

Segmentação de imagens de propagação de chamas durante a combustão em um motor

Eduardo Moraes Coraça

7 de julho de 2021

Resumo

A otimização de motores a combustão interna é um assunto de grande relevância, principalmente no contexto de geração de energia. Ao se propor novos combustíveis é importante a realização de uma análise morfológica da chama de combustão, o que requer processamento de fotos da chama em propagação. Porém, as técnicas normalmente utilizadas são manuais, implicando em alto tempo de processamento. Neste relatório, é proposto o uso de redes profundas para a segmentação semântica das imagens. As redes *U-Net*, *HRNet* e *DeepLab v3* e o efeito da resolução da imagem de entrada são analisados. Conclui-se que o uso de modelos de aprendizado profundo trouxeram resultados superiores a um método tradicional, reduzindo o tempo de processamento e permitindo uma maior quantidade de imagens processadas, o que representa um grande potencial de aprimoramento da confiabilidade dos resultados de pesquisas na área.

1 Introdução

Para se atender as metas de emissões de gases do efeito estufa, a difusão de veículos com motorização elétrica vem sendo incentivada. Entretanto, o impacto na geração de eletricidade será grande, podendo levar a aumentos na emissão indireta de gases. Neste contexto, a otimização de motores a combustão para a geração de energia é de grande relevância. Como exemplo de pesquisa nesse contexto cita-se o uso de novos combustíveis e novas estratégias de operação para melhorar a combustão. A análise óptica, através de câmaras de combustão adaptadas de forma que se possa utilizar câmeras de alta resolução, possibilita a extração de características morfológicas da chama que não são possíveis de se obter via técnicas exclusivamente termodinâmicas. Para tanto, filma-se o processo de combustão e utilizam-se algoritmos de processamento de imagem para a detecção da região da chama a cada instante. Normalmente realiza-se um pré-processamento na imagem e se aplica um método de *threshold* para a distinção entre chama e não-chama. Muitas vezes, o uso de *thresholds* automáticos não é o suficiente para a identificação correta da chama, podendo negligenciar regiões menos luminosas ou identificar regiões espúrias que na verdade são reflexos da chama nas partes metálicas do motor. Para contornar este problemas, os pesquisadores recorrem a *thresholds* manuais, o que exige maior tempo para o processamento. A Figura 1 indica um exemplo de foto original de uma chama se propagando, o mapa de segmentação desejado, onde a cor branca simboliza chama, e o mapa produzido por um método tradicional sem ajuste dos parâmetros de pré-processamento. Após a segmentação da chama, é possível determinar informações como a velocidade aparente de chama, área, perímetro, circularidade, *wrinkling*, *stretching* entre outros parâmetros fundamentais para caracterizar a combustão.

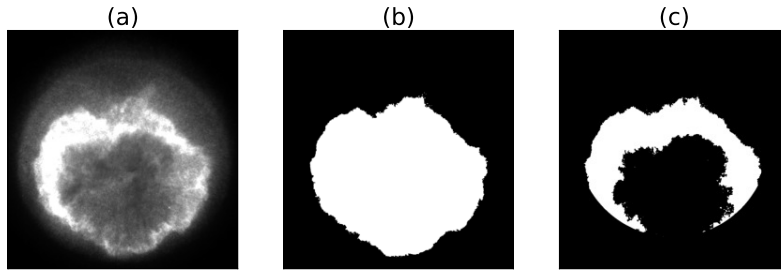


Figura 1: Exemplo de imagem de uma chama em propagação. (a) imagem original, (b) segmentação desejada, (c) imagem segmentada por um método tradicional (*threshold* de Otsu).

Analisando trabalhos anteriores da área, nota-se o uso de métodos de binarização [2], *threshold* de Otsu [1], [6] e o uso de pré-processamento de imagens, como o ajuste de contraste, brilho e recortes nas imagens [7], [12], não havendo um procedimento padrão que seja conveniente para todas as condições. Uma ferramenta de segmentação automática tornaria o processo mais rápido, possibilitando o processamento de um grande volume de dados em um tempo relativamente inferior, o que traz a possibilidade de avanços neste tipo de pesquisa. O uso de modelos de redes neurais de segmentação se mostra promissor. Nota-se também que não se tem conhecimento de um trabalho prévio em que se aplicou alguma ferramenta de aprendizado profundo para esta atividade, portanto a sua aplicação seria inovadora na área.

1.1 Objetivo

O objetivo final deste projeto é obter um modelo de rede de segmentação semântica que consiga, com boa precisão, distinguir uma chama de combustão e o seu interior do restante, de forma que seja possível levantar indicadores morfológicos para o seu estudo.

1.2 Metodologia

A seguinte metodologia será utilizada para o desenvolvimento deste trabalho:

- Preparação do *dataset* com imagens de diferentes combustíveis e pontos de operação;
- Definição de um modelo *baseline* baseado em uma técnica tradicional;
- Treinamento de modelos de aprendizado profundo para a tarefa de segmentação;
- Comparação dos resultados utilizando o coeficiente de Dice como métrica.

1.3 Revisão Bibliográfica

A segmentação de uma imagem envolve a sua separação em diferentes segmentos, os quais apresentam algum significado em relação a um contexto, e a aplicação deste tipo de técnica é relevante

em diferentes campos, como na área médica, em veículos autônomos e sistemas de vigilância. *Thresholding* é um conjunto de métodos tradicionais e de simples funcionamento baseados na aplicação de um limiar (*threshold*) para a separação das classes de interesse, podendo ser aplicado a nível global ou local [11]. O método de Otsu [10] é uma técnica de seleção automática de *threshold* de forma a maximizar a separação entre as classes em níveis de cinza.

O uso de técnicas baseadas em aprendizado profundo vem sendo empregado nos últimos anos, o que trouxe desempenho muito superior e representa uma mudança de paradigma na área de segmentação [9]. A tarefa de segmentação semântica a ser aprendida pelos modelos é a de classificação de cada *pixel* da imagem com rótulos semânticos, o que é um problema consideravelmente mais difícil do que tarefas de classificação de imagens, as quais atribuem um rótulo a uma imagem como um todo. Como exemplos de arquiteturas bem-sucedidas propostas pela comunidade de visão computacional, citam-se as redes convolucionais (CNNs), as redes recorrentes (RNN), as LSTM e as baseadas em *encoder-decoder*. Nota-se que, para muitas aplicações práticas, não há um número adequado de dados rotulados para o treinamento dos modelos. Neste contexto, o uso de *transfer learning*, definido como a utilização de um modelo previamente treinado em um outro conjunto de dados para o seu ajuste-fino com o *dataset* reduzido, e o uso de *data augmentation*, definido como uma geração artificial de dados a partir de transformações nas imagens originais, são técnicas de extrema importância. A extensão da segmentação semântica é a segmentação por instância, a qual tem como objetivo a detecção de diferentes objetos de uma mesma classe, para as diferentes classes presentes na imagem.

A rede *U-Net* [13] é um modelo do tipo *encoder-decoder* que gerou excelentes resultados, sendo vencedora em competições da área. Este modelo utiliza apenas operações de convolução e é dividido em duas partes: um caminho de contração e um de expansão. Durante a contração, camadas de convolução e ativação são aplicadas sucessivamente à imagem de entrada, a qual é reduzida por uma operação de *pooling*. Este procedimento é repetido sucessivamente, cada qual aplicando uma redução nas dimensões da imagem, porém aumentando o número de canais de saída. Ao atingir um gargalo definido, a rede passa a operar no modo de expansão, o qual consiste em camadas de *up-sampling* aproximadamente simétricas às da contração, promovendo o aumento na dimensão da imagem de forma que, na saída, a dimensão seja igual à da entrada. Nota-se que os mapas de ativação de cada bloco da contração são concatenados aos mapas da expansão. A Figura 2 ilustra a rede proposta pelos autores.

Apesar de ter sido proposta para aplicações na área médica, a rede *U-Net* vem sendo utilizada em aplicações fora deste domínio. Os autores alcançaram bons resultados fazendo uso extensivo de *data augmentation*, o que se mostrou essencial na situação de poucas imagens rotuladas, e notaram que a rede é rápida, levando menos de um segundo para a segmentação de uma imagem 512 x 512. O modelo *V-Net* [8] foi proposto para a segmentação de imagens médicas 3D. Os autores utilizaram uma função de perda baseada no coeficiente de Dice, o que permitiu o treinamento utilizando um *dataset* desbalanceado.

O modelo *HRNet* (*high-resolution network*) [14] foi proposto inicialmente para aplicações de segmentação semântica em estimativa de pose humana. Diferente da *U-Net*, o modelo proposto mantém representações de alta resolução ao longo das convoluções e produziu resultados expressivos em *benchmarks* de segmentação. A arquitetura apresenta um caminho de convoluções em alta resolução e caminhos paralelos de convoluções em menor resolução, conforme indicado na Figura 3. O modelo está dividido em quatro estágios, sendo que, para cada um, um novo caminho paralelo de convoluções é adicionado, cada um com uma resolução inferior aos caminhos antecedentes. Como saída, o modelo gera um mapa com a mesma dimensão da entrada. Os autores testaram diferentes

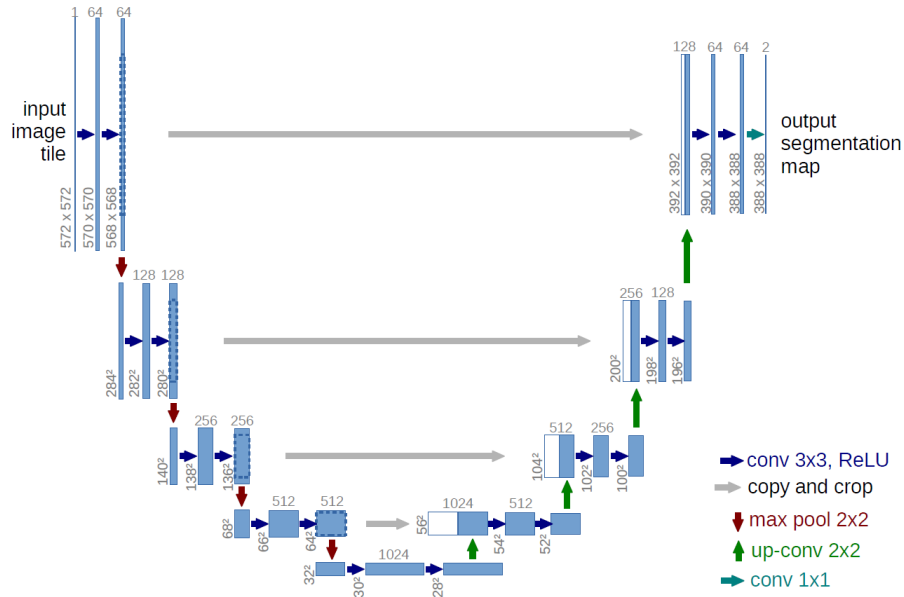


Figura 2: Arquitetura da rede *U-Net*, conforme apresentado em [13].

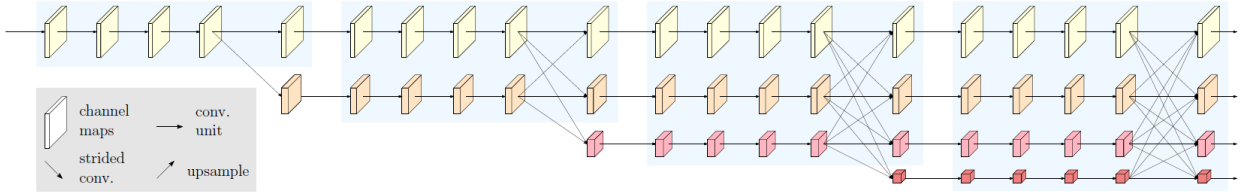


Figura 3: Arquitetura da rede *HRNet*, conforme apresentado em [14].

formas para a geração deste mapa e observaram, empiricamente, que a melhor arquitetura foi a que aplicava uma convolução nos mapas concatenados de cada resolução na última camada, conforme indicado na Figura 4.

Por fim, a família de modelos *DeepLab* [4] utiliza redes neurais convolucionais com dilatação, podendo se aproveitar de arquiteturas bem estabelecidas como a *VGG-16* ou *ResNet-50*. Um *kernel* de convolução com dilatação apresenta um espaçamento entre os seus pesos, de forma que o campo receptivo seja aumentado sem aumento do custo computacional. A rede faz uso de *pooling* piramidal para segmentar objetos de diferentes escalas. A saída da rede gera um mapa de ativações denso, o qual é interpolado de forma a retornar um mapa de resolução igual à entrada. A este mapa é aplicado o modelo probabilístico CRF (*Conditional Random Field*), o qual promove um refinamento do mapa de segmentação, permitindo uma melhor captura das bordas do objeto. Esta camada utiliza informação de todos os *pixels* da imagem de forma a favorecer o rotulamento da mesma classe para *pixels* vizinhos, o que promove a suavização do mapa gerado. Durante o treinamento da CRF, minimiza-se uma função de energia que contém um termo relacionado à reconstrução do

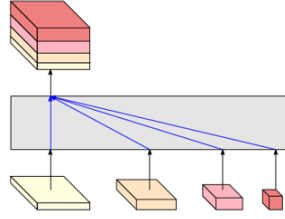


Figura 4: Arquitetura de saída da rede *HRNet* que apresentou o melhor desempenho, conforme apresentado em [14].

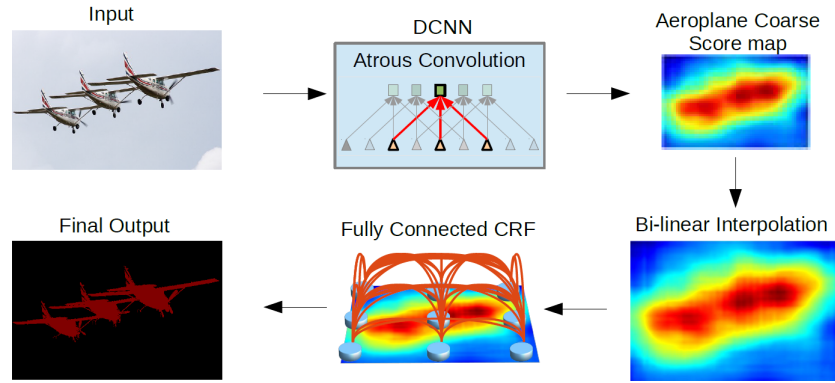


Figura 5: Arquitetura de saída da rede *DeepLab*, conforme apresentado em [4].

mapa de ativação e um termo relacionado à relação entre um par de *pixels*, para todos os *pixels* do mapa gerado. Esta arquitetura está indicada na Figura 5. É importante notar que o processo de treinamento da camada CRF é desacoplado da camada convolucional, podendo ser considerado um modelo de pós-processamento. Portanto, o treinamento do modelo não é feito de ponta a ponta.

Para contornar o problema do treinamento da CRF, foi proposto o modelo *DeepLab v3* [5], o qual pode ser treinado de ponta-a-ponta. Neste modelo, a imagem de entrada passa por camadas de convolução com dilatação, chamadas pelos autores de *Atrous Convolution*. O mapa de saída é alimentado a uma camada de *pooling* piramidal com dilatação, chamada de *Atrous Spatial Pyramid Pooling* (ASPP), a qual atua em quatro etapas. Em primeiro lugar, aplica-se *pooling* global. Em seguida, aplica-se uma convolução dilatada de *kernel* (1,1) seguida de três convoluções dilatadas com *kernel* (3,3). Em seguida, aplica-se uma camada de *image pooling* e, por fim, os resultados são concatenados e alimentados a uma última camada de convolução, a qual gera o mapa de segmentação. Estas etapas estão ilustradas na Figura 6.

2 Dataset

O *dataset* utilizado neste trabalho é composto por imagens de chama de combustão durante sua propagação, referentes a testes com etanol (164 imagens), gasolina (124 imagens) e gás hidrogênio

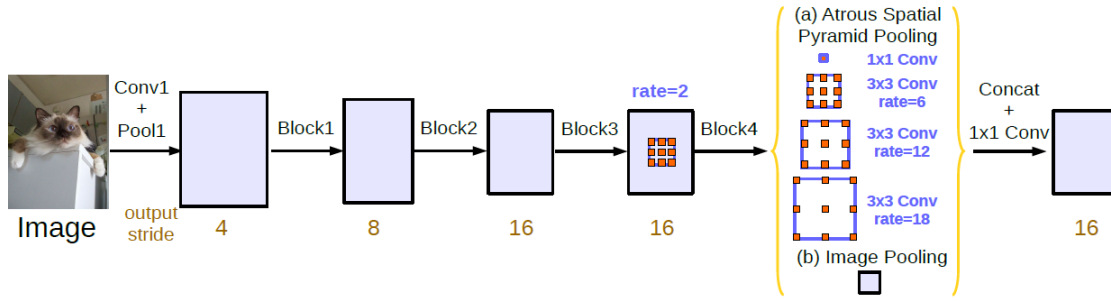


Figura 6: Arquitetura de saída da rede *DeepLab v3*, conforme apresentado em [5].

(100 imagens) como combustíveis, utilizados em trabalhos que já foram divulgados em artigos de revistas e congressos. Cada imagem conta com um respectivo mapa de segmentação, o qual foi gerado caso a caso definindo manualmente um *threshold* local e utilizando funções do pacote OpenCV em Python.

O equipamento utilizado para se obter as imagens é um motor de acesso óptico da empresa AVL, localizado no Laboratório de Combustão, Propulsão e Energia (LCPE) do Instituto Tecnológico de Aeronáutica (ITA). O motor possui um pistão com uma janela de quartzo. Abaixo do pistão é posicionado um espelho a 45° de forma a refletir a imagem para uma câmera de alta velocidade PCO Dimax posicionada lateralmente (Fig. 7). Conforme o motor opera com combustão, a câmera é sincronizada com a centelha e assim torna-se possível a aquisição de sucessivas imagens, mostrando a evolução da chama, conforme indicado na Figura 8.

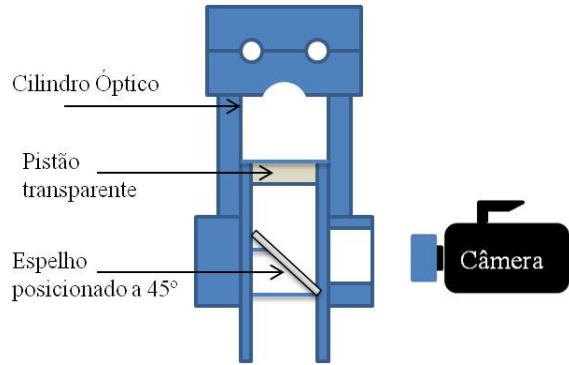


Figura 7: Configuração experimental utilizada na aquisição das imagens.

3 Metodologia

3.1 Arquiteturas

As redes utilizadas foram a *U-Net*, *HRNet* e *DeepLab v3*. Fora as redes, também foi utilizado o método do *threshold* de Otsu como um modelo *baseline* de comparação. No caso da *U-Net*, foi

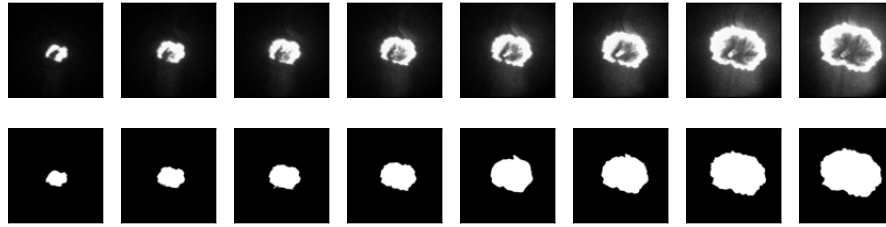


Figura 8: Exemplos de imagens capturadas para a propagação da chama da combustão de gás hidrogênio, na linha superior, e seu respectivo mapa de segmentação gerado manualmente na linha inferior.

utilizado um modelo proposto para segmentação de anomalias em imagens de ressonância magnética cerebral, conforme apresentado por [3]. Foram utilizados quatro blocos no caminho de contrações, os quais apresentavam 64, 128, 256 e 512 canais, enquanto no gargalo foram utilizados 1024 canais. Os blocos são compostos de duas convoluções com *kernel* (3,3) com ativação *ReLU* e uma camada de *Max Pooling* com *kernel* (2,2). Esta arquitetura está disponível em https://pytorch.org/hub/mateuszbuda_brain-segmentation-pytorch_unet/ e o modelo utilizado neste trabalho não foi pré-treinado.

A arquitetura *HRNet* utilizada foi adaptada do modelo apresentado em [14], disponível em <https://github.com/HRNet/HRNet-Semantic-Segmentation>. As convoluções apresentam *kernel* (3,3), as quais estão agrupadas em quatro estágios. De um estágio para o outro, um novo caminho em paralelo de convoluções é criado, o qual é aplicado à saída do estágio anterior após a aplicação de *downsampling* da imagem. O modelo utilizado não foi pré-treinado.

Por fim, foi utilizada a arquitetura *DeepLab v3* disponível em https://pytorch.org/vision/stable/_modules/torchvision/models/segmentation/segmentation.html, tendo como base a rede *ResNet-50*. Foi utilizada na saída da rede uma camada *DeepLabHead*, a qual consiste na aplicação de *pooling* piramidal, convolução, *batch normalization*, ativação *ReLU* e uma última camada de convolução, a qual apresenta como canais de saída o número de classes do problema. O modelo utilizado foi pré-treinado no conjunto *COCO*.

3.2 Configurações do treinamento

Com o objetivo de tornar a rede robusta a variações nas imagens de entrada, foi empregado o uso de *data augmentation* na fase de treinamento, contando com as seguintes transformações, as quais são aplicadas aleatoriamente:

- Espelhamento vertical;
- Espelhamento horizontal;
- Recorte de segmentos da imagem;
- Rotação da imagem original entre -30° e $+30^\circ$.

Fora as transformações citadas, também aplica-se normalização *z-score* em relação ao conjunto de treinamento. A Figura 9 ilustra diferentes imagens geradas artificialmente a partir dos aumentos aplicados à imagem original. Cada imagem conta com um respectivo mapa de segmentação, ao qual são aplicadas as mesmas transformações de aumento. Este mapa é uma matriz binária com dimensão igual à da imagem, sendo que o valor 0 corresponde à classe “não-chama” e o valor 1 à classe “chama”. Após a aplicação dos aumentos, a imagem é redimensionada de forma a apresentar mesma largura e altura, as quais serão variadas neste projeto de forma a se investigar o efeito do redimensionamento.

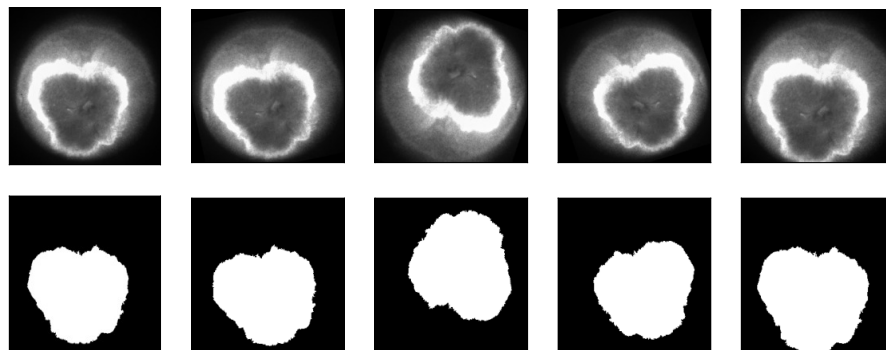


Figura 9: Exemplos de aumentos artificiais aplicados na imagem original (a primeira da esquerda para direita) e seu mapa de segmentação.

A métrica oficial para a avaliação e comparação dos resultados foi o coeficiente de Dice, o qual é dado abaixo, onde A é o mapa de referência, B é o mapa gerado pelo algoritmo e $|A \cap B|$ é o número de elementos na intersecção dos dois mapas. A função de perda a ser minimizada foi a *Dice Loss*, a qual utiliza os logits do mapa predito e retorna o complemento do coeficiente de Dice, de forma que a sua minimização leve a um aumento do coeficiente de Dice.

$$D(A, B) = \frac{2|A \cap B|}{|A| + |B|}. \quad (1)$$

Foram realizados experimentos em duas configurações de treinamento e teste. Os resultados foram avaliados para dimensões de entrada de $(256, 256)$, $(384, 384)$ e $(512, 512)$. Estas configurações foram:

- Configuração misturada: treinamento e teste realizados em conjuntos diferentes, cada um contendo dados dos três combustíveis;
- Configuração *Zero-Shot*: treinamento realizado com os dados de gasolina e etanol, teste realizado no conjunto de hidrogênio. Esta configuração é mais próxima do caso real em que um novo combustível deve ser analisado, porém é mais desafiadora do ponto de vista de aprendizado supervisionado.

O *batch size* utilizado foi 10, exceto nos casos de resolução $(512,512)$ onde se utilizaram duas imagens por *batch*. Os modelos foram treinados por 50 épocas, sendo que em alguns casos o treinamento foi interrompido antes pois a *loss* de treino e validação já havia se estabilizado.

4 Experimentos

4.1 Configuração misturada

A Tabela 1 resume o coeficiente de Dice médio para todas as amostras do conjunto de testes desta configuração, para as diferentes resoluções da imagem de entrada. Como referência, o valor médio utilizando o *threshold* de Otsu no conjunto de testes foi de 0.8270. Analisando a tabela, percebe-se que todos os modelos apresentaram alto valor da métrica de interesse. A Figura 10 indica o coeficiente de Dice para cada amostra, para a resolução de $(384,384)$. Nota-se que os resultados do modelo *DeepLab v3* apresentaram menor variação entre amostras e que, para todos os modelos, algumas amostras apresentaram coeficiente próximo de zero, as quais se referem a imagens do início da combustão.

Tabela 1: Coeficiente de Dice para a configuração misturada.

Resolução	<i>U-Net</i>	<i>HRNet</i>	<i>DeepLab v3</i>
$(256,256)$	0.9218	0.9092	0.9576
$(384,384)$	0.9096	0.9301	0.9545
$(512,512)$	0.9221	0.9238	0.9682

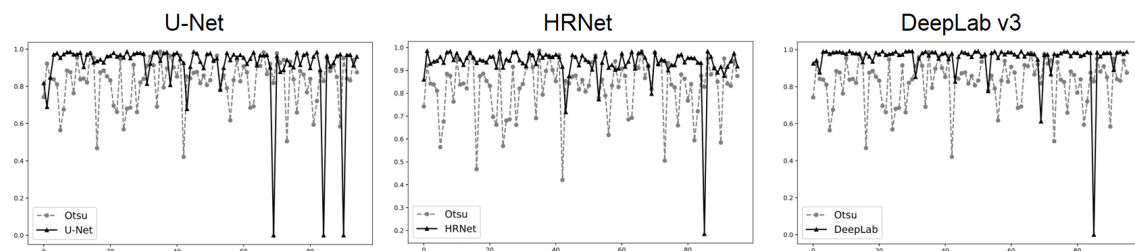
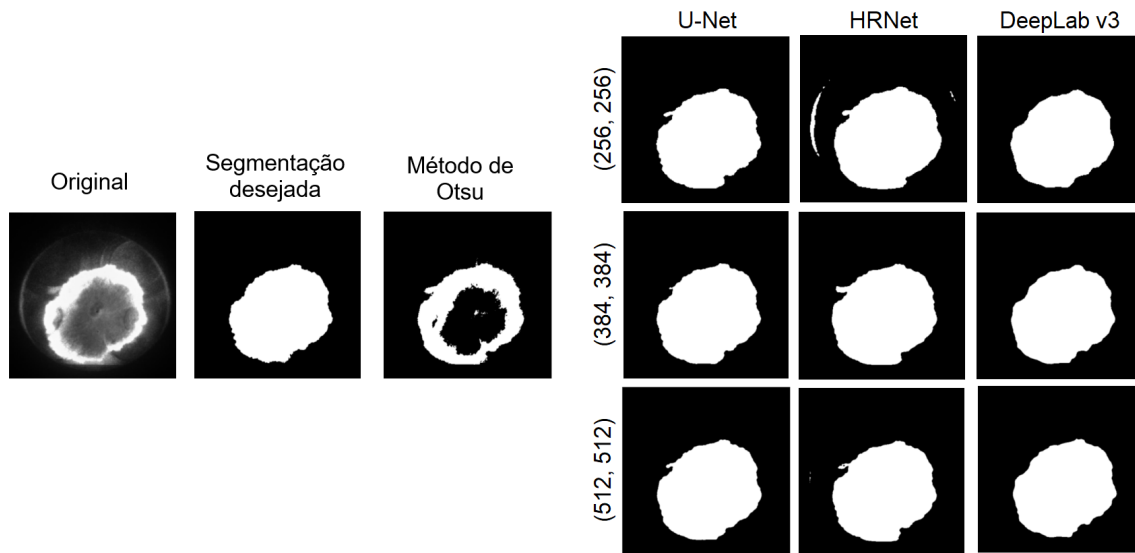
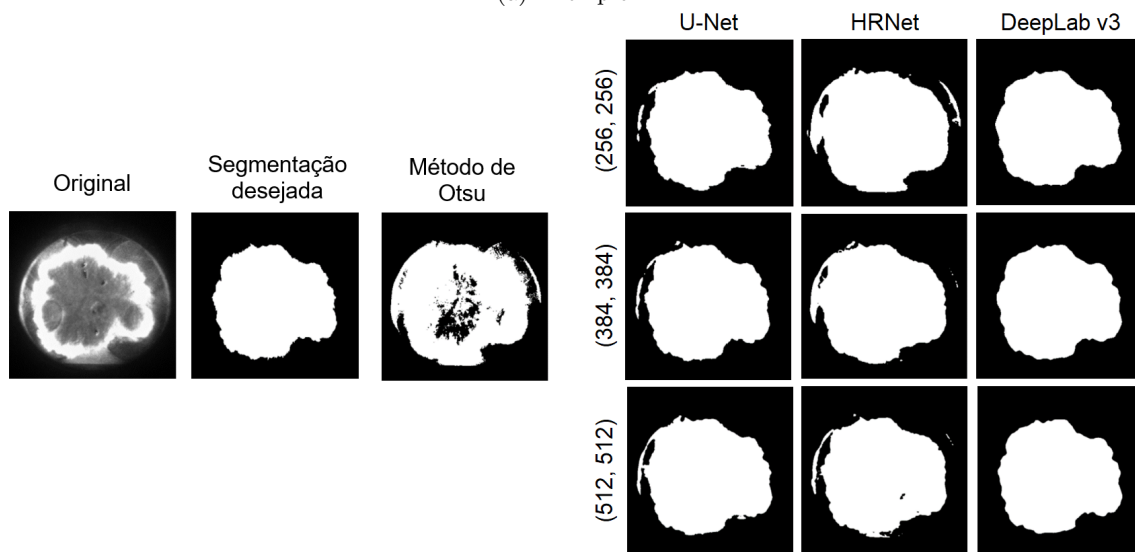


Figura 10: Coeficiente de Dice para cada amostra do conjunto de testes da configuração misturada utilizando resolução de $(384,384)$.

A Figura 11 indica dois exemplos desafiadores para a tarefa de segmentação, para os quais o método de Otsu não apresentou resultados satisfatórios. O Exemplo 1 apresenta a dificuldade na parte inferior da chama, enquanto o Exemplo 2 apresenta reflexões de luz na parede do cilindro. O modelo com melhor desempenho, independente da resolução da imagem, foi o *DeepLab v3*. Nota-se que os outros modelos destes exemplos apresentaram coeficiente de Dice elevado, próximo a 0.9, porém os mapas gerados apresentam erros relacionados às dificuldades relatadas. Isto mostra que a comparação dos resultados pelo coeficiente de Dice, por si só, é insuficiente para esta aplicação.



(a) Exemplo 1.



(b) Exemplo 2.

Figura 11: Exemplos de segmentações geradas para duas imagens selecionadas no conjunto de testes da configuração misturada.

4.2 Configuração *Zero-Shot*

A Tabela 2 resume o coeficiente médio para todas as amostras da configuração *Zero-Shot*, para as diferentes resoluções da imagem de entrada. Como referência, o valor médio utilizando o *threshold* de Otsu no conjunto de testes foi de 0.8306. Nota-se que o aumento da resolução da imagem impactou apenas o modelo *HRNet*, o qual apresentou melhor desempenho para maiores resoluções.

Tabela 2: Coeficiente de Dice para a configuração *Zero-Shot*.

Resolução	<i>U-Net</i>	<i>HRNet</i>	<i>DeepLab v3</i>
$(256, 256)$	0.9296	0.7614	0.9278
$(384, 384)$	0.8994	0.8639	0.9207
$(512, 512)$	0.8237	0.8855	0.9242

É possível observar que o modelo *DeepLab v3* apresentou resultados mais consistentes do que as demais arquiteturas no caso do Exemplo 1, enquanto que a rede *U-Net* foi mais próxima da segmentação desejada no Exemplo 2 ao se utilizar a menor resolução. A Figura 12 indica os coeficientes de Dice para cada amostra utilizando resolução de $(384, 384)$. Pode-se notar a presença de mínimos, os quais correspondem a imagens do início da combustão em que a chama ainda é muito pequena. Outro fato interessante é que, a partir da amostra 55, o desempenho médio cai. Isto ocorreu pois a segunda metade do *dataset* corresponde a uma mistura pobre de hidrogênio e ar, o que gerou chamas com maior dificuldade de segmentação do que a primeira metade, a qual corresponde a uma mistura estequiométrica.

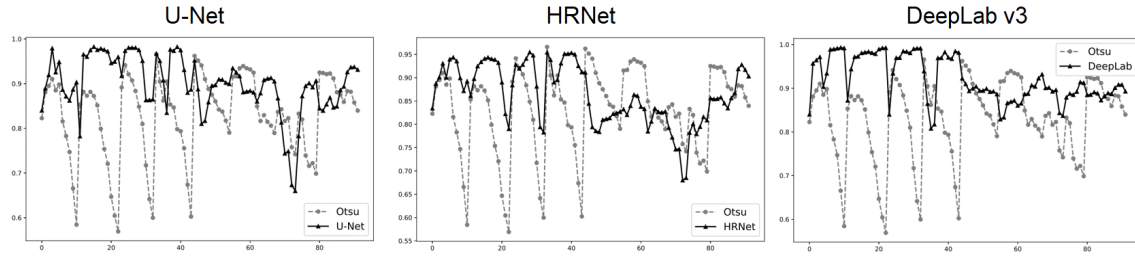
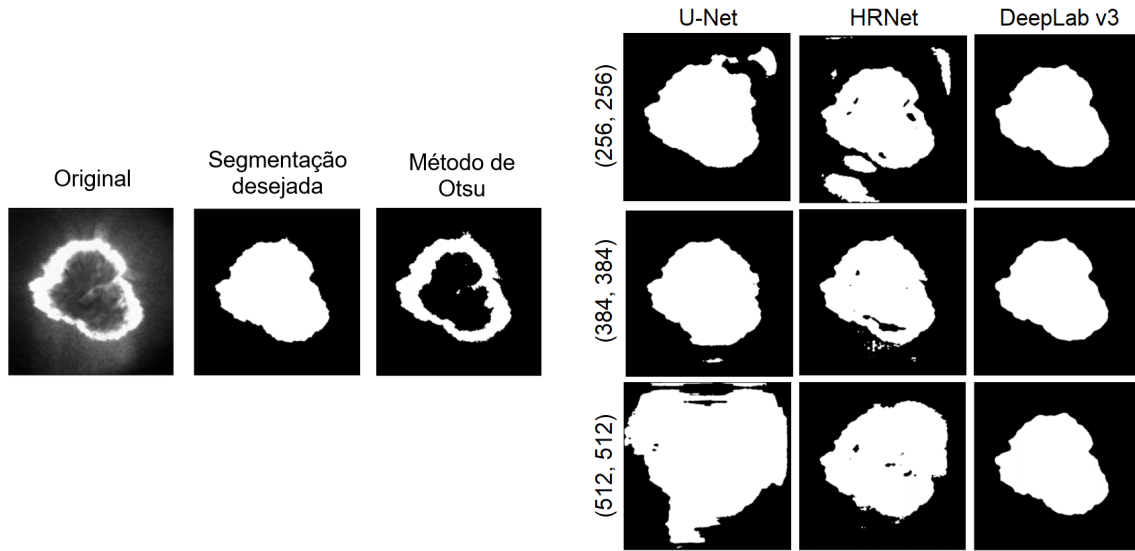
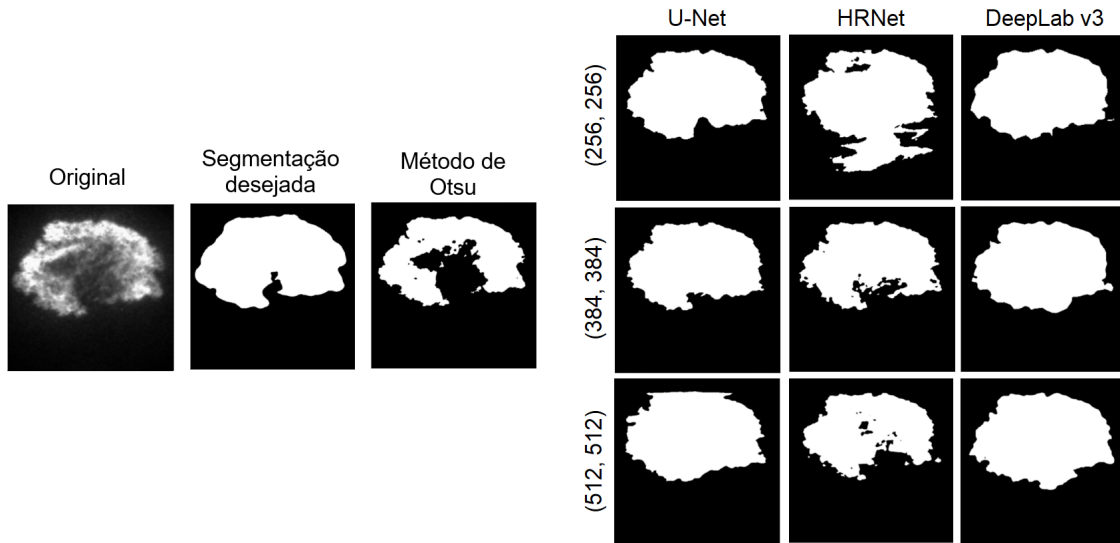


Figura 12: Coeficiente de Dice para cada amostra do conjunto de testes da configuração *Zero-Shot* utilizando resolução de $(384, 384)$.

A Figura 13 traz dois exemplos de imagens para diferentes resoluções, os quais representam casos desafiadores para a tarefa de segmentação. A dificuldade do Exemplo 1 é a de preencher o interior da chama com a classe correta, enquanto que para o Exemplo 2 o desafio é o correto fechamento da chama na parte inferior.



(a) Exemplo 1.



(b) Exemplo 2.

Figura 13: Exemplos de segmentações geradas para duas imagens selecionadas no conjunto de testes da configuração *Zero-Shot*.

5 Conclusões

Neste projeto, investigou-se o uso de modelos de aprendizado profundo para a tarefa de segmentação semântica de imagens de chama durante a combustão em um motor. Os resultados obtidos através do modelo *DeepLab v3* foram os mais consistentes com os mapas desejados, para diferentes resoluções de imagem, sendo capaz de superar algumas dificuldades inerentes deste tipo de imagem, como reflexões de luz, interior da chama com baixa intensidade e contorno indefinido em determinadas regiões da chama. Dados os bons resultados alcançados, o uso destes modelos se mostra promissor na área de análise óptica de combustão.

6 Trabalhos Futuros

Como possibilidades de trabalhos futuros, citam-se:

- Comparar os resultados através de outras métricas de segmentação, como acurácia e IoU, ou propor novas métricas que penalizem predições incorretas relacionadas aos desafios inerentes desta aplicação;
- Analisar modelos que utilizem mecanismo de atenção, os quais podem trazer bons resultados em relação aos desafios citados;
- Propor novas técnicas e *data-augmentation* que sejam capazes de tornar o modelo treinado mais robusto a reflexões nas paredes do cilindro;
- Comparar resultados em termos de parâmetros morfológicos da chama, o que está mais alinhado à aplicação final.

Referências

- [1] Pavlos G Aleiferis and Markus K Behringer. Flame front analysis of ethanol, butanol, iso-octane and gasoline in a spark-ignition engine using laser tomography and integral length scale measurements. *Combustion and Flame*, 162(12):4533–4552, 2015.
- [2] PG Aleiferis, J Serras-Pereira, and D Richardson. Characterisation of flame development with ethanol, butanol, iso-octane, gasoline and methane in a direct-injection spark-ignition engine. *Fuel*, 109:256–278, 2013.
- [3] Mateusz Buda, Ashirbani Saha, and Maciej A Mazurowski. Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm. *Computers in biology and medicine*, 109:218–225, 2019.
- [4] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. 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, 2017.
- [5] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.

- [6] Simona Silvia Merola, Adrian Irimescu, Luca Marchitto, Cinzia Tornatore, and Gerardo Valentino. Flame contour analysis through uv-visible imaging during regular and abnormal combustion in a disi engine. Technical report, SAE Technical Paper, 2015.
- [7] Simona Silvia Merola, Cinzia Tornatore, and Adrian Irimescu. Cycle-resolved visualization of pre-ignition and abnormal combustion phenomena in a gdi engine. *Energy Conversion and Management*, 127:380–391, 2016.
- [8] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. IEEE, 2016.
- [9] Shervin Minaee, Yuri Y Boykov, Fatih Porikli, Antonio J Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.
- [10] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.
- [11] Nikhil R Pal and Sankar K Pal. A review on image segmentation techniques. *Pattern recognition*, 26(9):1277–1294, 1993.
- [12] A Peñaranda, SD Martinez Boggio, PT Lacava, S Merola, and A Irimescu. Characterization of flame front propagation during early and late combustion for methane-hydrogen fueling of an optically accessible si engine. *International Journal of Hydrogen Energy*, 43(52):23538–23557, 2018.
- [13] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [14] Jingdong Wang, Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao, Dong Liu, Yadong Mu, Mingkui Tan, Xinggang Wang, et al. Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 2020.