Revista de Informática Teórica e Aplicada - RITA - ISSN 2175-2745
Vol. 32, Num. 1 (2025) 83-90

RESEARCH ARTICLE

Deep Learning-Based Segmentation of Nanomaterial Images Acquired via Scanning Electron Microscopy

Aprendizado Profundo para Segmentação de Imagens de Nanomateriais Obtidas por Microscopia Eletrônica de Varredura

Ivânia M. L. De Donato[1]*, Fernanda D. Marques[2], Juliana N. Lunz[2], Bráulio S. Archanjo[2], Francisco J. P. Lopes[1], Gilson A. Giraldi[3]

Resumo: This work investigates the application of deep learning techniques to analyze nanomaterial images acquired by scanning electron microscopy (SEM). The main objective is to improve the segmentation of zinc oxide (ZnO) nanoparticles and graphene oxide (GO) nanosheets. For this purpose, we developed a dedicated database for the supervised training of deep neural networks, using the U-Net architecture. Given the high spatial resolution of the original images, we investigated two approaches for image subdivision during the training and testing stages. Due to the high cost of acquiring and annotating SEM images of nanomaterials, we applied transfer learning through fine-tuning, adapting the network previously trained on ZnO images to segment GO images. The results were promising, with accuracies and recalls exceeding 95% for both ZnO and GO images.
Keywords: Deep learning — U-Net — Nanomaterial — Fine-tuning

Resumo: Este trabalho investiga a aplicação de técnicas de aprendizado profundo para analisar imagens de nanomateriais adquiridas por microscopia eletrônica de varredura (MEV). O objetivo principal é aprimorar a segmentação de nanopartículas de óxido de zinco (ZnO) e nanofolhas de óxido de grafeno (GO). Para este propósito, desenvolvemos um banco de dados específico para o treinamento supervisionado de redes neurais profundas, utilizando a arquitetura U-Net. Dada a alta resolução espacial das imagens originais, investigamos duas abordagens de subdivisão para as etapas de treinamento e teste. Devido ao alto custo de aquisição e anotação de micrografias de nanomateriais obtidas por MEV, aplicamos a transferência de aprendizado por meio de ajuste fino, adaptando a rede previamente treinada em imagens de ZnO para segmentar imagens GO. Os resultados foram promissores, com acurácias e *recalls* superiores a 95% para ambos os casos.
Palavras-Chave: Aprendizado Profundo — U-Net — Nanomaterial — Ajuste Fino

[1] *Universidade Federal do Rio de Janeiro (UFRJ), Brasil*
[2] *Divisão de Metrologia de Materiais – INMETRO, Brasil*
[2] *Laboratório Nacional de Computação Científica (LNCC), Brasil*
***Corresponding author**: ivania@ufrj.br
DOI: http://dx.doi.org/10.22456/2175-2745.143516 • **Received:** 24/10/2024 • **Accepted:** 28/11/2024

1. Introdução

A caracterização do tamanho e da proporção de partículas distribuídas em um sistema é fundamental para a correta classificação de um nanomaterial. Definições normativas relacionadas a essa classificação têm sido amplamente discutidas por comitês técnicos, evidenciando sua importância. Métodos não automatizados de medição de partículas estão sujeitos a inconsistências que podem comprometer a análise e a rotulagem do material. A adoção de métodos automáticos é imprescindível para garantir reprodutibilidade nos resultados.

O primeiro passo para a caracterização do tamanho e da porcentagem de nanopartículas em um material é a segmentação. Nossa proposta é utilizar o aprendizado profundo para essa tarefa, o que depende de grandes volumes de dados para o treinamento de redes. Entretanto, considerando o alto custo de aquisição e anotação de imagens de microscopia eletrônica de varredura (MEV), adotamos estratégias de *Few-Shot learning* (FSL) para lidar com a limitação de imagens rotuladas [1]. Realizamos a segmentação de nanomateriais presentes em imagens obtidas por MEV, por meio da arquitetura U-Net, amplamente utilizada em tarefas de segmentação de imagens [2]. Para o treinamento, validação e teste da rede, desenvolvemos um banco de dados constituído por amostras de imagens de nanopartículas de óxido de zinco (ZnO) e nano-

folhas de óxido de grafeno (GO) e suas versões segmentadas usando o software ImageJ [3]. Essas imagens apresentam alta resolução espacial e por isso foi necessário subdividi-las antes de serem processadas pela rede. Testamos duas estratégias de subdivisão, uma regular e outra aleatória. Após essa etapa, realizamos a transferência de aprendizado por meio do ajuste fino (*fine-tuning)*, uma técnica tradicional em FSL, para adaptar a U-Net previamente treinada em imagens de ZnO em um modelo para segmentar imagens de nanofolhas de GO.

Em relação ao estado da arte, algoritmos de aprendizado profundo que utilizam a U-Net têm se mostrado uma abordagem eficiente para a segmentação de imagens de MEV e outros tipos de micrografias. Estudos recentes, como [4, 5], exploraram o uso de redes neurais profundas para análise morfológica de nanopartículas e materiais similares.

No trabalho de [6], foi utilizada uma abordagem de aprendizado com poucos exemplos rotulados para segmentar nanopartículas de óxido de molibdênio e outros tipos de materiais, alcançando segmentação precisa mesmo diante de variações de contraste e tamanho de partículas. O estudo de [7] propôs um método automático de detecção de flocos de grafeno em imagens de microscopia óptica, utilizando uma U-Net modificada, que melhorou significativamente a detecção. Já [8] aplicou uma U-Net pré-treinada em imagens de vírus Chikungunya, Parapoxvírus e Marburg para segmentar o vírus SARS-CoV-2 em imagens de microscopia eletrônica de transmissão, resultando em um modelo eficiente e de baixo custo. O trabalho de [9] demonstrou a eficácia da técnica FSL ao treinar uma U-Net com imagens sintéticas, transferindo esse aprendizado para dados reais de células de glioblastoma. O estudo de [10] integrou FSL e U-Net para segmentar lesões de câncer de pulmão em tomografias, ajustando o modelo com feedback supervisionado on-line, proporcionando uma acurácia satisfatória.

Destacamos ainda o trabalho de [11], que empregou um sistema de reconhecimento de imagens MEV por aprendizado profundo para segmentar nanofios de ZnO, correlacionando a área de superfície com o desempenho do material. O estudo de [12] que comparou oito modelos de segmentação de micrografias com nanopartículas de diferentes tipos, e também o estudo de [13] que focou na detecção de nanopartículas de prata, destacando a importância de uma inspeção rigorosa dos parâmetros dimensionais dos nanomateriais.

Embora o presente artigo compartilhe similaridades com outros estudos que apresentaram avanços principalmente nas áreas de biologia e de medicina, nossa pesquisa traz um avanço ao aplicar técnicas de FSL, via ajuste fino, para a segmentação de nanomateriais, como uma forma eficaz de transferir conhecimento entre diferentes conjuntos de dados.

As principais contribuições do presente trabalho são:

- Disponibilização de um banco de imagens de nanomateriais obtidas por MEV, permitindo a reprodutibilidade científica.
- Análise da influência de diferentes estratégias para subdivisão de imagens de alta resolução, em recortes para o treinamento, validação e teste dos modelos.
- Uso de técnicas de FSL, no caso, transferência de aprendizado, na segmentação de nanomateriais, demonstrando eficácia em cenários com dados limitados.

O texto está organizado como segue: a Seção 2 descreve o banco de dados criado para o treinamento da rede, em seguida detalha o método adotado. Por fim, o artigo apresenta, na Seção 3, a discussão dos resultados obtidos no experimento computacional e, na Seção 4, as conclusões do estudo.

2. Método

2.1 Banco De Dados e Pré-processamento

Nosso banco de dados é composto por imagens de alta resolução de nanopartículas de ZnO e nanofolhas de GO, com dimensões de 1770×2048 pixels e 1752×2016 pixels, respectivamente. As imagens foram adquiridas por um MEV modelo Magellan 450, da FEI Company. No caso das nanopartículas de ZnO, as imagens foram capturadas com uma magnificação de 50.000x e um campo de visão horizontal (HFW) de 5,12 μm, e as imagens de GO foram capturadas com uma magnificação de 10.000x e um HFW de 25,6 μm.

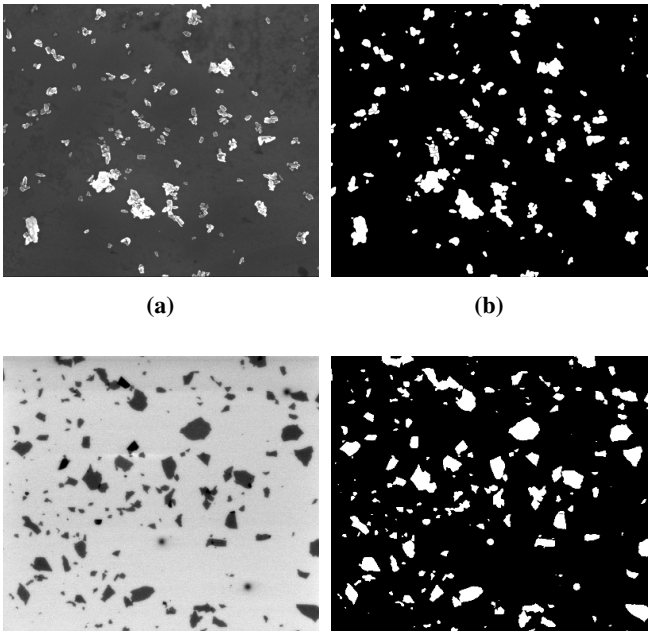
Cada imagem possui seu respectivo padrão-ouro, gerado de forma semiautomática no software ImageJ [3], pela aplicação do método de limiarização de Otsu, seguido por correções manuais para refinar as bordas das partículas e garantir maior precisão na segmentação. Os padrões-ouro destacam os nanomateriais em branco, servindo como referência para o treinamento e validação dos modelos de segmentação. Os dois pares da Figura 1 ilustram as imagens originais e os seus respectivos padrões-ouro.

2.2 Estratégias de Recorte

O tamanho 256×256 pixels é amplamente adotado em tarefas de segmentação de imagens com redes neurais, pois proporciona um equilíbrio adequado entre a resolução das imagens e o uso de memória durante o processamento [14, 15, 16, 17]. Recortes muito grandes podem resultar em maior consumo de memória e tempo de processamento, enquanto recortes menores podem comprometer a captura de informações relevantes sobre os objetos a serem segmentados [18].

No entanto, o recorte sem sobreposição poderia resultar em perda de informações nas bordas da imagem. Essa perda ocorre porque o filtro de convolução (*kernel*), que percorre a imagem extraindo características, possui um tamanho fixo e, ao processar pixels próximos à borda de um recorte, não tem acesso aos pixeis que estão presentes somente no recorte vizinho. Para minimizar esse problema, adotamos uma sobreposição de 4 pixels (equivalente a 1,56%) na largura e altura dos recortes. Esse valor de sobreposição foi definido empiricamente com o objetivo de reduzir a imprecisão na segmentação de pixels próximos às bordas.

Comparamos duas estratégias de recorte; na estratégia de recorte I, as amostras para o conjunto de treinamento e



(a) **(b)** **(c)** **(d)**

Figura 1. (a) Imagem original de ZnO e (b) Imagem padrão-ouro de ZnO, (c) Imagem original de GO e (d) Imagem padrão-ouro de GO.

validação foram recortadas com dimensões 256×256 pixels, com sobreposição de 4 pixels na altura e na largura (Figura 2), como já comentado acima.

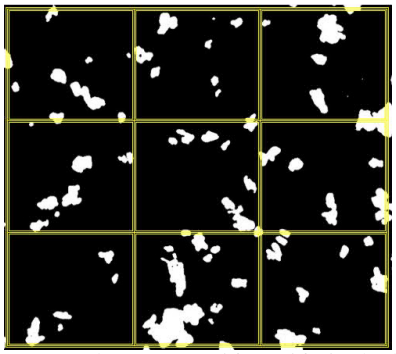


Figura 2. Esquema de recorte 256×256 pixels das imagens de ZnO, com sobreposição de 4 pixels na altura e na largura.

Na estratégia de recorte II para cada par (imagem/padrão-ouro), os recortes de 256×256 pixels foram realizados aleatoriamente em cada época durante as etapas de treinamento e validação.

Para o conjunto de teste, optamos por uma subdivisão em recortes 256×256 pixels, sem sobreposição, de modo a evitar possíveis efeitos de suavização no resultado das predições, nos pixeis próximos as bordas dos recortes.

2.3 Validação Cruzada

As estratégias I e II foram aplicadas pelo método de validação cruzada [19] com K = 4 (K-*fold*), onde o conjunto de treinamento e validação é dividido em 4 partes (ou *folds*). A cada iteração, um desses *folds* é usado para validação, enquanto os outros 3 *folds* são usados para treinamento. Esse processo é repetido 4 vezes, alternando o *fold* de validação, de forma que, ao final, cada parte do conjunto de dados foi utilizada uma unica vez para validação e três vezes para treinamento. Essa técnica é usada para analisar a variação nas estimativas de desempenho quando alteramos o conjunto de treinamento e para evitar que o modelo seja avaliado com base em uma única divisão arbitrária dos dados. Dentre os modelos gerados durante a execução do K-*fold*, selecionamos o melhor modelo de cada estratégia como sendo aquele que obteve a menor perda durante a validação. Dessa forma, temos dois modelos vencedores, denominados M_I e M_{II}, para as estratégias de recorte I e II, respectivamente.

2.4 Arquitetura U-Net

Selecionamos a arquitetura U-Net [2], representada na Figura 3, implementada em TensorFlow 2.15.0 para segmentar as nanopartículas de ZnO e as nanofolhas de GO. A rede é composta por camadas convolucionais para *downsampling* e camadas deconvolucionais para *upsampling*. Durante o *downsampling*, são aplicadas convoluções com um *kernel* de tamanho 4×4, *stride* de 2, com *padding*, função de ativação LeakyReLU e *batch normalization* [2]. A imagem de entrada, com resolução 256×256 é processada por uma camada convolucional com 64 filtros gerando um bloco (tensor) com dimensão 128×128 e profundidade 64. Esse bloco segue pelo codificador, onde a quantidade de filtros dobra a cada nova camada convolucional, até o valor 512, enquanto a resolução vai sendo reduzida. Após sete camadas de *downsampling*, a rede realiza *upsampling* através de sete camadas deconvolucionais. Cada camada do decodificador realiza *upsampling* seguido por deconvoluções com um *kernel* de tamanho 4×4, *batch normalization* e concatenação com as correspondentes camadas de *downsampling* via *skip connections*.

Após a última camada de *upsampling*, a rede retorna a dimensão de 256×256 e a convolução final realiza a ativação de saída, dada pela função sigmoide.

Os pesos da U-Net foram inicializados conforme o método Glorot Uniform [20] e para otimização, empregamos o algoritmo Adam [21].

A entropia cruzada binária [22], denotada por L_{BCE}, foi selecionada como função de perda, sendo representada por:

$$L_{BCE} = -\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{c}\left[y_{ij}\log(p_{ij}) + (1-y_{ij})\log(1-p_{ij})\right], \quad (1)$$

onde n é o total de imagens usadas para treinamento, c é o número de pixels de uma image e y_{ij} é o rótulo verdadeiro para o pixel j da imagem i e p_{ij} é a predição da rede para cada pixel j da imagem i.

Assim, dada uma imagem x_i, se a predição para o pixel j é um $p_{ij} > 0.5$ consideramos que o pixel j pertença à classe '1', caso contrário, ele será atribuído à classe '0'.

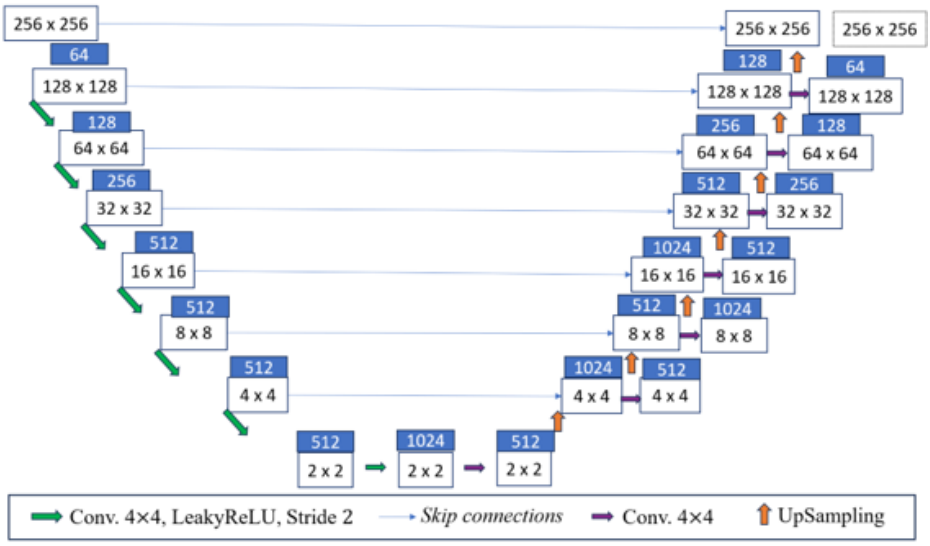


Figura 3. Arquitetura de aprendizado profundo, U-Net 256.

O desempenho dos modelos M_I e M_{II} sobre o conjunto de teste foi avaliado com base nas medidas de Precisão $P(i)$, *Recall* $R(i)$ e *F1-Score* $F1_{Score}(i)$, as quais são calculadas a partir do número de pixels verdadeiros positivos da classe i, denotado por TP_i, falso positivo (FP_i) e falso negativo (FN_i), onde $i = 0$ para o fundo e $i = 1$ para o objeto de interesse [23].

$$P(i) = \frac{TP_i}{TP_i + FP_i}, \tag{2}$$

$$R(i) = \frac{TP_i}{TP_i + FN_i}, \tag{3}$$

$$F1_{Score}(i) = 2 \cdot \frac{P(i) \cdot R(i)}{P(i) + R(i)}. \tag{4}$$

Utilizamos também a Intersecção sobre a União e a Acurácia. A primeira, denotada por IoU:

$$IoU = \frac{|y_{true} \cap y_{pred}|}{|y_{true} \cup y_{pred}|}, \tag{5}$$

onde y_{true} é o padrão-ouro e y_{pred} é a segmentação gerada pela rede. A Acurácia é calculada pela expressão:

$$Accuracy = \frac{\sum_{i=1}^{N} TP_i}{\sum_{i,j=1}^{N} C_{i,j}}, \tag{6}$$

onde $C_{i,j}$ é um elemento da matriz de confusão, sendo o número de amostras que pertencem à classe i e foram previstas como pertencentes à classe j. O valor de N é o número de classes, sendo $N = 2$ nesse caso.

2.5 Transferência de aprendizado

Implementamos o FSL por meio da transferência de aprendizado via ajuste fino, para investigar o comportamento da U-Net em cenários com disponibilidade limitada de dados. A metodologia é aplicada sobre os modelos M_I e M_{II} pré-treinados com as imagens de ZnO; ou seja, os modelos M_I e M_{II}, cujos pesos foram inicializados pelo método Glorot Uniform [20] e treinados a partir das imagens de ZnO, foram posteriormente retreinados com as imagens de GO, de dimensões 1752×2016 pixels, que foram pré-processadas conforme as estratégias de recortes descritas na Seção 2.2.

Nesta etapa, as redes foram inicializadas com os pesos obtidos no treinamento anterior. A taxa de aprendizado do otimizador Adam foi reduzida para 0,0001 para o ajuste fino do treinamento com as imagens de GO. Os demais hiperparâmetros do otimizador não foram alterados. Os dados foram normalizados dividindo os valores dos pixels pela intensidade máxima (255).

Além disso, como estamos simulando um cenário com número reduzido de imagens para explorar a técnica de FSL, optamos por realizar um único treinamento para cada modelo, substituindo o K-*fold* pela divisão dos dados de treinamento e validação na proporção 75:25, respectivamente. Para o conjunto de teste selecionamos 2 imagens.

3. Resultados e Discussão

3.1 Treinamento e avaliação dos modelos

Esta seção apresenta os resultados obtidos utilizando um subconjunto formado por 10 pares de imagens de alta resolução (original e padrão-ouro) de ZnO selecionados do banco de dados descrito na Seção 2.1. Deste subconjunto, 2 imagens foram escolhidas aleatoriamente para compor o conjunto de teste, enquanto as 8 restantes foram utilizadas em validação cruzada com K=4, descrita na Seção 2.3. As estratégias I e II foram aplicadas para recortar as imagens em retalhos de tamanho 256×256. Por razões de eficiência de processamento, todas as etapas computacionais foram realizadas na plataforma online gratuita Google Colaboratory, utilizando a

GPU Nvidia K80s disponível durante os treinamentos. Para assegurar a reprodutibilidade, o valor 42 foi escolhido como semente randômica do gerador de números aleatórios utilizado na inicialização das redes. A U-Net empregada possui um total de 41.830.401 parâmetros treináveis.

Os modelos foram treinados ao longo de 30 épocas. O otimizador Adam foi parametrizado com taxa de aprendizagem de 0,001 e tamanho de *batch* igual a 16. Utilizamos a entropia cruzada binária (1) como função de perda. Obtivemos os modelos M_I e M_{II}, correspondentes às estratégias de recorte I e II, respectivamente. As Figuras 4 e 5 ilustram a evolução dos treinamentos dos modelos M_I e M_{II}.

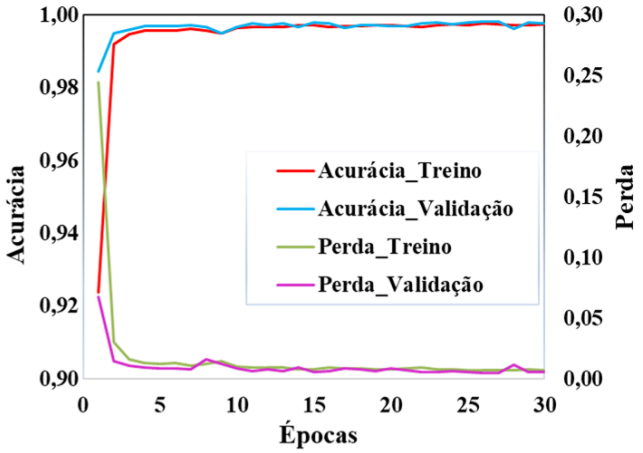


Figura 4. Curvas de acurácia e perda, da estratégia de recorte I, durante o treinamento e validação com as imagens de ZnO.

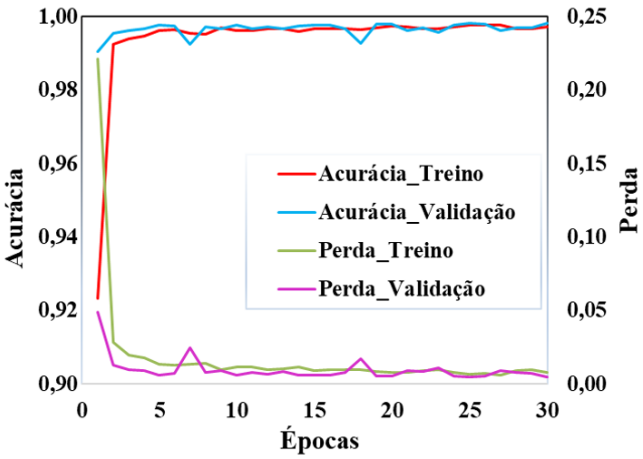


Figura 5. Curvas de acurácia e perda, da estratégia de recorte II, durante o treinamento e validação com as imagens de ZnO.

Ao comparar as acurácias para o treinamento e validação usando as estratégias de recorte I e II, Figuras 4 e 5, respectivamente, observamos o mesmo padrão, seguindo a mesma tendência nos dois casos.

Para avaliar o desempenho no conjunto de teste, composto por 112 recortes, utilizamos matrizes de confusão juntamente com as métricas de IoU, *recall* e *F1-Score* (Figura 6), cujos valores foram obtidos pelas médias das equações (5), (3) e (4), respectivamente.

M_I		M_{II}	
IoU:	95,69 % ± 3,62 %	IoU:	95,63 % ± 3,61 %
Recall:	97,92 % ± 1,89 %	Recall:	98,43 % ± 1,55 %
F1-Score:	97,76 % ± 2,05 %	F1-Score:	97,73 % ± 2,02 %

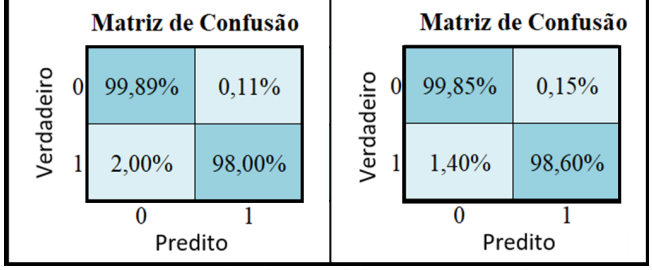


Figura 6. Desempenho dos modelos M_I e M_{II} sobre o conjunto de teste das imagens de ZnO.

As matrizes de confusão de cada estratégia de recorte, Figura 6, mostram que os modelos M_I e M_{II} acertaram pelo menos 98% das predições, evidenciando um bom desempenho para dados não vistos durante o treinamento.

Considerando o IoU, em média, para os diferentes recortes do conjunto de teste, o modelo M_I obteve 95,69% variando dentro do intervalo [92,07; 99,31]. O modelo M_{II} atingiu em média um IoU de 95,63% e um intervalo de [92,02; 99,24]. Com relação ao *F1-Score*, que mede o equilíbrio entre precisão e o *recall*, os intervalos estatísticos foram de [95,71; 99,81] e [95,71; 99,75] para os modelos M_I e M_{II}, respectivamente. Para o *recall*, que indica quão bem o modelo identifica corretamente os pixels que realmente fazem parte do objeto de interesse, os recortes do conjunto de teste apresentaram uma performance entre os intervalos [96,03; 99,81] e [96,88; 99,98] para os modelos M_I e M_{II}, respectivamente.

Com base na análise apresentada em [24] para comparar intervalos estatísticos, concluímos que os modelos M_I e M_{II} são estatisticamente equivalentes em relação às medidas de IoU e *F1-Score*. Com relação ao *recall*, ao considerar o critério fornecido por [24], segundo o qual, dados dois intervalos estatísticos $[\mu_1 - \sigma_1; \mu_1 + \sigma_1]$ e $[\mu_2 - \sigma_2; \mu_2 + \sigma_2]$, referentes aos modelos M_I e M_{II}, respectivamente, dizemos que M_I é inferior a M_{II} se:

$$\mu_1 < \mu_2 \rightarrow \rho = \frac{|(\mu_1 + \sigma_1) - (\mu_2 - \sigma_2)|}{2\sigma_1}. \qquad (7)$$

No caso, temos: $\mu_1 = Recall(M_I)$ e $\mu_2 = Recall(M_{II})$, sendo os desvios-padrão calculados de forma análoga. Com base nos valores apresentados anteriormente, obtemos $\rho \approx 0,77$. Como $\rho < 1$, podemos inferir que o modelo M_{II} é ligeiramente superior a M_I em termos de *recall*, uma vez que apresenta uma média maior e um valor de ρ relativamente baixo.

No contexto de análise de nanomateriais, a escolha do modelo ideal depende da aplicação específica. Um aspecto importante é a contagem precisa de nanopartículas com base nas

regiões segmentadas. A métrica utilizada para quantificação do tamanho das partículas, seja diâmetro, área ou volume, pode impactar diretamente os resultados dessa contagem. Considerando esta questão, é importante observar que o *recall* diminui com o aumento dos falsos negativos, que correspondem a pixels do objeto incorretamente classificados como fundo. Isso leva à subestimação do tamanho real das nanopartículas e pode, em certos casos, fragmentar partículas maiores em várias menores, potencialmente resultando em uma contagem superestimada. Por outro lado, a precisão é sensível a falsos positivos, ou seja, pixels do fundo da imagem que foram segmentados incorretamente como partículas, o que pode aumentar o tamanho estimado das partículas e causar uma contagem incorreta. Portanto, alcançar um equilíbrio entre *recall* e precisão é fundamental para garantir a acurácia na análise e contagem de nanopartículas.

3.2 Ajuste Fino

Com o objetivo de analisar o comportamento dos modelos em situações de poucas amostras no contexto de FSL, seguimos a metodologia descrita na Seção 2.5, utilizando imagens de GO com resolução de 1752×2016 pixels e sobreposição de 4 pixels, conforme a Seção 2.2. Neste caso, não foi empregada validação cruzada. Realizamos três treinamentos para cada tipo de recorte, gerando os seguintes modelos: recorte I - 6MI, 4MI e 2MI; recorte II - 6MII, 4MII e 2MII.

No treinamento com 6 imagens, obtivemos 252 recortes de 256×256 pixels, sendo 189 utilizados para treinamento e 63 para validação. Com 4 imagens, geramos 168 recortes, dos quais 126 foram usados para treinamento e 42 para validação. Finalmente, com 2 imagens, obtivemos 84 recortes, dos quais 63 foram destinados ao treinamento e 21 à validação.

A etapa de teste, em todos os casos, foi realizada com duas imagens de 1752×2016 pixels, que geraram 84 recortes de 256×256 pixels, sem sobreposição.

O processo de treinamento e validação foi conduzido ao longo de 30 épocas. Para o ajuste fino dos modelos, utilizamos uma taxa de aprendizado reduzida igual a 0,0001, visando otimizar os resultados. A função de perda adotada foi a entropia cruzada binária (1) e as medidas de desempenho foram avaliadas utilizando o conjunto de teste.

A Figura 7 e a Tabela 1 ilustram os resultados sobre o conjunto de teste das imagens de GO para a acurácia (6) e para as médias de precisão, *recall*, *F1-Score* e IoU computadas pelas equações (2), (3), (4) e (5), respectivamente.

Os resultados apresentados na Fig. 7 e Tabela 1 indicam que, utilizando 6 imagens, o M_{II} foi superior em quase todas as métricas, exceto no *recall*. Utilizando 4 imagens, os modelos tiveram desempenho muito semelhante, com pequenas vantagens para o M_{II}. No entanto, o M_I se destacou em situações de menor quantidade de dados (2 imagens de GO), o que pode ser uma consideração importante em aplicações práticas onde a obtenção de grandes volumes de dados é desafiadora.

Tabela 1. Desempenho (%) sobre o conjunto de teste dos modelos gerados com imagens de GO.

Medida	**6 MI**	**4 MI**	**2 MI**
Acurácia	96,99 ± 7,51	96,91 ± 7,80	97,94 ± 4,80
Precisão	88,22 ± 15,55	89,27 ± 16,46	91,86 ± 12,17
Recall	99,75 ± 0,29	98,53 ± 1,41	97,60 ± 1,99
IoU	88,00 ± 15,46	87,91 ± 15,94	89,66 ± 11,54
F1-Score	92,60 ± 12,46	92,48 ± 12,92	94,05 ± 8,21
Medida	**6 MII**	**4 MII**	**2 MII**
Acurácia	97,73 ± 6,03	96,98 ± 7,78	97,48 ± 4,16
Precisão	92,30 ± 15,04	89,96 ± 16,38	90,20 ± 12,87
Recall	97,62 ± 2,03	98,18 ± 2,08	95,02 ± 4,19
IoU	90,02 ± 14,28	88,26 ± 15,77	86,24 ± 12,31
F1-Score	93,93 ± 10,90	92,70 ± 12,79	92,02 ± 9,03

4. Conclusão

Os modelos M_I e M_{II} alcançaram altos valores para as medidas de desempenho consideradas na segmentação de ZnO, com uma superioridade do M_{II} em termos de *recall*. O ajuste fino indicou que o M_I, que utiliza recortes regulares com sobreposição, obteve um desempenho superior para o novo conjunto de dados formado por imagens GO, especialmente em situações críticas com dados limitados para o treinamento. Dessa forma, observamos que a estratégia de recorte II foi mais eficaz no treinamento tradicional, enquanto a estratégia de recorte I resultou em um modelo mais eficiente para o cenário de poucas amostras.

A pesquisa apresenta impactos em diversos contextos. Na nanotecnologia, a abordagem pode ser ampliada para análise de diferentes nanomateriais, contribuindo para a reprodutibilidade nos processos de fabricação e para o desenvolvimento de novos materiais. No campo das ciências biomédicas, pode ser aplicado na detecção de células ou tecidos em imagens microscópicas, auxiliando no diagnóstico precoce de doenças e no acompanhamento de tratamentos, o que representa um impacto significativo na área médica.

O método demonstrou potencial para cenários com poucas amostras. No entanto, observamos na Tabela 1 que os desvios padrão são altos para as métricas analisadas, com exceção do *recall*. Esse fato compromete a significância estatística dos resultados obtidos por ajuste fino nas imagens de GO. Para mitigar esses problemas, em trabalhos futuros, iremos explorar outras arquiteturas de redes neurais e novas técnicas de FSL [1]. Além disso, será desenvolvido um método para análise dimensional e contagem dos nanomateriais, juntamente com uma comparação entre diferentes tamanhos de recortes.

Author contributions

According to the submission guidelines, here is the CRediT authorship statement: Renato de Avila Lopes: Conceptualization, Investigation, Methodology, Resources, Validation, Visualization, Writing - original draft, Writing - review &

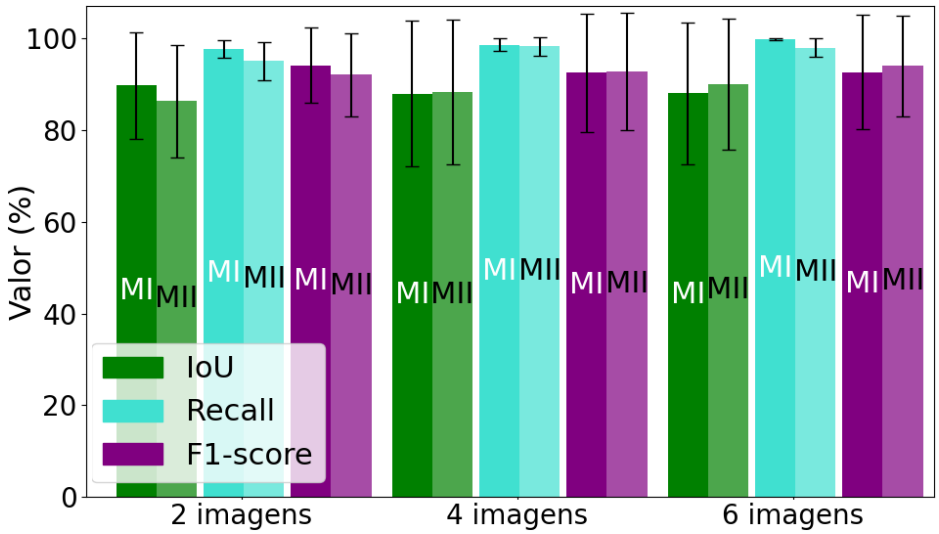


Figura 7. Médias de IoU, *recall* e *F1-Score* obtidas sobre o conjunto de teste dos modelos gerados com imagens de GO.

editing; Marcus Vinicius Leal de Carvalho: Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing; Edson Kitani, Francisco de Assis Zampirolli, Leopoldo Yoshioka, Luiz Antonio Celiberto Junior, and Ugo Ibusuki: Investigation, Validation, Writing - original draft, writing - review & editing.

Referências

[1] WANG, Y. et al. Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Curveys (csur)*, ACM New York, NY, USA, v. 53, n. 3, p. 1–34, 2020. Disponível em: ⟨https://doi.org/10.1145/3386252⟩.

[2] DU, G. et al. Medical image segmentation based on u-net: A review. *Journal of Imaging Science & Technology*, Society for Imaging Science and Technology, v. 64, n. 2, 2020. Disponível em: ⟨10.2352/J.ImagingSci.Technol.2020.64.2.020508⟩.

[3] SCHNEIDER, C. A.; RASBAND, W. S.; ELICEIRI, K. W. Nih image to imagej: 25 years of image analysis. *Nature Methods*, Nature Publishing Group, v. 9, n. 7, p. 671–675, 2012. Disponível em: ⟨https://doi.org/10.1021/acs.jpca.0c01513⟩.

[4] BALS, J.; EPPLE, M. Deep learning for automated size and shape analysis of nanoparticles in scanning electron microscopy. *RSC Advances*, Royal Society of Chemistry, v. 13, n. 5, p. 2795–2802, 2023. Disponível em: ⟨10.1039/D2RA07812K⟩.

[5] SUN, Z. et al. A deep learning-based framework for automatic analysis of the nanoparticle morphology in sem/tem images. *Nanoscale*, Royal Society of Chemistry, v. 14, n. 30, p. 10761–10772, 2022. Disponível em: ⟨https://pubs.rsc.org/en/content/articlelanding/2022/nr/d2nr01029a/unauth⟩.

[6] AKERS, S. et al. Rapid and flexible segmentation of electron microscopy data using few-shot machine learning. *npj Computational Materials*, Nature Publishing Group, v. 7, n. 1, p. 187, 2021. Disponível em: ⟨https://doi.org/10.1038/s41524-021-00652-z⟩.

[7] SIAO, H.-Y. et al. Machine learning-based automatic graphene detection with color correction for optical microscope images. *arXiv preprint arXiv:2103.13495*, 2021. Disponível em: ⟨https://doi.org/10.48550/arXiv.2103.13495⟩.

[8] XIAO, C. et al. Automatic sars-cov-2 segmentation in electron microscopy based on few-shot learning. *International Journal of Wavelets, Multiresolution and Information Processing*, World Scientific Publishing Company, v. 22, n. 02, p. 2350047, 2024. Disponível em: ⟨https://doi.org/10.1142/S0219691323500479⟩.

[9] DELERUYELLE, A.; VERSARI, C.; KLEIN, J. Self-mentoring: A new deep learning pipeline to train a self-supervised u-net for few-shot learning of bio-artificial capsule segmentation. *Computers in Biology and Medicine*, Elsevier, v. 152, p. 106454, 2023. Disponível em: ⟨https://doi.org/10.1016/j.compbiomed.2022.106454⟩.

[10] PROTONOTARIOS, N. E. et al. A few-shot u-net deep learning model for lung cancer lesion segmentation via pet/ct imaging. *Biomedical Physics & Engineering Express*, v. 8, n. 2, p. 025019, 2022. Disponível em: ⟨https://iopscience.iop.org/article/10.1088/2057-1976/ac53bd/meta⟩.

[11] LI, L. Zno sem image segmentation based on deep learning. In: *IOP Conference Series: Materials Science and Engineering*. IOP Publishing, 2020. v. 022035. Disponível em: ⟨https://iopscience.iop.org/article/10.1088/1757-899X/782/2/022035/meta⟩.

[12] SAAIM, K. M. et al. In search of best automated model: Explaining nanoparticle tem image segmentation. *Ultramicroscopy*, v. 233, p. 113437, 2022. Disponível em: ⟨https://doi.org/10.1016/j.ultramic.2021.113437⟩.

[13] GUTIéRREZ, J.; BARRERA, I.; GóMEZ, N. Nanoparticle detection on sem images using a neural network and semi-synthetic training data. *Nanomaterials*, v. 12, n. 11, p. 1818, 2022. Disponível em: ⟨https://doi.org/10.3390/nano12111818⟩.

[14] IEEE. *Few-Shot Learning for Segmentation of Yeast Cell Microscopy Images*. 1–4 p. Disponível em: ⟨https://doi.org/10.1109/SIU53274.2021.9477988⟩.

[15] CHEN, Z. et al. Deep learning-based method for sem image segmentation in mineral characterization, an example from duvernay shale samples in western canada sedimentary basin. *Computers & Geosciences*, v. 138, p. 104450, 2020. Disponível em: ⟨https://doi.org/10.1016/j.cageo.2020.104450⟩.

[16] ELIASSON, H.; ERNI, R. Localization and segmentation of atomic columns in supported nanoparticles for fast scanning transmission electron microscopy. *npj Computational Materials*, v. 10, n. 1, p. 168, 2024. Disponível em: ⟨https://doi.org/10.1038/s41524-024-01360-0⟩.

[17] FARAZ, K. et al. Deep learning detection of nanoparticles and multiple object tracking of their dynamic evolution during in situ etem studies. *Scientific Reports*, v. 12, n. 1, p. 2484, 2022. Disponível em: ⟨https://doi.org/10.1038/s41524-024-01360-0⟩.

[18] OZTURK, O.; SARITÜRK, B.; SEKER, D. Z. Comparison of fully convolutional networks (fcn) and u-net for road segmentation from high resolution imageries. *International Journal of Environment and Geoinformatics*, v. 7, n. 3, p. 272–279, 2020.

[19] MARSLAND, S. *Machine Learning: An Algorithmic Perspective*. Chapman & Hall/CRC, 2009. Disponível em: ⟨https://doi.org/10.1201/9781420067194⟩.

[20] GLOROT, X.; BENGIO, Y. Understanding the difficulty of training deep feedforward neural networks. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Sardinia, Italy: PMLR, 2010. v. 9, p. 249–256. Disponível em: ⟨https://proceedings.mlr.press/v9/glorot10a⟩.

[21] KINGMA, D. P.; BA, J. A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2017. Disponível em: ⟨https://arxiv.org/abs/1412.6980⟩.

[22] ZHANG, Z.; SABUNCU, M. Generalized cross entropy loss for training deep neural networks with noisy labels. In: *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2018. v. 31. Disponível em: ⟨https://proceedings.neurips.cc/paper/2018/hash/f2925f97bc13ad2852a7a551802feea0-Abstract.html⟩.

[23] HOSSIN, M.; SULAIMAN, M. N. A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, v. 5, n. 2, p. 1–11, 2019. Disponível em: ⟨https://doi.org/10.5121/ijdkp.2015.5201⟩.

[24] ALMEIDA, L. R. de et al. Pairwise rigid registration based on riemannian geometry and lie structures of orientation tensors. *Journal of Mathematical Imaging and Vision*, Springer, v. 63, n. 7, p. 894–916, 2021. Disponível em: ⟨https://doi.org/10.1007/s10851-021-01037-z⟩.