

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL

MÉTODO BASEADO EM *MACHINE LEARNING* E *DEEP LEARNING* PARA SEGMENTAÇÃO DA MATÉRIA ORGÂNICA PARTICULADA EM IMAGENS DE MICROTOMOGRAFIA DE RAIOS X DE AGREGADOS DE SOLO

Aline Barbosa de Oliveira

Geóloga

2023

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**MÉTODO BASEADO EM *MACHINE LEARNING* E *DEEP LEARNING* PARA SEGMENTAÇÃO DA MATÉRIA
ORGÂNICA PARTICULADA EM IMAGENS DE
MICROTOMOGRÁFIA DE RAIOS X DE AGREGADOS DE
SOLO**

Aline Barbosa de Oliveira

Orientador: Dr. Ricardo de Oliveira Bordonal

Coorientadora: Dra. Talita Rosas Ferreira

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia (Ciência do Solo)

2023

O48m

Oliveira, Aline Barbosa de

Método baseado em machine learning e deep learning para segmentação da matéria orgânica particulada em imagens de microtomografia de raios X de agregados de solo / Aline Barbosa de Oliveira. -- Jaboticabal, 2023
77 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Ricardo de Oliveira Bordonal

Coorientadora: Talita Rosas Ferreira

1. Agronomia. 2. Estrutura do solo. 3. Matéria orgânica. 4. Processamento de imagens. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

REGISTRO DE IMPACTO

Nesse estudo é proposto um protocolo de segmentação para a fração da matéria orgânica particulada que considera estratégias de aprendizado de máquina, principalmente por meio da aplicação de redes neurais de convolução. Além de proporcionar uma segmentação de boa qualidade, com maior detalhamento e ganhos metodológicos, permitiu uma maior rapidez quando comparada a técnicas tradicionais de segmentação.

IMPACT REGISTER

This study proposes a segmentation protocol for the particulate organic matter fraction that considers machine learning strategies, mainly through the application of convolution neural networks. Besides providing a good quality segmentation, with greater detail and methodological gains, it allowed for greater speed when compared to traditional segmentation techniques.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal



CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: MÉTODO BASEADO EM *Machine Learning* E *Deep Learning* PARA SEGMENTAÇÃO DA MATÉRIA ORGÂNICA PARTICULADA EM IMAGENS DE MICROTOMOGRAFIA DE RAIOS X DE AGREGADOS DE SOLO

AUTORA: ALINE BARBOSA DE OLIVEIRA

ORIENTADOR: RICARDO DE OLIVEIRA BORDONAL

COORIENTADORA: TALITA ROSAS FERREIRA

Aprovada como parte das exigências para obtenção do Título de Mestra em Agronomia (Ciência do Solo), pela Comissão Examinadora:



Documento assinado digitalmente

RICARDO DE OLIVEIRA BORDONAL

Data: 01/03/2023 18:59:58-0300

Verifique em <https://verificador.iti.br>

Pesquisador Dr. RICARDO DE OLIVEIRA BORDONAL (Participação Virtual)

Laboratório Nacional de Biorrenováveis / Centro Nacional de Pesquisa em Energia e Materiais Campinas/SP



Documento assinado digitalmente

LUIZ FERNANDO PIRES

Data: 01/03/2023 20:16:28-0300

Verifique em <https://verificador.iti.br>

Prof. Dr. LUIZ FERNANDO PIRES (Participação Virtual)

Departamento de Física / Universidade Estadual de Ponta Grossa - UEPG

Prof. Dr. NEWTON LA SCALA JUNIOR (Participação Virtual)

Departamento de Engenharia e Ciências Exatas DECEX / FCAV UNESP Jaboticabal

Jaboticabal, 01 de março de 2023

DADOS CURRICULARES DO AUTOR

ALINE BARBOSA DE OLIVEIRA – Filha de Nilson Barbosa de Oliveira e Rosileide Cassimiro de Oliveira, nascida em 22 de setembro de 1994, em Rio Claro, estado de São Paulo, Brasil. Em 2015, ingressou no curso de graduação em Geologia na Universidade Estadual de Campinas (UNICAMP), com conclusão em 2021. Atuou como bolsista ou voluntária no Programa de Apoio Didático (PAD), na condição de monitora, nas disciplinas “Introdução ao Sensoriamento Remoto”, em 2016, “Fundamentos de Paleontologia”, em 2019, e “Elementos de Paleontologia”, em 2020. Foi bolsista de Iniciação Científica pelo Conselho Nacional de Desenvolvimento Tecnológico (CNPq) nos anos de 2018-2019 com o projeto “Tafonômia e análise de *charcoals* como ferramenta para caracterização paleoambiental da Formação São Carlos, SP (Cretáceo Superior, Bacia Bauru, Brasil)” e 2020-2021 com o projeto “Fracionamento físico do carbono em solos agrícolas cultivados com cana-de-açúcar sob diferentes níveis de remoção de palha” com vínculo de pesquisadora colaboradora no Laboratório Nacional de Biorrenováveis do Centro Nacional de Pesquisa em Energia e Materiais (LNBR/CNPEM). Em 2021, ingressou no curso de mestrado em Agronomia (Ciência do Solo), como bolsista pela Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES), na Faculdade de Ciências Agrárias e Veterinárias (FCAV/UNESP), Câmpus de Jaboticabal-SP, com vínculo como pesquisadora colaboradora no LNBR/CNPEM, participando do grupo de pesquisa “Biotecnologia para Agricultura”.

(...)

*O que importa se o pneu furou
O que importa é se vai decolar
O que importa se escorregou
O que importa é se levantar
O que importa se a noite esfriou
O que importa é saber amar*

*O estado contente da mente
Depende somente
De acreditar*

(...)

*Sou feliz, alegre e forte
Tenho amor e sorte
Aonde quer que eu vá*

Marisa de Azevedo Monte

AGRADECIMENTOS

A Deus, por me capacitar e iluminar nos desafios de cada dia.

Aos meus pais por cada palavra de motivação e incentivo, por estarem sempre ao meu lado e acreditarem em mim.

Ao meu companheiro Felipe, por aguentar todas as minhas crises de ansiedade e angústias. Não é possível colocar em palavras quão importante você foi durante todos esses anos e como se tornou minha inspiração para almejar sempre mais. Obrigada pelas comidas gostosas e motivadoras!

Ao meu orientador Dr. Ricardo de Oliveira Bordonal pela orientação desde a iniciação científica e auxílio para o desenvolvimento deste trabalho.

A minha coorientadora Dra. Talita Rosas Ferreira pelo compartilhamento de muitos dos seus conhecimentos, por me ensinar “o caminho das pedras”, pela paciência e dedicação no desenvolvimento de cada etapa deste trabalho, além de ser uma inspiração de cientista mulher conquistando o seu espaço.

Ao pesquisador Dr. João Luis Nunes Carvalho pelas dicas e palavras de motivação.

Ao professor Dr. Alan Rodrigo Panosso, vice-coordenador do Programa de Pós-Graduação em Agronomia (Ciência do Solo), pelo auxílio nos protocolos, e a todos os professores do programa que contribuíram para o meu desenvolvimento e formação.

Aos meus amigos de laboratório Fernanda, Sarah e Marcelo, agradeço pelos momentos de café (e docinhos), conversas e desabafos, vocês deixaram os meus dias muito mais leves e felizes!

As minhas amigas irmãs Helena e Ieda, que me acompanham desde a graduação. Obrigada por sempre me incentivarem e estarem presentes em cada momento da minha vida.

Ao Laboratório Nacional de Biorrenováveis (LNLS) e Laboratório Nacional de Luz Síncrotron (LNLS) do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM) por fornecerem a estrutura para desenvolvimento deste estudo, além da aquisição das imagens (proposta IMX20180205) e processamento (proposta 20210070).

Ao grupo Tepuí, Grupo de Computação Científica (GCC) e a Kerolyn da Silva Araújo do Serviço de Apoio ao Usuário (SAU) do LNLS, pelo suporte e apoio com as máquinas (HPC), softwares e agendamentos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

SUMÁRIO

	Página
1 INTRODUÇÃO	1
2 REVISÃO DE LITERATURA	5
2.1 Frações da matéria orgânica do solo	5
2.2 Papel da estrutura do solo para a proteção e estabilização da MOP no solo	6
2.3 Avaliação da MOS por μ TCX	8
2.3.1 Segmentação de imagens	10
3 MATERIAL E MÉTODOS	13
3.1 Descrição da área experimental e coleta de amostras	13
3.2 Aquisição das imagens por μ TCX.....	14
3.3 Processamento das imagens de μ TCX.....	15
3.3.1 Remoção do <i>background</i>	18
3.3.2 Construção da base de dados de segmentação da MOP.....	21
3.3.3 <i>Deep learning: dataset</i>	25
3.3.4 <i>Deep learning: network</i>	30
3.3.5 Inferência para a classificação da MOP.....	31
3.4 Proposta de avaliação para o resultado dos treinamentos (estimativa de acurácia)	32
4 RESULTADOS.....	33
4.1 Remoção do background	33
4.2 Treinamento das redes neurais e inferências para a classificação da MOP (via <i>deep learning</i>)	33
4.3 Estimativa de acurácia das redes neurais.....	41
5 DISCUSSÃO	42
6 CONCLUSÃO.....	50
7 PERSPECTIVAS FUTURAS	50
8 REFERÊNCIAS.....	51
APÊNDICE	57

MÉTODO BASEADO EM MACHINE LEARNING E DEEP LEARNING PARA SEGMENTAÇÃO DA MATÉRIA ORGÂNICA PARTICULADA EM IMAGENS DE MICROTOMOGRAFIA DE RAIOS X DE AGREGADOS DE SOLO

RESUMO – A identificação da distribuição espacial da fração da matéria orgânica particulada (MOP) em agregados de solo é de fundamental importância para decifrar os mecanismos que determinam o aumento da estocagem e estabilidade do carbono no solo. A localização do carbono no agregado de solo pode afetar se ele permanecerá protegido ou será consumido por microrganismos que habitam o solo, por isso a necessidade de buscar compreender o papel da estrutura do solo na proteção física do carbono. O uso da técnica de Microtomografia Computadorizada de Raios X (μ TCX) tem proporcionado avanços na investigação *in situ* de processos relacionados à decomposição e proteção física da fração MOP no solo. No entanto, uma dificuldade que limita o progresso do uso dessa técnica está na identificação (ou segmentação) da MOP em imagens de μ TCX, especialmente quando não são realizados procedimentos de marcação química dessa fração. Por este motivo, o presente estudo tem como objetivo propor um método de segmentação da fração MOP. Buscou-se construir uma base de dados de fragmentos orgânicos, contidos em imagens de μ TCX, segmentados manualmente, utilizando o software Avizo, ou por meio de estratégias de *machine learning*, no software Annotat3D, para treinar uma rede neural de segmentação da MOP via *deep learning*. As imagens 3D consideradas para desenvolvimento do método foram adquiridas na linha de luz de μ TCX do Laboratório Nacional de Luz Síncrotron, com tamanho de pixel de 0,82 μ m. Para o treinamento das redes, considerou-se: o uso de um recurso de criação de dados artificiais (*augmentation*) e a mistura de dois conjuntos de imagens diferentes para diversificação dos dados de entrada. As duas melhores redes neurais treinadas obtiveram identificação parcial ou total dos fragmentos orgânicos nas imagens de μ TCX com uma estimativa de acurácia de 57% e 69%. Os resultados mostraram que o método foi bem sucedido para a segmentação da fração MOP em 17 agregados de solo, sem a utilização de marcação química. O recurso *augmentation* proporcionou melhor continuidade dos fragmentos orgânicos, e a mistura de imagens de diferentes conjuntos resultou em maior número de fragmentos identificados. Portanto, o método apresenta potencial para ser empregado em diferentes propostas experimentais com objetivos semelhantes, uma vez que a rede neural treinada pode ser adaptada e otimizada para a sua aplicação em diferentes conjuntos de imagens.

Palavras-chave: Processamento de imagem; rede neural de convolução; anotação; Annotat3D; estabilidade do carbono no solo

MACHINE LEARNING AND DEEP LEARNING BASED METHOD FOR SEGMENTATION OF PARTICULATE ORGANIC MATTER IN X-RAY MICROTOMOGRAPHY IMAGES OF SOIL AGGREGATES

ABSTRACT – Identifying the spatial distribution of particulate organic matter (POM) fraction in soil aggregates is fundamental to deciphering the mechanisms that determine the increased of carbon storage and stability in soils. The location of carbon in the soil aggregate can affect whether it remains protected or is consumed by soil-inhabiting microorganisms, hence the need to seek an understanding of the role of soil structure in the physical protection of carbon. The use of X-ray Computed Microtomography (μ CT) has provided advances in the in situ investigation of processes related to the decomposition and physical protection of POM in soil. However, one difficulty limiting progress in the use of this technique is the identification (or segmentation) of POM in μ CT images, especially in the absence of chemical staining procedures. For this reason, the present study aims to propose a method for direct segmentation of POM fraction, without chemical staining. We sought to build a database of organic fragments, contained in μ CT images, segmented manually using the Avizo software, or by machine learning strategies, in the Annotat3D software, in order to train a neural network for POM segmentation through deep learning. The 3D images considered for method development were acquired on the μ CT IMX beamline of the Brazilian Synchrotron Light Laboratory, with a pixel size of 0.82 μ m. For the training of the networks, we considered: the use of an artificial data creation feature (augmentation) and the mixture of two different image sets for diversification of the input data. The two best trained networks obtained partial or total identification of organic fragments in the μ CT images, with an estimated accuracy of 57% and 69%. The results showed that the method was successful in segmenting the POM fraction in 17 soil aggregates without the use of chemical staining procedure. The *augmentation* feature provided better continuity of organic fragments, and the mixing of images from different sets resulted in a greater number of identified fragments. Therefore, the method has the potential to be employed in different experimental proposals with similar objectives, since the trained neural network can be adapted and optimized for its application on different sets of images.

Keywords: Image processing; convolution neural network; annotation; Annotat3D; carbon stability in soil

LISTA DE TABELAS

Tabela 1. Parâmetros de treinamento definidos para a segmentação via deep learning no software Annotat3D para cada rede neural (A1, A2, B1 e B2).	31
Tabela 2. Número de fatias anotadas por amostra para remoção do background e tamanho do training data associado.	33
Tabela 3. Valores da função total loss relacionados ao treinamento e a validação para o número de iterações escolhido para cada rede neural (A1, A2, B1 e B2).	40
Tabela 4. Estimativa de acurácia obtida para cada rede neural (A1, A2, B1 e B2) considerada no estudo.	42

LISTA DE FIGURAS

Figura 1. Exemplo de um histograma genérico para ilustrar uma fase com tons de cinza intermediários como a MOP. Os picos A e B representam a matriz e os poros de uma imagem de μ TCX em tons de cinza. É possível observar que as curvas correspondentes a esses picos se sobrepõem e é nessa região de sobreposição que se encontram os tons de cinza intermediários que representam a MOP, por isso a dificuldade em propor um método de segmentação para essa fase baseado apenas em tons de cinza. Adaptado de Perry (2013).	10
Figura 2. Esquema do fluxo de trabalho para segmentação da MOP. A filtragem consiste em uma etapa opcional (setas tracejadas).	16
Figura 3. Sequência de trabalho para a remoção do background ("Sample 2", fatia 500). (a) Exemplo de roi box para corte das fatias vazias com eixos X, Y e Z da imagem, onde Z representa o eixo de rotação da amostra na medida de μ TCX. (b) Imagem grayscale após a reconstrução, com o fundo e a porção correspondente ao agregado de solo indicados na imagem. (c) As linhas rosadas definem a malha de superpixels gerada (interessante conferir se os limites respeitam as bordas do objeto que se pretende segmentar, retângulo em destaque na imagem). (d) Anotações para definir as diferentes classes da imagem: anotações em rosa representam o fundo e anotações em verde o agregado. (e) Resultado da segmentação via anotação após a aplicação do classificador para todo o volume e salvamento do arquivo training data. (f) Resultado da correção feita a partir dos recursos de segmentação do software Avizo 3D, após aplicação do fill holes para todas as fatias e do watershed. (g) Resultado da máscara para remoção do background.	20
Figura 4. Sequência de trabalho para segmentação manual da MOP. (a) Exemplo de seleção de um subvolume após a remoção do background. (b) Imagem grayscale ("Sample 2", subvolume com dimensão de 413×290×257). Nessa imagem são identificados visualmente 3 fragmentos de MOP devido a sua morfologia diferente em relação a matriz. (c) Desenho do contorno/limite para cada um dos fragmentos em cada fatia através da ferramenta magic wand (draw limit line). (d) Definição do melhor intervalo de seleção (representado na cor roxa) para separar a parte orgânica dos poros contidos dentro do limite previamente definido. (e) Seleção adicionada à classe MOP (representada em verde) em uma imagem label – a imagem label contendo a MOP está sobreposta à imagem grayscale.	22
Figura 5. Sequência de trabalho para segmentação via anotação da MOP. (a) Imagem grayscale ("Sample 17", subvolume com dimensão de 883×508×700) com os desenhos dos contornos para cada um dos fragmentos em cada fatia através da ferramenta magic wand	

(draw limit line). (b) Seleção da região correspondente ao contorno (MOP e poros pertencentes a essa estrutura). (c) Criação de uma imagem label da região selecionada (em amarelo). (d) Resultado da máscara do contorno (remoção do background e matriz). (e) No Annotat3D, após definir os parâmetros do classificador para segmentação a nível de pixel, são feitas anotações para identificar as diferentes classes. Anotações em rosa correspondem ao background, em verde correspondem à região orgânica da MOP, e em roxo correspondem aos poros. (f) Resultado da segmentação após a aplicação do classificador para toda a imagem. (g) Label resultante da segmentação da MOP sobreposta à imagem grayscale, sem a máscara do contorno. 24

Figura 6. Esquema representando a criação do dataset para segmentação via deep learning. (a) Agregado de solo ("Sample 17") com subvolumes definidos a partir da escolha de regiões com maior densidade de MOP e diferentes morfologias para diversificar os dados de entrada. Imagens que compõem o dataset: (b) imagem grayscale (subvolume com dimensão de 883×508×700), (c) label com definição de valores para cada classe (o valor 1 representando o background em toda a região azul clara da imagem, 0 a matriz em toda a região preta e 2 a MOP em toda a região azul escura) e (d) mapa de pesos com definição de um peso para cada área específica da imagem (valor 1 para a região pertencente ao background, 20 à matriz e 255 à MOP). 26

Figura 7. Renderizações 3D da MOP correspondente a base de dados construída a partir da segmentação manual, indicados pela letra "M", ou via anotação, indicados pela letra "A", para treinamento das redes neurais. 27

Figura 8. Efeito da aplicação do augmentation nas imagens durante a criação do dataset. Agregado de solo ("Sample 17") representando as fatias selecionadas para ilustrar a mudança de textura entre a (a) porção superior da imagem com textura mais nítida e a (b) porção inferior com textura menos nítida (embaçada). (c) Aplicação do augmentation Gaussian Blur (interface do Annotat3D para criação do dataset) e definição dos melhores intervalos de desfoque para o conjunto de dados, de modo que o máximo se aproximasse com o observado nas imagens (a partir do resultado na imagem de preview, subvolume "Sample 2"). 29

Figura 9. Comparação dos resultados para cada uma das redes treinadas (A1, A2, B1 e B2). Esquema com agregado de solo ("Sample 19", inferência de "teste cego") e fatias exemplificando a (a) porção superior da imagem (fatia 1512) com textura mais nítida e a (b) porção inferior da imagem, com textura mais embaçada (fatia 537). MOP segmentada nas fatias das porções superior e inferior, respectivamente, pelas redes (c, d) A1, (e, f) A2, (g, h) B1 e (i, j) B2. 36

Figura 10. Renderização 3D da MOP, para a amostra "Sample 19", segmentada a partir das inferências das redes (a) A1, (b) A2, (c) B1 e (d) B2. Linha vermelha tracejada indicando a região superior e inferior da imagem. 37

Figura 11. Gráficos de acompanhamento da função total loss no decorrer das iterações para o treinamento e validação das redes neurais (a) A1, (b) A2, (c) B1 e (d) B2. 39

Figura 12. Renderizações 3D da MOP resultantes das inferências da rede neural B2 para as 17 imagens do conjunto. 41

1 INTRODUÇÃO

Atualmente, busca-se promover a rápida ampliação e implementação do uso de práticas agrícolas que permitam a proteção do (C) no solo, com o objetivo de contribuir para a mitigação das mudanças climáticas (Amelung et al., 2020). Os solos são o maior e mais dinâmico reservatório de C terrestre, armazenando mais C na forma de matéria orgânica do solo (MOS) comparado aos reservatórios atmosféricos ou vegetal combinados (Lehmann e Kleber, 2015), podendo remover entre 0,79 e 1,54 Gt C ano⁻¹ da atmosfera. Por essa razão, aumentar a estocagem de C no solo tem se tornado parte de uma agenda global para a mitigação e adaptação às mudanças climáticas. Entre as importantes iniciativas neste contexto estão o "4p1000", lançado na COP21 pela UNFCCC em 2015, os seminários de Koronivia sobre agricultura, iniciados na COP23 em 2018, e o programa RECSOIL lançado pela FAO (Baveye et al., 2020; Rumpel et al., 2020).

A matéria orgânica particulada (MOP) é a fração ativa da MOS, podendo ser definida como partículas orgânicas que variam entre 0,053 e 2 mm, e constituída por fragmentos orgânicos (como folhas ou raízes) parcialmente decompostos. Encontra-se facilmente disponível para os microrganismos, sendo decomposta mais rapidamente e, por isso, é considerada um estoque transitório da MOS. Portanto, por estar menos protegida e ser mais susceptível às alterações no ambiente, o C associado à MOP se torna mais vulnerável às perdas durante o uso e manejo do solo (Lavalley et al., 2020; Cotrufo et al., 2013).

A MOP é importante por propiciar muitos benefícios agrícolas e ecológicos, servindo como fonte de alimento e energia para microrganismos e animais do solo, além de fornecer nutrientes para o crescimento das plantas e melhorar a estabilidade dos agregados, a infiltração de água e aeração do solo (Gosling et al., 2013). A proteção física da MOP dentro dos agregados é um dos principais mecanismos de sequestro de C pelo solo. Portanto, aprofundar o entendimento dos mecanismos que determinam a dinâmica e a distribuição espacial da MOP em agregados de solo é fundamental para assegurar as diversas funções que essa fração da MOS exerce na saúde do solo, incluindo seu potencial para o sequestro de C.

O recente desenvolvimento de técnicas sofisticadas de imagem, atingindo escalas de resolução espacial microscópica, tem levado a um progresso substancial na descrição da arquitetura do solo, tanto em termos de espaço poroso quanto da configuração espacial da fração mineral ou materiais orgânicos de forma direta (Schlüter et al., 2022). No entanto, análises em grandes volumes de solo e em número de repetições adequadas de amostras ainda são limitadas pela complexidade dos procedimentos propostos (Steffens et al., 2017). Essas técnicas permitiram o avanço no entendimento dos processos-chave (de proteção e/ou perda do C) para aumentar a persistência de C no solo, tais como o papel da proteção física da MOS (Lehmann e Kleber, 2015; Kravchenko e Guber, 2017). Na escala micrométrica, a estrutura do solo estabiliza o C dentro de inúmeros arranjos espaciais e canais de poros do solo, os quais governam diversos processos, incluindo a difusão de gases, água, compostos químicos dissolvidos, produtos de decomposição, enzimas e microrganismos na solução do solo (Six et al., 2004). Schlüter et al., 2022 mostraram que as frações da MOS derivadas de imagens de Microtomografia Computadorizada de Raios X (μ TCX) refletiram com precisão as frações de C obtidas através do fracionamento físico convencional do solo, como a MOP, com o benefício adicional de quantificar sua distribuição tridimensional em relação ao espaço poroso não perturbado.

Imagens de μ TCX voltadas para pesquisa científica podem ser obtidas a partir de tomógrafos industriais, de bancada ou baseados em luz síncrotron. No Brasil, está sendo desenvolvido o Sirius, uma nova fonte de luz síncrotron de 4ª geração no Laboratório Nacional de Luz Síncrotron do Centro Nacional de Pesquisa em Energia e Materiais (LNLS/CNPEM). No Sirius, haverá uma linha de luz de micro e nanotomografia (chamada MOGNO) cujo design proporcionará ganhos significativos em estudos de amostras de solo, como as possibilidades de se realizar experimentos com grande número de repetições de amostra (por ser uma linha de aquisição rápida de imagens 3D), resolvidos no tempo (4D) (Ferreira et al., 2022a) e em multiescala (Claro et al., 2023), com resolução espacial máxima de 100 nm. Tendo em vista o seu potencial para aplicações na área da agricultura, um dos desafios atualmente consiste em propor uma metodologia que seja precisa e eficiente (isto é, automatizada/rápida)

para a segmentação de imagens de amostras de solo e que seja acessível aos futuros usuários da linha de luz.

As técnicas de segmentação (classificação) de imagens são fundamentais para a análise de imagens em diversas áreas do conhecimento e visam particionar uma imagem digital em segmentos, a fim de definir com precisão os limites espaciais de objetos de interesse nas imagens. No contexto de imagens de μ TCX, a segmentação consiste em atribuir classificações coerentes e precisas a determinados grupos de tons de cinza na imagem, os quais devem representar uma determinada fase (ou classe) de interesse. No entanto, diferentes fases podem apresentar tons de cinza similares. Por exemplo, a MOP apresenta uma atenuação intermediária dos raios X entre a fase de poros e da matriz sólida, resultando em tons de cinza similares a estas duas fases, mas predominantemente similares aos da matriz sólida. Nesse caso, outras características da MOP nas imagens de tomografia, como a forma, tamanho e textura, podem auxiliar na sua identificação visual. Porém, o processo de segmentação com critérios que levem em consideração essas características é mais complexo e desafiador do que o processo já bem conhecido baseado em separação (*threshold*) simples de tons de cinza (Ferreira et al., 2022b).

Existem formas de tornar a segmentação da MOP direta por *threshold* através de técnicas que possibilitem a sua marcação química. Peth et al. (2014) desenvolveram uma abordagem para localizar a MOS por meio de uma combinação entre a μ TCX baseada em luz síncrotron e o ósmio (utilizando a fase vapor) como agente de marcação química da MOS (Quigley et al., 2018; Rawlins et al., 2016). Embora seja um método que permita a identificação da MOP, o tetróxido de ósmio (OsO_4) é extremamente tóxico, reativo e altamente volátil, trazendo riscos à saúde do operador e ao ambiente, por isso há estudos que buscam encontrar outros agentes promissores para marcação da MOP (Maenhout et al., 2021). Uma tentativa proposta por (Arai et al., 2019) é o uso do OsO_4 em estado líquido (sob baixa temperatura para não ocorrer a volatilização) para buscar reduzir os riscos relacionados à sua natureza altamente tóxica. Ainda é possível pontuar limitações associadas ao método que impedem sua aplicação para um grande conjunto de amostras ou para uma ampla variedade de condições de solo (Zheng et al., 2020) e até mesmo em relação à

precisão da marcação química da MOS, pois o OsO_4 tem mostrado uma contribuição e afinidade com a fração mineral do solo (Schlüter et al., 2022).

Mais recentemente estão sendo propostos avanços para a segmentação de imagens por meio de estratégias de *machine learning* e *deep learning*. Essas estratégias objetivam proporcionar a segmentação de uma imagem 3D a partir do treinamento de um modelo para classificar pixels como pertencentes a uma classe (como em Rippner et al., 2022; Schlüter et al., 2022; Soltaninejad et al., 2020). Os algoritmos de *machine learning* e *deep learning* realizam a extração de recursos e a classificação de imagens através de modelos de menor ou maior grau de complexidade, respectivamente. Estes algoritmos se diferenciam ainda pelo tipo e quantidade de dados de entrada para o treinamento do modelo, bem como pela forma de entrada dos dados. Mais especificamente, os algoritmos de *machine learning* são mais dependentes da interação com o operador, exigindo dedicação contínua para o desenvolvimento do modelo. Por outro lado, os algoritmos de *deep learning* recebem dados prontos para o desenvolvimento do modelo, isto é, são menos interativos. Ambos os tipos de modelos, quando bem ajustados, podem ser capazes de segmentar grandes volumes de imagens com alto grau de precisão. Entre os softwares que proporcionam abordagens de segmentação baseada em *machine learning* e/ou *deep learning* estão o Annotat3D, Weka Segmentation (plugin do ImageJ/Fiji), Ilastik (Berg et al., 2019), Dragonfly, SuRVoS (Pennington et al., 2022) e Avizo. No caso da MOP, essas estratégias podem ser utilizadas com ou sem o uso de marcação química.

No presente trabalho, será dedicada maior atenção ao software Annotat3D, o qual foi utilizado anteriormente em Ferreira et al. (2022b) para o desenvolvimento de um protocolo de segmentação da fase de poros em imagens tomográficas de agregados de solo. Portanto, o objetivo é propor um método de segmentação da MOP, sem marcação química, a partir da: i) elaboração de um banco de dados de fragmentos orgânicos segmentados via *machine learning* e manualmente, que servirá como referência (dados observados) para o ii) treinamento de uma rede neural (modelo) capaz de iii) reproduzir a classificação de referência e vi) classificar novos dados semelhantes. Com esse protocolo pretende-se propor uma forma eficiente e precisa de segmentar a MOP, sem a necessidade de marcação química, a fim de

diminuir o tempo de operação necessário e consequentemente, permitir um aumento de repetição amostral em pesquisas futuras.

2 REVISÃO DE LITERATURA

2.1 Frações da matéria orgânica do solo

O C pode ser encontrado em diferentes frações no solo que desempenham distintas funções para o desenvolvimento das plantas (fornecendo nutrientes para microbiota e plantas), na estrutura do solo e no sequestro de C. Existem diversos tipos de fracionamento da MOS que buscam separá-la em frações que sejam homogêneas quanto a sua natureza, dinâmica e função no solo. No fracionamento químico, que constituíram os primeiros esforços de separação da MOS em componentes significativos, buscou-se uma melhor caracterização e identificação química dos componentes da MOS a partir do estudo de substâncias húmicas isoladas por extrações alcalinas. Já o fracionamento físico permite estudos sobre a função e estrutura da MOS a partir de métodos relacionados a diferenças de tamanho e densidade. Assim, a MOS pode ser separada em MOP e matéria orgânica associada a minerais (MOAM) que podem ser diferenciadas pela sua formação e persistência no solo.

A MOP é a fração constituída por fragmentos orgânicos como restos de plantas e animais parcialmente decompostos pelos microrganismos do solo, resultando em pequenas partículas orgânicas (variando entre 0,053 e 2 mm) e, por esse motivo, acabam ficando facilmente disponíveis para os microrganismos (Lavallee et al., 2020; Cotrufo et al., 2013). A MOP é um componente chave dos estoques de C do solo podendo compreender de 6 a 37% do C orgânico total do solo (SOC) em solos agrícolas (Gregorich et al., 2006). A MOAM é a fração derivada principalmente de corpos e subprodutos dos microrganismos presentes no solo (quimicamente transformados pela biota do solo) e de certos compostos vegetais (lixiviados diretamente do material vegetal), e se encontra em um estágio avançado de decomposição pela comunidade microbiana na forma de moléculas orgânicas microscópicas que possuem uma alta afinidade com a superfície dos minerais

(Lavallee et al., 2020; Cotrufo et al., 2013). Dentre as principais diferenças entre a MOP e a MOAM, destaca-se a permanência de cada uma delas no solo. A persistência da MOP no solo é controlada principalmente pela proteção física em agregados, enquanto a da MOAM é controlada principalmente pela associação mineral.

A MOP entra no solo a partir da serrapilheira (camada orgânica superficial) e da rizosfera, por meio da fragmentação de raízes. Seu acúmulo é geralmente considerado um estágio inicial para o sequestro de C nos solos, assim conhecer as concentrações da MOP pode ser útil para entender o estado de fertilidade e a estabilidade do C nos solos agrícolas (Kravchenko et al., 2014). Além disso, a MOP pode ser um indicador sensível das mudanças na qualidade do solo resultantes das práticas de manejo e uso da terra.

2.2 Papel da estrutura do solo para a proteção e estabilização da MOP no solo

A estrutura do solo é o arranjo tridimensional dos constituintes sólidos do solo e vazios em diferentes escalas (Rabot et al., 2018), resultante de interações de fatores bióticos e abióticos, incluindo clima, composição mineral, matéria orgânica, raízes, hifas fúngicas, fauna do solo e diferentes sistemas de manejo. As funções ecossistêmicas essenciais do solo, como armazenamento e transporte de água, ciclagem de nutrientes, armazenamento de C, crescimento vegetal e manutenção da biodiversidade, são em grande parte governadas pela estrutura do solo (Schlüter et al., 2020). Nas últimas décadas, tornou-se mais evidente a importância dos mecanismos relacionados à estrutura do solo, sendo reconhecida a importância da proteção física como um dos principais mecanismos de sequestro de C em longo prazo juntamente à proteção via associações organo-minerais, em relação a recalcitrância química dos compostos orgânicos (Lehmann e Kleber, 2015). Assim, modificações na rede de poros do solo podem beneficiar o acúmulo de C ou intensificar a sua decomposição.

A proteção física do C no solo ocorre quando a matéria orgânica não é decomposta por estar inacessível aos microrganismos dentro dos agregados do solo, que atuam como uma barreira física para a atividade microbiana possibilitando o estoque da MOS. Kravchenko e Guber (2017) ressaltam a importância da arquitetura de poros (isto é, tamanho dos poros, formas, conectividades, tortuosidades e outras

características morfológicas) para os processos que influenciam a estabilidade ou decomposição do C, pois consideram a arquitetura de poros como o principal regulador dos fluxos de gases (difusão de gás em poros preenchidos por ar) e líquidos no solo (transporte de compostos químicos dissolvidos como enzimas, nutrientes e produtos de decomposição do material orgânico, além de microrganismos presentes na solução do solo). Por isso, os poros são os locais de troca de água, gases e outros recursos, ao mesmo tempo que fornecem suporte estrutural e abrigo para os organismos presentes no solo, que irão direta ou indiretamente, estar relacionados a diferentes processos que determinam a proteção física da MOS.

Pesquisas avançaram no conceito de proteção do C via oclusão intra-agregado, por meio da distribuição de agregados de diferentes tamanhos, especialmente menores que 250 μm (Six et al., 2000), e seus possíveis teores de C. Trata-se de um mecanismo de proteção muito importante para o sequestro de C, principalmente quando se considera o impacto dos sistemas agrícolas na estrutura do solo. Porém é necessário avançar no entendimento dos mecanismos que governam a acessibilidade de C pela atividade microbiana, que é impulsionada pelas características da rede de poros, para uma melhor compreensão de como o C pode estar protegido no solo. Assim, para modelar a presença, ausência ou a acessibilidade física de um substrato orgânico aos decompositores é necessário quantificar as rotas pelas quais esse substrato poderia ser acessado pelos microrganismos – por isso a importância de se conhecer a rede de poros (Kravchenko e Guber, 2017). É implícito que as distribuições de tamanho dos agregados refletem a presença e as características dos poros circundantes (Ashman et al., 2003; Dexter, 1988; Elliott e Cambardella, 1991; Young et al., 2001) e, já é conhecido que a oclusão do C intra-agregado minimiza as vias pelas quais ele pode ser acessível. No entanto, nem a descrição quantitativa dessas rotas, nem a avaliação de sua eficácia pode ser obtida a partir de informações sobre as distribuições de tamanho e teores totais de C dos agregados do solo, evidenciando a necessidade de uma quantificação mais específica. Desse modo, a caracterização dos atributos do espaço poroso do solo não perturbado é uma abordagem promissora para relacioná-lo com as funções do solo (Rabot et al., 2018).

A proteção física do C em uma escala de 5-1000 μm pode estar associada à presença de diversos microambientes que geram uma ampla gama de condições

limitantes para a acessibilidade de C por decompositores microbianos e que podem influenciar na atividade microbiana (Kuz'yakov e Blagodatskaya, 2015). É nessa escala que a proteção do C no solo torna-se, em grande parte, uma função das barreiras físicas entre os decompositores e as substâncias orgânicas e/ou uma função das condições locais criadas, em cada tipo de poro, que incapacitam os decompositores (Ekschmitt et al., 2008). Já em escalas menores que 5 μm , as propriedades dos microambientes são determinadas pelas características mineralógicas e superficiais das partículas do solo.

Kravchenko et al. (2019) sugerem que o processamento e a proteção do C aconteçam em diferentes locais da matriz do solo, sendo necessário que o C seja transportado para os locais de proteção antes de ocorrer de fato a sua proteção e, que esse mecanismo, está associado a poros na faixa de 30-150 μm – locais preferenciais de novos inputs de C e de comunidades microbianas ativas, facilitando o processamento dos novos inputs de C no solo. Assim, quanto maior for o volume da matriz do solo em contato com esses poros, maior será o potencial para os produtos da decomposição microbiana serem transportados e protegidos dentro da matriz do solo.

Portanto, o ganho ou perda de C no solo está relacionado ao equilíbrio entre a decomposição microbiana de novos aportes de C, seja como resíduos de plantas ou compostos que foram processados pelos organismos localizados na matriz do solo, e a sua proteção. O material orgânico que não for perdido como CO_2 para a atmosfera depois de ser consumido pela comunidade microbiana, pode escapar de novos processos de decomposição por meio de associações organo-minerais (relacionado também com a textura do solo), por aprisionamento em pequenos poros do solo ou em regiões da matriz que estejam protegidas da ação de microgarnismos.

2.3 Avaliação da MOS por μTCX

Os métodos analíticos convencionais para o estudo da estabilidade da MOS são realizados a partir de amostras pré-processadas, ou seja, amostras que são moídas e posteriormente peneiradas. Apesar de a moagem e o peneiramento permitirem uma maior uniformidade das características gerais do solo, esse processo destrói completamente a sua estrutura, que de acordo com Kravchenko e Guber

(2017), é um componente essencial para promover a proteção física do C no solo. Desse modo, é importante o uso de técnicas que permitam a utilização de amostras indeformadas, as quais possibilitam a investigação *in situ* dos mecanismos envolvidos na decomposição e proteção da MOP com a estrutura intacta do agregado, ou seja, como naturalmente ocorre a proteção física do C. Entre estas técnicas estão a Tomografia de Raios X, Tomografia de Raios Gama, Tomografia de Nêutrons, a Ressonância Magnética Nuclear (com ressalvas para uso em solos com óxido de ferro) e a Microscopia Eletrônica de Varredura (mas requer a preparação das amostras em seções finas).

A Microtomografia Computadorizada de Raios X (μ TCX) é uma técnica que permite quantificar a heterogeneidade nos agregados de solo, possibilitando modelar os processos envolvidos na proteção física do C devido ao imageamento de amostras indeformadas. A μ TCX se baseia na transmissão de raios X por uma determinada amostra que possuem diferentes componentes (fases) para os quais podem ser associados diferentes valores de coeficiente de atenuação a partir da interação com a radiação. No caso do solo, de maneira geral, estas fases são: poros, água, matriz sólida e matéria orgânica.

Os diferentes coeficientes de atenuação que constituem a amostra de interesse aparecem na sua respectiva imagem de tomografia como diferentes tons de cinza, os quais representam, portanto, as diversas fases da amostra. Em uma imagem de tomografia, os tons de cinza mais escuros estão associados às fases que menos absorvem os raios X (espaço poroso do solo) e os tons de cinza mais claros estão associados às fases que mais os absorvem (matriz sólida do solo). Entretanto, as fases que apresentam coeficientes de atenuação intermediários, como é o caso da matéria orgânica no solo, usualmente apresentam tons de cinza similares às outras fases (por vezes similares aos tons relacionados aos poros e por vezes aos tons relacionados à matriz sólida), tornando difícil a distinção ao olho humano (como exemplificado pela área de sobreposição entre os picos A e B na Figura 1). Portanto, o desafio em aplicar a μ TCX em estudos da MOP está na segmentação (classificação) desta fase em relação às demais.

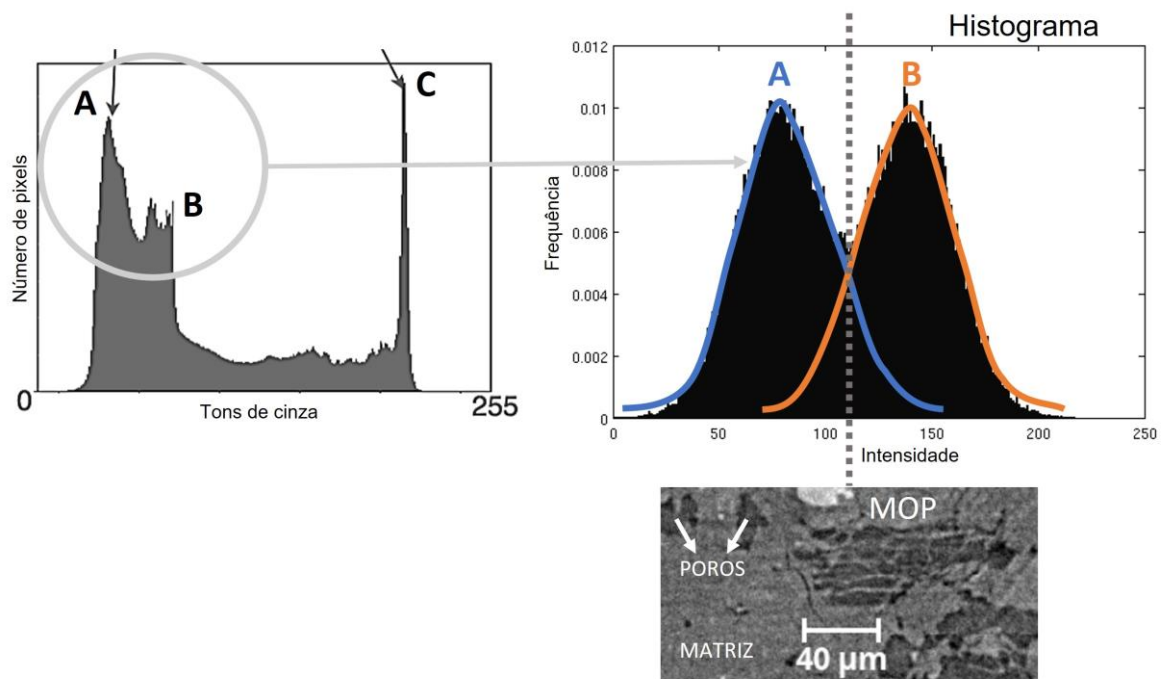


Figura 1. Exemplo de um histograma genérico para ilustrar uma fase com tons de cinza intermediários como a MOP. Os picos A e B representam a matriz e os poros de uma imagem de μ TCX em tons de cinza. É possível observar que as curvas correspondentes a esses picos se sobrepõem e é nessa região de sobreposição que se encontram os tons de cinza intermediários que representam a MOP, por isso a dificuldade em propor um método de segmentação para essa fase baseado apenas em tons de cinza. Adaptado de Perry (2013).

2.3.1 Segmentação de imagens

A segmentação é o processo de particionar uma imagem digital em vários segmentos (ou classes), a fim de definir a localização espacial ou identificação precisa dos objetos contidos na imagem (Vasconcelos et al. 2018). No contexto de imagens de μ TCX, a segmentação consiste em atribuir uma classificação coerente e precisa a determinados grupos de tons de cinza na imagem que representam fases de interesse. Esse processo é importante, porque em muitos casos, não é possível realizar análises quantitativas diretamente da imagem obtida em tons de cinza pela μ TCX.

Entre os métodos de segmentação tradicionalmente utilizados estão: o método manual, que consiste em usar ferramentas de “pintura” (*pincel*) diretamente na imagem e, apesar de ser bastante preciso a depender da dedicação do operador, é demasiado lento (no caso de uma imagem 3D, precisa ser realizado fatia a fatia da

imagem), útil para pequenas correções de imagens já segmentadas ou em casos em que nenhum outro método se aplique; o método semi-manual inclui ferramentas que visam automatizar o processo, como *thresholding* (um corte simples no histograma de tons de cinza da imagem que separa duas fases) e *watershed* (baseado em crescimento de sementes para delimitação entre diferentes fases) (Tuller et al., 2015), as quais podem proporcionar bons resultados para fases que apresentam alto contraste de tons de cinza (como poros e matriz sólida) em imagens de boa qualidade (com baixo ruído e artefatos) (Ferreira et al., 2018; Ferreira et al., 2019; Pires et al., 2020).

Por outro lado, o processamento de grandes conjuntos de dados usando métodos de segmentação tradicionais é muitas vezes limitado por desafios relacionados com a qualidade e reprodutibilidade da segmentação, bem como devido ao elevado consumo de tempo do operador nesta etapa (Ferreira et al., 2022b). Uma limitação atual é a falta de ferramentas eficazes para a segmentação da MOP em imagens de amostras de solo, devido à dificuldade de diferenciação entre as fases. Mais especificamente, os métodos tradicionais descritos não são boas alternativas para a segmentação da MOP em grandes conjuntos de amostras: o método manual é inviável por ser lento e depender completamente do operador, o que limita o número de amostras a serem investigadas e, conseqüentemente, a interpretação estatística dos resultados; os métodos semi-manuais são falhos por se basearem unicamente nos tons de cinza, os quais, como descrito anteriormente, se confundem com os tons associados a poros e matriz sólida.

Alguns fragmentos orgânicos podem ser segmentados a partir da sua localização e, principalmente, pela sua morfologia, permitindo a sua diferenciação da fase sólida da amostra através de uma avaliação visual. A partir da identificação visual dos fragmentos orgânicos, Kravchenko et al. (2014) propõem um teste para a separação da matéria orgânica com base em geoestatística, por meio de uma análise estatística complexa de todas as estruturas da matéria orgânica identificadas e necessita de grande input do operador para o sucesso da segmentação. Trata-se de um procedimento laborioso para aplicação em um grande número de amostras, restringindo o número de amostras que poderá ser utilizada para a elucidação do questionamento científico proposto.

Avanços recentes em aprendizado de máquina, especificamente a aplicação de redes neurais convolucionais para análise de imagens, estão permitindo rápidos progressos para a obtenção de uma segmentação mais rápida e precisa de dados de imagem (Rippner et al., 2022). No entanto, os desafios permanecem na aplicação destas redes neurais para a análise de imagens para estudos agrícolas relevantes (com comparação de tratamentos, por exemplo). Assim, novos protocolos de processamento de imagem vêm sendo desenvolvidos para melhorar a robustez, objetividade ou precisão na caracterização da microestrutura do solo. Nesse sentido, as estratégias baseadas em *machine* e *deep learning*, mais especialmente a última, são identificadas como promissoras para se estabelecer uma metodologia que forneça resultados bons e rápidos de segmentação da MOP.

O software Annotat3D, desenvolvido pelo GCC/LNLS, é baseado em técnicas de aprendizado de máquina (*machine learning*) e computação de alto desempenho (HPC – *high performance computing*) equipados com GPUs (*graphics processing unit*) (mais informações em Pinto et al. (2022)). Atualmente, esse software pode ser utilizado por todos os usuários do LNLS, internos e externos, e deve ser disponibilizado em breve com a comunidade em geral via Github/Gitlab (Pinto et al., 2022). Sua elaboração e aprimoramento visa atender futuros usuários das diversas linhas de luz do Sirius, buscando inclusive avançar para uma interface web.

O Annotat3D é baseado em uma classificação de superpixel (Vasconcelos et al., 2018). Um superpixel é formado por um grupo de pixels vizinhos que possuem intensidades semelhantes em relação aos seus tons de cinza (isto é, um certo número de pixels, elementos constituintes da imagem, que compartilham características em comum), resultando em uma redução do dado que precisa ser processado para a segmentação final. Essa abordagem requer menor capacidade computacional e, portanto, menor tempo de processamento do que uma classificação realizada a nível de pixel. O plugin Weka Segmentation, desenvolvido para o software ImajeJ/FIJI, é um exemplo de ferramenta de segmentação de imagens 2D e 3D (processamento em Java), também baseada em *machine learning*, porém em nível de pixel (Arganda-Carreras et al., 2017). Da mesma forma, o Avizo é um software comercial de uso geral para visualização e análise de dados científicos e industriais e tem sido utilizado para processamento e análise de imagens tomográficas em uma ampla gama de

aplicações. No entanto, ambos não se beneficiam da capacidade de processamento paralelo de um sistema de computação heterogêneo equipado com várias GPUs e apresentam dificuldade em suportar grandes dados volumétricos.

No software Annotat3D existem duas vias de segmentação, sendo uma via anotação (associado a algoritmos de *machine learning*) e outra via *deep learning*. A via anotação envolve um fluxo de trabalho interativo para treinamento de um classificador (modelo) e requer a definição de alguns parâmetros, como superpixel, extração de recursos e tipo de classificação. Na via de segmentação *deep learning* são utilizadas redes neurais profundas, em que um conjunto de algoritmos é usado em aprendizado de máquina para modelar os dados utilizando várias camadas de neurônios. Após obter um conjunto de dados bem segmentados, por exemplo na via de anotação, é possível utilizá-lo como dado de entrada para uma rede neural profunda que possa aprender e aplicar a segmentação a outras imagens semelhantes. No software Annotat3D, são utilizadas Redes Neurais de Convolução (RNC), redes neurais artificiais que estão entre as abordagens mais bem sucedidas na extração de características de imagens (Peixinho, 2017) e estão possibilitando uma segmentação mais rápida e precisa de dados de imagem (Rippner et al., 2022). São oferecidas duas arquiteturas neurais diferentes: a Unet (abordagens 2D e 3D) e a Vnet (apenas para abordagem 3D).

3 MATERIAL E MÉTODOS

3.1 Descrição da área experimental e coleta de amostras

A coleta das amostras ocorreu em junho de 2018 em um experimento de campo de longa duração, localizado no município de Iracemápolis, São Paulo, sudeste do Brasil (22°36' S-47°34' W), com altitude de 614 m acima do nível do mar. O solo da área experimental foi classificado como Latossolo Vermelho eutrófico de textura argilosa (Santos et al., 2018), e o experimento foi instalado em setembro de 2013 sob o delineamento de blocos casualizados com quatro repetições. Cada parcela foi composta por 10 linhas de cana-de-açúcar, com comprimento de 10 m e espaçamento entre linhas de 1,5 m. Neste trabalho, as amostras de solo foram provenientes de dois

tratamentos contrastantes de adição de palha de cana-de-açúcar anual na superfície do solo por um período de 5 anos (0 e 15 Mg ha⁻¹ de massa seca).

Para cada tratamento, uma amostra indeformada (10 cm comprimento × 10 cm largura × 5 cm altura) da camada superficial (5 cm) foi coletada de cada bloco (repetição), totalizando 8 amostras intactas de solo para ambos os tratamentos. No laboratório, as amostras foram desagregadas manualmente aplicando-se uma força de tração suficiente para rompê-las em seus planos naturais de fraqueza (fissuras naturais), e em seguida, foram secas ao ar. Esse procedimento resultou em agregados de tamanhos diversos, os quais passaram pelo peneiramento à seco utilizando um agitador de peneiras para obtenção de agregados com diâmetros entre 0,5 e 1 mm, de modo a atender o campo de visão das medidas de μ TCX. Assim, foram selecionados 3 agregados correspondentes a cada repetição (bloco experimental), resultando em 10 agregados por tratamento e 20 no total. A coleta de amostras e seleção dos agregados foram realizadas previamente ao desenvolvimento do presente trabalho.

3.2 Aquisição das imagens por μ TCX

Imagens 3D dos 20 agregados de solo selecionados foram geradas, em 2018 (proposta: IMX20180205), na linha de luz de Microtomografia de Raios X (IMX), no antigo acelerador de elétrons (UVX), pertencente ao LNLS/CNPEM. Destas 20 imagens, três foram perdidas, totalizando 17 imagens (10 para o tratamento sem palha e 7 para o tratamento com palha). As amostras foram escaneadas usando um feixe policromático com filtros de Si de 550 μ m de espessura, sendo adquiridas 2048 projeções de 2048×2048 pixels ao longo de 180° de rotação da amostra. Uma objetiva de 10x de aumento foi utilizada, com distância entre a amostra e o detector de 40 mm e com tamanho de pixel resultante de 0,82 μ m. As imagens foram reconstruídas a partir de um algoritmo de retroprojeção filtrada, desenvolvido pelo Grupo de Computação Científica (GCC) do LNLS (Miqueles et al., 2014; Miqueles et al., 2018), resultando em imagens tridimensionais (2048×2048×2048 pixels) de 16 bits. A resolução horizontal x vertical da imagem foi estimada em 2,35 × 0,98 μ m² (o cálculo encontra-se no apêndice deste trabalho). A aquisição das imagens foi realizada previamente ao desenvolvimento do presente trabalho.

3.3 Processamento das imagens de μ TCX

O fluxograma apresentado na Figura 2 resume as etapas desenvolvidas no presente trabalho, com o uso do software Annotat3D para as principais tarefas e suporte do software Avizo em etapas intermediárias. Tipicamente, a filtragem da imagem é um primeiro passo importante para suavizar possíveis ruídos e realçar as bordas das fases contidas na imagem, tal como foi realizado em Ferreira et al. (2022b). No entanto, considerando que o ruído das presentes imagens não era pronunciado e buscando-se manter as características da MOP para o melhor resultado possível, optou-se por trabalhar com as imagens originais, ou seja, sem filtragem preliminar. Por este motivo, a filtragem aparece como uma etapa opcional na Figura 2.

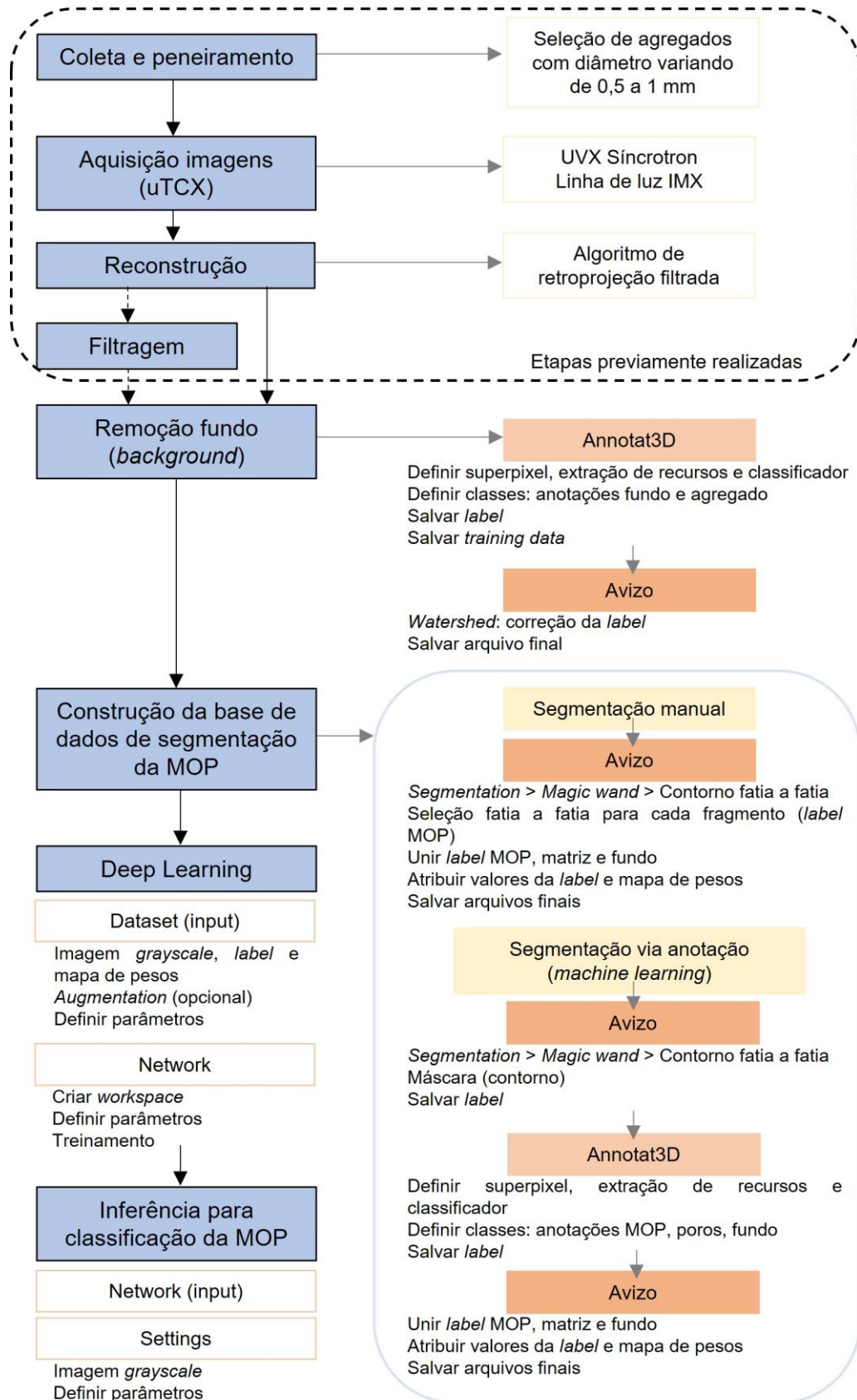


Figura 2. Esquema do fluxo de trabalho para segmentação da MOP. A filtragem consiste em uma etapa opcional (setas tracejadas).

Nas próximas seções, serão descritas em detalhes as metodologias empregadas nas etapas de remoção do *background* (fundo), construção da base de dados de segmentação da MOP pela via anotação no Annotat3D e etapas intermediárias no Avizo, treinamento de rede neural pela via *deep learning*, e aplicação desta rede neural para a segmentação da MOP (inferência) em 17 amostras. No apêndice do presente trabalho consta uma descrição geral do processo de segmentação utilizando o Annotat3D, incluindo definições de parâmetros de configuração necessários para cada via de processamento, com base em Pinto et al. (2022), Ferreira et al. (2022b) e na experiência adquirida durante o desenvolvimento deste trabalho.

Neste momento, é importante apresentar as definições de alguns termos que serão recorrentes ao longo do texto:

- 1) Imagem *grayscale*: representa a imagem 3D original (pós reconstrução) obtida pela técnica de μ TCX, a qual constitui um mapa de tons de cinza associados com os valores dos coeficientes de atenuação de cada fase presente na amostra; no presente trabalho, são imagens do tipo *raw* em 16 bit e, portanto, os valores de tons de cinza podem variar de 0 a 2^{16} (65.536);
- 2) Classe: representa alguma das fases definidas pelo operador no processo de segmentação da imagem *grayscale* (por exemplo, fundo, matriz sólida, e MOP);
- 3) Imagem *label*: representa uma imagem 3D resultante de algum procedimento de segmentação e, portanto, contém uma ou mais classes; no presente estudo, são consideradas poucas classes, e por isso utilizamos imagens de 8 bit para essa categoria, em que os valores das classes podem variar de 0 a 2^8 (256);
- 4) Fatia: imagem 2D que compõe uma imagem 3D (*grayscale* ou *label*). Uma imagem 3D consiste em uma sequência de imagens 2D (comumente chamadas de fatias) empilhadas na direção vertical da amostra (equivalente ao seu eixo de rotação na medida, Z, Figura 2a), de modo que é possível acessar fatias 2D nas três direções espaciais, formadas pelos planos XY, XZ e YZ;

- 5) Inferência: aplicação da rede neural (modelo) treinada nas mesmas imagens usadas como input do treinamento, as quais servem como validação do modelo, ou em imagens não usadas no treinamento, isto é, usadas para teste do modelo (também chamado aqui de “teste cego”); o resultado da inferência é um mapa de probabilidade das diferentes classes definidas.

3.3.1 Remoção do *background*

Foi realizada a definição de um subvolume das imagens, no software Avizo, contendo apenas fatias da imagem que continham informação do agregado, buscando eliminar as fatias vazias que gerariam um processamento desnecessário (Figura 3a). Para isso, os módulos *roi box* e *extract subvolume* foram utilizados.

Em seguida, o software Annotat3D foi utilizado para criar uma imagem *label* contendo as classes “agregado” e “fundo” (Figura 3b). Na extração de recursos (*feature extraction*) foram selecionados todos os filtros disponíveis (Tabela 1A do apêndice). Os superpixels (grupo de pixels vizinhos que possuem intensidades semelhantes em relação aos seus tons de cinza, compartilhando características em comum) foram definidos como *Waterpixels 3D* (Pinto et al., 2022), com o *seed spacing* de 4 e *compactness* de 100, e o *feature pooling* foi configurado com as funções mínimo, média e máximo (Figura 3c). Os valores correspondentes ao *seed spacing* e ao *compactness*, que definem a grade referente a malha de superpixels, são escolhidos a partir de testes e da observação dos melhores intervalos que possibilitem respeitar os limites de cada classe (Figura 3c, retângulo em destaque), esses valores podem mudar de acordo com a imagem e as classes que se pretendem segmentar. O classificador (modelo) escolhido foi o *random forest* com os demais parâmetros no modo *default* (Tabela 1A do apêndice). Foram feitas as anotações nas regiões correspondentes ao agregado (em verde) e ao *background* (em rosa) (Figura 3d). As anotações são inicialmente realizadas em fatias espaçadas, buscando ter anotações no início, meio e fim das imagens e, em sequência, são realizadas anotações em fatias intermediárias e/ou regiões que ainda apresentam mais erros.

Para otimizar esta etapa de trabalho no Annotat3D, foi utilizada a função *training data*. Esta função permite salvar as informações de um treinamento de

classificador para uma determinada imagem e aplicar então na imagem seguinte. O arquivo *training data* carrega informações vetoriais sobre as classes agregado e fundo, reduzindo assim a quantidade de anotações necessárias para atingir um bom resultado na sequência de imagens. Um novo *training data* foi salvo para cada imagem, tornando-se mais robusto com a diversidade de informações fornecidas. Assim, foram realizadas anotações inicialmente na amostra “Sample 2”, para a qual o arquivo *training data* foi salvo e aplicado para a amostra “Sample 3”, com a adição de novas anotações e um novo *training data* foi salvo, sendo aplicado sucessivamente nas amostras “Sample 5”, “Sample 7” e “Sample 8”. O *training data* salvo na amostra “Sample 7” foi então utilizado nas demais imagens do conjunto.

A imagem *label* resultante do processo descrito no Annotat3D, para cada amostra, contendo duas classes (Figura 3e), foi utilizada no Avizo para a elaboração da máscara de remoção do *background*. Para isso, foi aplicada a ferramenta *fill holes* em todas as fatias, para melhor preenchimento das classes. As classes fundo e agregado foram então ajustadas e corrigidas, utilizando-se a operação *shrink* para reduzir o tamanho das classes do fundo e agregado. As classes ajustadas foram empregadas como sementes de um processo de *watershed* para a obtenção da *label* final (Figura 3f) e aplicação da máscara de remoção do *background* (Figura 3g). Se o resultado não fosse satisfatório, eram feitas novas anotações no Annotat3D a fim de melhorar as regiões que precisavam de ajustes e, o processo descrito anteriormente no Avizo era realizado novamente, para aprimorar as sementes e diminuir o número de correções necessárias fatia a fatia por meio de ferramentas manuais disponíveis no Avizo. Desta forma, as correções manuais necessárias pós *watershed* (fatia a fatia) foram mínimas, reduzindo o tempo de dedicação do operador. Esse procedimento foi adotado para as 17 imagens consideradas nesse estudo. Uma descrição do processo de *watershed* pode ser encontrada em Ferreira et al. (2022b).

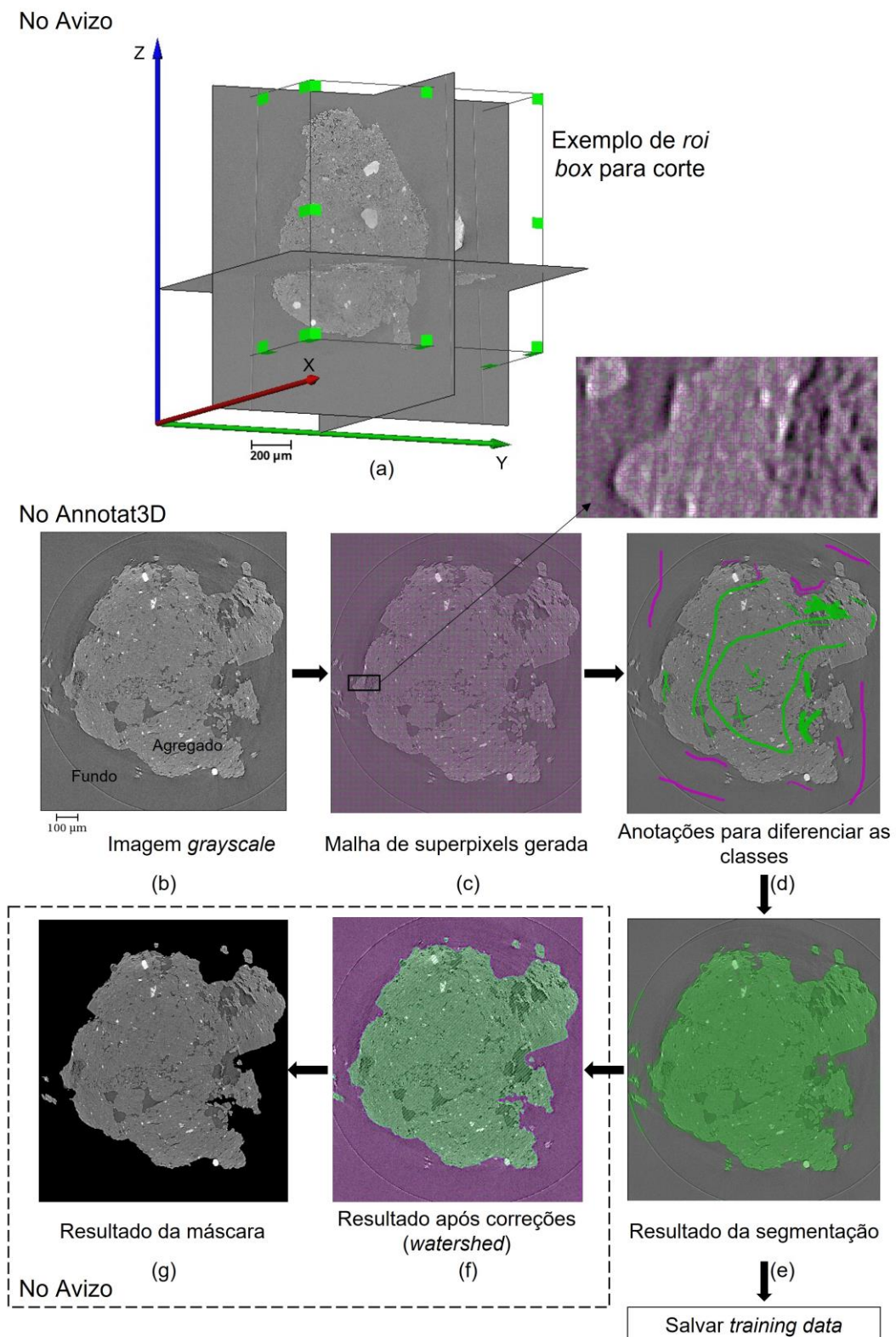


Figura 3. Sequência de trabalho para a remoção do background (“Sample 2”, fatia 500). (a) Exemplo de *roi box* para corte das fatias vazias com eixos X, Y e Z da imagem, onde Z representa o eixo de rotação da amostra na medida de μTCX . (b) Imagem *grayscale* após a reconstrução, com o fundo e a porção

correspondente ao agregado de solo indicados na imagem. (c) As linhas rosadas definem a malha de superpixels gerada (interessante conferir se os limites respeitam as bordas do objeto que se pretende segmentar, retângulo em destaque na imagem). (d) Anotações para definir as diferentes classes da imagem: anotações em rosa representam o fundo e anotações em verde o agregado. (e) Resultado da segmentação via anotação após a aplicação do classificador para todo o volume e salvamento do arquivo *training data*. (f) Resultado da correção feita a partir dos recursos de segmentação do software Avizo 3D, após aplicação do *fill holes* para todas as fatias e do *watershed*. (g) Resultado da máscara para remoção do *background*.

3.3.2 Construção da base de dados de segmentação da MOP

Foram selecionadas regiões de interesse, isto é, 20 subvolumes, dentro de 3 imagens diferentes, contendo maior concentração e/ou variedade de morfologias de fragmentos orgânicos que pudessem ser identificados visualmente. Um exemplo de seleção de subvolume em uma das imagens de agregados é apresentado na Figura 4a. O tamanho dos subvolumes respeitava sempre o tamanho pré-definido para o *patch size* do *dataset* na via *deep learning* (explicado mais à frente no texto). No presente caso, os subvolumes tinham no mínimo 256 fatias em cada uma das três direções da imagem. Para a construção dos dados de entrada do treinamento da rede neural via *deep learning*, a MOP foi segmentada de duas formas: integralmente manual ou via anotação (parcialmente manual).

A segmentação manual (Figura 4b-e) foi realizada inteiramente no Avizo, no ambiente *segmentation*. Inicialmente, foi definido o desenho do contorno de cada fragmento orgânico utilizando-se a ferramenta *magic wand* e a operação *draw limit line*, fatia a fatia (Figura 4c), com o intuito de selecionar apenas a MOP na região interna desse limite. Dentro dos contornos, definia-se o intervalo de tons de cinza pelo parâmetro *masking*, a partir de um histograma variando de 0 (tons mais escuros) até 65535 que melhor se adequava à morfologia da MOP. Esse processo também foi realizado fatia a fatia (entre 40 e 540 fatias) para cada fragmento individualmente, onde era definido o melhor valor de início para desconsiderar os poros pertencentes a esses fragmentos (correspondente à seleção em rosa na Figura 4d), e a seleção era adicionada à classe de MOP (Figura 4e) em uma imagem *label* no Avizo. Esta abordagem foi adotada para 14 dos 20 subvolumes totais segmentados, para

familiarização com as ferramentas de segmentação manual, até que se optou por testar a abordagem via anotação, descrita a seguir.

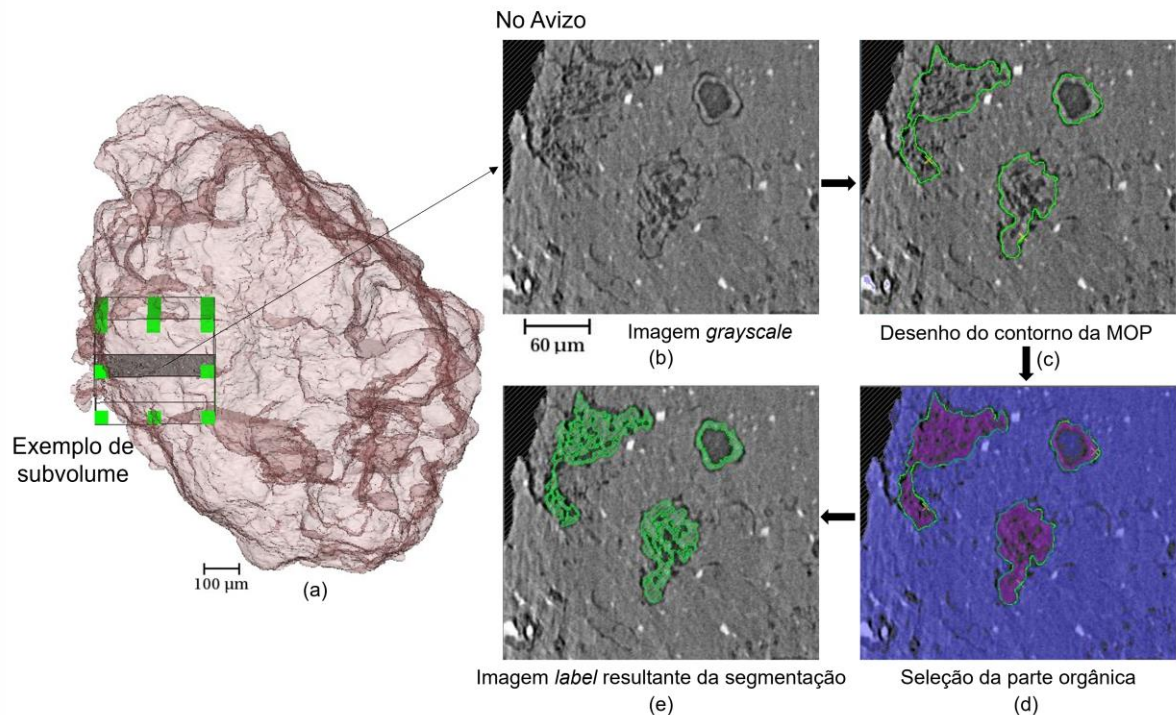


Figura 4. Sequência de trabalho para segmentação manual da MOP. (a) Exemplo de seleção de um subvolume após a remoção do *background*. (b) Imagem *grayscale* ("Sample 2", subvolume com dimensão de 413×290×257). Nessa imagem são identificados visualmente 3 fragmentos de MOP devido a sua morfologia diferente em relação a matriz. (c) Desenho do contorno/limite para cada um dos fragmentos em cada fatia através da ferramenta *magic wand* (*draw limit line*). (d) Definição do melhor intervalo de seleção (representado na cor roxa) para separar a parte orgânica dos poros contidos dentro do limite previamente definido. (e) Seleção adicionada à classe MOP (representada em verde) em uma imagem *label* – a imagem *label* contendo a MOP está sobreposta à imagem *grayscale*.

Na segmentação via anotação (usando algoritmos de *machine learning*) (Figura 5a-g), também foi preciso desenhar os contornos para cada um dos fragmentos orgânicos contidos em cada subvolume definido (Figura 5a), mas, nesse caso, a seleção foi de toda a região interna ao contorno (Figura 5b). A imagem *label* resultante do contorno (Figura 5c) foi usada para mascarar a matriz e o fundo na imagem *grayscale* (Figura 5d). Nessa imagem *grayscale*, contendo apenas os fragmentos isolados, foi realizada a segmentação via anotação no software Annotat3D (Figura 5e), para diferenciar a porção realmente orgânica dos poros que se encontram dentro

dessa estrutura (classes definidas: fundo, poros internos e MOP). Em seguida, todos os filtros foram selecionados para a extração de recursos. Porém, nesse caso, os melhores resultados de segmentação foram obtidos com a utilização do recurso de segmentação a nível de pixel (*pixel segmentation*) ao invés de superpixels (Figura 5f). É importante ressaltar que, enquanto o processo de anotação para a remoção do *background* possibilitou a definição de uma malha de superpixels mais grosseira e menos adaptada às bordas (parâmetros mais relaxados), o processo de anotação da MOP demandou uma classificação mais criteriosa em nível de pixel. Primeiro foram feitas anotações mais espaçadas, no início, meio e fim das imagens e, depois, nas fatias intermediárias ou em regiões que demandassem mais anotações para uma melhor identificação das classes. A Figura 5g mostra a classe de MOP resultante da segmentação na via anotação sobreposta à imagem *grayscale*. A abordagem de segmentação da MOP via anotação foi aplicada para 6 subvolumes.

Diferentemente da segmentação integralmente manual, a segmentação via anotação não requer anotações em todas as fatias de cada fragmento dentro dos subvolumes. Foram necessárias em média 6 fatias anotadas (no plano XY) por subvolume (compostos por um número total de fatias em Z variando entre 257 e 365) e, no máximo, 22 fatias para o maior subvolume em Z (com 700 fatias), para que o classificador estimasse a segmentação de uma só vez para todos os fragmentos contidos no subvolume. É importante mencionar que a estratégia de mascarar o fragmento via anotação foi adotada para diminuir o número de fatias anotadas necessárias para atingir um resultado de qualidade similar ao alcançado com a segmentação integralmente manual, principalmente relacionado à morfologia de cada fragmento.

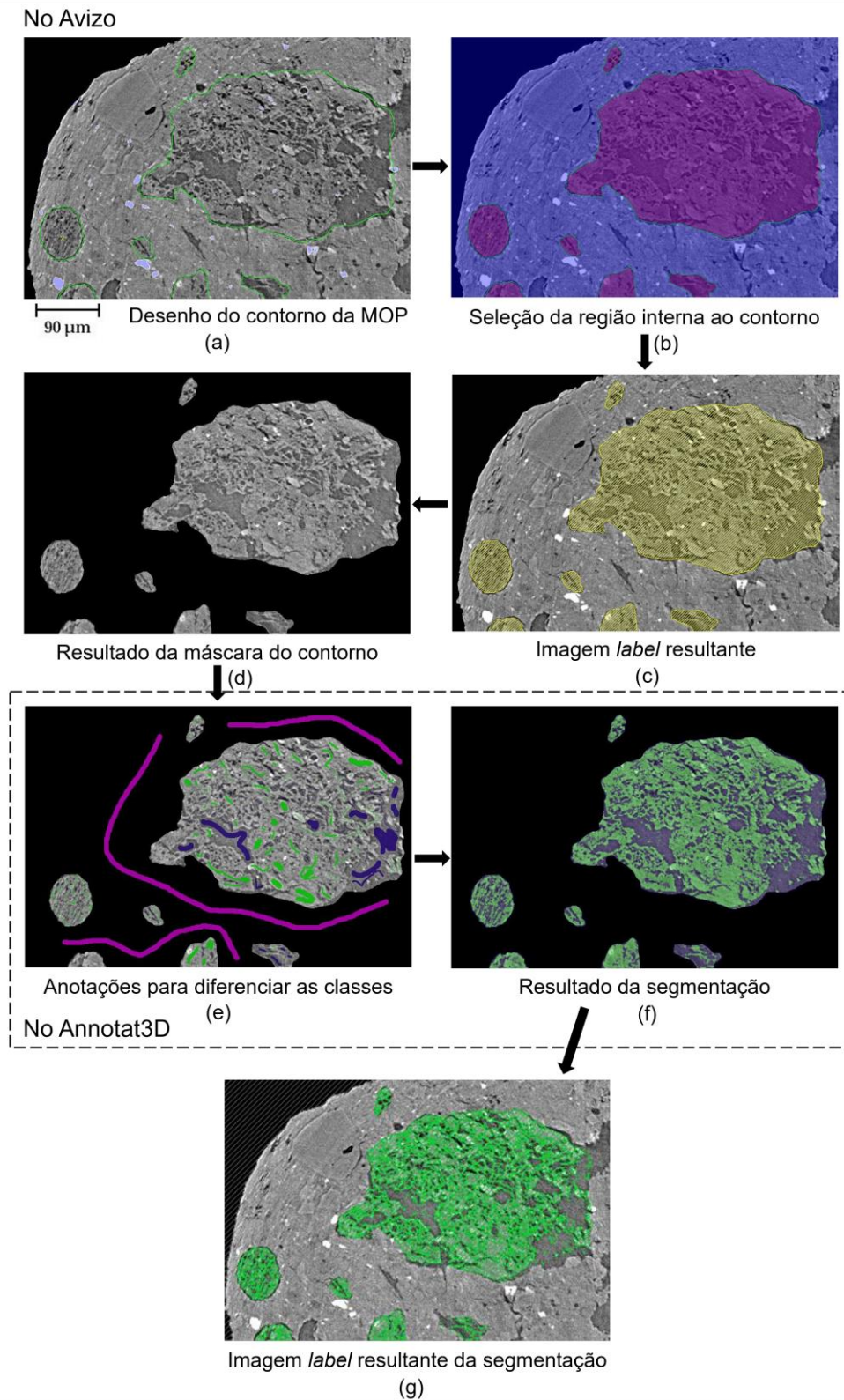


Figura 5. Sequência de trabalho para segmentação via anotação da MOP. (a) Imagem *grayscale* ("Sample 17", subvolume com dimensão de $883 \times 508 \times 700$) com os desenhos dos contornos para cada um dos fragmentos em cada fatia através da ferramenta *magic wand* (*draw limit line*). (b) Seleção da região

correspondente ao contorno (MOP e poros pertencentes a essa estrutura). (c) Criação de uma imagem *label* da região selecionada (em amarelo). (d) Resultado da máscara do contorno (remoção do *background* e matriz). (e) No Annotat3D, após definir os parâmetros do classificador para segmentação a nível de pixel, são feitas anotações para identificar as diferentes classes. Anotações em rosa correspondem ao *background*, em verde correspondem à região orgânica da MOP, e em roxo correspondem aos poros. (f) Resultado da segmentação após a aplicação do classificador para toda a imagem. (g) *Label* resultante da segmentação da MOP sobreposta à imagem *grayscale*, sem a máscara do contorno.

3.3.3 Deep learning: dataset

Na via *deep learning*, é necessário inicialmente criar um *dataset*, composto por uma ou mais imagens *grayscale* e suas respectivas imagens *labels* e mapas de peso. Sendo assim, para cada subvolume selecionado (Figura 6a, b), as classes de MOP (segmentadas manualmente ou pela via anotação), matriz e fundo foram unidos em uma imagem *label*, onde foi arbitrariamente atribuído um valor diferente para cada classe: 1 ao *background*, 0 à matriz e 2 à MOP (Figura 6c). Além da imagem *label* para cada subvolume, foi preparada uma imagem representando o mapa de pesos, na qual são atribuídos pesos diferentes a determinadas classes da imagem *label*: peso 1 para o *background*, 20 para a matriz e 255 para a MOP (Figura 6d). Esses pesos foram atribuídos de modo a estabelecer uma hierarquia de regiões prioritárias durante o treinamento da rede neural na via *deep learning*, isto é, quanto maior o peso associado a uma classe, maior a prioridade de atenção nesta região. Sendo assim, o maior peso possível (255, considerando uma imagem 8 bits) foi associado à MOP porque esta classe consiste em pequenos fragmentos com menor abundância em relação a outros elementos da imagem (como matriz e fundo). Assim, ao final do trabalho, formou-se um *dataset* com 20 subvolumes de imagens *grayscale*, *labels* e mapas de peso; as renderizações 3D da MOP para cada subvolume (Figura 7, amostras “Sample 2”, “Sample 13”, e “Sample 17”).

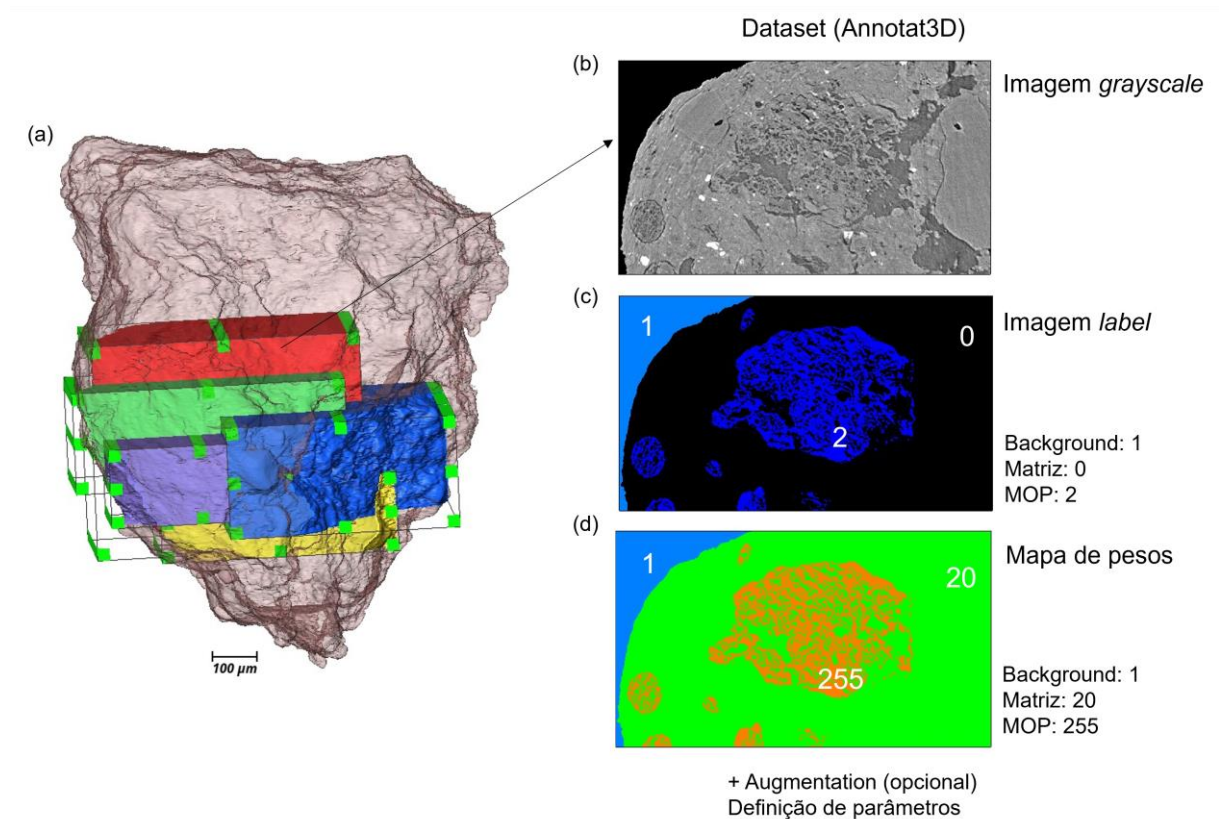


Figura 6. Esquema representando a criação do *dataset* para segmentação via *deep learning*. (a) Agregado de solo ("Sample 17") com subvolumes definidos a partir da escolha de regiões com maior densidade de MOP e diferentes morfologias para diversificar os dados de entrada. Imagens que compõem o *dataset*: (b) imagem *grayscale* (subvolume com dimensão de 883×508×700), (c) *label* com definição de valores para cada classe (o valor 1 representando o *background* em toda a região azul clara da imagem, 0 a matriz em toda a região preta e 2 a MOP em toda a região azul escura) e (d) mapa de pesos com definição de um peso para cada área específica da imagem (valor 1 para a região pertencente ao *background*, 20 à matriz e 255 à MOP).

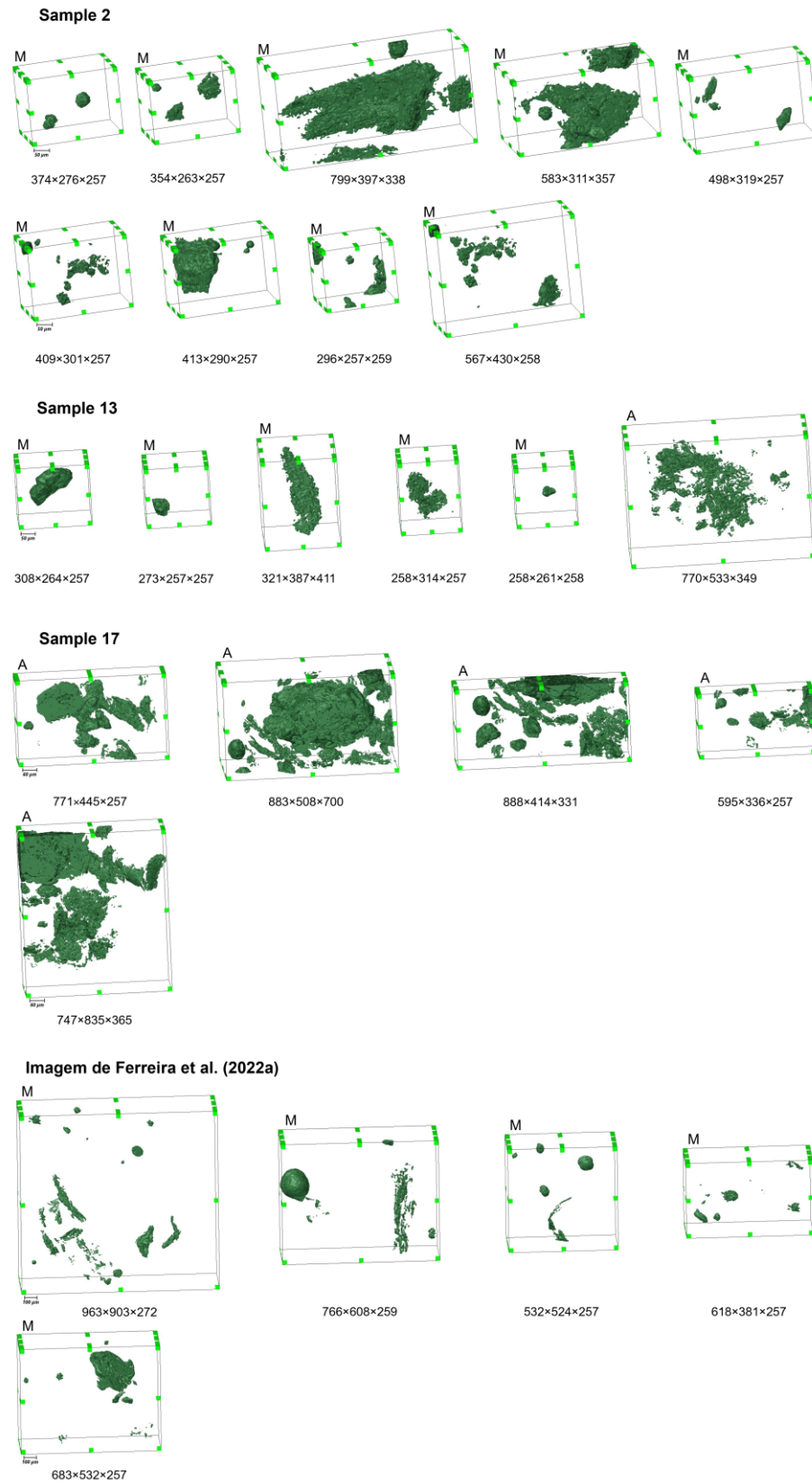


Figura 7. Renderizações 3D da MOP correspondente a base de dados construída a partir da segmentação manual, indicados pela letra “M”, ou via anotação, indicados pela letra “A”, para treinamento das redes neurais.

Observou-se que em todas as imagens de μ TCX dos agregados, a porção inferior (fatias iniciais na direção Z) apresentava um aspecto mais embaçado (Figura 8b), enquanto a porção superior (fatias finais na direção Z) apresentava uma textura mais nítida, com melhor diferenciação dos elementos que constituem a imagem, principalmente da matriz (Figura 8a). Por esse motivo, considerou-se a utilização de estratégias de *augmentation* (Tabela 1A do apêndice). Estas estratégias são capazes de gerar dados artificiais a partir de deformações que são aplicadas aos dados originais, resultando em um aumento do número de dados rotulados, além de criar problemas ainda mais desafiadores para a rede neural (busca por uma rede mais generalizada). A estratégia de *augmentation* do tipo *Gaussian Blur* foi a que apresentou melhores contribuições.

A definição dos valores utilizados para aplicar o *augmentation Gaussian Blur* ocorreu pela comparação do resultado de desfoque provocado na imagem *grayscale* original a partir da alteração dos valores de sigma que melhor representassem a textura da porção inferior da imagem. O parâmetro sigma irá definir o nível de desfoque (efeito *blur*) (Figura 8c) provocado na imagem, assim, quanto maior o valor de sigma mais desfocada ficará a imagem (sigma mínimo com menor desfoque, Figura 8c, *Augmented – Min*, e sigma máximo com maior desfoque, Figura 8c, *Augmented - Max* em relação a imagem *grayscale* original). É importante ressaltar que a implementação do *augmentation* resulta no aumento do tamanho do *dataset* e isso pode impactar diretamente no tempo necessário de treinamento.

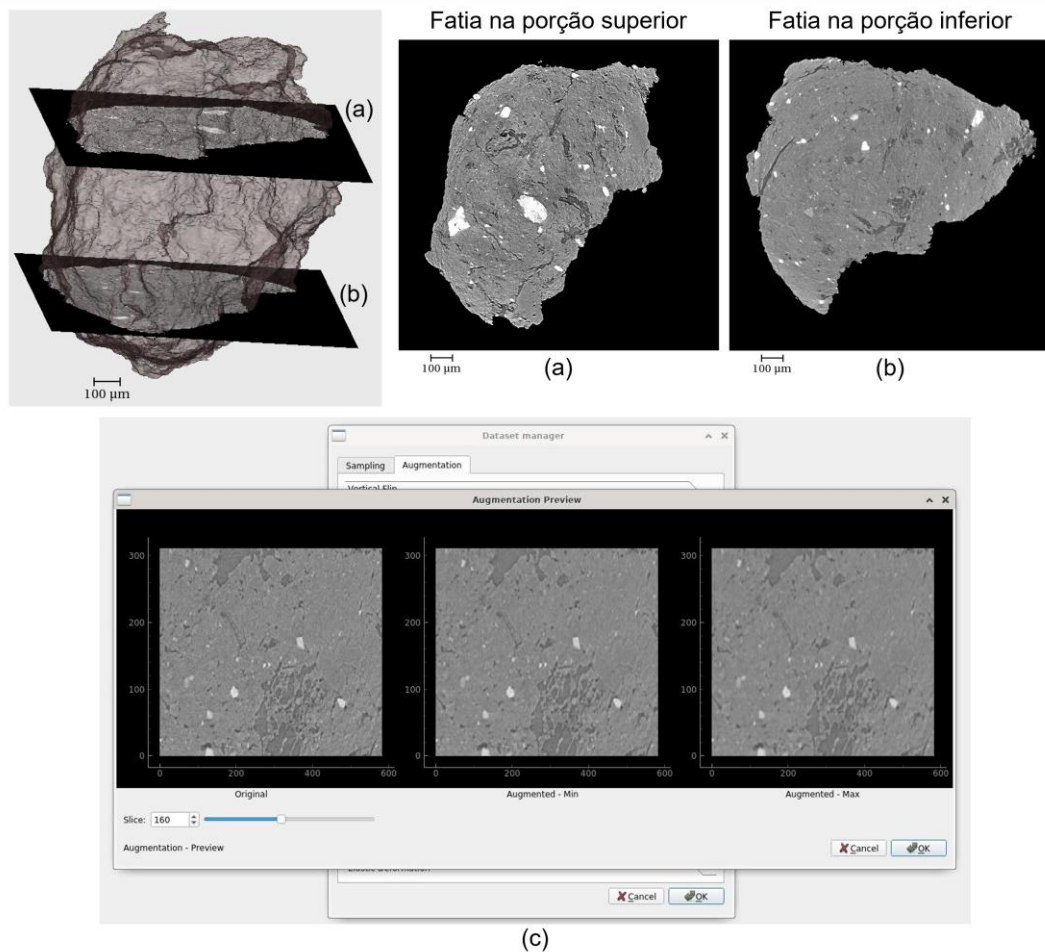


Figura 8. Efeito da aplicação do *augmentation* nas imagens durante a criação do *dataset*. Agregado de solo (“Sample 17”) representando as fatias selecionadas para ilustrar a mudança de textura entre a (a) porção superior da imagem com textura mais nítida e a (b) porção inferior com textura menos nítida (embaçada). (c) Aplicação do *augmentation Gaussian Blur* (interface do Annotat3D para criação do *dataset*) e definição dos melhores intervalos de desfoque para o conjunto de dados, de modo que o máximo se aproximasse com o observado nas imagens (a partir do resultado na imagem de *preview*, subvolume “Sample 2”).

Nesse estudo foram definidos quatro *datasets* diferentes:

- *Dataset A1*: composto por 20 subvolumes pertencentes a imagens *grayscale* de 3 agregados diferentes (Figura 7), suas respectivas *labels* e mapas de peso;
- *Dataset A2*: composto pelos mesmos subvolumes do *dataset A1* com utilização do recurso de *augmentation* de desfoque gaussiano (*Gaussian Blur*) com os parâmetros de sigma: mínimo de 00.71 e máximo de 01.20 (Figura 8);
- *Dataset B1*: composto pela combinação de dois conjuntos de subvolumes diferentes, sendo os mesmos subvolumes do *dataset A1* somados a 5

subvolumes provenientes de uma imagem grayscale de um agregado de Ferreira et al. (2022b) (Figura 7). Os dados relacionados a MOP não foram abordados no trabalho de Ferreira et al. (2022b), mas os autores realizaram a segmentação manual da MOP nesta imagem e disponibilizaram para o presente estudo;

- *Dataset B2*: composto pelos mesmos subvolumes do *dataset B1* com utilização do recurso de *augmentation* de desfoque gaussiano (*Gaussian Blur*) com os parâmetros de sigma: mínimo de 00.71 e máximo de 01.20 (Figura 8).

3.3.4 *Deep learning: network*

Foi escolhida a rede com arquitetura neural do tipo V-net (3D) com base no melhor desempenho de classificação para o protocolo de segmentação de poros (Ferreira et al., 2022b). Neste estudo foram realizados quatro treinamentos da rede neural que diferem entre si pelo *dataset* utilizado e, portanto, são identificados pelos mesmos nomes definidos anteriormente: A1, A2, B1 e B2 (Tabela 1).

Cada treinamento inicia-se pelo carregamento de um dos *datasets*. No menu de configurações foram selecionadas quatro GPUs (NVIDIA Tesla V100 32 Gb). O *batch size* foi definido como um, mas o valor total é a multiplicação deste número pelo número de GPUs utilizadas ($1 \times 4 = 4$). A taxa de aprendizagem (*learning rate*), a função *loss* e o otimizador (*optimizer*) foram mantidas em 0,00001, *Cross Entropy* e Adam (geralmente mais rápido), respectivamente. O *patch size* foi definido como 128 (parâmetros apresentados na Tabela 2A do apêndice). Para a realização de cada treinamento foram utilizados 128 núcleos de CPU e memória de 1TB. As iterações (*iterations*) atingiram os números de 210.000 (treinamento A1), 240.000 (treinamentos A2 e B1) e 270.000 (treinamento B2), avançando por meio de etapas de *finetune* a cada 30.000 iterações, demorando ~ 6 a 8 horas cada *finetune*.

Tabela 1. Parâmetros de treinamento definidos para a segmentação via deep learning no software Annotat3D para cada rede neural (A1, A2, B1 e B2).

Workspace		A1	A2	B1	B2
Número de imagens		20	20	25	25
Dataset	Label	0: matriz	0: matriz	0: matriz	0: matriz
		1: fundo	1: fundo	1: fundo	1: fundo
		2: MOP	2: MOP	2: MOP	2: MOP
	Weigth map	1: fundo	1: fundo	1: fundo	1: fundo
		20: matriz	20: matriz	20: matriz	20: matriz
		255: MOP	255: MOP	255: MOP	255: MOP
	Classes	3	3	3	3
	Sample size (configured)	20	20	20	20
	Sample size (total)	20	40	20	40
	Size (GB)	16	38	19.1	45.4
Network	Patch size	256 ³	256 ³	256 ³	256 ³
	Augmentation	-	Gaussian blur	-	Gaussian blur
	Network	Vnet	Vnet	Vnet	Vnet
	GPUs	4	4	4	4
	Batch	1 (x4 = 4)	1 (x4 = 4)	1 (x4 = 4)	1 (x4 = 4)
	Iterations	240k	210k	240k	270k
	Learning rate	0,00001	0,00001	0,00001	0,00001
	Loss function	Cross Entropy	Cross Entropy	Cross Entropy	Cross Entropy
	Optimizer	Adam	Adam	Adam	Adam
	Patch Size	128	128	128	128
Inference	Network	Input	Input	Input	Input
	GPUs	4	4	4	4
	Volume padding	60	60	60	60
	Patch border	60	60	60	60

3.3.5 Inferência para a classificação da MOP

Todos os treinamentos foram seguidos da etapa de inferência, em que a rede neural treinada é aplicada em um volume de imagem para efetuar a sua segmentação (demorando em média 10 minutos por imagem). A sequência realizada neste trabalho foi: 1) inferências nos mesmos subvolumes fornecidos no treinamento para validação da rede; 2) “testes cegos” de inferência nas imagens inteiras das quais esses

subvolumes foram selecionados (“Sample 2”, “13” e “17”); 3) “testes cegos” de inferência nas demais imagens do experimento (isto é, imagens não usadas no treinamento) (total de 14 imagens).

A partir das inferências, para se obter a imagem *label* final da segmentação de cada imagem e gerar as renderizações 3D, foi necessário realizar um *threshold* no mapa de probabilidade resultante com o intuito de definir um valor (entre 0,4 e 0,5) que eliminasse possíveis ruídos associados ao processo de segmentação, mantendo-se apenas a MOP de fato presente na imagem.

3.4 Proposta de avaliação para o resultado dos treinamentos (estimativa de acurácia)

Com intuito de estabelecer critérios de comparação entre as redes treinadas (A1, A2, B1 e B2) e entender o potencial de segmentação (eficácia e precisão) de um modo quantitativo, foi proposta uma comparação entre as imagens *grayscale* e os resultados das inferências para as principais redes. Para cada imagem foi realizada uma contagem dos fragmentos orgânicos visualmente identificados em cada um dos planos (XY, YZ e XZ). O valor médio dessas contagens foi considerado como uma referência do número total de fragmentos orgânicos em cada imagem.

No caso do resultado das inferências, buscou-se automatizar a contagem para propiciar um método viável de comparação de diferentes redes, quando aplicadas para todas as imagens, sem grande consumo de tempo. Desse modo, primeiramente definiu-se o menor volume de pixels contidos nos menores fragmentos identificados na contagem visual/manual. O valor da mediana do volume de pixels (4.819 pixels³) entre 34 fragmentos considerados visualmente menores e segmentados manualmente foi considerado como valor de corte para a contagem automática de fragmentos. Isso foi feito para excluir possíveis ruídos relacionados ao resultado da segmentação e/ou fragmentos que não poderiam ser facilmente identificados de forma visual/manual.

Nas imagens *labels* resultantes da inferência, foi aplicada a seguinte sequência no software Avizo para remover volumes inferiores ao limite estabelecido: no ambiente *segmentation*, a operação *fill holes* foi aplicada em todas as fatias, seguida pela operação *remove islands* para remoção de volumes inferiores a 4.819 pixels³, correspondente a 13,85 µm, considerando todo volume 3D da imagem. Em seguida,

a ferramenta *label analysis* foi utilizada para contar automaticamente o número de fragmentos de MOP segmentados pela via *deep learning* nas imagens. Portanto, a partir das contagens visual/manual nas imagens *grayscale* e automática nas imagens resultantes das inferências, foi determinada a porcentagem de acertos das redes neurais. Essa porcentagem foi considerada então como uma estimativa de acurácia das redes.

4 RESULTADOS

4.1 Remoção do background

Com o uso do *training data* foi possível diminuir o número de fatias anotadas na sequência de imagens processadas para obter uma imagem *label* com melhor definição dos limites entre agregado e fundo (Tabela 2). Foi necessário maior número de fatias anotadas nas primeiras 5 imagens para que o classificador retornasse um resultado que melhor respeitasse as regiões correspondentes às classes agregado e fundo e que permitisse seguir para o uso do *watershed*. Com a aplicação do arquivo *training data* gerado para a amostra “Sample 7” nas 14 imagens restantes, foi necessário uma média de 20 fatias anotadas por imagem.

Tabela 2. Número de fatias anotadas por amostra para remoção do background e tamanho do training data associado.

Amostra	Número de fatias anotadas	Tamanho do <i>training data</i>
Sample 2	71	114.3 MB
Sample 3	49	561.3 MB
Sample 5	49	1.2 GB
Sample 7	20	2.4 GB
Sample 8	40	5 GB

4.2 Treinamento das redes neurais e inferências para a classificação da MOP (via *deep learning*)

Os testes iniciais auxiliaram na definição dos melhores parâmetros de treinamento da rede neural. A avaliação do resultado nestes casos foi visual,

comparando qualitativamente a imagem *label* usada no treinamento da rede (parte do *dataset*) e a imagem *label* resultante das inferências nas mesmas e em novas imagens. Os parâmetros mais relevantes foram os tamanhos do *patch size* definido para o *dataset* e para o treinamento. Percebeu-se que os melhores resultados eram obtidos ao atribuir um maior valor ao *patch size* do *dataset* em relação ao do treinamento (256 e 128, respectivamente). Outros testes foram realizados para determinar o melhor valor para o *sample size* do *dataset*, considerando os valores de 10, 20 e 40, entre os quais o valor 20 apresentou o melhor resultado. Com relação aos valores atribuídos à matriz no mapa de pesos, variou-se entre 20 e 130, mas o valor 20 apresentou o melhor resultado. Também é possível definir pesos diferentes em sub-regiões de uma mesma classe, como em Ferreira et al. (2022b), mas no presente estudo isso não foi necessário.

A partir de uma comparação visual das inferências e imagens *label* do *dataset*, foram obtidos melhores resultados com o aumento do número de combinações de imagens no *dataset*, cada uma formada por um subvolume de imagem *grayscale*, *label* e mapa de pesos. Além disso, a inclusão de fragmentos que apresentavam diferentes morfologias também favoreceu os resultados. Foi notável a necessidade de garantir a segmentação de todos os fragmentos presentes nos subvolumes do *dataset*, pois quando existiam fragmentos não segmentados nestes subvolumes, os mesmos não eram identificados posteriormente na etapa de inferência. Com o avanço dos resultados das inferências nos “testes cegos” (ou seja, imagens não usadas no treinamento), foi possível perceber as diferenças entre os resultados de segmentação nas porções inferior e superior das imagens (Figura 9a e b). Nota-se que entre todas as redes testadas, a rede A1 identificou menor contribuição de fragmentos na fatia da porção inferior da imagem, com fragmentos parcialmente segmentados ou não identificados (Figura 9d) comparado a fatia da porção superior (Figura 9c). O uso do *augmentation* na rede A2 proporcionou melhor preenchimento dos fragmentos justamente na fatia da porção inferior (Figura 9f), porém foram perdidos integralmente ou parcialmente alguns fragmentos anteriormente identificados na fatia da porção mais nítida da imagem (Figura 9e).

As redes B1 e B2 demonstram alguns benefícios do uso de diferentes conjuntos de imagens, devido à contribuição de diferentes qualidades de imagem (ruídos), filtros,

entre outras características. A rede B1, cujo *dataset* continha cinco subvolumes de uma imagem diferente do conjunto desse estudo, resultou em um maior preenchimento dos fragmentos para a fatia da porção inferior da amostra, mesmo sem a utilização do *augmentation* (Figura 9h) (resultado similar ao da rede A2). A rede B2 auxiliou em completar ainda mais os fragmentos que foram identificados na porção inferior (Figura 9j). Em relação à porção inferior, ambas as redes B1 e B2 obtiveram resultados parecidos (como pode ser observados nas Figuras 9g e i), diminuindo o número de fragmentos que foram perdidos com o uso do *augmentation* (como observado na rede A2, Figura 9e).

Essas diferenças podem ser observadas com maior clareza na Figura 10, onde são apresentadas renderizações 3D da MOP resultantes das inferências realizadas a partir de cada rede neural treinada, considerando uma mesma imagem de “teste cego”. Entre a rede A1 (Figura 10a) e A2 (Figura 10b) é possível notar diferenças em relação ao uso do *augmentation* na região inferior do agregado, principalmente relacionado ao melhor preenchimento da forma correspondente à MOP, porém notou-se a diminuição de alguns fragmentos na porção superior anteriormente identificados na rede A1 (principalmente dos fragmentos menores). Entre as redes A2 e B1 (Figuras 10b e c) chamou a atenção a semelhança entre os resultados mesmo sem o uso do *augmentation* na rede B1, mas com a inclusão no *dataset* de imagens de outro conjunto de dados com características diferentes. E em relação às redes B1 e B2 (Figura 10c e d) é possível observar fragmentos mais completos na porção inferior da imagem, mostrando que o *augmentation* intensificou a melhora obtida pela mistura de imagens.

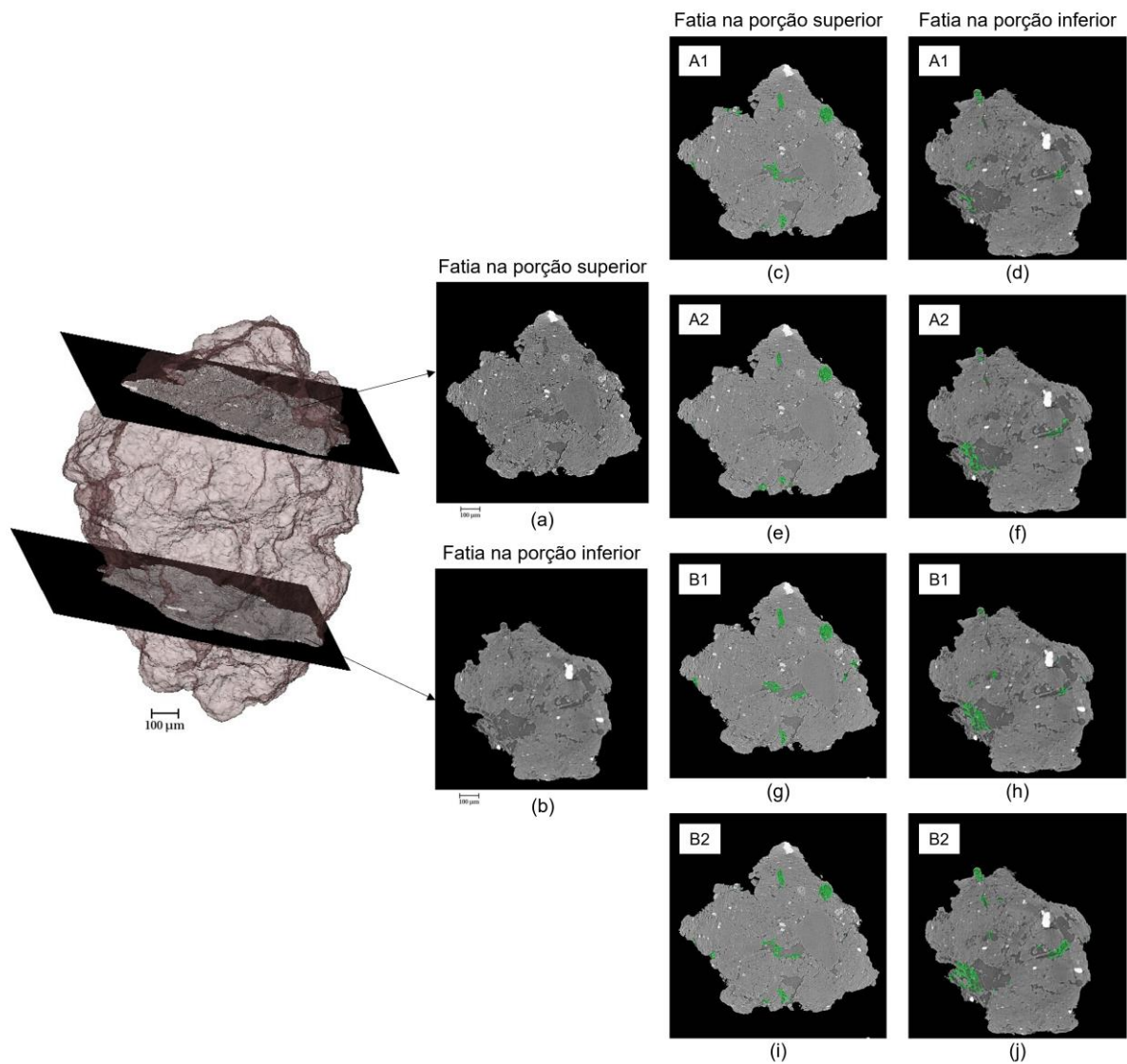


Figura 9. Comparação dos resultados para cada uma das redes treinadas (A1, A2, B1 e B2). Esquema com agregado de solo (“Sample 19”, inferência de “teste cego”) e fatias exemplificando a (a) porção superior da imagem (fatia 1512) com textura mais nítida e a (b) porção inferior da imagem, com textura mais embaçada (fatia 537). MOP segmentada nas fatias das porções superior e inferior, respectivamente, pelas redes (c, d) A1, (e, f) A2, (g, h) B1 e (i, j) B2.

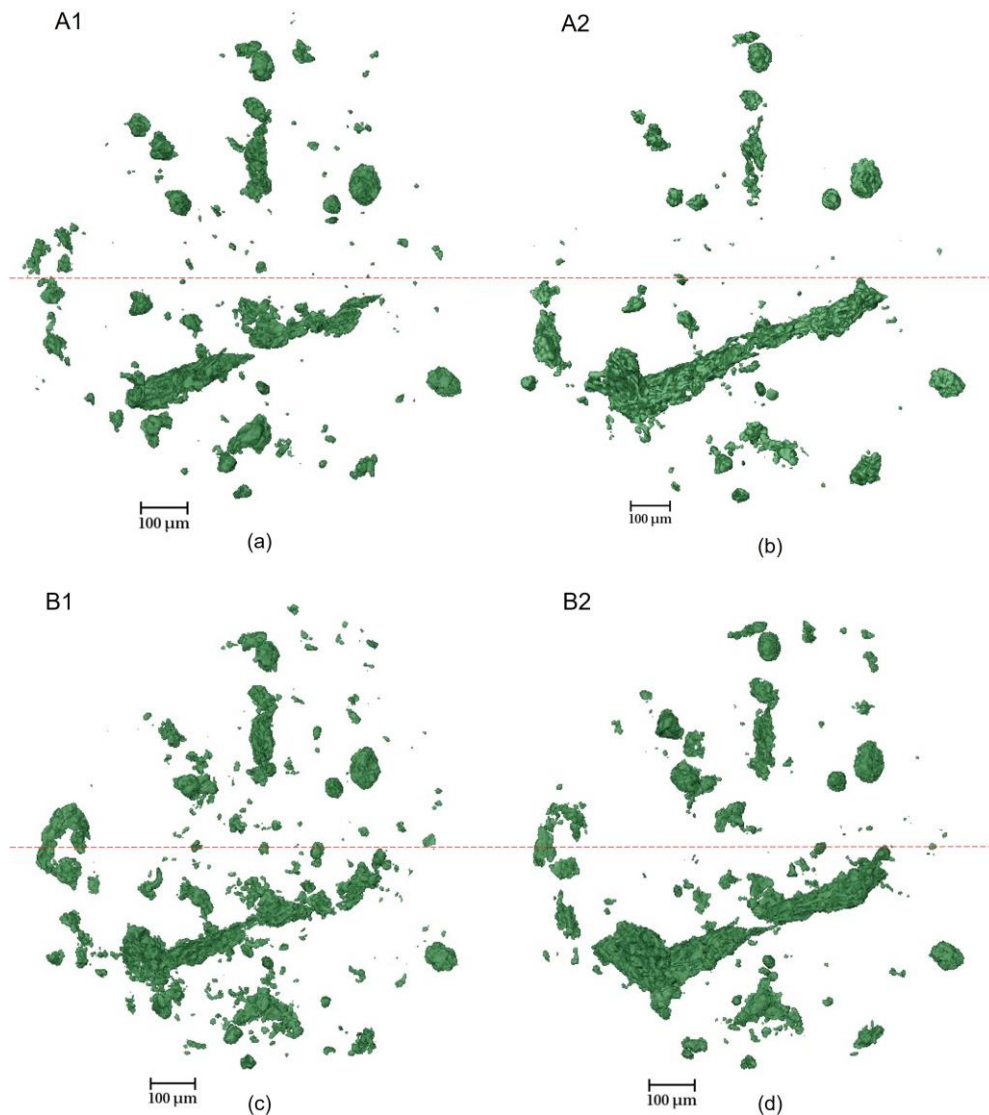


Figura 10. Renderização 3D da MOP, para a amostra “Sample 19”, segmentada a partir das inferências das redes (a) A1, (b) A2, (c) B1 e (d) B2. Linha vermelha tracejada indicando a região superior e inferior da imagem.

A Fig. 11 mostra a evolução da função *total loss* (*Cross Entropy*) durante o treinamento no decorrer das iterações, sendo que este comportamento está associado com a qualidade dos treinamentos. Neste ponto é importante explicar que, para o treinamento das redes neurais, ocorre uma separação dos *patches* definidos no *dataset*. Uma parte destes *patches* é usada no treinamento da rede e uma parte é reservada para a validação da rede a cada intervalo de iterações. Sendo assim, a cada intervalo de iterações do treinamento, a função *total loss* é computada de duas formas, a partir de uma comparação entre: 1) os *patches* das imagens *labels* que

compõe o *dataset* usadas no treinamento e o resultado da rede nestes mesmos patches (curva de treinamento); e entre 2) os *patches* das imagens *labels* que compõe o *dataset* não usadas no treinamento e o resultado da rede nestes mesmos patches (curva de validação). Portanto, para fins de comparação, assumiu-se que o melhor estado de um treinamento ocorre quando a curva de validação apresenta valor de *loss* menor ou similar à curva de treinamento. No entanto, este critério foi complementado por uma comparação visual das alterações dos resultados da inferência para cada etapa de *finetune* (intervalo de 30.000 iterações) do treinamento, até que se identificavam melhorias com novas etapas de *finetune* ou até que se notasse diminuição dos fragmentos identificados devido a um efeito de *overfitting*. Para a rede A1 (Figura 11a) o melhor resultado visual ocorreu com 240k, enquanto na rede A2 (Figura 11b) foi em 210k, uma vez que no último caso o maior número de iterações não resultou em ganhos visuais e promoveu o aumento da função *loss* associada à validação. Para a rede B1 (Figura 11c), o melhor resultado também ocorreu com 240k iterações e na rede B2 (Figura 11d) com 270k. Nota-se que o número de iterações escolhido ocorreu quando a função *loss* relacionada à validação foi menor que a do treinamento, para todas as redes, sendo os menores valores resultantes na seguinte sequência: B2, B1, A2 e A1, respectivamente (Tabela 3).

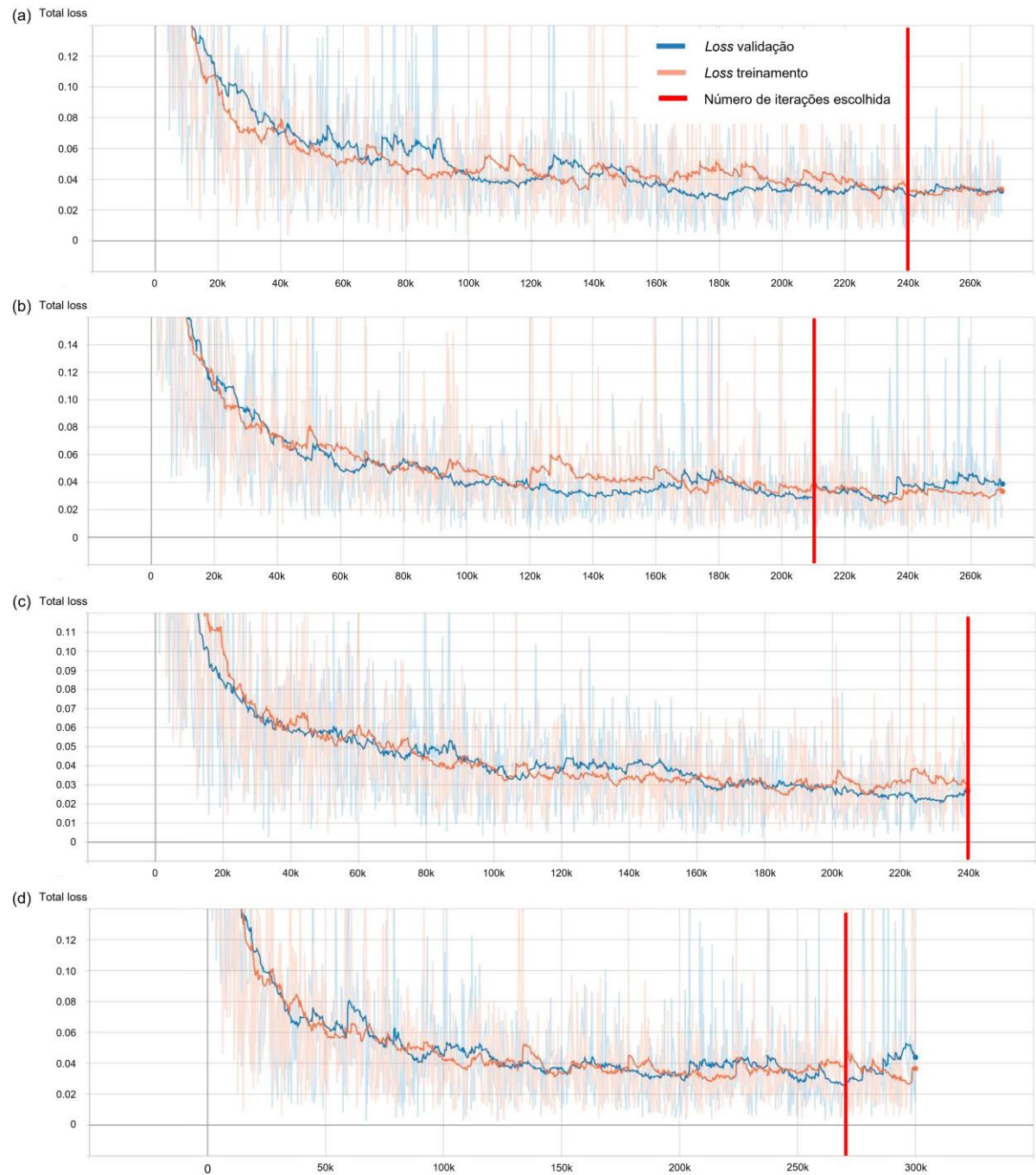


Figura 11. Gráficos de acompanhamento da função *total loss* no decorrer das iterações para o treinamento e validação das redes neurais (a) A1, (b) A2, (c) B1 e (d) B2.

Tabela 3. Valores da função total loss relacionados ao treinamento e a validação para o número de iterações escolhido para cada rede neural (A1, A2, B1 e B2).

	A1	A2	B1	B2
Número de iterações	240k	210k	240k	270k
<i>Loss</i> treinamento	0,03483	0,03726	0,02773	0,03732
<i>Loss</i> validação	0,02960	0,02737	0,02659	0,02623

A Figura 12 mostra as renderizações 3D resultantes da rede B2 para as 17 imagens do conjunto, junto com o limite/contorno de cada agregado. O resultado da segmentação permitiu observar diferentes morfologias da MOP, mais alongadas (fibrosas) ou mais arredondadas, além de diversos tamanhos. Schlüter et al. (2022) associaram as formas mais alongadas e fibrosas da MOP com a serapilheira mais fresca e raízes que estão localizadas principalmente dentro do sistema de poros. Já a MOP mais compacta, com menor quantidade de vazios internos (poros) e com formas mais arredondadas, como resultado de uma fragmentação e translocação repetidas.

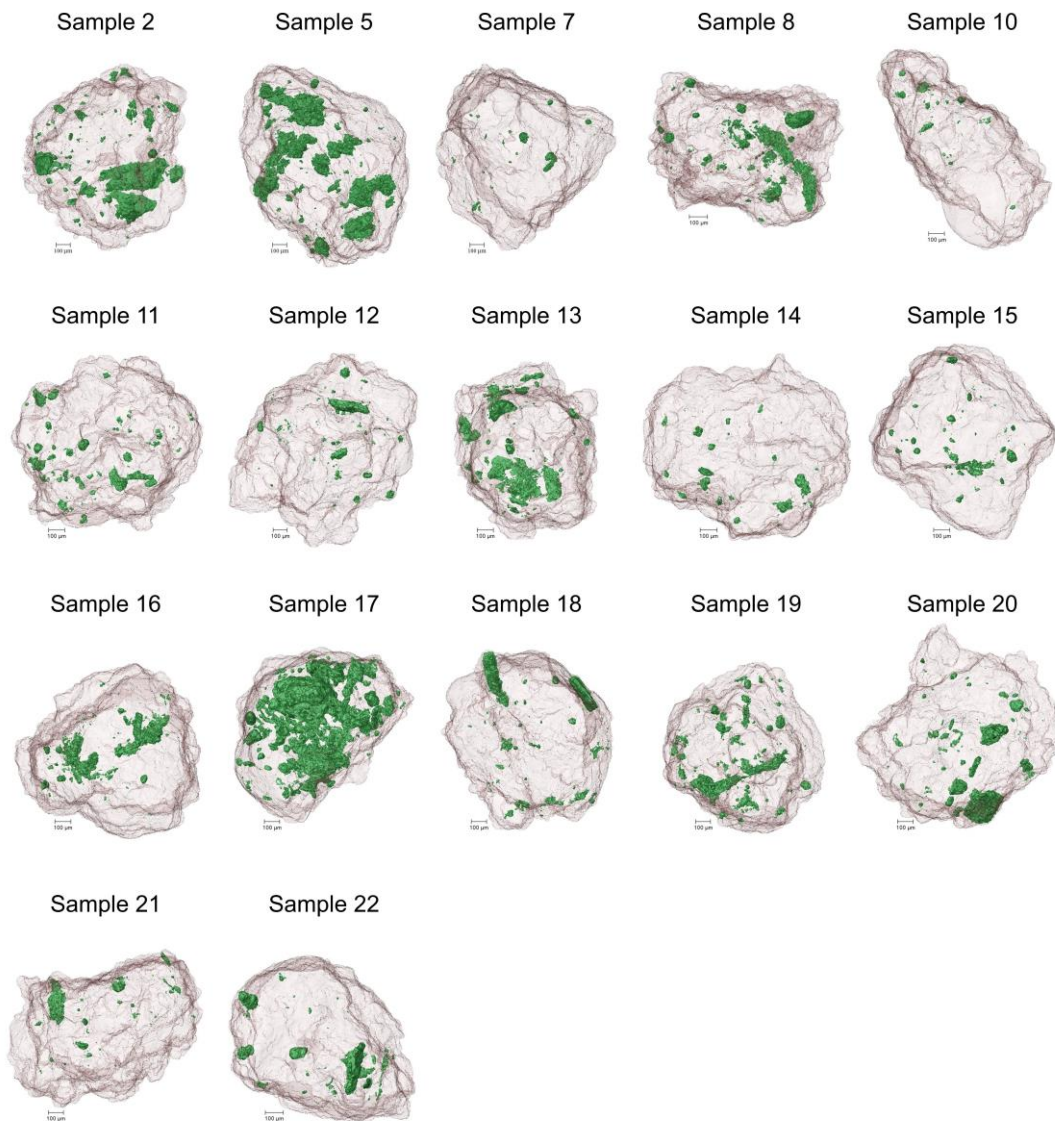


Figura 12. Renderizações 3D da MOP resultantes das inferências da rede neural B2 para as 17 imagens do conjunto.

4.3 Estimativa de acurácia das redes neurais

As redes A1 e B1 resultaram em maiores estimativas de acurácia comparado com as redes A2 e B2, respectivamente (Tabela 4). A mistura de dois conjuntos diferentes de imagens, na rede B1, atingiu as maiores porcentagens. Já a utilização do *augmentation* para a criação do *dataset* não refletiu em um aumento da estimativa de acurácia, ou seja, em um aumento no número total de fragmentos orgânicos identificados. Em alguns casos, o número de fragmentos determinado a partir dos

resultados de segmentação das redes neurais foi maior que o valor adotado de referência, tornando a estimativa de acurácia maior que 100%.

Tabela 4. Estimativa de acurácia obtida para cada rede neural (A1, A2, B1 e B2) considerada no estudo.

Amostra	Número de fragmentos adotado como referência	A1	A2	B1	B2
		Estimativa de acurácia (%)			
Sample2	116	63	50	74	67
Sample5	91	135	40	132	85
Sample7	81	22	35	31	25
Sample8	64	47	39	61	49
Sample10	37	54	31	90	46
Sample11	83	53	31	66	58
Sample12	51	49	99	67	61
Sample13	109	43	38	56	48
Sample14	61	46	28	61	43
Sample15	81	39	25	31	34
Sample16	74	52	79	52	43
Sample17	177	112	19	108	106
Sample18	63	51	25	60	54
Sample19	111	61	19	88	62
Sample20	77	58	34	69	53
Sample21	63	40	36	51	43
Sample22	53	42	39	68	42
Valor médio		57	39	69	54

5 DISCUSSÃO

Tipicamente, a filtragem da imagem é um primeiro passo importante para suavizar possíveis ruídos e realçar as bordas das fases contidas na imagem. Por outro lado, a filtragem inicial do dado original pode mascarar informações relevantes para os algoritmos de *machine* e *deep* learning. Portanto, considerando que o ruído das presentes imagens não era pronunciado, a opção de se trabalhar com as imagens

originais, sem filtragem preliminar, foi eficiente para preservar as características da MOP. Até o momento, os testes de uso do filtro *non-local means* mostraram resultados inferiores aos apresentados neste trabalho, demonstrando a importância da avaliação inicial sobre a necessidade da etapa de filtragem.

A utilização de superpixels na via anotação (algoritmos de *machine learning*) para geração de sementes das classes agregado e fundo, destinadas ao procedimento de *watershed*, foi mais eficiente do que a geração manual utilizando somente operações morfológicas, conforme descrito em Ferreira et al. (2022b). Nessa etapa, foi possível definir valores tipicamente grandes para os parâmetros *seeds* e *compactness*, que definem as dimensões dos superpixels (Figura 3c), acelerando assim o resultado do classificador. A estratégia de utilização do *training data* possibilitou ainda uma otimização no processo, diminuindo o número necessário de fatias anotadas por imagem. No entanto, justamente por acumular informações, este arquivo tende a aumentar em tamanho e consequentemente aumentar o tempo de aplicação do classificador (Tabela 2). Portanto, o melhor arquivo *training* gerado contou com informações de 4 amostras no total (salvo na “Sample 7”), pois a partir deste número o ganho em precisão do resultado não mais compensava o tempo necessário para realizar a aplicação do classificador no Annotat3D, inviabilizando o ritmo do trabalho.

A segmentação integralmente manual da MOP contida nos subvolumes foi considerada impraticável para a preparação do *dataset* destinado ao treinamento da rede neural. Trata-se de um procedimento que consome demasiado tempo do operador, em operações mecânicas e repetitivas (foram dedicados ~5 meses de trabalho para a segmentação manual de 14 subvolumes, Figura 7). Por isso, a construção de um *dataset* apropriado como o que foi construído neste trabalho, de maneira manual, demandaria vários meses de dedicação do operador, sendo essa uma tarefa desmotivadora. Assim, considera-se que a segmentação da MOP pela via anotação no Annotat3D foi crucial (~3 semanas para 6 subvolumes, Figura 7), mesmo que essa estratégia ainda necessite que os contornos da MOP sejam definidos manualmente no Avizo (Figura 5b).

A presença da máscara que limitava os contornos da MOP, na via anotação, foi essencial para que esta estratégia fosse de fato eficiente (Figura 5d-f), por facilitar

a separação entre a parte maciça do fragmento orgânico dos poros de dentro dessa estrutura. Vale ressaltar que, diferente da segmentação manual, a segmentação via anotação não requer anotações em todas as fatias do subvolume. Pelo contrário, poucas anotações em poucas fatias do subvolume são suficientes para que o classificador possa estimar a segmentação no restante deste subvolume. Esta grande vantagem da segmentação via anotação em relação à manual permitiu o aumento considerável do banco de dados de entrada (número de fragmentos segmentados) para a construção do *dataset* destinado à via *deep learning*. Houve inicialmente uma tentativa de segmentar a MOP nos subvolumes, pela via anotação, sem a definição da máscara dos contornos da MOP, o que facilitaria ainda mais este processo. No entanto, o resultado dessa segmentação era bom apenas localmente, isto é, na fatia anotada, com a presença de muitos erros em fatias não anotadas, exigindo assim um maior número de anotações e tornando o processo tão demorado quanto a segmentação integralmente manual (por exemplo, foram realizadas anotações em 44 fatias de um subvolume de $217 \times 190 \times 109$ pixels³ contendo apenas um fragmento orgânico e ainda havia erros em relação a MOP e às bordas dos poros).

Observou-se também que, na via anotação, a segmentação em nível de pixel (com a máscara dos contornos) proporcionou melhor definição da MOP em relação à segmentação com superpixels, mesmo com uma malha bastante fechada (*seed spacing*: 2 e *compactness*: 5). A segmentação em nível de pixel foi mais adequada para preservar a morfologia do fragmento, com qualidade semelhante à apresentada pela segmentação manual, e proporcionou uma diferenciação satisfatória entre MOP (associado a tons de cinza mais claros) e poros da MOP (associado a tons de cinza mais escuros). A utilização do classificador em nível de pixel ao invés de superpixels não demandou maior número de anotações, reforçando a vantagem dessa condição.

Percebeu-se que com o aumento do banco de dados, ou seja, aumento no número e variedade morfológica de fragmentos orgânicos utilizados para o treinamento da rede neural na via *deep learning*, o erro de diferenciação entre MOP e poros reduzia gradativamente. Isso demonstra que a metodologia empregada foi eficiente para alcançar o objetivo de segmentar a MOP, com a preservação de suas estruturas e definição satisfatória entre seus limites com os poros. Mas é importante ressaltar a dificuldade para produção desses conjuntos de dados que permitam o

treinamento de modelos (como apontado por Rippner et al. (2022)), pois além da demora para realizar as anotações ou outros processos/etapas que necessitam serem feitos à mão, é necessário adquirir um conhecimento especializado para a sua realização de maneira adequada nas imagens de treinamento (Kamilaris and Prenafeta-Boldú 2018).

Um dos fatores que colaboraram com a evolução dos resultados de segmentação da MOP pela rede neural foi a definição de subvolumes com maior densidade de fragmentos de MOP. Isso proporcionou maior interação da rede neural com as estruturas que se desejava segmentar, em relação ao restante da imagem. Ou seja, em subvolumes que contêm fragmentos mais dispersos, há uma maior probabilidade de que os *patches* do treinamento sejam posicionados aleatoriamente em regiões da imagem sem nenhuma contribuição da MOP. Além disso, percebeu-se a importância de segmentar todos os fragmentos orgânicos contidos no subvolume determinado. Isso porque, quando alguns fragmentos estão segmentados e outros não, a rede neural pode aprender padrões contraditórios, isto é, regiões de MOP ora são aprendidos como sendo pertencentes a essa classe, ora não. Como consequência, nos resultados da inferência, a MOP não segmentada anteriormente, próxima a um fragmento segmentado como dado de entrada, geralmente não era identificada. Outro fator identificado como relevante no treinamento da rede neural foi a atribuição de um mapa de peso, com peso máximo para a MOP (255) em relação à matriz (20) e fundo (1). Ressalta-se que foi também testado um peso de 130 para a matriz, mas apesar de pequenas melhoras, notou-se uma diminuição nos fragmentos corretamente identificados na inferência. Assim, o valor de 20 consolidou-se como uma boa opção neste caso.

Para a definição dos parâmetros, com os treinamentos iniciais, notou-se que os melhores resultados para a classificação da MOP ocorriam quando o *patch size* do *dataset* era maior que o *patch size* do treinamento. Particularmente, a combinação escolhida para o *patch size* do *dataset* e do treinamento foi de 256 e 128, a qual foi mais eficiente do que as combinações de 128 e 64 ou 64 e 32, respectivamente. As combinações de *patch size* do *dataset* e treinamento equivalentes (mesmo tamanho: 64 e 64 ou 128 e 128, respectivamente) também apresentaram resultados inferiores. Esses testes alternativos foram realizados com o intuito de definir uma configuração

que propiciasse melhor segmentação da MOP e diminuição do tempo necessário para o treinamento. É importante ressaltar que a combinação do tamanho definido para os diferentes parâmetros pode auxiliar na diminuição ou aumento do tempo de treinamento. Quando é necessário aumentar o tamanho do *patch size* pode-se diminuir o tamanho do *sample size* para ter um tamanho de *dataset* razoável e, dessa forma, não inviabilizar o tempo de treinamento.

Em relação ao desenvolvimento dos treinamentos, observou-se que a utilização de *datasets* contendo novas imagens isoladamente a cada finetune não resultava diretamente em uma melhora no resultado da inferência. O efeito desta operação era concentrado na região que continha os novos fragmentos, apresentando uma melhora na definição e completude dos fragmentos inseridos, porém ocorria uma diminuição no número de fragmentos que eram identificados anteriormente ou na sua continuidade no decorrer das fatias. Para tentar corrigir esse efeito, a cada finetune era criado um novo *dataset* com todas as imagens segmentadas até o momento (novas segmentações e antigas), essa medida proporcionou uma melhora nos resultados das inferências. Por isso, quando finalizada a segmentação de todos os fragmentos que compõem o *dataset* deste trabalho, um treinamento foi reiniciado com todos os subvolumes segmentados (imagens *grayscale*, *labels* e mapa de pesos correspondentes com ou sem *augmentation*).

As diferenças de textura e nitidez observadas entre as porções superior e inferior das imagens se devem à condição de medida de μTCX . Sabe-se que o feixe de raio X transmitido pela amostra incide no detector com uma distribuição de fótons nem sempre simétrica e homogênea, pois depende de um alinhamento prévio a cada experimento. É provável que o alinhamento realizado no experimento desenvolvido em Ferreira et al. (2022b) tenha sido ligeiramente diferente do alinhamento realizado para as imagens deste estudo. Com isso, houve maior concentração de fótons na região superior do detector, correspondente à região mais nítida das imagens consideradas neste trabalho. Por este motivo, priorizou-se a definição dos subvolumes na região mais nítida das imagens (superior). Essa decisão se deve ao fato de que a mistura de subvolumes entre as diferentes regiões reduzia a qualidade da segmentação da MOP de forma geral, devido à confusão de dados de entrada para o treinamento da rede neural. É possível que, se fossem fornecidos subvolumes em

maiores quantidades de ambas as regiões, a rede neural se adaptasse naturalmente. No entanto, não havia tempo hábil para preparar este *dataset* mais completo.

Desta forma, apesar de as imagens consideradas no presente estudo apresentarem menor ruído do que as imagens estudadas em Ferreira et al. (2022b), eliminando a necessidade de filtragem preliminar, a característica de não homogeneidade na textura da imagem levou à necessidade de um recurso (*augmentation*) disponível no Annotat3D, porém não utilizado em Ferreira et al. (2022b). O *augmentation* do tipo *gaussian blur* foi importante para alcançar um resultado de segmentação da MOP de forma homogênea ao longo da direção Z da imagem (definida pelo eixo de rotação da amostra na medida de μTCX , exemplificado na Figura 3a). Este recurso permitiu simular dados de entrada artificiais, com base nos dados reais, que representavam de forma apropriada a variação na textura da imagem. O *augmentation* poupou o esforço de gerar subvolumes para o *dataset* em grandes quantidades em ambas as regiões das imagens. É válido mencionar que também foram testados outros recursos de *augmentation* disponíveis no Annotat3D, como a rotação vertical, rotação horizontal, contraste e deformação elástica, sem observação de otimização dos resultados de segmentação.

Outro fator crítico para o treinamento da rede neural foi a adição de subvolumes provenientes de um outro experimento de μTCX na construção do *dataset*, mostrando a importância na diversificação da rede, seja buscando enriquecer o banco de dados com diferentes tipos de morfologias da MOP, texturas e propriedades de imagem. Neste caso, também houve a preocupação de que a adição destas imagens pudesse prejudicar a performance da rede neural, uma vez que poderia trazer um fator de confusão para o treinamento. No entanto, a performance da rede foi surpreendentemente superior, mesmo com apenas 5 subvolumes representativos deste outro experimento.

A partir de uma análise visual comparativa dos resultados das inferências para cada uma das redes percebeu-se que o uso do *augmentation* contribuiu positivamente na definição dos fragmentos orgânicos, proporcionando, na maioria dos casos, melhor preenchimento dos fragmentos, principalmente da porção inferior das imagens (Figura 10b,d comparadas com Figura 10a,c, respectivamente). A mistura dos conjuntos de imagens resultou em uma maior contribuição na identificação de novos fragmentos

(particularmente dos menores) nas porções inferior e superior das imagens (Figura 10c,d comparadas com as Figura 10a,b, respectivamente). Por outro lado, as redes A1 e B1 (ambas sem aplicação do *augmentation*) resultaram em valores médios de estimativa de acurácia mais altos (Tabela 4) indicando um maior número de fragmentos identificados, onde a B1 (com mistura de imagens) culminou no maior valor. As redes A2 e B2 (ambas com aplicação do *augmentation*) resultaram em valores médios mais baixos, porém a B2 (com mistura de imagens) apresentou um valor melhor. A redução na estimativa de acurácia com o uso do *augmentation* é decorrente de um menor número de fragmentos identificados pelas redes A2 e B2 (Tabela 4). No entanto, esta redução no número de fragmentos com o uso do *augmentation* pode estar relacionado com a maior conectividade proporcionada por este recurso entre as fatias de cada fragmento (Figura 10b,d comparadas com Figura 10a,c). Essas diferenças também foram observadas em outras imagens do conjunto, além disso, foi possível constatar fragmentos com morfologias semelhantes em todas as amostras e que o método definido permitirá comparar os diferentes tratamentos em um trabalho futuro.

É importante ressaltar, que mesmo adotando a média do número de fragmentos contados nas imagens *grayscale* como valor de referência para a estimativa de acurácia, é esperado que exista variações entre os valores estabelecidos e reais devido à dificuldade de realizar a diferenciação e contagem de cada fragmento visualmente no decorrer das fatias. Por este motivo, em alguns casos a estimativa de acurácia foi maior que 100%. Uma outra proposta de avaliação da acurácia das redes seria aumentar o banco de dados de fragmentos de MOP segmentados pela via anotação destinando um conjunto destes fragmentos exclusivamente para a comparação com os resultados das inferências em amostras não usadas no treinamento das redes (não viável no presente trabalho pela limitação de tempo). Assim, espera-se aprimorar o modo de validação dos resultados com o avanço dos estudos e esforços para otimização da rede. Porém, os valores obtidos a partir dessa análise quantitativa inicial unida à inspeção visual demonstraram o potencial de segmentação da MOP pelo método proposto sem realizar a marcação química das amostras. Portanto, a definição da melhor rede pode ser feita a partir da questão científica que se pretende aplicar, onde é preciso avaliar se a maior identificação em

número de fragmentos é mais interessante que a melhor definição da forma do mesmo, caso um efeito parecido seja observado nos resultados. Uma proposta, ainda não testada neste trabalho, é a possibilidade de aplicar redes diferentes para cada uma das porções com diferentes texturas da imagem, considerando separadamente a rede B1 para a porção superior e a rede B2 para a região inferior das imagens.

Com esse protocolo foi possível obter bons resultados para a segmentação da MOP sem a necessidade da marcação química por ósmio (como em Schlüter et al., (2022)). Arai et al. (2019) e Zheng et al. (2020) apontam três aspectos a serem observados ao se aplicar a abordagem atual de marcação química em agregados de solo. Primeiro, a marcação por ósmio funciona predominantemente para C insaturado, enquanto o C pode existir de outras formas no solo, além de ocorrer ligações preferenciais, que podem resultar em variações relacionadas ao método na capacidade de marcação da MOS. Em segundo lugar, a acessibilidade do OsO_4 à MOS dentro da estrutura agregada, principalmente quando o fragmento orgânico está coberto por materiais inorgânicos do solo (como em precipitações de ferro), ou seja, a estrutura do solo pode afetar na eficiência da marcação química. E, em terceiro lugar, a dificuldade em detectar fragmentos orgânicos muito pequenos. Ainda é necessário avançar em testes para verificar a eficiência da marcação da MOS na presença de diferentes minerais (inclusive na fração que está associada a minerais) e na avaliação da seletividade do Os para diferentes tipos de C insaturado (Peth et al., 2014; Belazi et al., 2009; Schlüter et al., 2022).

Assim, a contribuição deste trabalho está em propor um protocolo de segmentação que considera estratégias de aprendizado de máquina, principalmente por meio da aplicação de RNC, para a diferenciação da MOP de modo direto. Além de proporcionar uma segmentação de boa qualidade, buscando uma identificação mais detalhada para caracterização da MOP, permitiu uma segmentação mais rápida de um conjunto composto por 17 imagens de agregados de solo comparada a técnicas tradicionais de segmentação. Com a evolução deste estudo, espera-se que a partir de pequenas adaptações na rede treinada, seja possível a sua aplicação em diferentes conjuntos de imagens, otimizando o tempo necessário de processamento.

6 CONCLUSÃO

Foi apresentado neste trabalho um método de segmentação da MOP em imagens de μ TCX de agregados de solo. Trata-se de um método baseado em estratégias de *machine learning* para a construção de uma base de dados de fragmentos segmentados de MOP que alimentam o treinamento de redes neurais via estratégias de *deep learning*, sendo tais estratégias implementadas no software Annotat3D. Foram estabelecidos protocolos mais eficientes de criação do banco de dados de fragmentos, os quais podem ser reproduzidos por outros operadores para aumentar ainda mais este banco de dados. O banco de fragmentos segmentados de MOP permitiu o treinamento de redes neurais para a generalização da segmentação da MOP em diferentes amostras. As inferências das redes neurais (obtidas em alguns minutos por imagem) identificaram parcial ou totalmente os fragmentos orgânicos contidos nas imagens de μ TCX, onde as melhores redes atingiram valores médios de acurácia de 57 e 69%. O uso do recurso *augmentation* (redes A2 e B2 em relação às redes A1 e B1) proporcionou melhor continuidade dos fragmentos orgânicos identificados, principalmente na porção menos nítida das imagens, mostrando ser um artifício interessante para o aumento na diversificação dos dados de entrada para treinamento. A mistura de imagens de diferentes conjuntos (redes B1 e B2) resultou na identificação de um maior número de fragmentos, o que pode significar futuras melhorias a partir do uso e adaptação da rede para outras imagens.

Estes resultados mostraram o potencial do método proposto para a identificação e segmentação da MOP em imagens de μ CTX. Espera-se que esse protocolo possibilite avanços no ponto de vista metodológico para estudos científicos relacionados ao entendimento do papel da estrutura do solo na proteção física do carbono do solo.

7 PERSPECTIVAS FUTURAS

Como continuidade deste trabalho, a segmentação da MOP proposta aqui será integrada à da rede de poros (segmentação resultante da adaptação da rede treinada

para poros em Ferreira et al. (2022b)). O intuito será avançar na questão científica com relação aos impactos da remoção ou manutenção da palha de cana-de-açúcar em microescala. No caso, serão utilizados os resultados obtidos a partir da rede B2, pois acredita-se que a morfologia será mais importante que um número maior de fragmentos (geralmente pequenos) identificados. Será proposta quantificações que correlacione a MOP e a sua morfologia com a rede de poros, com o benefício de quantificar a sua posição em relação ao espaço poroso não perturbado.

8 REFERÊNCIAS

Amelung W, Bossio D, et al. (2020) Towards a global-scale soil climate mitigation strategy. **Nature Communications** 11:5427.

Arai M, Uramoto GI, Asano M, Uematsu K, Uesugi K, Takeuchi A, Morono Y, Wagai R (2019) An improved method to identify osmium-stained organic matter within soil aggregate structure by electron microscopy and synchrotron X-ray micro-computed tomography. **Soil and Tillage Research** 191:275-281.

Arganda-Carreras I, Kaynig V, Rueden C, Eliceiri KW, Schindelin J, Cardona A, Seung HS (2017) Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. **Bioinformatics** 33:2424-2426.

Ashman MR, Hallett PD, Brookes PC (2003) Are the links between soil aggregate size class, soil organic matter and respiration rate artefacts of the fractionation procedure? **Soil Biology and Biochemistry** 35:435-444.

Bartels M (2013) Cone-beam X-ray phase contrast tomography of biological samples: optimization of contrast, resolution and field of view. Universitätsverlag Göttingen.

Baveye, PC (2006) Comment on? Soil structure and management: A review? by CJ Bronick and R. Lal. **Geoderma** 134:231-232.

Baveye PC, White RE, Amiraslani F, Chenu C, Cardenas MG, Kaonga M, Koutika L-S, Ladha JK (2020) The “4p1000” initiative: A new name should be adopted Comment to: Rumpel. **Ambio** 49:361-362.

Belazi D, Solé-Domènech S, Johansson B, Schalling M, Sjövall P (2009) Chemical analysis of osmium tetroxide staining in adipose tissue using imaging ToF-SIMS. **Histochemistry Cell Biology** 132:105-115.

Berg S, Kutra D, et al. (2019) ilastik: interactive machine learning for (bio)image analysis. **Nature Methods** 16:1226-1232.

Claro PIC, Borges EPBS, Schleder GR, Archilha NL, Pinto A, Carvalho M, Driemeier CE, Fazzio A, Gouveia RF (2023) From micro- to nano- and time-resolved x-ray computed tomography: Bio-based applications, synchrotron capabilities, and data-driven processing. **Applied Physics Reviews** 10:021302.

Cotrufo MF, Wallenstein MD, Boot CM, Denef K, Paul E (2013) The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: Do labile plant inputs form stable soil organic matter? **Global Change Biology** 19:988-995.

Dexter AR (1988) Advances in characterization of soil structure. **Soil Tillage Research** 11:199-238.

Ekschmitt K, Kandeler E, Poll C, Brune A, Buscot F, Friedrich M, Gleixner G, Hartmann A, Kästner M, Marhan S, Miltner A, Scheu S, Wolters V (2008) Soil-carbon preservation through habitat constraints and biological limitations on decomposer activity. **Journal of Plant Nutrition and Soil Science** 171:27-35.

Elliott ET, Cambardella CA (1991) Physical separation of soil organic matter. **Agriculture, Ecosystems & Environment** 34:407-419.

Ferreira TR, Cássaro FAM, Zhou H, Pires LF (2022b) X-ray Computed Tomography Image Processing & Segmentation: A Case Study Applying Machine Learning and Deep Learning-Based Strategies. **X-ray Imaging of the Soil Porous Architecture**. Springer International Publishing, p. 57-80.

Ferreira TR, Pires LF, Reichardt K (2022a) 4D X-Ray Computed Tomography in Soil Science: an Overview and Future Perspectives at Mogno/Sirius. **Brazilian Journal of Physics** 52:33.

Ferreira TR, Archilha NL, Pires LF (2022c) An analysis of three XCT-based methods to determine the intrinsic permeability of soil aggregates. **Journal of Hydrology** 612:128024.

Ferreira TR, Pires LF, Wildenschild D, Brinatti AM, Borges JAR, Auler AC, dos Reis AMH (2019) Lime application effects on soil aggregate properties: Use of the mean

weight diameter and synchrotron-based X-ray μ CT techniques. **Geoderma** 338:585-596.

Ferreira TR, Pires LF, Wildenschild D, Heck RJ, Antonino ACD (2018) X-ray microtomography analysis of lime application effects on soil porous system. **Geoderma** 324:119-130.

Gosling P, Parsons N, Bending GD (2013) What are the primary factors controlling the light fraction and particulate soil organic matter content of agricultural soils? **Biology and fertility of soils** 49:1001-1014.

Gregorich EG, Beare MH, McKim UF, Skjemstad JO (2006) Chemical and Biological Characteristics of Physically Uncomplexed Organic Matter. **Soil Science Society of America Journal** 70:975-985.

Gureyev TE, Nesterets YI, Stevenson AW, Miller PR, Pogany A, Wilkins SW (2008) Some simple rules for contrast, signal-to-noise and resolution in in-line x-ray phase-contrast imaging. **Optics express** 16:3223-3241.

Kamilaris A, Prenafeta-Boldú FX (2018) Deep learning in agriculture: A survey. **Computers and Electronics in Agriculture** 147:70-90.

Kravchenko AN, Guber AK (2017) Soil pores and their contributions to soil carbon processes. **Geoderma** 287:31–39.

Kravchenko NA, Guber AK, Razavi BS, Koestel J, Quigley MY, Robertson GP, Kuzyakov Y (2019) Microbial spatial footprint as a driver of soil carbon stabilization. **Nature Communications** 10:3121.

Kravchenko AN, Negassa W, Guber AK, Schmidt S (2014) New Approach to Measure Soil Particulate Organic Matter in Intact Samples Using X-ray Computed Microtomography. **Soil Science Society of America Journal** 78:1177-1185.

Kuzyakov Y, Blagodatskaya E (2015) Microbial hotspots and hot moments in soil: Concept & review. **Soil Biology and Biochemistry** 83:184-199.

Lavallee JM, Soong JL, Cotrufo MF (2020) Conceptualizing soil organic matter into particulate and mineral-associated forms to address global change in the 21st century. **Global Change Biology** 26:261-273.

Lehmann J, Kleber M (2015) The contentious nature of soil organic matter. **Nature** 528:60-68.

Maenhout P, De Neve S, Wragg J, Rawlins B, De Pue J, Hoorebeke L Van, Cnudde V, Sleutel S (2021) Chemical staining of particulate organic matter for improved contrast in soil X-ray μ CT images. **Scientific Reports** 11:370

Miqueles E, Koshev N, Helou ES (2018) A Backprojection Slice Theorem for Tomographic Reconstruction. **IEEE Transactions on Image Processing** 27:894-906

Miqueles EX, Rinkel J, O'dowd F, Bermú Dez B A JS v (2014) Synchrotron Radiation Generalized Tatarski's algorithm for ring artefacts reduction. **Journal of Synchrotron Radiation** 21:1333-1346.

Pennington A, King ONF, Tun WM, Ho EML, Luengo I, Darrow MC, Basham M (2022) SuRVoS 2: Accelerating Annotation and Segmentation for Large Volumetric Bioimage Workflows Across Modalities and Scales. **Frontiers in Cell and Developmental Biology** 10:842342.

Perry D (2013) Image processing/ Histograms. Disponível em: http://daniel-perry.github.io/class/cs6640/histograms/dperry_cs6640_project1_report.html

Peth S, Chenu C, Leblond N, Mordhorst A, Garnier P, Nunan N, Pot V, Ogurreck M, Beckmann F (2014) Localization of soil organic matter in soil aggregates using synchrotron-based X-ray microtomography. **Soil Biology and Biochemistry** 78:189-194.

Pinto A, Borin G, Carlos B, Bernardi ML, Sarmento MF, Peixinho AZ, Spina T, Miqueles EX (2022) Annotat3D: A Modern Web Application for Interactive Segmentation of Volumetric Images at Sirius/LNLS. **Synchrotron Radiation News** 35:36-43.

Pires LF, Auler AC, Roque WL, Mooney SJ (2020) X-ray microtomography analysis of soil pore structure dynamics under wetting and drying cycles. **Geoderma** 362:114103.

Quigley MY, Rivers ML, Kravchenko AN (2018) Patterns and sources of spatial heterogeneity in soil matrix from contrasting long term management practices. **Frontiers in Environmental Science** 6:28

Rabot E, Wiesmeier M, Schlüter S, Vogel HJ (2018) Soil structure as an indicator of soil functions: A review. **Geoderma** 314:122-137.

Rawlins BG, Wragg J, Reinhard C, Atwood RC, Houston A, Lark RM, Rudolph S (2016) Three-dimensional soil organic matter distribution, accessibility and microbial respiration in macroaggregates using osmium staining and synchrotron X-ray computed tomography. **SOIL** 2:659-671.

Rippner DA, Raja P V., Earles JM, Momayyezi M, Buchko A, Duong F V., Forrestel EJ, Parkinson DY, Shackel KA, Neyhart JL, McElrone AJ (2022) A workflow for segmenting soil and plant X-ray computed tomography images with deep learning in Google's Colaboratory. **Frontiers in Plant Science** 13:3225.

Rumpel C, Amiraslani F, Chenu C, Garcia Cardenas M, Kaonga M, Koutika LS, Ladha J, Madari B, Shirato Y, Smith P, Soudi B, Soussana JF, Whitehead D, Wollenberg E (2020) The 4p1000 initiative: Opportunities, limitations and challenges for implementing soil organic carbon sequestration as a sustainable development strategy. **Ambio** 49:350-360.

Santos HG, Jakomine PKT, Anjos LHC, Oliveira VA, Lumberras JF, Coelho MR, Almeida JA, Araujo-Filho JC, Oliveira JB, Cunha TJF (5 Eds.) (2018). Sistema brasileiro de classificação de solos. Embrapa, 350p.

Schlüter S, Leuther F, Albrecht L, Hoeschen C, Kilian R, Surey R, Mikutta R, Kaiser K, Mueller CW, Vogel H-J (2022) Microscale carbon distribution around pores and particulate organic matter varies with soil moisture regime. **Nature Communications** 13: 2098.

Schlüter S, Sammartino S, Koestel J (2020) Exploring the relationship between soil structure and soil functions via pore-scale Imaging. **Geoderma** 370:114370.

Six J, Bossuyt H, Degryze S, Denef K (2004) A history of research on the link between (micro)aggregates, soil biota, and soil organic matter dynamics. **Soil and tillage research** 79:7-31.

Six J, Elliott ET, Paustian K (2000) Soil macroaggregate turnover and microaggregate formation: A mechanism for C sequestration under no-tillage agriculture. **Soil Biology and Biochemistry** 32:2099-2103.

Soltaninejad M, Sturrock CJ, Griffiths M, Pridmore TP, Pound MP (2020) Three Dimensional Root CT Segmentation Using Multi-Resolution Encoder-Decoder Networks. **IEEE Transactions on Image Processing** 29:6667-6679

Steffens M, Rogge DM, Mueller CW, Hö C, Lugmeier J, Kö A, Kö Gel-Knabner I (2017) Identification of Distinct Functional Microstructural Domains Controlling C Storage in Soil. **Environmental science & technology** 51:12182-12189.

Tuller M, Kulkarni R, Fink W (2015) Segmentation of X-ray CT data of porous materials: A review of global and locally adaptive algorithms. **Soil-Water-Root Processes: Advances in Tomography and Imaging** 61:157-182.

Vasconcelos GJQ, Antonietti G, Libel GC, Rosa PRR, Archilha NL, Carvalho TJ, Pedrini H, Spina T V (2018) Evaluation of Segmentation Methods Based on Classification Patterns for Micro-Tomography Applications in Rock Analysis. In: **Conference on Graphics, Patterns and Images, 31.(SIBGRAPI)**. Sociedade Brasileira de Computação, Foz do Iguaçu, PR, Brazil.

Young IM, Crawford JW, Rappoldt C (2001) New methods and models for characterising structural heterogeneity of soil. **Soil and Tillage Research** 61:33-45.

Zheng H, Kim K, Kravchenko A, Rivers M, Guber A (2020) Testing Os Staining Approach for Visualizing Soil Organic Matter Patterns in Intact Samples via X-ray Dual-Energy Tomography Scanning. **Environmental Science Technology** 54:8980-8989

APÊNDICE

1. Resolução vs. tamanho de pixel

A resolução geométrica (σ_R) da imagem gerada pode ser definida como a variância da função de espalhamento pontual (*point spread function* - PSF) do sistema de imagem, conforme a Eq. (1) (Bartels, 2013; Gureyev et al., 2008):

$$\sigma_R = \sqrt{\left(1 - \frac{1}{m}\right)^2 \sigma_S^2 + \frac{1}{m^2} \sigma_D^2} \quad (1)$$

onde σ_S e σ_D foram considerados neste trabalho como o tamanho nominal da fonte ($\sigma_S = 919 \times 228 \mu\text{m}^2$ – horizontal $\times$ vertical) e o tamanho do pixel do detector ($\sigma_D = 0,82 \mu\text{m}$), respectivamente; m é a magnificação geométrica ($m = 1 + \frac{Z_1}{Z_2}$) que, por sua vez, depende das distâncias fonte-amostra ($Z_1 = 16.700 \text{ mm}$) e amostra-detector ($Z_2 = 40 \text{ mm}$). A ampliação é ≈ 1 devido à distância Z_1 muito grande e Z_2 muito pequena. Portanto, a influência do tamanho da fonte na resolução geométrica é menor e o tamanho do pixel do detector é a limitação crítica (como discutido em Ferreira et al., 2022c). Assim, a resolução horizontal $\times$ vertical da imagem foi estimada em $2,35 \times 0,98 \mu\text{m}^2$.

2. Descrição geral do processo de segmentação utilizando o software Annotat3D

Conforme descrito no texto principal, a filtragem é uma etapa opcional e depende da avaliação inicial de sua necessidade. Após a decisão pelo uso ou não de filtragem preliminar, se a amostra é menor do que o campo de visão do detector, é necessário atribuir um rótulo (*classe*) ao fundo (*background*) da imagem (entorno do agregado preenchido por ar). O próximo passo é construir os dados de entrada (imagens segmentadas), por exemplo na via anotação do software Annotat3D, para então avançar para a via *deep learning*. Pelo Annotat3D é necessário configurar os parâmetros que serão utilizados, como a extração de recursos, superpixels e classificador, além de declarar as fases que serão consideradas.

Os superpixels são definidos a partir do espaçamento de sementes (*waterpixels seed spacing*), compactação (*waterpixels compactness*) e agrupamento de superpixels (*superpixel feature pooling*). A função extração de recursos (*feature extraction*) extrai características numéricas ou simbólicas de uma imagem. Em outras palavras, esta função transforma a imagem em um vetor, por meio da aplicação de

filtros na imagem de entrada, com o intuito de enriquecer as informações concentradas nos superpixels e garantir as propriedades que sejam relevantes aos objetos de interesse (trata-se de uma versão filtrada da imagem de entrada). As opções de filtro disponíveis são apresentadas na Tabela 1A. E, por fim, são definidos o tipo de classificador e o número de árvores (*random forest n. trees*). Assim, é possível iniciar as anotações, ou seja, a marcação de regiões na imagem para identificar elementos específicos que serão a base para o classificador. Com base nessas informações e na extração de recursos/filtros, o classificador aplica a segmentação para toda a imagem, sendo necessárias novas anotações até atingir a segmentação esperada.

Tabela 1A. Parâmetros considerados para definir a extração de recursos, superpixels e classificador na segmentação via anotação do Annotat3D.

Change classifier parameters		
Feature extraction	Superpixels	Classification
FFT Gaus	Waterpixels	Classifier:
None (Original Img)	Waterpixels 3D	Random Forest
Sobel	Pixel Segmentation	Linear SVM
Minimum	Waterpixels Seed Spacing	Multi Layer Perceptron
Average	Waterpixels Compactness	KNN
Median	<i>Superpixel feature pooling:</i>	AdaBoost
FFT Gabor	Minimum	<i>Find best parameters</i>
FFT DoG	Mean	<i>Random forest N. trees</i>
Membrane Projections	Maximum	
Maximum		
Variance		
LBP		
<i>Multi-scale filtering windows</i>		
<i>Feature selection:</i>		
Importance thershold		

Assim, são feitas as anotações para as fases de interesse declaradas, em diferentes fatias da imagem obtida pela μ TCX e o classificador é executado à medida que as fases são extrapoladas para as demais regiões da imagem que não contém anotações (segmentação via anotação). A partir das classificações resultantes deste processo são definidas as sementes (*seeds*) para segmentação baseada em

watershed usando o software Avizo (quando necessárias correções e/ou ajustes nos limites, contatos e preenchimentos da *label* resultante no Annotat3D), além de possíveis correções manuais. Essa etapa pode ser realizada em qualquer outro software de processamento de imagens que o operador julgue mais interessante (por exemplo, o Avizo 3D) ou uma combinação entre programas/ferramentas de segmentação.

Para o *deep Learning*, primeiro é necessário designar um espaço de trabalho, que consiste na criação de um diretório estruturado automaticamente pelo Annotat3D que irá salvar todas as decisões do *pipeline*. Em seguida, carrega-se o conjunto de dados (*dataset*) que será utilizado no treinamento (imagem *grayscale*, *label* correspondente resultante da segmentação descrita no passo anterior, e mapa de pesos). É necessário que as imagens do *dataset* apresentem um campo de visão suficiente para capturar a forma e a textura do elemento/objeto de interesse, ou seja, que contenham informações necessárias para a sua diferenciação das outras fases da imagem. Em seguida é possível acrescentar um mapa de pesos para cada uma das imagens utilizadas (opcional), para que o classificador entenda que deve priorizar áreas específicas, com a finalidade de induzir a rede a focar na fase/objeto de interesse, principalmente quando essa fase corresponde a uma parcela bem pequena da imagem. É necessário ainda, definir o número de classes, o tamanho da amostra (*sample size*) e o tamanho do *patch* (*patch size*). O tamanho da amostra caracteriza o número de subimagens (*patches*) a serem selecionadas das imagens de entrada que irão realmente alimentar o treinamento da rede neural a partir de uma estratégia uniforme, ou seja, os elementos são extraídos aleatoriamente sobre uma distribuição uniforme. O tamanho do *patch* representa as dimensões em 3D de cada uma dessas subimagens (que define o campo de visão do treinamento).

No Annotat3D ainda há um recurso, durante a criação do *dataset*, com estratégias de aumento do conjunto de dados (*augmentation*) (Tabela 1A) por gerar dados artificiais a partir de deformações que são aplicadas aos dados originais. Uma observação é que o tamanho do conjunto de dados pode aumentar drasticamente a cada estratégia incrementada no conjunto, por isso é necessária uma avaliação dos ganhos gerados e o tempo de processamento necessário.

O próximo passo, após a finalização do *dataset*, é definir a arquitetura de rede neural que será utilizada e a função de perda a ser minimizada por um algoritmo de aprendizagem, calculada a partir da imagem de referência (*dataset*) e a imagem de saída (imagem que está sendo classificada como resultado do treinamento da rede neural). A função de perda dependerá de um equilíbrio entre a minimização dos erros, o tempo de treinamento da rede e a precisão visual a ser obtida (entre a classificação resultante e a imagem de referência, conhecida como “padrão ouro”) e mede essencialmente o erro obtido durante a otimização para modelar a função de classificação. Dentre as configurações (Tabela 2A) estão a definição do *batch size*, que irá definir quantos *patches* (recortes) da imagem serão considerados ao mesmo tempo durante o treinamento da rede; o parâmetro iterações (*iterations*), que irá determinar quantas vezes a rede testará o conjunto de dados, sendo que cada iteração promove uma atualização no treinamento (este número define quantas vezes um lote é processado por todas as camadas da rede e são reajustadas); a taxa de aprendizagem (*learning rate*); a função de perda (*loss*); otimizador (*optimizer*). Por último, são definidos o tamanho do *patch size* do treinamento (tamanho da entrada de rede) e o *drop classifier*.

O operador pode acompanhar o desenvolvimento do treinamento por meio da variação da perda (erro) de acordo com o número de iterações ou por uma análise de visualização, através da aba *Board*, enquanto o treinamento está acontecendo. É possível ainda realizar um refinamento (*finetune*) no intuito de tentar melhorar o resultado do treinamento, alterando parâmetros ou aumentando o número de iterações para continuar com o treinamento. Com o término do treinamento são inicialmente feitas inferências para classificação do objeto segmentado na mesma imagem *grayscale* 3D utilizada no treinamento, para validação ou refutação da sua aplicabilidade. Quando o resultado da inferência for satisfatório, é realizada uma inferência “às cegas”, ou seja, usando imagens *grayscale*s 3D de outro agregado de solo, que não foi usada para o treinamento. Novamente o resultado é avaliado, e caso sejam identificadas regiões de erro, novos dados de entrada podem ser propostos (que abrangem a região com erros ou a mesma morfologia/textura da estrutura não identificada corretamente), para então proceder com *finetune*, até atingir a otimização final da RNC.

Tabela 2A. Parâmetros considerados para a segmentação via *deep learning* no Annotat3D.

Dataset manager		Network manager			Inference	
Sampling	Augmentation	Network	Dataset	Settings	Inference	Settings
Data	Vertical Flip	Unet2d	Load dataset (input):	<i>System:</i>	Network (input)	<i>System:</i>
Label	Horizontal Flip	Unet3d	Dimensions	GPUs	Input images	GPUs
Weight	Rotate 90 degrees	Vnet	Data info	<i>Training:</i>	Output folder	<i>Patches:</i>
<i>Sampling:</i>	Rotate -90 degrees		Label info	Batch size		Volume padding
Classes	Contrast		Weight info	Iterations		Patch border
Sample size	Linear Contrast			Learning rate		
Patch size	Dropout			Loss		
	Gaussian Blur			Optimizer		
	Average Blur			<i>Network:</i>		
	Additive Poisson Noise			Patch size		
	Elastic Deformation			Drop classifier		