

PPM-DeepLab: Módulo de Pirâmide de Pooling como Codificador da rede DeepLabV3+ para Segmentação de Rins, Cistos e Tumores Renais

Caio Eduardo Falcão Matos¹, Marcus Vinicius Silva Lima Oliveira¹,
João Otávio Bandeira Diniz², Arthur Guilherme Santos Fernandes¹,
Geraldo Braz Junior¹, Anselmo Cardoso de Paiva¹

¹Núcleo de Computação Aplicada – NCA, Universidade Federal do Maranhão (UFMA)
Caixa Postal 65.080-805 – São Luís – MA – Brasil

²Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA)
Caixa Postal 65075-441 – São Luís – MA – Brasil

{caioefalcao, arthurgsf, geraldo, paiva}@nca.ufma.br,
marcus.oliveira@discente.ufma.br, joao.bandeira@ifma.edu.br

Abstract. *Kidney cancer ranks among the leading causes of cancer-related deaths worldwide. Early detection and diagnosis are crucial in the fight against this disease. Recently, convolutional neural networks (CNNs) have demonstrated their effectiveness in semantic segmentation tasks. In this study, we introduce PPM-DeepLab, a novel architectural model designed for the segmentation of kidneys, cysts, and tumors in computed tomography (CT) images. Specifically, we explored the Pyramid Pooling Module (PPM) to enhance the DeepLabv3+ network by incorporating contextual information from various scales. Our proposed model achieved promising results, with Dice indices of 94.89% for kidneys, 83.95% for cysts, and 84.62% for renal tumors.*

Resumo. *O câncer renal está entre as principais causas de mortes relacionadas ao câncer em todo o mundo. A detecção e o diagnóstico precoces são cruciais na luta contra esta doença. Recentemente, redes neurais convolucionais (CNNs) demonstraram sua eficácia em tarefas de segmentação semântica. Neste estudo, apresentamos o PPM-DeepLab, um novo modelo arquitetônico projetado para a segmentação de rins, cistos e tumores em imagens de tomografia computadorizada (TC). Especificamente, exploramos o Pyramid Pooling Module (PPM) para aprimorar a rede DeepLabv3+ incorporando informações contextuais de várias escalas. Nosso modelo proposto alcançou resultados promissores, com índices de Dice de 94,89% para rins, 83,95% para cistos e 84,62% para tumores renais.*

1. Introdução

O câncer consiste em um termo genérico associado a um vasto grupo de patologias que podem afetar qualquer parte do corpo. Em geral, o câncer possui como característica a rápida proliferação de células anormais que podem invadir e comprometer órgãos e tecidos adjacentes [Smith et al. 2005]. Segundo o relatório da Organização Mundial de Saúde (OMS), o câncer é uma das principais causas de morte em todo o mundo, sendo responsável por quase 10 milhões de mortes em 2020 [WHO - World Health Organization 2023].

O câncer renal, também conhecido como carcinoma de células renais, está entre os 14 tipos de câncer mais comuns no mundo, com mais de 400.000 diagnósticos e mais de 170.000 mortes apenas em 2020 [WCRF - World Cancer Research Fund Internacional 2023]. Estima-se

que cerca de 81.800 novos casos de câncer renal (52.360 em homens e 29.440 em mulheres) serão diagnosticados e cerca de 14.890 pessoas (9.920 homens e 4.970 mulheres) morrerão desta em 2023 nos Estados Unidos. Esses números incluem todos os tipos de câncer de rim e pelve renal. Esta patologia, apresenta sua incidência com maior probabilidade em homens em relação as mulheres [ACS - American Cancer Society 2023b]. Em nível nacional, pesquisas relacionadas ao câncer renal na população brasileira evidenciam o aumento de óbitos. Segundo o estudo, entre os anos 2000 e 2019, o total de óbitos passou de 1.528 para 3.720 correspondendo a um aumento de aproximadamente 143% em pouco mais de 19 anos. Desta forma, este câncer possui taxa de mortalidade de aproximadamente 1,35 óbitos a cada 100 mil homens e mulheres, somente em 2019 [INCA - Instituto Nacional do Câncer 2022].

Desta forma, o diagnóstico precoce desta patologia torna-se ferramenta essencial para o prognóstico e tratamento da doença elevando assim, as chances de cura do paciente. Pacientes diagnosticados em estágio inicial da doença, possuem cerca de 93% de taxa de sobrevivência relativa em cinco anos. Por outro lado, quando o diagnóstico é efetuado em estágio avançado (metástase da doença) esta taxa é reduzida para 12%, demonstrando assim, a importância deste diagnóstico precoce [ACS - American Cancer Society 2023a].

Nos últimos anos, a medicina em geral tem feito grandes avanços na prevenção, detecção e tratamento. Exames de imagem, como a tomografia computadorizada (TC), representam uma alternativa não invasiva para obtenção de informações adicionais para auxiliar o médico. Dessa forma, os avanços nas metodologias de detecção e diagnóstico auxiliado por computador (CAD) têm contribuído para a detecção precoce do câncer renal, indicando áreas suspeitas e diagnosticando com precisão as anormalidades.

Abordagens baseadas em aprendizado profundo têm sido amplamente aplicadas com sucesso nessa tarefa, com destaque para as redes neurais convolucionais que demonstram excelentes resultados na tarefa de segmentação semântica em imagens médicas [Fonseca et al. 2022, Teixeira et al. 2022, Diniz et al. 2021, Baffa et al. 2021]. Destaca-se ainda, a modalidade de redes convolucionais baseadas em codificadores e decodificadores com convoluções dilatadas, quando aplicada para segmentação dos rins, cistos e tumores renais têm obtido ótimos resultados [Abdelrahman and Viriri 2022, Oliveira et al. 2022].

Portanto, a correta segmentação dos rins, cistos e tumores renais proporcionam ao médico especialista uma melhor análise e diagnóstico de lesões suspeitas, contribuindo para o planejamento do tratamento. Essa tarefa normalmente é realizada de forma manual, tornando o processo suscetível ao cansaço (físico e visual) e à distração. Desta forma, este trabalho propõe o PPM-DeepLab, uma nova arquitetura baseada no DeepLabv3+ usando blocos *Pyramid Pooling Module* (PPM). Especificamente, explorou-se o PPM para estender o DeepLabv3+ como composição de codificador para fornecer informações em contextos locais em diferentes resoluções.

De modo geral, pode-se destacar como contribuição do método proposto os seguintes pontos:

- Proposição de nova arquitetura de rede convolucional capaz de segmentar rins, cistos e tumores renais em imagens de TC com alta precisão;
- Adaptação e integração de modelos arquiteturais (PPM + DeepLab) não desenvolvidos na literatura;
- Um método robusto para segmentação com alta precisão dos rins, cistos e tumores renais. Desta forma, acredita-se que o método proposto fornece uma segmentação eficiente podendo ser integrado aos sistemas CAD, com o intuito de elevar a produtividade e contribuir para o diagnóstico do câncer renal.

2. Trabalhos Relacionados

A segmentação dos rins, cistos e tumores renais em imagens de TC visa auxiliar os especialistas no diagnóstico do câncer renal. Assim, a aplicação e desenvolvimento de técnicas voltadas ao processamento de imagens, bem como, a aprendizagem profunda são investigadas por diversos trabalhos. São apresentados a seguir trabalhos na literatura relacionados com a temática desta pesquisa.

Abordagens que aplicam sucessivas redes de segmentação semântica para a localização de objetos em imagens são denominadas como arquitetura em cascata ou *emsemble*. Os trabalhos [Zhao et al. 2022, Golts et al. 2022, George 2022] desenvolvem esta abordagem para a segmentação de rins, cistos e tumores renais. O método proposto no primeiro trabalho é estruturado duas etapas, segmentação geral e específica, que aplicam uma versão modificada da arquitetura U-Net, denominada nnUNet, para efetuar a segmentação semântica. Este trabalho foi validado na base de imagens KiTS21, atingindo como melhores resultados 97,52% para segmentação de rins e massas, 88,51% para massas renais e 86,93% para tumores renais.

Por outro lado, em [Golts et al. 2022], implementa uma cascata baseada nas redes *3D U-Net* aplicada em dois estágios, que buscam informações contextuais espaciais em baixa e alta resoluções, respectivamente. e o segundo é aplicado em dados de entrada de alta resolução pela segmentação dos rins e tumores renais. Esta abordagem foi aplicada no conjunto de dados KiTS21 e obteve como melhores resultados, 97,15%, 87,90% e 86,38% de coeficientes Dice para segmentação de rins e massas, massas renais e tumores renais, respectivamente.

A metodologia apresentada por [George 2022] também baseia-se no *emsemble* aplicando o modelo de rede neural com arquitetura de rede 3D U-Net. A rede é projetada para seguir uma abordagem em rede em cascata de dois estágios. A base KiTS21 foi usada para avaliação do método, que obteve como melhores resultados para segmentação de rins e massas, massas renais e tumores renais, respectivamente 97,60%, 87,60% e 83,10%.

Em [Yang et al. 2022] é proposto uma abordagem para a segmentação de rim, cistos e tumores renais em tomografias computadorizadas (TC) usando a técnica de transferência de aprendizado (*transfer learning*). Para tanto, utilizam a rede Residual 3D U-Net com pesos pré-treinados e módulo Atrous Spatial Pyramid Pooling associado ao decodificador. Como resultados, esta metodologia atingiu, 97,30%, 87,40% e 82,20% de coeficientes Dice para segmentação de rins e massas, massas renais e tumores renais, respectivamente.

Adaptações da arquitetura padrão da rede U-Net são amplamente exploradas para tarefa de segmentação em imagens médicas. Os trabalhos [Sun et al. 2022, Wu and Liu 2022] propõem variações desta rede utilizando mecanismos de atenção para a tarefa de segmentação de rins, cistos e tumores renais. O primeiro trabalho é baseado na variação da rede U-Net denominada DenseU-Net adaptada com o mecanismo de atenção, *Attention Gate*, como *skip connections*. Por outro lado, o segundo trabalho propõem a utilização de outro mecanismo denominado *Contrast Attention* para efetuar *skip connections* entre codificador e decodificador da rede.

Observando os trabalhos relacionados, destaca-se a utilização de variações na arquitetura da rede convolucional U-Net [Ronneberger et al. 2015] em todos os métodos propostos demonstrando assim, o potencial das redes neurais convolucionais baseadas em codificadores-decodificadores. Contudo, métodos baseados em redes convolucionais, como U-Net, trabalham em suas camadas com dados de entrada com diferentes tamanhos o que em arquiteturas com alta profundidade pode interferir na identificação. Desta forma, o método proposto neste trabalho, objetiva melhorar a capacidade da rede convolucional Deeplabv3+ em lidar com imagens

de diferentes tamanhos identificação informações contextuais em diferentes escalas.

3. Metodologia Proposta

Nesta seção, apresenta-se a metodologia proposta para segmentação dos rins, cistos e tumores renais em imagens de tomografia computadorizada (TC). A primeira etapa efetuou-se a aquisição das imagens através da utilização da base de imagens KiTS21 [Heller et al. 2021]. Em seguida, aplicou-se uma etapa de pré-processamento onde efetuou-se o redimensionamento e melhoramento das imagens. Na etapa seguinte, é proposto um novo modelo arquitetural, PPM-Deeplab, baseado na integração de blocos de Piramide de Pooling (Pyramid Pooling Module - PPM) como composição do codificador da rede convolucional Deeplabv3+. Em seguida, o modelo arquitetural desenvolvido é aplicado a base de imagem pré-processada e os resultados obtidos são avaliados na última etapa da metodologia. As quatro etapas desenvolvidas neste trabalho são descritas na Figura 1.

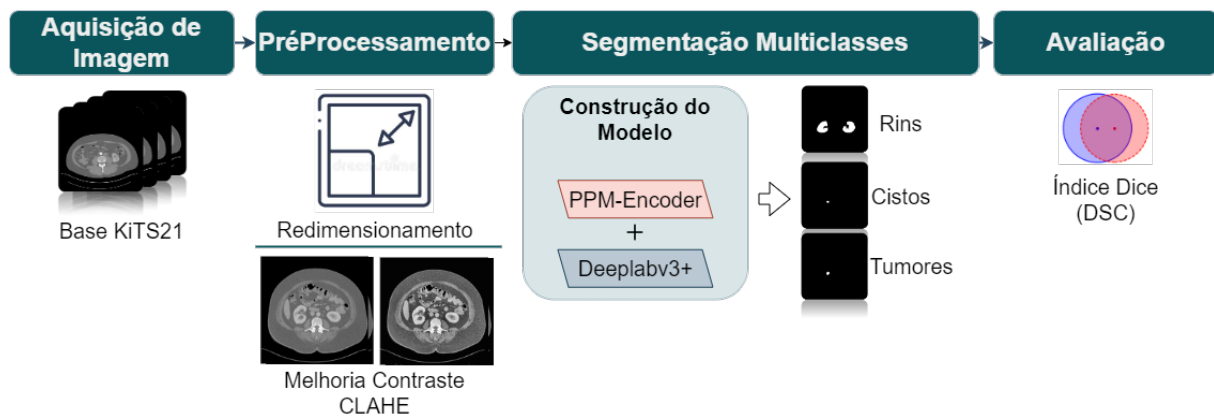


Figura 1. Metodologia Proposta

3.1. Base de Imagem

Neste trabalho, utilizou-se a base de imagens fornecida pelo desafio KiTS21 (*Kidney Tumor Segmentation 2021*) [Heller et al. 2021]. O objetivo do desafio consiste no avanço das pesquisas na área de segmentação de tumores renais, com a finalidade de melhorar o diagnóstico e tratamento de pacientes com câncer renal. A competição fornece uma grande base de dados de pacientes com tumores renais, incluindo anotações manuais das regiões dos rins, cistos e dos tumores em imagens de tomografia computadorizada. Este conjunto de dados, são provenientes de pacientes submetidos à nefrectomia por tumores renais no centro médico da Universidade de Minnesota (EUA) entre 2010 e 2021. Um total de 300 exames foram coletados neste levantamento, onde 210 foram disponibilizados para treinamento e validação e 90 para teste segundo o desafio. Além dos exames de imagem, anotações manuais das regiões dos rins e tumores é disponibilizada para cada paciente.

As imagens de TC e anotações (segmentação fornecida), disponibilizadas em formato NIFTI [INITIATIVE 2000] para cada volume são convertidas em fatias, gerando uma imagem em escala de cinza com resolução 512x512, sendo que a quantidade de fatias variam entre pacientes. Na Figura 2, pode-se observar um exemplo de uma fatia do volume, onde (a) representa uma fatia do exame de TC e (b) a marcação dos rins e tumores nas cores verde e vermelho, respectivamente [Heller et al. 2021].

3.2. Pré-Processamento

O objetivo desta etapa de pré-processamento consiste na redução da dimensionalidade das fatias e no aumento do contraste dos objetos de interesse (rins, cistos e tumores) em relação ao demais

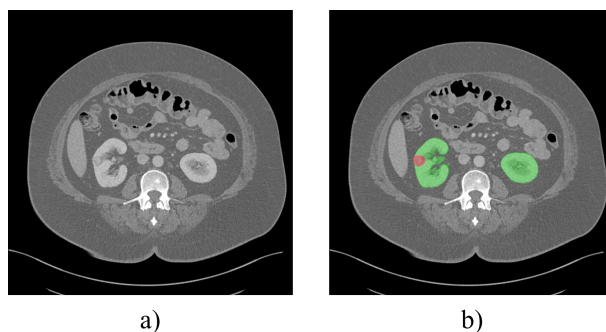


Figura 2. Exemplo da base KiTS21. (a) Fatia do exame de TC e (b) marcação dos rins, cistos e tumores disponibilizada. Fonte: [Heller et al. 2021]

órgãos presentes no exame. Com isso, é possível fornecer melhores características as demais etapas da metodologia.

Devido à limitações computacionais, foi necessário efetuar o redimensionamento das fatias (imagens) de todos os exames fornecidos da base. As imagens foram reduzidas de 512x512 *pixels* para 256x256 *pixels* tornando assim possível a aplicação do modelo arquitetural proposto neste trabalho.

Em seguida, cada fatia passou por realce de contraste usando CLAHE. A técnica de pré-processamento Equalização do Histograma Adaptativa Limitada ao Contraste (do inglês, Contrast Limited Adaptive Histogram Equalization - CLAHE) consiste em uma variante da equalização de histograma adaptável (AHE) que regula a amplificação excessiva do contraste. Esta técnica, opera em pequenas regiões da imagem, onde cada *pixel* da imagem original está no centro da região contextual. Como resultado deste processo, o histograma gerado pela nova imagem apresenta melhores contrastes [Pizer et al. 1987].

Nesta metodologia, o CLAHE foi aplicado com configuração do limite de contraste para 2.0 e tamanho da grade para equalização do histograma de 8x8. A Figura 3 ilustra a aplicação da técnica na base KiTS21.

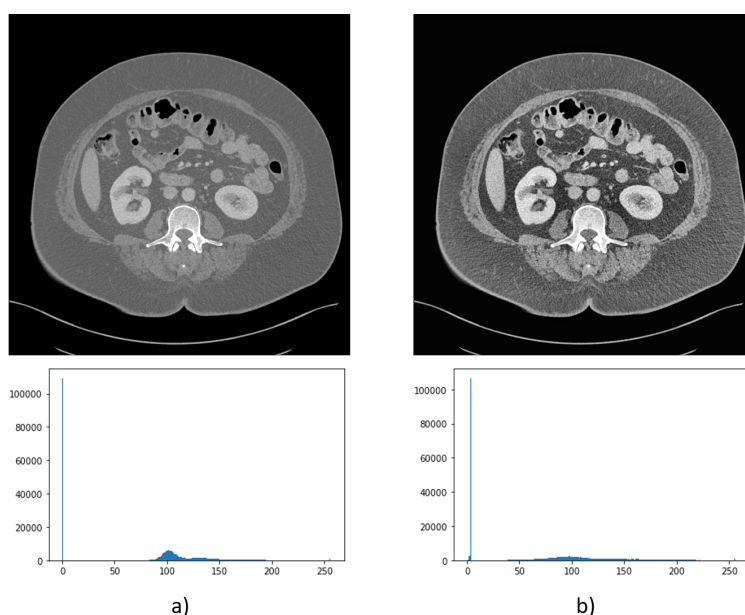


Figura 3. Resultado da aplicação do CLAHE com os respectivos histogramas. (a) imagem original e (b) imagem pré-processada.

3.3. Segmentação Multi-classe

Esta etapa visa o desenvolvimento de uma nova arquitetura de rede convolucional baseada em codificadores e decodificadores voltada para a segmentação de rins, cistos e tumores renais em imagem de TC. Para tanto, propomos um novo modelo arquitetural baseado na adaptação das redes convolucionais DeepLabv3+ [Chen et al. 2018] e *Pyramid Pooling Module* (PPM) [Zhao et al. 2017].

3.3.1. Módulo de Pirâmide de Pooling - PPM

O Módulo de Pirâmide de Pooling (PPM) é um módulo de processamento de imagem usado em algoritmos de segmentação semântica. O objetivo do módulo de pooling de pirâmide é extrair informações contextuais de diferentes escalas espaciais da imagem. Isso é importante porque diferentes objetos na imagem podem ter tamanhos diferentes e, portanto, requerem diferentes níveis de contexto para serem adequadamente segmentados [Zhao et al. 2017].

Neste trabalho, utilizou-se a arquitetura PPM apresentada na Figura 4 onde a partir do mapa de características de entrada é aplicado o módulo de pirâmide de *pooling* que divide a imagem em várias grades com diferentes resoluções (1x1, 3x3, 5x5 e 7x7). As operações de convolução, *Batch Normalization*, e ativação ReLU são aplicadas para cada resolução. Por fim, um *upsample* é utilizado para que todas as saídas da pirâmide de *pooling* sejam padronizadas e concatenadas juntamente com o mapa de características de entrada gerando assim o mapa de característica de saída. O resultado é uma representação contexto de várias escalas da imagem que contém observações localizadas, que são combinada para gerar uma saída final.

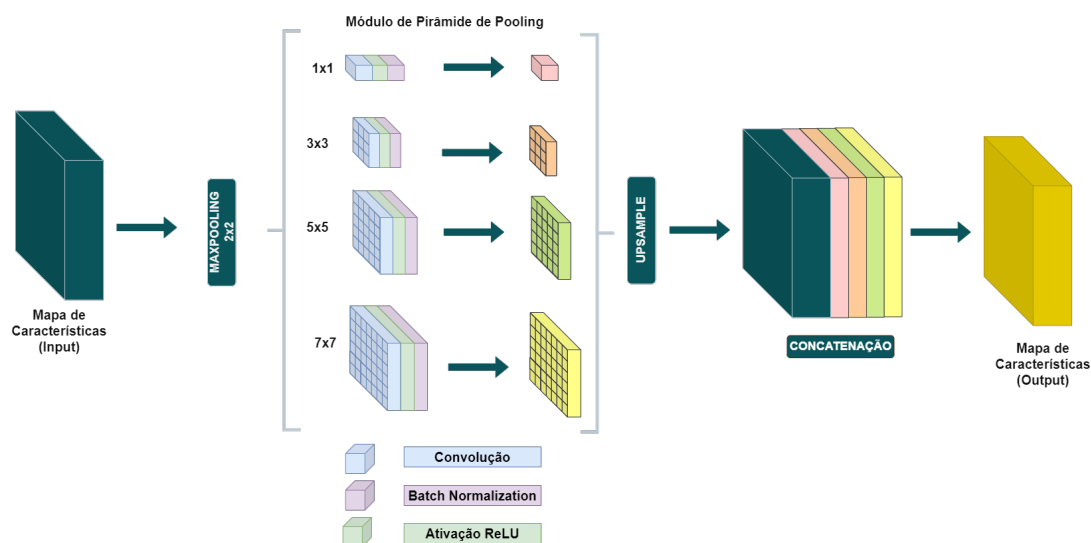


Figura 4. Arquitetura PPM

3.3.2. Arquitetura DeepLabv3+

O DeepLabv3+ é um modelo de rede neural convolucional voltado para segmentação semântica de imagens desenvolvido em 2018 e utilizado em diversas aplicações de visão computacional, como detecção de objetos, reconhecimento de cenas e segmentação de imagens. Este modelo de aprendizado profundo é baseado em uma arquitetura estruturada em codificador-decodificador para a tarefa de segmentação. O codificador objetiva a extração de características contextuais em

várias escalas, enquanto o decodificador é voltado para os limites dos objetos extraídos, recuperando gradativamente as informações espaciais para reconstrução da saída [Chen et al. 2018].

A arquitetura da DeepLabv3+ é baseada em redes ResNet [He et al. 2016] ou Xception [Chollet 2017] para extrair características de alto nível das imagens e uma técnica chamada "Atrous Spatial Pyramid Pooling" (ASPP) [Chen et al. 2017] para capturar informações de diferentes escalas espaciais e de contexto. Além disso, ele utiliza um decodificador de alta resolução que recria a resolução da imagem original para melhorar a precisão da segmentação.

3.3.3. Arquitetura Proposta - PPM-DeepLabv3+

Neste trabalho, foi proposto uma adaptação da arquitetura padrão DeepLabv3+, onde efetuou-se substituição da rede extratora de característica presente codificador padrão da DeepLabv3+, pela arquitetura de Pirâmides de Pooling (PPM). Para tanto, um total de quatro blocos de PPM foram utilizados como composição do PPM Encoder. Os blocos são aplicados sequencialmente, onde a saída de um bloco corresponde a entrada do bloco seguinte. As imagens de entrada da rede são submetidas como mapa de característica de entrada do primeiro bloco PPM. Esta adaptação, tem como objetivo melhorar a capacidade da rede DeepLab em obter informações contextuais em diferentes escalas, permitindo a análise sobretudo de pequenas regiões da imagem tornando a rede mais robusta e precisa tarefa de segmentação do câncer renal.

Após a execução do PPM Encoder, o mapa de característica final é encaminhado ao módulo ASPP. Este módulo é formado uma piramide de pooling que utiliza convoluções atrosas (Atrous Convolution) sendo necessário para a captura características e observação de objetos em múltiplas escalas.

Sendo voltado principalmente a identificação de objetos em contextos globais nas imagens [Chen et al. 2018]. Neste trabalho utilizou-se os parâmetros padrão definido pela rede DeepLabv3+, $rate = 6$, $rate = 12$ e $rate = 18$. A Figura 5 demonstra todas as etapas do modelo arquitetural proposto.

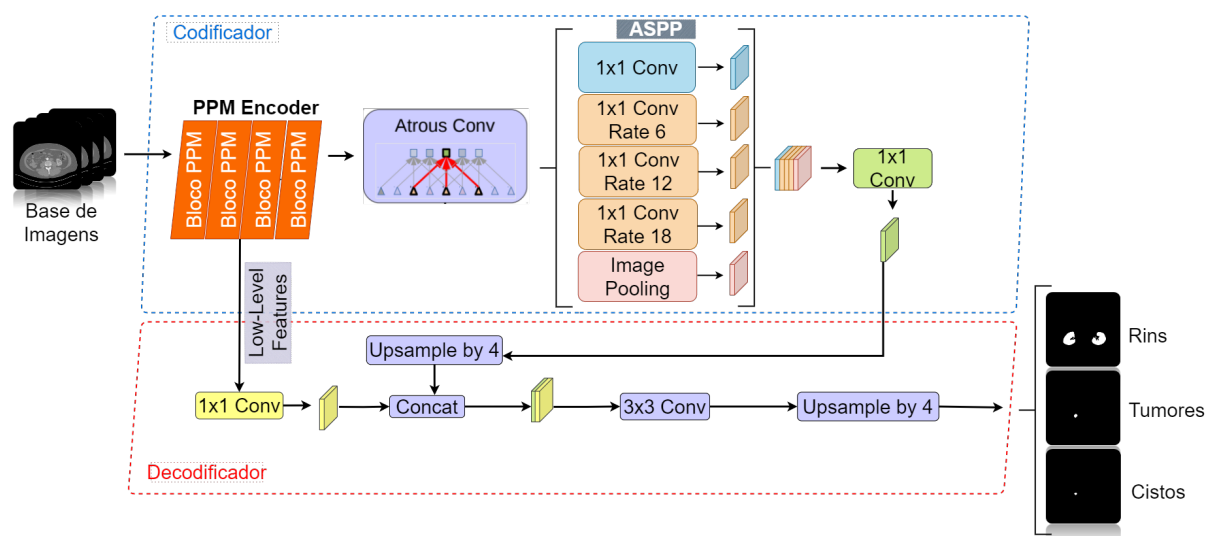


Figura 5. Arquitetura PPM-DeepLabv3+

4. Resultados e Discussão

Nesta seção, apresenta-se em detalhes os resultados dos experimentos realizados na avaliação da metodologia proposta. Inicialmente, demonstra-se as métricas aplicadas para validação de desempenho, em seguida uma discussão dos resultados obtidos. Além disso, realizou-se uma análise comparativa com trabalhos relacionados.

Para medir a precisão desta metodologia, utilizou-se métrica amplamente aplicada em sistemas CADx, e em diversos trabalhos presentes na literatura voltados para segmentação de imagens médicas. O coeficiente de similaridade Dice (DSC) [Fleiss et al. 1981] foi adotado como métrica de avaliação e comparação com os trabalhos relacionados.

Em nossos experimentos, dividiu-se os 300 exames disponibilizados pela desafio KiTS21 em três conjuntos de dados respeitando as proporções 70% (210 exames) para treinamento, 20% (60 exames) para teste e 10% (30 exames) validação. Este processo de separação dos conjuntos de dados foi desenvolvido por paciente (exame) evitando assim que fatias do mesmo paciente não sejam compartilhadas em diferentes conjuntos de dados. A escolha dos exames para formação dos conjuntos é realizada de forma aleatória.

A arquitetura PPM-DeepLabv3+, proposta neste trabalho, foi treinadas por 100 épocas, usando um tamanho de *batch* igual a 4, otimizador Adam com taxa de aprendizagem de 0,0001 e *DiceLoss + BinaryFocalLoss* como função de perda. A definição deste conjunto de parâmetros foi encontrada após vários experimentos, sendo este conjunto, produziu os melhores resultados.

A Tabela 1 demonstra que a arquitetura proposta obteve melhores resultados para a segmentação de rins, cistos e tumores renais quando comparados a aplicação do modelo padrão DeepLabv3+ na base KiTS21. Estes resultados, apontam que a inserção do PPM como codificador da rede DeepLabv3+ resultou em um modelo arquitetural codificador-decodificador robusto, com alto desempenho para segmentação do rins, cistos e tumores renais.

Tabela 1. Comparação da arquitetura proposta (PPM-Deeplabv3+) com modelo Deeplabv3+ padrão.

Métodos	DSC - Rins	DSC - Cisto	DSC - Tumor
DeepLabv3+	94,48%	64,44%	83,65%
PPM-DeepLabv3+	94,89%	83,95%	84,62%

Exemplos da segmentação dos rins, cistos e tumores renais efetuada pela PPM-DeepLabv3+ podem ser visualizados na Figura 6. Observa-se que, para as três fatias apresentadas o modelo proposto conseguiu segmentar de maneira eficiente os rins, cisto e tumor renal, quando comparado a marcação do especialista.

4.1. Comparação com Trabalhos Relacionados

Na Tabela 2, comparou-se o desempenho da metodologia proposta com os trabalhos que ocupam o estado da arte na base de dados KiTS21. Segundo o desafio, os resultados obtidos para as classes de rins, cistos e tumores devem ser combinadas para fins de computação de uma métrica para o superconjunto. Desta forma, as macroclasses são definidas como, Rins e Massas (Rim + Tumor + Cisto), Massa Renal (Tumor + Cisto) e Tumor.

Observando de forma comparativa, o método proposto neste trabalho obteve resultados comparáveis em relação ao estado da arte para segmentação de rins, cistos e tumores renais. Destaca-se ainda, que obteve-se excelente resultado para segmentação de tumores renais atingindo 84.62% correspondendo a uma diferença de apenas 2.31% em relação ao método proposto

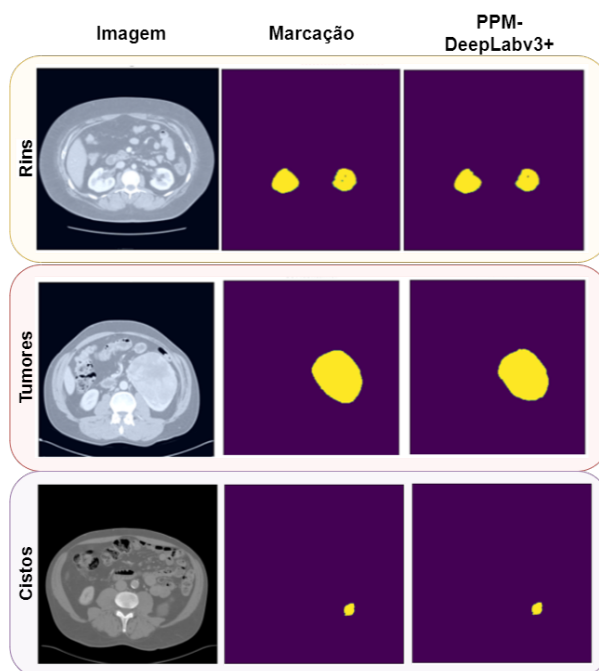


Figura 6. Exemplo de segmentação de rins, cisto e tumor renal utilizando a PPM-DeepLabv3+

por [Zhao et al. 2022] vencedor do desafio. Estes resultados, demonstram a potencialidade da utilização do PPM em conjunto a rede DeepLabv3+ que possibilitam a construção de um modelo arquitetural robusto que através da aplicação do conjunto de pirâmides de *pooling* (PPM e ASPP) realiza a análise em várias resoluções obtendo assim, características tanto locais quanto globais voltado para o contexto de segmentação do câncer renal.

Tabela 2. Comparação dos resultados da metodologia proposta com os trabalhos relacionados

Trabalho	Método	Rins e Massas	Massas Renais	Tumor	DSC-Médio
[Zhao et al. 2022]	Cascaded nnU-Net	97,52%	88,51%	86,93%	90,99%
[Golts et al. 2022]	Ensemble of 3D U-Net	97,15%	87,90%	86,38%	90,48%
[George 2022]	Cascaded 3D U-Net	97,60%	87,60%	83,10%	89,43%
[Yang et al. 2022]	Transfer Learning U-Net	97,30%	87,40%	82,20%	88,97%
[Sun et al. 2022]	Attention gate DenseU-Net	97,10%	81,20%	81,50%	88,70%
[Wu and Liu 2022]	CAU-UNet	97,00%	86,30%	81,10%	88,10%
Metódo Proposto	PPM-DeepLabv3+	90,81%	84,29%	84,62%	86,57%

5. Conclusão

A construção de metodologias e arquiteturas voltadas a segmentação semântica de órgãos e lesões tumorais em exames de tomografia computadorizada apresenta-se como uma tarefa desafiadora devido a heterogeneidade das estruturas presentes na região do abdominal.

A metodologia proposta neste trabalho buscou desenvolver e avaliar um novo modelo arquitetural baseado em redes convolucionais. Para tanto, efetuou-se integração da arquitetura PPM como codificador e extrator de características da rede DeepLabv3+ para segmentação de rins, cisto e tumores renais em imagens de TC.

O método proposto alcançou índice Dice de 90.81% para Rins e Massas, 84.29% para Massas renais e 84.62% para Tumores Renais, alcançando um lugar de destaque entre os melhores trabalhos encontrados em a literatura. Além disso, a utilização do PPM integrado ao

codificador da rede DeepLab surge como área de pesquisa que possibilita a construção de novos modelos baseados no PPM para a segmentação em imagens médicas. Como trabalhos futuros, recomendamos a investigação de técnicas de otimização de parâmetros para determinar a quantidade de blocos PPM a serem utilizados, bem como, a profundidade da pirâmide de *pooling* e as resoluções a serem utilizadas. De maneira geral, os resultados obtidos apresentam evidências sólidas de que o método proposto pode ser incorporado a um sistema CAD, a fim de fornecer suporte adicional aos médicos especialistas no diagnóstico do câncer renal.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, Fundação de Amparo a Pesquisa do Maranhão (FAPEMA), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Empresa Brasileira de Serviços Hospitalares (Ebserh) Brazil (Proc. 409593/2021-4).

Referências

- Abdelrahman, A. and Viriri, S. (2022). Kidney tumor semantic segmentation using deep learning: A survey of state-of-the-art. *Journal of Imaging*, 8(3):55.
- ACS - American Cancer Society (2023a). Cancer staging. Disponível em <https://www.cancer.org/treatment/understanding-your-diagnosis/staging.html>.
- ACS - American Cancer Society (2023b). Key statistics. Disponível em <https://www.cancer.org/cancer/kidney-cancer/about/key-statistics.html>.
- Baffa, M., Coelho, A., and Conci, A. (2021). Segmentação de imagens infravermelhas para detecção do câncer de mama utilizando u-net cnn. In *Anais do XXI Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 119–128, Porto Alegre, RS, Brasil. SBC.
- Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848.
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 801–818.
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258.
- Diniz, J., Ferreira, J., Silva, G., Quintanilha, D., Silva, A., and Paiva, A. (2021). Segmentação de coração em tomografias computadorizadas utilizando atlas probabilístico e redes neurais convolucionais. In *Anais do XXI Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 83–94, Porto Alegre, RS, Brasil. SBC.
- Fleiss, J. L., Levin, B., Paik, M. C., et al. (1981). The measurement of interrater agreement. *Statistical methods for rates and proportions*, 2(212-236):22–23.
- Fonseca, P., Araújo, A., Almeida, J., and Júnior, G. B. (2022). Treinando rede neural profunda com divisão proporcional de imagens para segmentação de estruturas da retina. In *Anais do XXII Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 1–12, Porto Alegre, RS, Brasil. SBC.

- George, Y. (2022). A coarse-to-fine 3d u-net network for semantic segmentation of kidney ct scans. In *Kidney and Kidney Tumor Segmentation: MICCAI 2021 Challenge, KiTS 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings*, pages 137–142. Springer.
- Golts, A., Khapun, D., Shats, D., Shoshan, Y., and Gilboa-Solomon, F. (2022). An ensemble of 3d u-net based models for segmentation of kidney and masses in ct scans. In *Kidney and Kidney Tumor Segmentation: MICCAI 2021 Challenge, KiTS 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings*, pages 103–115. Springer.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- Heller, N., Isensee, F., Maier-Hein, K. H., Hou, X., Xie, C., Li, F., Nan, Y., Mu, G., Lin, Z., Han, M., et al. (2021). The state of the art in kidney and kidney tumor segmentation in contrast-enhanced ct imaging: Results of the kits19 challenge. *Medical image analysis*, 67:101821.
- INCA - Instituto Nacional do Câncer (2022). Ministério da saúde amplia tratamento para câncer renal. Disponível em <https://www.inca.gov.br/noticias/ministerio-da-saude-amplia-tratamento-para-cancer-renal>.
- INITIATIVE, N. I. T. N. (2000). Neuroimaging informatics technology initiative. Disponível em: <https://nifti.nih.gov/>. Acesso em: 1 de dez. de 2022.
- Oliveira, M., Matos, C., Júnior, G. B., Paiva, A., Almeida, J., Costa, G., Levy, M., and Freitas, M. (2022). Ppm-unet: Uma rede neural convolucional para a segmentação de rins em imagens de tc. In *Anais do XXII Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 268–276, Porto Alegre, RS, Brasil. SBC.
- Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J. B., and Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3):355–368.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241, Munich, Germany. Springer.
- Smith, R. A., Cokkinides, V., and Eyre, H. J. (2005). American cancer society guidelines for the early detection of cancer, 2005. *CA: a cancer journal for clinicians*, 55(1):31–44.
- Sun, P., Mo, Z., Hu, F., Song, X., Mo, T., Yu, B., Zhang, Y., and Chen, Z. (2022). Segmentation of kidney mass using agdenseu-net 2.5 d model. *Computers in Biology and Medicine*, 150:106223.
- Teixeira, J., Rezende, M., Diniz, D., Carneiro, C., Luz, E., Souza, M., Ushizima, D., Medeiros, F., and Bianchi, A. (2022). Segmentação automática de núcleos cervicais em imagens de papanicolaou. In *Anais do XXII Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 346–357, Porto Alegre, RS, Brasil. SBC.
- WCRF - World Cancer Research Fund Internacional (2023). Kidney cancer statistics. Disponível em <https://www.wcrf.org/cancer-trends/kidney-cancer-statistics/>.
- WHO - World Health Organization (2023). Câncer - visão geral. Disponível em <https://www.who.int/news-room/fact-sheets/detail/cancer>.

- Wu, M. and Liu, Z. (2022). Less is more: Contrast attention assisted u-net for kidney, tumor and cyst segmentations. In *Kidney and Kidney Tumor Segmentation: MICCAI 2021 Challenge, KiTS 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings*, pages 46–52. Springer.
- Yang, X., Zhang, J., Zhang, J., and Xia, Y. (2022). Transfer learning for kits21 challenge. In *Kidney and Kidney Tumor Segmentation: MICCAI 2021 Challenge, KiTS 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings*, pages 158–163. Springer.
- Zhao, H., Shi, J., Qi, X., Wang, X., and Jia, J. (2017). Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2881–2890, Honolulu, HI, USA. IEEE.
- Zhao, Z., Chen, H., and Wang, L. (2022). A coarse-to-fine framework for the 2021 kidney and kidney tumor segmentation challenge. In *Kidney and Kidney Tumor Segmentation: MICCAI 2021 Challenge, KiTS 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings*, pages 53–58. Springer.