inferindo aproximadamente 45,56% mais rápido em comparação com o seu melhor competidor sobre 2874 nuvens de pontos, demonstrando eficiência computacional. A análise comparativa foi realizada considerando os dois modelos da APES, Local e Global, a rede PartField e a consolidada PointNet, demonstrando a eficácia da arquitetura proposta em manter alto desempenho com um tempo de inferência otimizado.

I. INTRODUÇÃO

Com o avanço tecnológico, sensores de alta precisão para escaneamento tridimensional, como o LiDAR, passaram a ser amplamente utilizados. Esses dispositivos vêm promovendo uma transformação significativa na representação e compreensão do mundo físico em ambientes digitais, ao possibilitar a captura de nuvens de pontos [1], conjuntos densos de coordenadas geométricas no espaço tridimensional (3D), que

oferecem uma descrição detalhada e precisa das superfícies de objetos e cenas [2].

O processamento e a análise desses dados têm como uma de suas principais finalidades a segmentação de partes, etapa fundamental para identificar as distintas instâncias funcionais que compõem um objeto representado tridimensionalmente. Trata-se de uma tarefa desafiadora, que exige identificar padrões para revelar a estrutura intrínseca de cada objeto [3].

A segmentação de nuvens de pontos por meio de técnicas de Aprendizado Profundo tem sido objeto de estudo por diversos autores. Destaca-se, por exemplo, o trabalho pioneiro de Qi et al. [4], que propuseram a PointNet, a primeira rede capaz de processar diretamente nuvens de pontos, sem a necessidade de converter os dados para formatos mais estruturados, como grades volumétricas (do inglês, *voxels*). Essa abordagem alcançou uma média de 83,7% na métrica intersecção sobre união média (do inglês, *mean intersection over union* – mIoU), descrita na Seção II. Mais recentemente, Liu et al. (2025) apresentaram a PartField [5], uma abordagem direta para o aprendizado de características 3D baseadas em partes, capaz de capturar o conceito geral das partes e sua hierarquia sem depender de modelos predefinidos ou informações textuais. A PartField obteve 79,18% de mIoU quando realizou-se a segmentação no conjunto de dados PartObjaverse-Tiny [6].

Em 2023, Wu et al. propuseram o método *Attention-Based Point Cloud Edge Sampling* (APES) [7], que combina camadas de atenção com seleção direta de pontos baseada em estatísticas matemáticas. A APES realiza a subamostragem de nuvens de pontos de duas formas distintas: local e global. Em testes realizados com o dataset ShapeNetPart [8], a APES Local atingiu 85,6% de mIoU na tarefa de segmentação, enquanto a APES Global obteve 85,8% de mIoU, ambas comparáveis a modelos consolidados na literatura.

À luz do exposto, e com o objetivo de contribuir para o avanço das pesquisas na área, este trabalho propõe uma rede neural compacta para a segmentação de partes em objetos 3D representados por nuvens de pontos. O método desenvolvido alcançou resultados expressivos, obtendo 84,26% de mIoU com desvio padrão de 15,96%, além de apresentar um tempo de inferência aproximadamente 45,56% inferior em relação ao melhor competidor.

O restante do trabalho está dividido como se segue. A Seção II apresenta a metodologia adotada, destacando a arquitetura proposta. Ademais, a Seção III contém detalhes sobre o processo de treinamento utilizado no desenvolvimento e apresentando o conjunto de dados utilizado. Além disso, as análises quantitativas, considerando a métrica de mIoU e seu desvio padrão; qualitativas, comparando visualmente os resultados de cada modelo; e a métrica de tempo de inferência. O método proposto é comparado com a PointNet, PartField e APES. Por fim, a Seção IV discute as conclusões derivadas dos resultados obtidos e propõe sugestões para trabalhos futuros.

II. METODOLOGIA

A proposição deste método objetivou treinar e testar uma nova arquitetura de rede neural para segmentação de objetos representados tridimensionalmente em nuvens de pontos. As métricas utilizadas visam avaliar a capacidade de segmentar corretamente partes de objetos 3D bem como também avaliar se o método proposto performa em tempo adequado para aplicações de tempo real. As etapas e processos adotados para a elaboração deste artigo constam nas subseções a seguir.

A. Compact Attention-based Point Segmentation via Ball Query

A arquitetura da rede proposta neste estudo, intitulada *Compact Attention-based Point Segmentation via Ball Query* (CAPS-BQ), foi concebida a partir da adaptação da rede APES, a qual emprega um mecanismo de atenção para gerar mapas de correlação e, a partir deles, realizar a amostragem das nuvens de pontos [7].

Ademais, a APES é composta por camadas de incorporação, atenção, operações de subamostragem e sobreamostragem, bem como por camadas convolucionais e faz uso do método de k -vizinhos mais próximos (do inglês, *k-nearest neighbors* – k -NN) [9] para seleção dos pontos das bordas. A Figura 1 mostra uma representação da arquitetura completa da rede original.

Entretanto, as operações de subamostragem utilizadas para a extração dos mapas de correlação, seguidas da etapa de sobreamostragem, produzem uma distribuição de pontos significativamente distinta daquela presente na nuvem original. Como consequência, o problema de segmentação passa a se assemelhar a uma tarefa de reconstrução de nuvem de pontos, uma vez que os pontos gerados não correspondem diretamente à distribuição espacial original, especialmente nas regiões afastadas das bordas [7].

Diante desse cenário, aplicou-se a técnica de ablação para a remoção das camadas intermediárias da rede. Originalmente empregada na neurociência, a análise de ablação consiste em danificar ou remover, de forma controlada, partes de um sistema — como tecidos neurais — a fim de avaliar o impacto dessa modificação na execução de determinadas tarefas [10].

No contexto deste estudo, a ablação foi conduzida removendo-se as camadas intermediárias da APES, preservando apenas a camada de incorporação, que processa as coordenadas 3D de entrada, transformando-as em um vetor de características em um espaço de dimensão superior; uma única

camada de atenção responsável por extrair mapas de correlação normalizados; e as camadas convolucionais que expandem a profundidade das características dos pontos, combinando ao fim informações locais, globais e de categoria.

Adicionalmente, o mecanismo de seleção de pontos foi substituído por uma abordagem baseada em raio. Essas modificações resultaram em uma rede mais compacta, cujos componentes serão apresentados em detalhe adiante. A Figura 1 ilustra a arquitetura do método proposto, obtida após a aplicação dessas alterações.

B. Mecanismo de Atenção R2P

Esta subseção trata do mecanismo de atenção utilizado neste trabalho, denominado raio-para-ponto (do inglês, *radius-to-point* – R2P). Ainda tomando como base a implementação da APES, alterou-se o modo de seleção dos pontos para análise, do k -NN para uma abordagem baseada em raio.

Então, considere uma nuvem de pontos $S = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N\}$ com $|S| = N$. Para cada ponto $\mathbf{p}_i \in S$, define-se um bloco local $S_i \subseteq S$ como o conjunto dos k vizinhos de \mathbf{p}_i , incluindo o próprio ponto. Então, para medir a correlação local entre as características dos pontos, utiliza-se um mecanismo de atenção que serve como mapa de correlação normalizado. A medida de correlação local h_l entre o ponto central \mathbf{p}_i e seu vizinho $\mathbf{p}_{ij} \in S_i$ é dada por

$$h_l(\mathbf{p}_i, \mathbf{p}_{ij}) = Q(\mathbf{p}_i)^\top K(\mathbf{p}_{ij} - \mathbf{p}_i), \quad (1)$$

onde Q e K correspondem a camadas lineares aplicadas às características latentes do ponto central \mathbf{p}_i e à diferença das características entre o vizinho \mathbf{p}_{ij} e o ponto central \mathbf{p}_i , respectivamente. O produto interno \top ocorre no espaço latente das características.

Aplica-se uma normalização escalonada pela raiz quadrada da dimensão d da característica, conforme o mecanismo de atenção do Transformer, obtendo-se o mapa de correlação normalizado

$$\mathbf{m}_i^l = \text{softmax} \left(\frac{\{h_l(\mathbf{p}_i, \mathbf{p}_{ij})\}_{\mathbf{p}_{ij} \in S_i}}{\sqrt{d}} \right). \quad (2)$$

Para cada mapa de correlação normalizado \mathbf{m}_i^l , calcula-se o desvio padrão dos valores

$$\sigma_i = \text{std}(\mathbf{m}_i^l). \quad (3)$$

Por fim, os pontos de borda são selecionados como aqueles que apresentam os maiores valores de σ_i , pois indicam maior variabilidade local na correlação e, consequentemente, potencial presença de bordas na nuvem de pontos.

1) *Seleção de Pontos Baseada em Raio*: A seleção de pontos baseada em raio (do inglês, *Ball Query*) é uma técnica de consulta espacial utilizada para identificar vizinhos locais ao redor de um ponto de referência (centróide). O procedimento consiste em selecionar todos os pontos contidos em uma esfera de raio fixo r , centrada no ponto de referência, extraindo

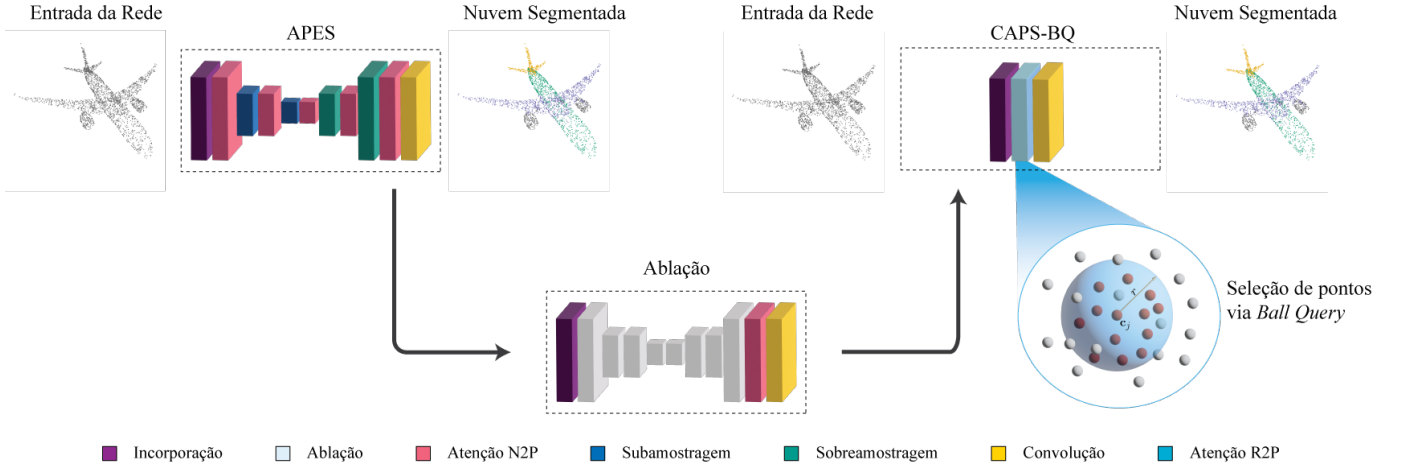


Figura 1. Arquitetura da CAPS-BQ após a ablação e alteração do método seletor de pontos do mecanismo de atenção.

regiões locais com diferentes distribuições de densidade e variações geométricas [11].

Seja uma nuvem de pontos tridimensional:

$$\mathcal{P} = \{\mathbf{p}_i \in \mathbb{R}^3 \mid i = 1, \dots, N\}, \quad (4)$$

sendo N o número total de pontos, e um conjunto de pontos de amostragem (centroides):

$$\mathcal{C} = \{\mathbf{c}_j \in \mathbb{R}^3 \mid j = 1, \dots, M\}, \quad M \ll N. \quad (5)$$

Para cada ponto de referência \mathbf{c}_j , o *Ball Query* retorna o conjunto de índices:

$$\mathcal{N}_j = \{i \mid \|\mathbf{p}_i - \mathbf{c}_j\|_2 \leq r\}, \quad (6)$$

onde $\|\cdot\|_2$ representa a distância Euclidiana e $r > 0$ é o raio pré-definido. Explicitamente, a condição de inclusão é dada por:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \leq r, \quad (7)$$

com $\mathbf{p}_i = (x_i, y_i, z_i)$ e $\mathbf{c}_j = (x_j, y_j, z_j)$.

Diferentemente do método k -NN, no qual o número de vizinhos é fixo, o *Ball Query* retorna uma quantidade variável de pontos, determinada pela densidade local da nuvem. Consequentemente, as características locais da região tornam-se mais generalizáveis no espaço, o que é preferível para tarefas que exigem reconhecimento de padrões locais, como rotulação semântica de pontos [11].

Na Figura 1 é possível visualizar a ilustração do procedimento, no qual \mathbf{c}_j representa ponto localizado no centroide de uma esfera de raio r utilizada como referência. Então, os pontos localizados dentro dessa esfera (em vermelho) são selecionados pelo método, enquanto aqueles situados fora dela (em cinza) são descartados.

III. RESULTADOS E DISCUSSÕES

Nesta seção será apresentada a métrica utilizada para avaliação dos modelos bem como o conjunto de dados considerado para tal e os parâmetros de treinamento do modelo proposto. Estão também apresentados e discutidos os resultados obtidos com o modelo proposto na Seção II, para a segmentação de partes em objetos 3D representados por de nuvens de pontos.

A. Base de Dados e Execução dos Experimentos

O conjunto de dados ShapeNetPart [8] é rotulado para segmentação de partes em objetos 3D. Também, é formado por 16.880 modelos, divididos em 16 categorias com o número de partes variando entre 2 e 6, totalizando 50 partes distintas.

Seguindo a abordagem de [7] para a execução dos experimentos, utilizou-se a versão subamostrada da ShapeNetPart produzida por [11], em que cada amostra é formada por 2.048 pontos. Ainda, dividiu-se o conjunto de dados em 14.006 amostras para treinamento e 2.874 para teste.

O treinamento e teste de cada um dos modelos foram realizados em ambiente local de execução, utilizando uma máquina com processador Intel Core i7-9750H de 9ª geração (2,60 GHz, 6 núcleos e 12 threads), 16GB de memória RAM DDR4 e uma GPU NVIDIA GeForce RTX 2060 com 6GB de memória dedicada.

B. Métrica de Avaliação

Para a avaliação quantitativa do desempenho do modelo na tarefa de segmentação de partes 3D, considerou-se como métrica a *Intersection-over-Union* (IoU), que calcula a razão entre a interseção e a união dos pontos preditos e dos pontos reais (*ground truth*) para uma única classe, métrica amplamente adotada na literatura de visão computacional e processamento de nuvens de pontos [12] [13]. Para uma classe específica, o IoU é definido pela seguinte Equação 8:

$$\text{IoU} = \frac{\text{Área de Sobreposição}}{\text{Área de União}} = \frac{|P \cap G|}{|P \cup G|}, \quad (8)$$

onde P representa o conjunto de pontos preditos como pertencentes à classe e G representa o conjunto de pontos do *ground truth* para a mesma classe.

Como a tarefa de segmentação de partes envolve múltiplas categorias, o IoU é calculado para cada classe de parte individualmente. O mIoU é, então, a média aritmética desses valores de IoU sobre todas as N classes de partes avaliadas, fornecendo uma medida agregada e balanceada da performance do modelo, como descrito na Equação 9:

$$\text{mIoU} = \frac{1}{N} \sum_{i=1}^N \text{IoU}_i. \quad (9)$$

A escolha do mIoU como métrica principal se deve à sua robustez, penalizando previsões incorretas e fornecendo uma avaliação mais fidedigna do que métricas baseadas puramente em acurácia.

C. Parâmetros de treinamento

A otimização do modelo foi realizada com o AdamW. A taxa de aprendizado foi inicializada em 1×10^{-4} e reduzida progressivamente até 1×10^{-8} ao longo do treinamento, seguindo uma estratégia de agendamento do tipo recozimento coseno (do inglês, *cosine annealing*).

Ainda, para a regularização dos pesos da rede, foi aplicado um decaimento de peso de 1×10^{-4} . Adicionalmente, uma regularização com probabilidade de 0,5 foi utilizado nas duas camadas finais totalmente conectadas. O treinamento foi conduzido por 200 épocas, com um tamanho de lote de 16.

A seguir, estão apresentados os resultados obtidos através dos experimentos realizados, bem como as comparações com outros métodos descritos na literatura.

D. Resultados Quantitativos

Nesta subseção, apresenta-se o desempenho das arquiteturas de redes neurais avaliadas em comparação com o método proposto. Conforme descrito na Seção II, a comparação baseia-se na métrica mIoU e em seu respectivo desvio padrão.

A Tabela I compara o método proposto com modelos amplamente consolidados na literatura. Ao analisar a CAPS-BQ em relação às variantes originais da APES, diferenciadas pelo método de subamostragem empregado, observa-se que a CAPS-BQ apresentou desempenho competitivo. Especificamente, o resultado de $84,26\% \pm 15,96\%$ obtido pelo método proposto superou ligeiramente a APES Local ($83,98\% \pm 16,55\%$), embora tenha permanecido 0,62% abaixo da APES Global.

Em contrapartida, quando comparada a um modelo mais recente, com abordagem distinta para a mesma problemática, como a PartField, que utiliza extração do campo de características combinada ao algoritmo *k-means* para nuvens de pontos, a CAPS-BQ apresentou desempenho superior, obtendo $84,26\% \pm 15,96\%$ contra $61,35\% \pm 23,68\%$ da PartField. De forma semelhante, frente a um modelo amplamente consolidado, como a PointNet, a CAPS-BQ também registrou métricas superiores, alcançando $84,26\% \pm 15,96\%$ em

Tabela I
COMPARAÇÃO QUANTITATIVA ENTRE A CAPS-BQ E MÉTODOS DO ESTADO DA ARTE.

Método	mIoU (%)	Desvio Padrão (%)
APES Local [7]	83,98	16,55
APES Global [7]	84,88	15,79
PartField [5]	61,35	23,68
PointNet [4]	83,05	14,11
CAPS-BQ (Proposto)	84,26	15,96

comparação a $83,05\% \pm 14,11\%$, o que representa um ganho absoluto de 1,21% nas inferências.

Esses resultados sustentam a hipótese de que as operações de ablação realizadas na arquitetura não comprometeram a capacidade do modelo em distinguir as partes dos objetos 3D, preservando sua efetividade na tarefa de segmentação. Além disso, essa simplificação reduziu substancialmente a profundidade e a complexidade computacional da rede, sem perda expressiva de desempenho, evidenciando que a extração de atributos discriminativos essenciais pôde ser mantida mesmo com uma estrutura mais enxuta, favorecendo cenários que demandam maior eficiência.

E. Análise Qualitativa

Embora as métricas quantitativas forneçam uma avaliação objetiva do desempenho dos modelos, a análise qualitativa é fundamental para identificar aspectos visuais que podem não ser totalmente capturados por essas métricas.

A Figura 2 apresenta uma comparação visual entre os resultados de segmentação obtidos pelo método proposto, seus métodos pares e o padrão-ouro. Ademais, a análise das amostras confirma as tendências observadas na avaliação quantitativa. Se comparada ao padrão-ouro, a CAPS-BQ demonstrou consistência na identificação do número correto de partes dos objetos, bem como na delimitação precisa de cada uma delas, mesmo em cenários mais complexos, com objetivos contendo cinco ou seis partes.

Em comparação aos demais métodos, observa-se elevada consistência visual na definição das fronteiras entre as diferentes partes dos objetos. Em determinados casos, a APES Local apresenta desempenho superior em regiões de transição com detalhes finos, desempenho este que não se mantém quando as transições são mais suaves. Além disso, o método PartField demonstrou limitações evidentes, apresentando dificuldades em separar adequadamente as partes dos objetos.

Novamente, essas observações reforçam a capacidade do método proposto de identificar corretamente a quantidade e os limites das partes, favorecendo a captura de vizinhos espacialmente relevantes e preservando a coerência estrutural, mantendo-se competitiva e demonstrando sua eficácia na tarefa de segmentação de partes em objetos 3D.

F. Comparativo dos Tempos de Inferência

Os resultados desta subseção correspondem ao comparativo do tempo de inferência de cada modelo. Esta métrica de tempo

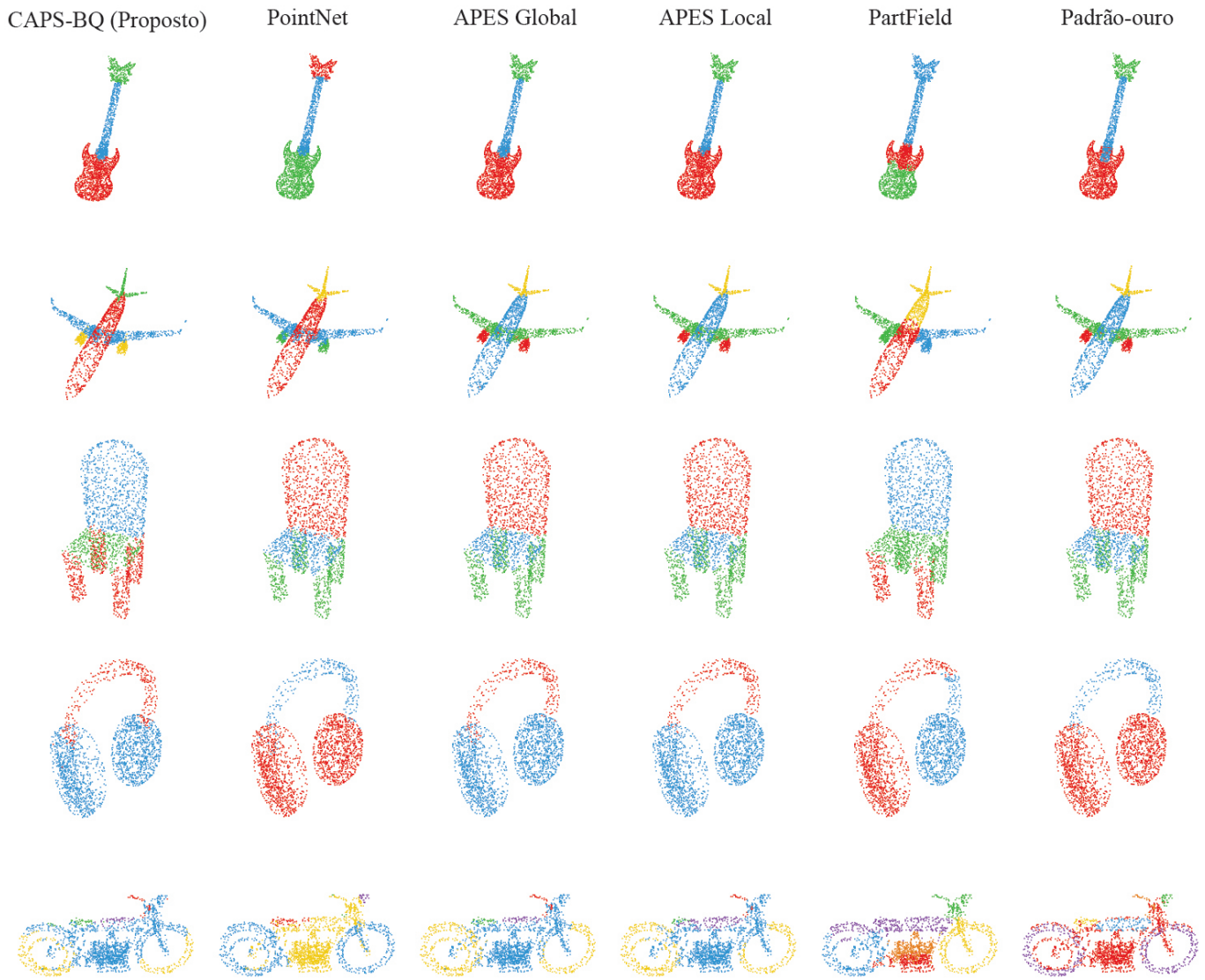


Figura 2. Comparação visual entre as inferências da CAPS-BQ e seus pares na tarefa de segmentação de partes em objetos 3D.

é considerada devido a sua importância prática: o tempo de inferência é crucial em aplicações que requerem desempenho em tempo real em ambientes críticos, como carros autônomos, por exemplo [14].

A Figura 3 evidencia o desempenho significativamente superior do método proposto em termos de tempo de inferência ao analisar o tempo total de inferência no conjunto de teste. Na segmentação das 2874 nuvens do conjunto de testes, descrito na Seção II, a CAPS-BQ completou a tarefa em 49 segundos no total. Isso se traduz em um tempo médio de segmentação de aproximadamente 17,05 milissegundos (ms) por nuvem, estabelecendo um novo padrão de desempenho entre os modelos avaliados.

Em contraste, as variantes APES Local e APES Global registraram um tempo médio de cerca de 32,71 ms por nuvem (94 segundos no total cada). De forma similar, a consolidada

PointNet apresentou um desempenho de aproximadamente 31,32 ms por inferência (90 segundos no total). Esses valores demonstram que a CAPS-BQ é quase duas vezes mais rápida que seus competidores diretos, realizando a mesma operação em uma fração do tempo.

No caso da PartField, a desvantagem em velocidade é ainda mais acentuada, com um tempo médio total de 52,54 ms por nuvem. Essa lentidão é explicada por sua metodologia em duas etapas: a extração de características, que consumiu em média 31,32 ms por nuvem (90 segundos no total), seguida pelo agrupamento com *k*-means, que adicionou outros 21,22 ms ao processo (61 segundos no total).

Portanto, a superioridade da CAPS-BQ em eficiência de inferência proporciona vantagens práticas relevantes, tais como menor latência em fluxos de processamento integrados; capacidade de operação em tempo real ou quase em tempo real,

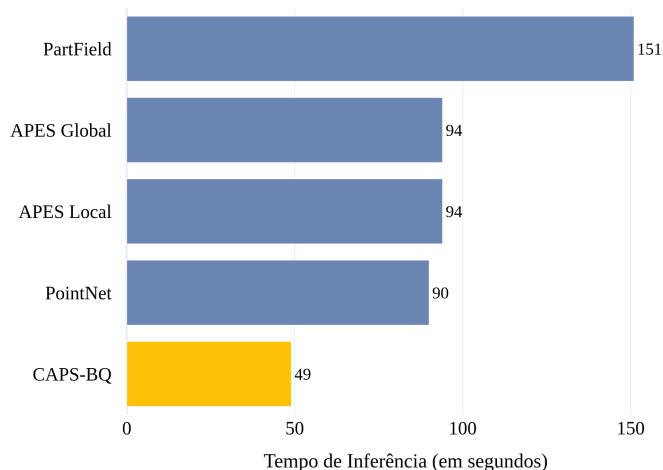


Figura 3. Comparativo dos tempos de inferência totais (em segundos) para a segmentação de 2874 nuvens de pontos.

essencial para aplicações interativas; e maior adequação para execução em dispositivos embarcados, ampliando o potencial de uso em cenários com restrições de recursos.

IV. CONCLUSÃO E TRABALHOS FUTUROS

Este estudo apresentou a CAPS-BQ, uma rede neural compacta desenvolvida para a segmentação de partes em objetos tridimensionais representados por nuvens de pontos. A arquitetura proposta utiliza um mecanismo de atenção especificamente projetado para capturar informações locais por meio da análise da vizinhança dos pontos, utilizando a seleção baseada em raio como critério de agrupamento.

O método proposto foi comparado a modelos consolidados na literatura por meio de três tipos de análise: quantitativa, qualitativa e tempo de inferência. Na avaliação quantitativa, a CAPS-BQ obteve um desempenho de $84,26\% \pm 15,96\%$ em mIoU, ficando apenas 0,62% abaixo da APES Global, que apresentou o melhor resultado entre os métodos avaliados.

Ademais, a análise qualitativa corroborou visualmente o potencial da CAPS-BQ na tarefa de segmentação de partes em objetos 3D, complementando as métricas numéricas. O método proposto demonstrou robustez tanto na identificação precisa do número de partes dos objetos quanto na delimitação de suas fronteiras, evidenciando que o processo de ablação e a seleção de pontos via *Ball Query* preservaram a elevada capacidade de segmentação da rede.

Outrossim, a comparação dos tempos de inferência dos modelos no conjunto de testes evidenciou a expressiva superioridade da CAPS-BQ em relação aos métodos concorrentes. Com um tempo total de predição de 49 segundos, o método proposto apresentou um desempenho aproximadamente 45,56% mais rápido que o melhor competidor, a PointNet (90 segundos) e de 47,87% se comparado às formulações Global e Local da APES (94 segundos cada).

Por fim, apesar dos resultados promissores, identificam-se oportunidades para aprimoramentos no método proposto.

Entre as possíveis direções para trabalhos futuros, destacam-se: a avaliação de seu desempenho em conjuntos de dados compostos por objetos reais; a investigação do impacto das operações de ablação sobre a complexidade computacional; e a exploração de diferentes estratégias de seleção de pontos. Essas investigações futuras poderão contribuir para validar de forma mais abrangente a versatilidade e a eficácia da abordagem proposta.

AGRADECIMENTOS

Os autores agradecem ao Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE) pelo apoio financeiro concedido por meio do EDITAL Nº 7/2024 PRPI/REITORIA-IFCE, que possibilitou a realização desta pesquisa.

REFERÊNCIAS

- [1] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu, and M. Bennamoun, "Deep learning for 3d point clouds: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 12, pp. 4338–4364, 2020.
- [2] H. Ran, J. Liu, and C. Wang, "Surface representation for point clouds," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 18 942–18 952.
- [3] R. Richter and J. Döllner, "Concepts and techniques for integration, analysis and visualization of massive 3d point clouds," *Computers, Environment and Urban Systems*, vol. 45, pp. 114–124, 2014.
- [4] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.
- [5] M. Liu, M. A. Uy, D. Xiang, H. Su, S. Fidler, N. Sharp, and J. Gao, "Partfield: Learning 3d feature fields for part segmentation and beyond," *arXiv preprint arXiv:2504.11451*, 2025.
- [6] Y. Yang, Y. Huang, Y.-C. Guo, L. Lu, X. Wu, E. Y. Lam, Y.-P. Cao, and X. Liu, "Sampart3d: Segment any part in 3d objects," *arXiv preprint arXiv:2411.07184*, 2024.
- [7] C. Wu, J. Zheng, J. Pfommer, and J. Beyerer, "Attention-based point cloud edge sampling," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 5333–5343.
- [8] L. Yi, V. G. Kim, D. Ceylan, I.-C. Shen, M. Yan, H. Su, C. Lu, Q. Huang, A. Sheffer, and L. Guibas, "A scalable active framework for region annotation in 3d shape collections," *ACM Transactions on Graphics (ToG)*, vol. 35, no. 6, pp. 1–12, 2016.
- [9] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.
- [10] R. Meyes, M. Lu, C. W. De Puiseau, and T. Meisen, "Ablation studies in artificial neural networks," *arXiv preprint arXiv:1901.08644*, 2019.
- [11] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," *Advances in neural information processing systems*, vol. 30, 2017.
- [12] J. Cen, Z. Zhou, J. Fang, W. Shen, L. Xie, D. Jiang, X. Zhang, Q. Tian *et al.*, "Segment anything in 3d with nerfs," *Advances in Neural Information Processing Systems*, vol. 36, pp. 25 971–25 990, 2023.
- [13] Y. He, H. Yu, X. Liu, Z. Yang, W. Sun, S. Anwar, and A. Mian, "Deep learning based 3d segmentation in computer vision: A survey," *Information Fusion*, vol. 115, p. 102722, 2025.
- [14] D. Parekh, N. Poddar, A. Rajpurkar, M. Chahal, N. Kumar, G. P. Joshi, and W. Cho, "A review on autonomous vehicles: Progress, methods and challenges," *Electronics*, vol. 11, no. 14, p. 2162, 2022.