



Bernardo Amaral Pascarelli Ferreira

**Super-Resolução em Imagens
Tomográficas de Briquetes de
Minério de Ferro Utilizando
Aprendizado Profundo**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Química, de Materiais e Processos Ambientais do Departamento de Engenharia Química e de Materiais da PUC-Rio.

Orientador: Prof. Sidnei Paciornik

Coorientadora: Dra. Karen Soares Augusto

Rio de Janeiro
Agosto de 2023



Bernardo Amaral Pascarelli Ferreira

**Super-Resolução em Imagens
Tomográficas de Briquetes de
Minério de Ferro Utilizando
Aprendizado Profundo**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Química, de Materiais e Processos Ambientais da PUC-Rio. Aprovada pela Comissão Examinadora abaixo.

Prof. Sidnei Paciornik

Orientador

Departamento de Engenharia Química e de Materiais - PUC-Rio

Dra. Karen Soares Augusto

Coorientadora

Departamento de Engenharia Química e de Materiais - PUC-Rio

Dr. Otávio da Fonseca Martins Gomes

Centro de Tecnologia Mineral - CETEM

Prof. Guilherme Lúcio Abelha Mota

Departamento de Informática e Ciência da Computação - UERJ

Rio de Janeiro, 30 de agosto de 2023

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Bernardo Amaral Pascarelli Ferreira

Graduado em Engenharia de Materiais e Nanotecnologia pela Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio – em 2021. Principais áreas de interesse: Caracterização Microestrutural, Processamento Digital de Imagens, Visão Computacional e Aprendizado Profundo.

Ficha Catalográfica

Ferreira, Bernardo Amaral Pascarelli

Super-resolução em imagens tomográficas de briquetes de minério de ferro utilizando aprendizado profundo / Bernardo Amaral Pascarelli Ferreira ; orientador: Sidnei Paciornik ; coorientadora: Karen Soares Augusto. – 2023.

99 f. : il. color. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Química e de Materiais, 2023.

Inclui bibliografia

1. Engenharia Química e de Materiais – Teses. 2. Caracterização mineral. 3. Microtomografia de raios-X. 4. Visão computacional. 5. Aprendizado profundo. 6. Super-resolução de única figura. I. Paciornik, Sidnei. II. Augusto, Karen Soares. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Química e de Materiais. IV. Título.

CDD: 620.11

Agradecimentos

Aos meus pais, Ivan e Eneida, e minha irmã, Fernanda, por todo o apoio e carinho.

Às orientações, acolhimento e paciência de meu orientador, Sidnei Paciornik, e de minha coorientadora, Karen Augusto, que muito me ensinam, incentivam e inspiram.

Ao Marcos Henrique pela disposição e ajuda nas diversas capturas e reconstruções das imagens em microCT, sem as quais este trabalho não poderia ser feito.

Ao Alei Domingues, por trazer este desafio, pela colaboração durante o projeto, e por disponibilizar todas as amostras necessárias.

Aos meus amigos Igor, Richard, Thalita, Laura e Francisco pela amizade e pelas inúmeras ajudas durante o período desse projeto.

Aos meus amigos Giovani, Breno e Thais, que considero irmãos.

À minha parceira Tatiana, por fazer parte da minha vida.

Ao CNPq, Vale e PUC-Rio, pelos auxílios concedidos, que possibilitaram o desenvolvimento deste projeto.

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

Resumo

Ferreira, Bernardo Amaral Pascarelli; Paciornik, Sidnei (Orientador); Augusto, Karen Soares (Coorientadora). **Super-Resolução em Imagens Tomográficas de Briquetes de Minério de Ferro Utilizando Aprendizado Profundo**. Rio de Janeiro, 2023. 99p. Dissertação de Mestrado – Departamento de Engenharia Química e de Materiais, Pontifícia Universidade Católica do Rio de Janeiro.

A indústria mineral vem presenciando, ao longo das últimas décadas, uma redução da qualidade de minério de ferro extraído e o surgimento de novas demandas ambientais. Esta conjuntura fortalece a busca por produtos provenientes do minério de ferro que atendam aos requisitos da indústria siderúrgica, como é o caso de novos aglomerados de minério de ferro. A Microtomografia de Raios-X (microCT) permite a caracterização da estrutura tridimensional de uma amostra, com resolução micrométrica, de forma não-destrutiva. Entretanto, tal técnica apresenta diversas limitações. Quanto melhor a resolução, maior o tempo de análise e menor o volume de amostra adquirido. Modelos de Super Resolução (SR), baseados em Deep Learning, são uma poderosa ferramenta para aprimorar digitalmente a resolução de imagens tomográficas adquiridas em pior resolução. Este trabalho propõe o desenvolvimento de uma metodologia para treinar três modelos de SR, baseados na arquitetura EDSR, a partir de imagens tomográficas de briquetes de redução direta: Um modelo para aumento de resolução de 16 μm para 6 μm , outro para aumento de 6 μm para 2 μm , e o terceiro para aumento de 4 μm para 2 μm . Esta proposta tem como objetivo mitigar as limitações do microCT, auxiliando o desenvolvimento de novas metodologias de Processamento Digital de Imagens para os aglomerados. A metodologia inclui diferentes propostas para avaliação do desempenho da SR, como comparação de PSNR e segmentação de poros. Os resultados apontam que a SR foi capaz de aprimorar a resolução das imagens tomográficas e mitigar ruídos habituais da tomografia.

Palavras-Chave

Caracterização mineral; Microtomografia de Raios-X; Visão Computacional; Aprendizado Profundo; Redes Neurais Convolucionais; Super-Resolução de Única Figura.

Abstract

Ferreira, Bernardo Amaral Pascarelli; Paciornik, Sidnei (Advisor); Augusto, Karen Soares (Co-advisor). **Super-Resolution in Tomographic Images of Iron Ore Briquettes Employing Deep Learning**. Rio de Janeiro, 2023. 99p. Dissertação de Mestrado – Departamento de Engenharia Química e de Materiais, Pontifícia Universidade Católica do Rio de Janeiro.

The mining industry has been witnessing a reduction of extracted iron ore's quality and the advent of new environmental demands. This situation reinforces a search for iron ore products that meet the requirements of the steel industry, such as new iron ore agglomerates. X-ray microtomography (microCT) allows the characterization of a sample's three-dimensional structure, with micrometer resolution, in a non-destructive analysis. However, this technique presents several limitations. Better resolutions greatly increase analysis time and decrease the acquired sample's volume. Super-Resolution (SR) models, based on Deep Learning, are a powerful tool to digitally enhance the resolution of tomographic images acquired at lower resolutions. This work proposes the development of a methodology to train three SR models, based on EDSR architecture, using tomographic images of direct reduction briquettes: A model for enhancing the resolution from 16 μm to 6 μm , another for enhancing from 6 μm to 2 μm , and the third for enhancing 4 μm to 2 μm . This proposal aims to mitigate the limitations of microCT, assisting the development and implementation of new Digital Image Processing methodologies for agglomerates. The methodology includes different proposals for SR's performance evaluation, such as PSNR comparison and pore segmentation. The results indicate that SR can improve the resolution of tomographic images and reduce common tomography noise.

Keywords

Mineral characterization; X-ray Microtomography; Computer Vision; Deep Learning; Convolutional Neural Networks; Single Image Super Resolution.

Sumário

1 Introdução	14
2 Objetivos	16
3 Revisão de literatura	17
3.1. Aglomerados de minério de ferro	17
3.2. Microtomografia de Raios-X	19
3.2.1. Funcionamento	20
3.2.2. Resolução da Imagem	21
3.2.3. Limitações	22
3.3. Processamento digital de imagens	23
3.4. Redes Neurais	24
3.4.1. Redes Neurais Convencionais	25
3.4.2. Aprendizagem	28
3.4.3. Redes Neurais Convolucionais	30
3.4.4. Redes Neurais Residuais	33
3.5. Super Resolução	35
3.5.1. Formulação	35
3.5.2. Métodos de Super Resolução	36
3.5.3. Métodos baseados em Aprendizado Profundo	37
3.5.4. Conjunto de dados	42
3.5.5. Métricas de Avaliação de Qualidade	43
3.5.6. Super Resolução em MicroCT	45
4 Metodologia	46
4.1. Preparo das Amostras	47
4.2. Captura em MicroCT	49
4.3. Reconstrução das Imagens Tomográficas	52
4.4. Criação do Conjunto de Dados	
4.4.1. Re-Sampling	53
4.4.2. Recorte da Região de interesse	54

4.4.3. Registro	55
4.4.4. Exportar sequência de imagens	56
4.4.5. Divisão do conjunto de dados	57
4.5. Pipeline de dados	58
4.6. Treinamento da Super Resolução	59
4.7. Etapa de Teste	61
4.7.1. Análise Visual	61
4.7.2. PSNR	61
4.7.3. Segmentação de poros	61
4.7.4. Precision, Recall e F1-Score	61
4.7.5. Segmentação de Volume 3D	63
5 Resultados e discussões	64
5.1. Resultados dos treinamentos	64
5.1.1. Modelo SR-I: 16 para 06 (3X)	64
5.1.2. Modelo SR-II: 06 para 02 (3X)	75
5.1.3. Modelo SR-III: 04 para 02 (2X)	84
5.2. Considerações finais	91
6 Conclusão	93
7 Referências bibliográficas	95

Lista de siglas e abreviaturas

ANN: Artificial Neural Networks

BN: Batch Normalization

CNN: Convolutional Neural Networks

DL: Deep Learning

DRCN: Deep Recursive Convolution Network

EDSR: Enhanced Deep Super Resolution

ESPCN: Efficient Sub-Pixel Convolutional Neural Network

FN: Falso Negativo

FP: Falso Positivo

GPU: Graphics Processing Unit

HR: High-Resolution

LR: Low-Resolution

MicroCT: Microtomografia computadorizada de Raios-X

PDI: Processamento Digital de Imagens

PReLU: Parametric Rectified Linear Unit

PSNR: Peak Signal-to-Noise Ratio

RSISR: Real-World Single Image Super Resolution

Roi: Region of Interest

SGD: Stochastic Gradient Descent

SISR: Single Image Super Resolution

SR: Super Resolução

SRCNN: Super-Resolution Convolutional Neural Network

SRResNet: Super-Resolution Residual Network

VDSR: Very Deep Super-Resolution

VP: Verdadeiro Positivo

Lista de figuras

Figura 1: Esquema das etapas que compõe o funcionamento da microCT. Adaptado de [10].	21
Figura 2: Geometria cônica. Retirado de [11].	21
Figura 3: Fluxograma clássico das etapas de PDI.	23
Figura 4: Esquema de um neurônio artificial. Retirado de [17].	25
Figura 5: Curvas das funções de ativação famosas. Adaptado de [20].	27
Figura 6: Exemplo de uma rede neural com arquitetura composta uma camada de entrada, duas camadas intermediária e uma camada de saída. Retirado de [24].	28
Figura 7: Exemplo da operação de convolução. Retirado de [26].	32
Figura 8: Exemplo do comportamento da degradação de acurácia. O gráfico da esquerda apresenta o erro de treinamento, enquanto o da esquerda apresenta o erro do teste. Retirado de [27].	33
Figura 9: Esquema de um bloco residual composto por duas camadas convolucionais com ativação ReLU. Retirado de [27].	34
Figura 10: Arquitetura da SRResNet. Retirado de [41].	38
Figura 11: Demonstração da operação de PixelShuffle para $r = 2$. Os quatro canais com dimensão de LR compõe os pixels da imagem de dimensões HR. Retirado de [40].	39
Figura 12: Comparação entre a composição dos blocos residuais na arquitetura SRResNet (esquerda) e na EDSR (direita). Adaptado de [42].	40
Figura 13: Esquema da arquitetura EDSR. Retirado de [42].	40
Figura 14: Gráfico Desempenho versus número de operações, para avaliação comparativa de diversos modelos de SR. Retirado de [34].	41
Figura 15: Comparação da qualidade visual de uma imagem de barco com diferentes tipos de distorções. Todas as imagens possuem MSE = 210. Retirado de [44].	44
Figura 16: Esquema das etapas presentes na metodologia desenvolvida.	47
Figura 17: Recortes da amostra de briquete F-33 C. Ao lado esquerdo, uma das faces foi cortada longitudinalmente em 3 fatias; ao lado direito, a outra face foi cortada transversalmente em 4 fatias.	48
Figura 18: Uma amostra com a pequena barra fixada em seu maior eixo.	48
Figura 19: Equipamento da ZEISS, modelo Xradia 510 versa, utilizado para captura tomográfica das amostras de briquete.	49

Figura 20: Interior do equipamento de microCT. Observa-se a fonte de raios-x à esquerda, a amostra a ser analisada, e do detector à direita.	51
Figura 21: Visualização da diferença de volumes, na mesma escala, entre a captura em 6 μm (esquerda) e 2 μm (direita) da amostra F33C-TA.	54
Figura 22: Visualização dos dois volumes cilíndricos, 6 μm (esquerda) e 2 μm (direita), gerados a partir da seleção de uma Região de Interesse.	55
Figura 23: Fatias de uma Rol em 6 μm (esquerda) e 2 μm (direita). É possível observar que diversas estruturas pertencentes ao briquete, como os poros destacados pelos retângulos vermelhos, não são idênticos, revelando que as fatias não estão alinhadas.	55
Figura 24: Fatias dos pares de imagens devidamente registradas. É possível observar que após o registro, as mesmas estruturas do briquete estão presentes nas fatias.	56
Figura 25: Fatia número 1 da sequência (direita) e a fatia seguinte (número 2). As imagens são muito semelhantes.	57
Figura 26: Fatia número tal da sequência (direita) e a fatia seguinte (número tal). As imagens já possuem características distintas.	57
Figura 27: Exemplo da operação de patching. A imagem original (a) possui 964x964 pixels de tamanho. O padding automático expande a imagem para uma dimensão de 1152x1152 pixels. Assim, o pipeline gera 9 patches de 384x384 pixels.	59
Figura 28: Visualização de uma imagem de baixa resolução (esquerda), imagem super resolvida pelo modelo SR-I (centro), e a imagem de alta resolução (direita). Corte da amostra F34B-TA.	65
Figura 29: Exemplo do efeito halo presente em regiões de fronteira de uma imagem LR e HR, para a região de interesse indicada.	66
Figura 30: Visualização das versões da Rol em baixa resolução, super resolvida e alta resolução	67
Figura 31: Processo de segmentação das regiões porosas. (a) e (b) se referem a imagem de LR. (c) e (d) se referem a imagem de HR. (e) e (f) se referem a imagem super resolvida.	68
Figura 32: Visualização das imagens binárias obtidas para segmentação de poros nas versões LR, HR e SR.	69
Figura 33: Comparação de poros segmentados. As imagens se referem à binária super resolvida com a seleção, em amarelo, de poros da imagem de baixa resolução (a) e com a seleção, em verde, de poros da imagem de alta resolução (b).	72
Figura 34: Resultados da operação AND entre LR e HR (a), e SR e HR (b).	73
Figura 35: Visualização dos poros segmentados em LR (esquerda), super resolução (centro) e HR (direita).	74

Figura 36: Comparação das imagens LR (esquerda), super resolvida pelo modelo SR-II (centro), e HR (direita). Corte da amostra F34B-TA.	76
Figura 37: Visualização das versões da Rol em baixa resolução, super resolvida e alta resolução.	77
Figura 38: Comparação das imagens super resolvida (esquerda), 4 μm (centro), e HR (direita).	78
Figura 39: Visualização das versões da Rol super resolvida, em 4 μm e em 2 μm .	79
Figura 40: Processo de segmentação das regiões porosas. A segmentação da LR foi feita na faixa tonal de 1-118; de 4 μm foi feita entre 1-115; da super resolvida foi feita entre 1-100; e da HR foi feita entre 1-101.	80
Figura 41: Visualização das imagens binárias obtidas para segmentação de poros nas versões LR, 4 μm , SR e HR.	81
Figura 42: Resultados da operação AND entre LR e HR (a), 4 μm e HR (b), e SR e HR (c).	82
Figura 43: Visualização dos poros segmentados em LR (esquerda), super resolução (centro) e HR (direita).	83
Figura 44: Comparação das imagens LR (esquerda), super resolvida pelo modelo SR-III (centro), e HR (direita). Corte da amostra F34B-TA.	85
Figura 45: Visualização das versões da Rol em baixa resolução (esquerda), super resolvida (centro) e alta resolução (direita).	86
Figura 46: Processo de segmentação das regiões porosas. A segmentação da LR foi feita na faixa tonal de 1-114; da super resolvida foi feita entre 1-104; e da HR foi feita entre 1-100.	87
Figura 47: Visualização das imagens binárias obtidas para segmentação de poros nas versões LR, HR e SR.	88
Figura 48: Resultados da operação AND entre LR e HR (a), e SR e HR (b).	89
Figura 49: Visualização dos poros segmentados em LR (esquerda), super resolução (centro) e HR (direita).	90

Lista de tabelas

Tabela 1: Comparação entre as especificações da SRResNet e EDSR. Adaptado de [42].	41
Tabela 2: Características principais do modelo Zeiss Xradia 510 Versa.	49
Tabela 3: Resumo das condições de captura para cada amostra.	52
Tabela 4: Tamanho dos conjuntos de dados para os diferentes modelos.	58
Tabela 5: Resumo dos conjuntos de treinamento e validação para os modelos propostos.	60
Tabela 6: Análises quantitativas feitas a partir da segmentação de poros.	70
Tabela 7: Valores de fração em área para diferentes faixas tonais.	71
Tabela 8: Medidas de Precision e Recall.	73
Tabela 9: Valores de porosidade das amostras 3D.	74
Tabela 10: Análises quantitativas feitas a partir da segmentação de poros.	82
Tabela 11: Medidas de Precision e Recall.	83
Tabela 12: Valores de porosidade das amostras 3D.	84
Tabela 13: Análises quantitativas feitas a partir da segmentação de poros.	88
Tabela 14: Medidas de Precision e Recall.	89
Tabela 15: Valores de porosidade das amostras 3D.	90
Tabela 16: Dados de tempo de análise e volume adquirido para a amostra F34B-TA.	91
Tabela 17: Dados de tempo para aplicação da SR na amostra F34B-TA.	91
Tabela 18: Resumo das comparações de tempo.	92

1

Introdução

A produção de aglomerados de minério de ferro tem sido cada vez mais relevante para a indústria minero-siderúrgica. Isso se deve a diversos fatores, como a falta de granulados de alta qualidade (alto teor de ferro), a necessidade em reaproveitar materiais oriundos do processamento mineral e da produção de aço, e a procura em reduzir emissões de poluentes. Além disso, os métodos de aglomeração permitem a geração de produtos com propriedades controladas para atender demandas específicas das rotas siderúrgicas.

Existe, também, uma procura por tecnologias de aglomeração que tragam maior eficiência econômica e térmica à fabricação de aglomerados, e que possam reduzir impactos ambientais.

Portanto, existe uma conjuntura que fortalece o desenvolvimento e aperfeiçoamento tanto das tecnologias de aglomeração, quanto da qualidade dos aglomerados. E para isso, a etapa de caracterização dos aglomerados torna-se fundamental.

A Microtomografia de Raios-X (microCT) é um método de caracterização não-destrutivo que permite a visualização tridimensional da estrutura interna de uma amostra. A microCT, junto ao Processamento Digital de Imagens (PDI), permite a elaboração de diversas metodologias para análise quantitativas, como discriminação automática entre poros e trincas em pelotas, e quantificação da porosidade de amostras de pelotas. [6][7]

Entretanto, a microCT possui certas limitações. O aumento de resolução da aquisição tomográfica só é possível com a redução do campo de visão do volume de amostra que será capturado. Além disso, a microCT é uma análise demorada, e os equipamentos possuem elevados custos financeiros.

Portanto, a existência de limitações na microCT, tendo em vista a demanda em estudar os aglomerados de minério de ferro, é uma motivação para investigar soluções que possam torná-la um método mais eficiente para a caracterização de aglomerados.

Neste contexto, a Super Resolução (SR) baseada em *Deep Learning* (DL) é uma técnica capaz de aumentar digitalmente a resolução de uma imagem, enquanto restaura detalhes de alta frequência espacial. A aplicação desse método para imagens tomográficas possibilitaria o aumento de resolução sem a necessidade de reduzir o volume de captura de uma amostra, assim como traria redução no tempo de aquisição de dados e, conseqüentemente, economia no uso do equipamento.

O presente trabalho propõe o desenvolvimento de uma metodologia para treinar modelos de SR a partir de imagens tomográficas de briquetes de redução direta capturadas em piores e melhores resoluções. Esta proposta tem como objetivo mitigar as limitações do microCT, auxiliando o desenvolvimento de novas metodologias de Processamento Digital de Imagens para os aglomerados.

2 Objetivos

O presente trabalho tem como principal objetivo adaptar e avaliar o uso da técnica de Super Resolução Enhanced Deep Super Resolution para imagens de briquete de redução direta capturadas por microCT. Dentro dos objetivos específicos, pode-se listar:

- Criar um conjunto de imagens tomográficas de briquetes de minério de ferro, conduzidos em capturas de pior e melhor resolução.
- Realizar o treinamento de três diferentes modelos de Super Resolução: Um capaz de melhorar a resolução de 16 μm para 6 μm ; outro capaz de melhorar a resolução de 6 μm para 2 μm , e o terceiro, capaz de melhorar a resolução de 4 μm para 2 μm .
- Compreender o desempenho dos modelos por meio de análises de imagens e por métricas amplamente utilizadas para avaliação de qualidade visual.

3

Revisão de literatura

3.1 Aglomerados de minério de ferro

O minério de ferro é uma matéria-prima fundamental para a indústria siderúrgica, movimentando cerca de 300 bilhões de dólares por ano, o que a torna um dos *commodities* mais importantes para a infraestrutura global.

Devido a sua relevância na produção de aço, os minérios de ferro são, principalmente, avaliados de acordo com seu teor de ferro. Como a produção eficiente de aço em alto-forno requer, em geral, uma matéria-prima com concentração de ferro superior a 58%, os minérios abaixo desse valor são considerados de baixa qualidade. Minérios de alta qualidade possuem concentração de ferro superior a 60%. [1][2]

A produção de aço segue uma tendência de crescimento desde os anos 50, fortemente acelerada a partir os anos 2000, o que ocasionou o esgotamento dos recursos de minérios de alta qualidade [2]. A exploração de reservas de minério com menores concentrações de ferro trouxe à tona a necessidade de processos de beneficiamento para aumentar o teor de ferro encontrado nesses minérios.

Tais operações, a fim de separar os minerais de interesse econômico dos minerais de ganga, incluem etapas de redução de tamanho das partículas dos minérios (cominuição), como a britagem e moagem. Assim, os processos geram produtos ricos em ferro, mas de granulometria fina.

Por exemplo, o minério de ferro bruto passa primeiramente por uma britagem, que gera os granulados (de 6,3 a 31 mm) e os finos de sinterização (*sinter feed*, de 0,15 a 6,3 mm). Os granulados, não possuindo alto teor de ferro, passam por mais processos de beneficiamento, onde se dá origem a um pó muito fino, chamado de finos de pelota (*pellet feed*, menos de 0,15 mm) [3]. Apesar de ricos em ferro, os finos gerados nas etapas de processamento mineral não possuem tamanho adequado para as rotas siderúrgicas. Por isso, as tecnologias de aglomeração são necessárias para transformar os finos em um produto aproveitável para a indústria minero-siderúrgica. Esse aproveitamento não se

torna vantajoso apenas pela questão de esgotamento do minério de alta qualidade, mas também devido à procura por reduzir impactos ambientais. [4]

Dentre os principais processos de aglomeração utilizados na indústria, tem-se a sinterização, a pelletização e a briquetagem.

A sinterização é um processo de fusão incipiente, em uma temperatura ótima entre a temperatura de calcinação e a temperatura de fusão, da mistura de finos de sinterização com fundentes. Esse método promove uma ligação rígida das partículas após o processo de solidificação da fase líquida, dando origem ao sínter. [4]

A pelletização tem como matéria-prima uma mistura úmida dos finos de pelletização com determinadas quantidades de aglomerante e óxidos básicos. Essa mistura é levada para os discos ou tambores de pelletização. Durante a rotação destes discos, a ação da tensão superficial de capilaridade da mistura úmida promove a formação de um aglomerado, nomeado de pelota crua. Essas são, então, submetidas a um tratamento térmico em atmosfera controlada, para adquirirem resistência mecânica, dando origem à pelota. [4]

A briquetagem é o processo mais simples de aglomeração. Ele consiste na aplicação de pressão em partículas finas, com adição de água, e auxílio ou não de aglutinantes. As forças de atração molecular de Van der Waals tem atuação na união das partículas, após a aplicação de pressão externa reduzir a distância entre partículas. Essa força externa pode ser aplicada por processos como prensas de rolos ou prensas hidráulicas. Para obterem resistência mecânica desejada, os briquetes passam por um processo de endurecimento, que pode ser realizado à temperatura ambiente, em estufas e secadores (aproximadamente 400°) ou fornos (acima de 1.000°). [4]

O aumento de tamanho e o aproveitamento dos finos não são os únicos aspectos que justificam a adoção das etapas de aglomeração. Durante a produção de aço, deseja-se uma reação estável e homogênea no alto-forno, ao mesmo tempo em que se alcança eficiência econômica e consumo mínimo de coque. Portanto, propriedades dos aglomerados, como redutibilidade, porosidade, permeabilidade, comportamento mecânico, entre outros, são essenciais para atender as demandas dos processos siderúrgicos [1][5]. Sendo assim, as etapas de aglomeração possibilitam a obtenção de produtos com características específicas que atendem as diferentes tecnologias empregadas na indústria siderúrgica.

Por outro lado, os processos de aglomeração estão associados a altos custos. Portanto, existe uma procura por tecnologias que tragam maior eficiência econômica e térmica, reduzindo também impactos ambientais. [4]

Em suma, o aperfeiçoamento das técnicas da aglomeração e o desenvolvimento de novos aglomerados é fundamental para o atual contexto da indústria minero-siderúrgica, tendo em vista a procura do aumento de eficiência na produção de aço, redução de custos de operação e redução de impactos ambientais. Por consequência, a caracterização desses produtos torna-se uma etapa essencial tanto para o estudo das relações entre as tecnologias de aglomeração e as propriedades do aglomerado produzido, quanto para o controle de qualidade durante a produção em escala industrial.

Devido às diversas propriedades que podem ser analisadas nos aglomerados, uma gama de técnicas de caracterização é importante para seu estudo. Em especial, este trabalho tem foco na utilização da Microtomografia de Raios-X, devido às recentes metodologias de processamento digital de imagens tomográficas desenvolvidas, como discriminação automática entre poros e trincas em pelotas, ou quantificar a porosidade de amostras de pelotas [6][7].

3.2 Microtomografia de Raios-X

A Microtomografia de Raios-x (microCT) é uma técnica de microscopia que permite a visualização tridimensional da estrutura interna de uma amostra sólida, com resolução típica da ordem de micrometros. Isso possibilita adquirir medidas precisas do volume tridimensional da amostra.

Além de requerer pouca ou sequer nenhuma preparação de amostra, a microCT é uma caracterização não-destrutiva. Portanto, é possível realizar uma captura tomográfica da amostra, sem interferir com sua estrutura original, possibilitando que uma mesma amostra seja investigada por outros métodos de caracterização ou que seja submetida posteriormente a processos físico-químicos. Esta característica é muito relevante na investigação de aglomerados, já que a microCT possibilita que uma mesma amostra de aglomerado seja tomografada em diversas etapas de um processo de redução ou de ensaios mecânicos. Também possibilita que uma análise quantitativa da captura tomográfica seja diretamente comparada ao resultado quantitativo de outra caracterização.

3.2.1 Funcionamento

A microCT é uma técnica que utiliza o mapeamento da atenuação do feixe de raios-x pela matéria para produzir imagens. Este processo pode ser explicado a partir de três constituintes básicos: a fonte de raios-X, um porta-amostra giratório e um detector. [8]

O feixe de Raios-X é gerado por meio da desaceleração brusca de elétrons altamente energizados durante a colisão com um alvo metálico (efeito *Bremsstrahlung*), dentro do tubo de raios-X [9]. O poder de penetração e a intensidade do feixe de raios-X, assim como sua faixa de comprimentos de onda, dependem da voltagem aplicada à fonte de raios-X.

O feixe que incide na amostra é atenuado devido às diferentes interações entre os raios-X e matéria (absorção fotoelétrica, espalhamento Compton e produção de pares). Assim, o feixe que atravessa a amostra sem interação passa pelo cintilador, que o transforma em luz visível. Esse feixe convertido alcança, finalmente, o fotodetector, que registra os dados de uma projeção bidimensional da amostra.

A partir desta configuração, é possível capturar um conjunto de projeções bidimensionais de diferentes ângulos da amostra, fazendo uso da rotação do porta-amostra. O passo angular desta rotação é um parâmetro controlável, o que possibilita o aumento ou redução do número de projeções. Um maior número de projeções promove a captura de mais informações da estrutura da amostra.

Após esse processo, um algoritmo de reconstrução é utilizado para gerar, a partir do conjunto de projeções, imagens que representam camadas bidimensionais do volume da amostra. Essas camadas, ou fatias, são empilhadas para formar o volume tridimensional da amostra, onde um pixel passa a ser um voxel.

A Figura 1 apresenta de forma resumida todas as etapas descritas.

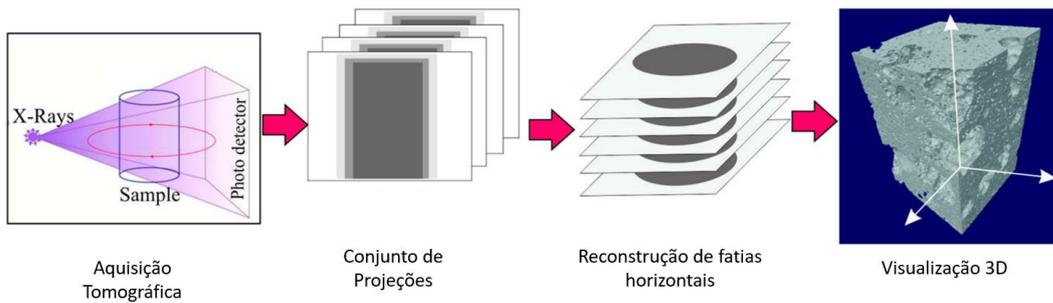


Figura 1: Esquema das etapas que compõe o funcionamento da microCT. Adaptado de [10].

3.2.2 Resolução da Imagem

Conceitualmente, a resolução espacial é a menor distância em que é possível distinguir duas estruturas distintas. Assim, quanto menor o valor de resolução, maior é a capacidade de um sistema identificar estruturas. Em um equipamento de microCT, o fator limitante do poder máximo de resolução do sistema está relacionado com o tamanho do ponto focal produzido pela fonte de raios-X.

Em um microtomógrafo de geometria cônica (Figura 2), como é o caso do equipamento utilizado neste trabalho, a resolução espacial da captura é definida pela magnificação geométrica, ou seja, pelas relações de distâncias entre fonte, amostra e detector.

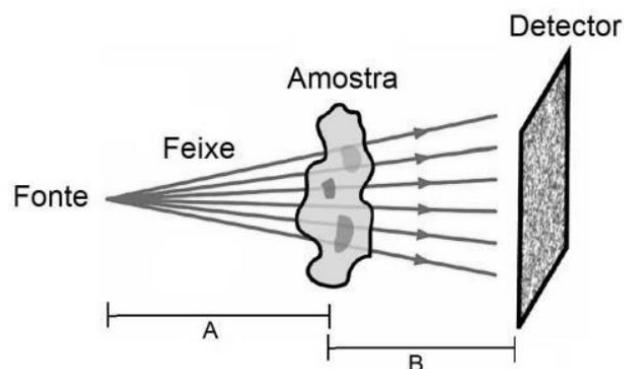


Figura 2: Geometria cônica. Retirado de [11].

Esta relação entre resolução (R) e magnificação geométrica (M) é dada por Equação (1).

$$R = \frac{k}{M} \quad (1)$$

onde k é o tamanho do pixel do detector e M , a magnificação geométrica, é uma relação entre distância fonte-amostra (A) e distância amostra-detector (B), segundo a Equação (2).

$$M = \frac{A + B}{A} \quad (2)$$

Alguns microtomógrafos também são equipados com um conjunto de lentes objetivas, como é o caso do equipamento utilizado neste trabalho. Isso permite que as condições de resolução sejam definidas pela magnificação geométrica ou pela magnificação ótica, ou a combinação das duas.

É essencial informar que, apesar de, conceitualmente, uma resolução menor ser sinônimo de uma resolução melhor, no sentido de discriminar estruturas menores, este trabalho utiliza termos em outra perspectiva. Este estudo utiliza os termos “baixa resolução” ou “menor resolução” como sinônimos de uma captura tomográfica em pior resolução espacial. Os termos “alta resolução” ou “maior resolução” são utilizadas no sentido de uma captura em melhor resolução espacial, ou seja, que discrimina estruturas menores.

3.2.3 Limitações

Algumas limitações presentes na técnica da microCT dificultam seu emprego e, conseqüentemente, sua popularização, principalmente na indústria.

A principal limitação está relacionada com a relação entre a resolução espacial da captura e o volume de amostra capturado. De acordo com a Equação (1), para realizar o aumento da resolução, é necessário reduzir a distância entre a fonte de raios-x e a amostra. Entretanto, essa aproximação faz com que o feixe de raios-x atravesse um volume de amostra menor. Ou seja, o aumento da resolução espacial resulta na captura de um volume menor da amostra.

Outra adversidade diz respeito ao longo tempo das aquisições tomográficas. Este problema é agravado ao se considerar que a análise de uma amostra de maiores dimensões em melhores resoluções necessita de várias capturas.

Além disso, equipamentos de microCT possuem elevado custo, assim como sua manutenção, tendo em vista que uma fonte de Raios-X possui vida útil.

3.3 Processamento digital de imagens

O PDI diz respeito a qualquer forma de manipulação de imagens digitais, a fim de corrigir e ressaltar regiões de interesse nas imagens para obter informações qualitativas e quantitativas durante a Análise Digital de Imagens (ADI). A Figura 3 apresenta as principais etapas presentes no PDI.

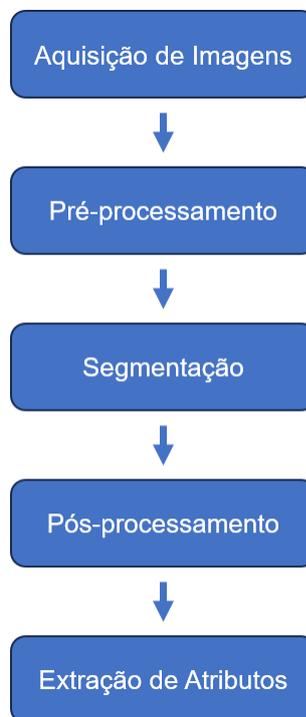


Figura 3: Fluxograma clássico das etapas de PDI e ADI.

Após a aquisição tomográfica, a imagem obtida passa por devidos tratamentos na etapa de pré-processamento. Nesta etapa, corrigem-se eventuais defeitos gerados no momento da captura da imagem, como redução de ruídos, correção de iluminação não-homogênea, melhora de contraste, entre outros. Outras técnicas para realce de detalhes ou regiões de interesse também podem ser aplicadas.

Após verificar que a imagem possui as qualidades desejadas, realiza-se a etapa de segmentação, cujo objetivo é isolar, ou separar, a região de interesse da

imagem. Dentre os diversos métodos de segmentação existentes, este trabalho emprega a limiarização (*threshold*), que discrimina uma região por meio da seleção de uma faixa tonal.

É comum que os métodos de segmentação não sejam capazes de discriminar perfeitamente todas as informações desejadas. Ou seja, as imagens resultantes do processo de segmentação normalmente contêm defeitos, como a presença de objetos indesejados ou a falta de determinados detalhes. Esses defeitos podem ser corrigidos na etapa de pós-processamento, como filtragens simples por tamanho e forma.

Após o pós-processamento, última etapa de PDI, obtém-se uma imagem com as regiões desejadas para realizar análises quantitativas. Assim, a etapa de extração de atributos fornece dados numéricos para a descrição das características de cada objeto e da imagem como um todo.

3.4 Redes Neurais

As Redes Neurais Artificiais (*Artificial Neural Networks* - ANN) são algoritmos computacionais de Aprendizado de Máquina (*Machine Learning*), possuindo capacidade de autoaprendizagem e adaptabilidade aos dados utilizados. [12]

Ao longo da última década, a área de ANN tem experimentado notável crescimento em sua relevância e na obtenção de conquistas, principalmente devido às significativas contribuições que conferiram ao estado-da-arte dos campos de visão computacional e linguística computacional.

As ANNs têm sido protagonistas na transformação de diversas indústrias. Obtiveram grande sucesso em resolver diversas tarefas, tais como classificação de imagens, reconhecimento de objetos, segmentação, reconhecimento de fala, tradução automática, criação de novas imagens e vídeos, entre outros. Sua presença é evidente em produtos e serviços populares, utilizados no dia a dia (mecanismos de busca, redes sociais, e-mail, e-commerce, câmeras fotográficas, entre outros), bem como em soluções tecnológicas avançadas que impactam positivamente áreas críticas, como medicina (diagnóstico de câncer e medicina de precisão) e transporte (carros autônomos). O constante e acelerado progresso nesses campos indica que as redes neurais continuarão a desempenhar um papel crucial no futuro da computação e inovação tecnológica. [13][14][15]

Embora a origem das redes neurais remeta à década de 40, o rápido avanço desse campo ocorreu apenas na última década, devido ao desenvolvimento de poderosas unidades de processamento gráfico (*GPUs*), à diminuição dos custos computacionais e à ascensão da cultura da *big data*. [14][16]

3.4.1 Redes Neurais Convencionais

Atualmente, as redes neurais se dividem em diferentes técnicas com características próprias, cada uma mais indicada para abordar determinados tipos de problemas. Dentre elas, têm-se as *Convolutional Neural Networks*, *Residual Neural Networks*, *Recurrent Neural Networks*, *Long Short-Term Memory Networks*, *Generative Adversarial Networks*, *Auto-Encoders*, entre outras.

Apesar de toda essa variedade de métodos, todas essas redes neurais compartilham dos mesmos conceitos básicos, elaborados em redes neurais mais simples, ou redes neurais *convencionais*.

3.4.1.1 Neurônios Artificiais

As ANNs são algoritmos computacionais que se inspiram na capacidade de processamento de dados e de aprendizado dos neurônios biológicos. Nesse sentido, adotam neurônios artificiais como unidades básicas de processamento de informações.

A Figura 4 apresenta o fluxograma com as etapas que compõem o algoritmo do neurônio artificial:

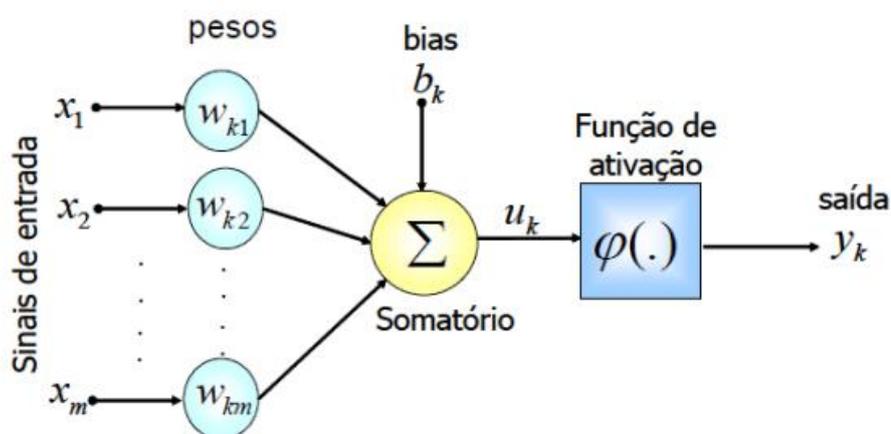


Figura 4: Esquema de um neurônio artificial. Retirado de [17].

O modelo de neurônio artificial inicia com a recepção de um conjunto (de tamanho n) de sinais de entrada x_j (inputs). Cada valor de entrada é associado a um peso w_j . A principal função do neurônio é calcular o produto de cada entrada x_j por seu respectivo peso w_j , para então realizar o somatório (Σ) de todos esses produtos. Adicionalmente, também é possível acrescentar ao somatório o termo bias b_k , cujo propósito é ajustar o valor do somatório por meio de uma transformação afim. [18]

Assim, o sinal produzido pelo somatório (u_k) pode ser descrito matematicamente como:

$$u_k = b_k + \sum_{j=1}^n x_j w_j \quad (3)$$

O sinal u_k é, então, aplicado na função de ativação, etapa fundamental para o funcionamento dos neurônios. De forma prática, ela determina o nível de ativação de um neurônio, pois ajusta ou limita os valores de saída dos neurônios de acordo com a função escolhida. Além disso, a função de ativação desempenha um papel crítico na eliminação da linearidade presente na Equação (3), uma vez que fenômenos naturais raramente apresentam comportamentos lineares. A utilização da função de ativação também é primordial para a correta execução do método de retropropagação (*backpropagation*), responsável pelo processo de aprendizagem das redes. Diversos tipos de funções matemáticas são empregados como função de ativação em um neurônio, cuja escolha sempre depende de características do sinal de entrada, qual é o objetivo da rede, entre outros conceitos. Existe um grande empenho, na literatura de redes neurais, em definir ou identificar funções de ativação que melhorem o desempenho de rede. [19][20][21]

A Figura 5 expõe as curvas de algumas funções de ativação amplamente reconhecidas na área, tais como as funções *sigmoid*, *tanh* e ReLU. Em especial, a função *Parametric ReLU* (PReLU) [22] é utilizada nas redes neurais deste trabalho.

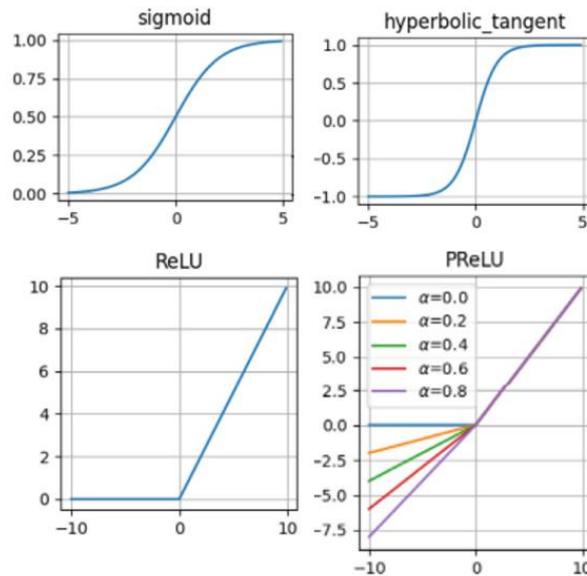


Figura 5: Curvas das funções de ativação famosas. Adaptado de [20].

A saída de um neurônio pode, então, ser modelada como:

$$y_k = \varphi(u_k) \quad (4)$$

Desta forma, entende-se que o neurônio artificial é capaz de produzir uma função matemática, a partir de seus pesos, do bias e da função de ativação. Neste sentido, a criação de uma rede de neurônios interligados, onde os sinais de saída de um neurônio y_k alimentam outros neurônios, é um método muito eficiente para gerar funções matemáticas extremamente complexas, sem a necessidade de modelagem humana. Todos os parâmetros presentes nessas funções são ajustados durante o processo de aprendizagem, que será discutido mais adiante, utilizando os dados de entrada como referência.

3.4.1.2 Arquitetura de Redes Neurais

A arquitetura de uma rede neural se refere, resumidamente, às formas de organizar as conexões entre neurônios e o fluxo de dados entre eles.

A disposição dos neurônios é feita por meio de camadas. A camada de neurônios que recebem os dados brutos é chamada de camada de entrada, enquanto a última camada é chamada de camada de saída. As saídas de cada neurônio de uma camada alimentam os neurônios presentes na camada seguinte. A rede neural mais básica é composta apenas por uma camada de entrada e uma camada de saída. Essa rede é historicamente conhecida como rede *Perceptron*, cuja simplicidade a habilita a resolver somente problemas lineares. [23]

Como solução para esta limitação, redes neurais com mais camadas entre a entrada e saída foram propostas. Essas novas camadas passaram a ser conhecidas como camadas escondidas (*Hidden Layers*) ou intermediárias, e tais redes foram denominadas de *Perceptrons* de múltiplas camadas. A Figura 6 traz um exemplo de uma rede neural com múltiplas camadas.

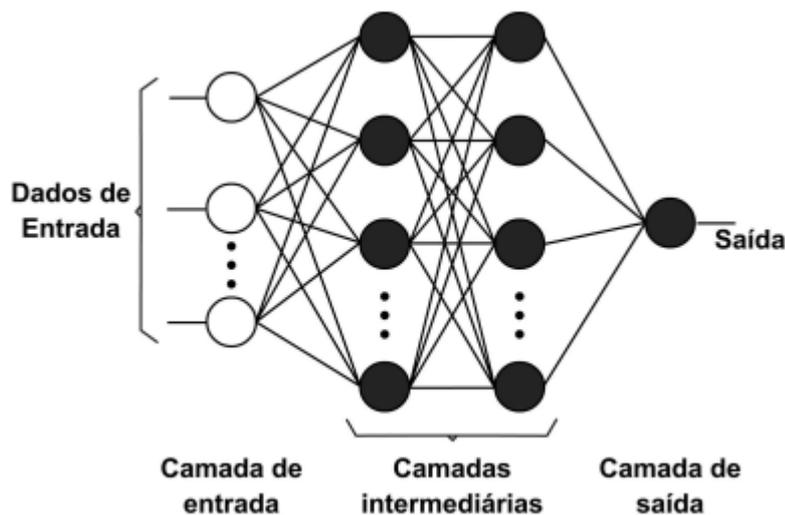


Figura 6: Exemplo de uma rede neural com arquitetura composta uma camada de entrada, duas camadas intermediária e uma camada de saída. Retirado de [24].

A quantidade de camadas presentes em uma arquitetura de rede neural é, comumente, designada pelo termo profundidade da rede. Assim, a denominação de Redes de Aprendizagem Profunda (*Deep Learning - DL*) surge para indicar redes que contam com muitas camadas escondidas.

Além disso, o sentido de fluxo de sinais também é uma característica importante na arquitetura das redes neurais. Nas redes *perceptron* de múltiplas camadas, o fluxo de dados ocorre em um único sentido, da camada de entrada para a camada de saída, conhecido como topologia *Feed-Forward*. Existem outras formas de fluxo de sinais, como por exemplo as topologias presentes em Redes Neurais Recorrentes e Redes Neurais Residuais.

3.4.2 Aprendizagem

A aprendizagem, ou treinamento, é uma etapa iterativa em que a rede neural ajusta seus pesos e bias, entendidos como parâmetros livres, repetidamente até alcançar um objetivo.

3.4.2.1 Tipos de Aprendizagem

Duas formas de aprendizado devem ser destacadas: Aprendizagem supervisionada e Aprendizagem não-supervisionada. [18]

A aprendizagem supervisionada utiliza um conjunto de dados que possui as entradas e as correspondentes saídas esperadas (alvos, ou *target*). Por meio deste par de entrada-saída, a rede neural consegue atualizar e ajustar seus pesos iniciais, até que as saídas da rede correspondam às saídas desejadas.

A aprendizagem não-supervisionada, por outro lado, utiliza apenas o conjunto de entrada, sem indicar a saída desejada. Este método é indicado quando se deseja que a rede descubra padrões nos dados.

No caso da aprendizagem supervisionada, que será utilizada neste trabalho, existem diversos métodos, conhecidos como algoritmos de aprendizagem, para modificar os pesos durante o treino, de forma que a saída gerada pela rede se torne o mais próximo possível da saída esperada.

3.4.2.2 Algoritmos de aprendizagem

Ao final de uma rede neural, após a propagação de um conjunto de dados de entrada x_j pelos neurônios, obtém-se uma saída y_k . Essa saída y_k pode, então, ser comparada à saída esperada t_k (*target*), fornecida pelo conjunto de dados do aprendizado supervisionado. Esta comparação é feita calculando a discrepância existente entre as duas saídas, por meio de uma função de erro (*Loss Function*). A partir do erro calculado, é possível ajustar os pesos da rede neural para que este erro seja diminuído. Ou seja, os algoritmos de aprendizagem supervisionada visam minimizar o erro entre a saída y_k e o alvo t_k .

Para realizar este processo de aprendizagem, são necessárias duas etapas. A primeira se refere ao cálculo dos gradientes da função de erro ao longo da rede neural, realizada pela classe de algoritmos conhecida como retropropagação (*Backpropagation*). A segunda etapa diz respeito a determinação do conjunto de variáveis que resultam na minimização da função de erro, a partir dos gradientes calculados pela retropropagação. Esta etapa é realizada pelos algoritmos de otimização, tais como *Stochastic Gradient Descent* (SGD) e *Adaptive Moment Estimation* (Adam). [25]

Assim, o algoritmo de otimização e a retropropagação são capazes de ajustar os parâmetros livres da rede neural. Esses processos são repetidos até que se atinja, idealmente, o mínimo global da função de erro.

3.4.2.3 Número de épocas

O número de épocas (*Epochs*) é outro parâmetro fundamental para o treinamento de redes neurais. Ele define o número de vezes em que o conjunto de dados completo será alimentado pela rede. Em redes DL, que costumam possuir extensos conjuntos de dados, são necessárias diversas iterações, ou seja, propagação das entradas e ajuste dos parâmetros livres, para que todo o conjunto de dados seja propagado pela rede. Por isso, costuma-se utilizar também o parâmetro de tamanho de lote (*Batch Size*), que representa o número de dados a ser utilizado em uma única iteração.

Desta forma, a relação entre número de épocas, número de iterações e tamanho de lote é dada por:

$$N^{\circ} \text{ de iterações/época} = \frac{\text{tamanho do conjunto de treinamento}}{\text{tamanho de lote}} \quad (5)$$

3.4.3 Redes Neurais Convolucionais

A aplicação das redes neurais em problemas de processamento e análise digital de imagens possui algumas peculiaridades. As redes neurais convencionais abordadas são extremamente deficientes para lidar com imagens, devido à sua incapacidade de capturar efetivamente as características essenciais de uma imagem, que dependem significativamente da posição e vizinhança dos pixels que a compõem. Uma rede neural convencional apenas recebe os valores de intensidade de cada pixel, sem os fatores de posicionamento e sem a compreensão das relações de vizinhança de pixels. Em outras palavras, as redes neurais convencionais são incapazes de extrair os atributos fundamentais de uma imagem.

Inspirado no funcionamento do córtex visual biológico, esse problema foi resolvido pela substituição da simples multiplicação realizada entre os sinais e os pesos por uma operação de convolução. [25]

Tais operações de convolução, feitas em diversas camadas ocultas, permitem que a rede realize etapas de pré-processamento de imagens e a extração de suas características essenciais (*feature extraction*).

Portanto, as Redes Neurais Convolucionais (CNNs) são altamente recomendadas para abordar os desafios do campo de visão computacional.

3.4.3.1 Convoluções

A convolução é definida como uma operação de multiplicação entre duas funções, de forma que: [25]

$$s(t) = x(t) * w(t) = \int x(a) \cdot w(t - a) da \quad (6)$$

Convoluções podem ser utilizadas em mais de um eixo ao mesmo tempo. Considerando a aplicação desta operação no processamento digital de imagens bidimensionais, esta operação pode ser objetivamente representada na seguinte forma:

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i - m, j - n) * K(m, n) \quad (7)$$

Na Equação (7), I representa uma imagem bidimensional, K é um *kernel* bidimensional, S representa uma saída denominada mapa de característica (*feature map*), enquanto i , j , m e n são os índices dos elementos matriciais. Esta equação demonstra que é feito um somatório da multiplicação entre o *kernel* K e a região resultante da interseção entre a imagem I e o *kernel* K . O *kernel* K é movido horizontalmente e verticalmente pela imagem I , de forma que o somatório é feito em cada uma das possíveis posições do *kernel* K dentro da imagem I . Os valores dos somatórios formam a matriz de saída, chamada de mapa de característica. A Figura 7 exemplifica uma etapa da operação de convolução entre a imagem I e o *kernel* K .

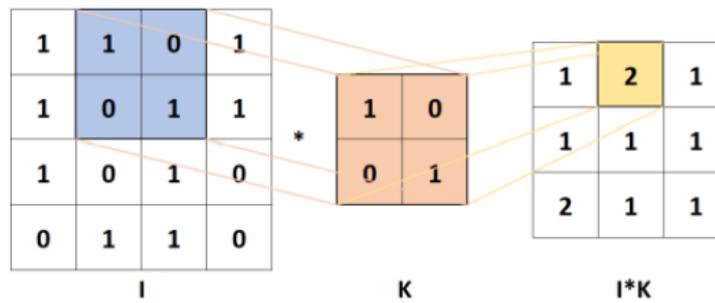


Figura 7: Exemplo da operação de convolução. Retirado de [26].

O uso da operação de convolução em redes neurais possui um paralelo com as Operações Locais, utilizadas na área de Processamento Digital de Imagens. As operações locais são comumente conhecidas como filtros de imagem. Esses filtros, definidos pelos coeficientes da matriz *kernel*, possibilitam a detecção de bordas, suavização ou realce das imagens originais, entre outras possibilidades. Assim, o emprego da convolução em diversas camadas ocultas possibilita que as redes de DL aprendam a decidir quais características essenciais de uma imagem são relevantes para resolver um determinado problema de visão computacional, uma vez que a rede irá modificar os filtros para minimizar o erro.

Convém destacar durante a especificação de uma arquitetura de *CNN*, sempre é descrita a quantidade de *kernels* utilizados em uma camada convolucional. O número de *kernels* normalmente é mencionado como número de filtros ou número de canais.

3.4.3.2 Batch Normalization

Embora as redes convolucionais profundas tenham proporcionado notáveis avanços no campo da visão computacional, mais especificamente em classificação de imagens, várias limitações de desempenho foram identificadas.

Até então, as evidências demonstravam que o aumento da profundidade da rede era uma característica positiva para algumas tarefas de reconhecimento visual triviais [27][28]. Os diversos experimentos de aumento de profundidade da rede ajudaram na detecção de obstáculos no treinamento de redes muito profundas.

Dentre esses obstáculos, tem-se o notório problema de dissipação/explosão do gradiente (*vanishing/exploding gradient*). A explosão de gradiente diz respeito ao aumento exponencial da norma do gradiente durante o treinamento. A dissipação, ou fuga, de gradiente se refere ao comportamento oposto, onde a

norma do gradiente durante o treinamento se aproxima rapidamente do valor zero. Estes comportamentos resultam em um aprendizado lento e maior dificuldade de alcançar a convergência do modelo. [29][30]

As propostas para solução deste problema envolvem a normalização inicial dos dados de entrada e a introdução de camadas intermediárias de normalização, conhecidas como normalização de lote (*Batch Normalization* - BN). O uso de BN permite a utilização de maiores taxas de aprendizado e menor número de iterações de treinamento. [31]

3.4.4 Redes Neurais Residuais

Embora as soluções para o problema de dissipação/explosão do gradiente tenham permitido o treinamento bem-sucedido de redes convolucionais mais profundas, outro obstáculo foi identificado.

Foi observado que conforme a profundidade de uma rede é aumentada, a acurácia atinge saturação a partir de um determinado número de camadas, e passa a degradar rapidamente com a adição de mais camadas. Este problema ficou conhecido como degradação da acurácia. Na Figura 8, que expõe o comportamento desta degradação, é possível notar que a rede com maior número de camadas apresenta uma curva com maiores valores de erro.

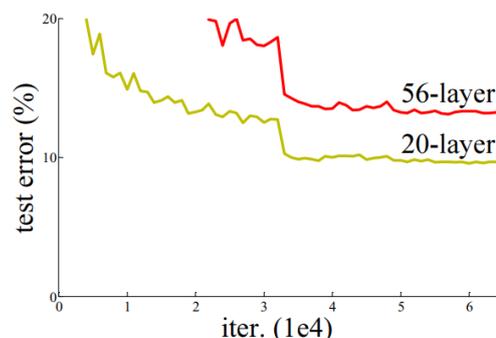


Figura 8: Exemplo do comportamento da degradação de acurácia. O gráfico erro do teste para duas arquiteturas com diferentes profundidades. O erro observado para a rede com 56 camadas é maior que o erro apresentado para a rede de apenas 20 camadas.

Retirado de [27].

Uma explicação mais intuitiva para a questão da degradação pode ser feita a partir do exemplo de um treinamento inicial uma rede muito profunda. Na primeira iteração de uma rede neural, todos os pesos iniciais são valores aleatórios. Assim, uma entrada é propagada por uma grande quantidade camadas

com pesos aleatórios até a saída, fazendo com que o sinal da entrada seja muito alterado ao longo da rede neural. Por meio dos algoritmos de aprendizado, os pesos são modificados a fim de minimizar o erro entre a saída e o alvo. Entretanto, os algoritmos de aprendizado acabam realizando modificações mais impactantes para os pesos das camadas mais próximas à saída, cujo sinal pode ser muito distante da informação original da entrada. E os pesos mais próximos da entrada acabam tendo atualizações menos significantes. Em outras palavras, redes neurais muito profundas são difíceis de treinar, porque as informações da entrada são negativamente afetadas quando propagadas em muitas camadas.

Para resolver este problema, He et al. [27] propuseram um método para que as informações da entrada alcancem camadas mais profundas sem serem negativamente afetadas, de forma que os algoritmos de aprendizagem sejam capazes de realizar um ajuste de pesos mais significativo nas camadas mais superficiais.

Isso é feito por meio da adição das conexões de atalho (*skip connections*). Essas conexões permitem que uma informação seja propagada diretamente para camadas posteriores, pulando uma ou mais camadas intermediárias, facilitando o cálculo dos gradientes durante o treinamento.

Na prática, isto é implementado por meio dos blocos residuais (*residual blocks*). Um exemplo de bloco residual é demonstrado na Figura 9. É possível notar que uma informação (x) é propagada em uma sequência de duas camadas convolucionais ($\mathcal{F}(x)$). Então, ao sinal de saída desta sequência de camadas ($\mathcal{F}(x)$), adiciona-se a informação original (x). A operação de adição ($\mathcal{F}(x) + x$) representa a conexão de atalho.

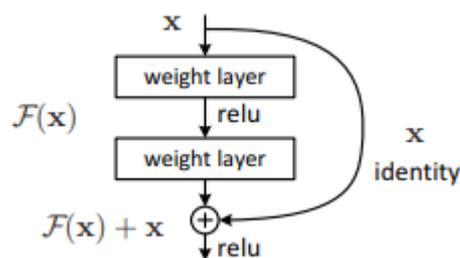


Figura 9: Esquema de um bloco residual composto por duas camadas convolucionais com ativação ReLU. Retirado de [27].

De forma mais genérica, os blocos residuais podem variar em sua composição, podendo incluir diferentes números de camadas e a presença de etapas de *batch normalization* e funções de ativação. Entretanto, os blocos

residuais são projetados para serem facilmente empilhados, ou seja, múltiplos blocos residuais podem ser postos em sequência para construir uma rede neural profunda. Portanto, o bloco residual é a unidade básica para a construção das redes residuais.

As redes construídas a partir dos blocos residuais são popularmente conhecidas como *ResNets*. As *ResNets* possibilitaram o treinamento bem-sucedido de redes com mais de 100 camadas de profundidade, além de apresentar melhor acurácia e alcançar a convergência mais rapidamente, quando comparadas à CNNs tradicionais. [27]

3.5 Super Resolução

A Super Resolução (SR) é um desafio clássico dos campos de Processamento Digital de Imagens (PDI) e Visão Computacional (Computer Vision), em que se deseja reconstruir, ou restaurar, uma imagem em alta resolução (*high-resolution - HR*) a partir de sua versão em baixa resolução (*low-resolution - LR*). A reconstrução de uma imagem HR por SR engloba tanto a obtenção de uma imagem com maior densidade de pixels, quanto a recuperação de detalhes de altas frequências indistinguíveis na LR. [32]

A SR é uma técnica almejada para várias aplicações que envolvem tarefas de reconhecimento ou detecção, como em imagiologia médica (exames de imagem), vigilância inteligente e sensoriamento remoto. [33]

A reconstrução pode ser realizada a partir de uma única imagem ou de várias imagens de entrada, o que permite categorizar os métodos de SR em duas famílias: *Single Image Super-Resolution (SISR)* e *Multi-Image Super-Resolution* [34]. Devido à natureza deste projeto, cuja intenção é a obtenção de uma imagem de alta resolução para cada imagem tomográfica capturada em baixa resolução, os métodos utilizados pertencem à classe *SISR*.

3.5.1 Formulação

A SR é um problema inverso ao processo de degradação. A relação entre a degradação de uma imagem desconhecida em alta-resolução X e sua versão correspondente em baixa-resolução Y é descrita em termos de uma função de degradação $D(\cdot)$, definida por um parâmetro desconhecido θ_D :

$$Y = D(X, \theta_D) \quad (8)$$

Portanto, como a SR tem como objetivo estimar a melhor correspondência para a imagem em alta-resolução X , ela é estabelecida pela contraversão do processo de degradação:

$$\hat{X} = SR(Y, \theta_{SR}) \quad (9)$$

Onde a função $SR(\cdot)$ representa o processo de restauração de imagem, definido por um parâmetro desconhecido θ_{SR} . Esse processo fornece em uma estimativa \hat{X} , chamada de imagem **super resolvida**, da imagem de alta resolução X . [34]

É possível encontrar, na literatura, simplificações para a função de degradação $D(X, \theta_D)$. Normalmente, essa simplificação é feita ao considerar que o processo de degradação de uma imagem é composto pela aplicação de um filtro passa-baixa, por uma etapa de *down-sampling* e pela adição de ruído. Assim, a Equação (8) é reescrita como:

$$Y = (X \otimes k) \downarrow_s + n \quad (10)$$

Onde a imagem em baixa-resolução Y é dada pelo *down-sampling* de fator s (conotado pelo operador \downarrow_s) feito sobre o resultado da convolução entre a imagem em alta-resolução X e um kernel passa-baixa k , além da adição de ruído n .

A Equação (9) e, conseqüentemente, a Equação (10) sugerem que diversas imagens de alta-resolução podem ser a solução de uma única imagem em baixa resolução. Sendo assim, a SISR é um problema matemático mal-posto (*ill-posed*). [33][35]

Portanto, os algoritmos de SISR utilizam métodos complexos para modelar uma função inversa à função de degradação, a fim de gerar as melhores estimativas de uma imagem HR a partir da imagem LR.

3.5.2 Métodos de Super Resolução

Os algoritmos de SR são divididos, de forma prática, em três categorias básicas: métodos baseados em interpolação, métodos baseados em reconstrução e métodos baseados em aprendizagem. [36][32]

Métodos baseados em interpolação compreendem os algoritmos clássicos de interpolação, como interpolação linear, bi-linear e bi-cúbica. Apesar de simples e rápidos, esses métodos não são capazes de recuperar detalhes refinados,

originando imagens com desfoque excessivo. Os métodos baseados em reconstrução determinam certas prioridades e restrições de reconstrução entre as imagens LR e HR, fazendo uso de técnicas de realce de bordas, regularizações e de-convoluções. [32][36]

Os métodos baseados em aprendizagem incluem as técnicas baseadas em *Machine Learning*. Esses métodos são capazes de analisar relações estatísticas entre pares de imagens correspondentes em baixa e alta resolução durante o aprendizado supervisionado. É possível categorizar os métodos baseados em aprendizagem em 5 grupos: métodos de *neighbor embedding*, métodos de *sparse coding*, métodos de *self-exemplar*, métodos de *locally linear regression* e métodos de DL. [36]

Com o recente avanço e sofisticação das CNNs, novos algoritmos baseados em DL vêm sendo desenvolvidos e explorados no âmbito da SR, ocupando o espaço de estado-da-arte atual da *SISR*.

3.5.3 Métodos baseados em Aprendizado Profundo

O uso de DL para SR teve início com a elaboração da *Super-Resolution Convolutional neural network (SRCNN)* [37]. Esse algoritmo apresentou resultados promissores que serviram de inspiração para o desenvolvimento de diversas redes convolucionais para SR, que incluíram técnicas mais avançadas como *residual learning* e *recurrent learning*. Dentre os modelos desenvolvidos, pode-se citar *Very Deep Super-Resolution (VDSR)* [38], *Deep Recursive Convolution Network (DRCN)* [39], *Efficient Sub-Pixel Convolutional Neural Network (ESPCN)* [40], entre outras.

Em especial, as redes *Super-Resolution Residual Network (SRResNet)* e *Enhanced Deep Super Resolution (EDSR)*, utilizada neste trabalho, devem ser discutidas com maior profundidade.

3.5.3.1 Arquitetura SRResNet

A arquitetura *Super-Resolution Residual Network (SRResNet)* foi proposta em 2017 por Ledig et al. [41]. A Figura 10 apresenta um modelo esquemático de sua arquitetura.

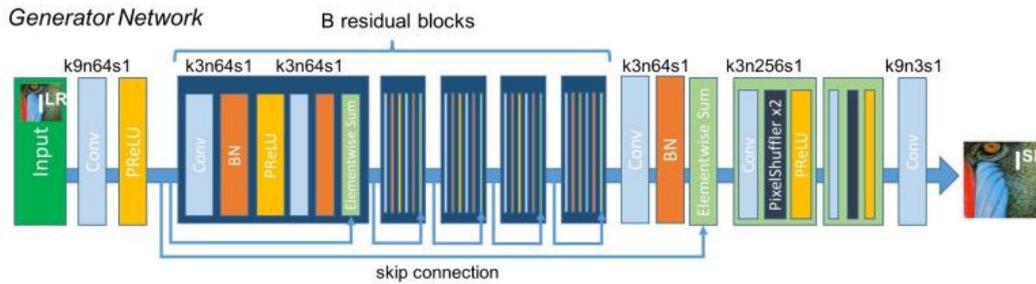


Figura 10: Arquitetura da SRResNet. Retirado de [41].

A primeira etapa do modelo consiste em uma camada de convolução (com 64 filtros de dimensões 9×9 pixels e *stride* igual a um), responsável pela extração de características, utilizando a *PReLU* como função de ativação. Em seguida, o sinal é propagado por uma quantidade B de blocos residuais, compostos por duas camadas de convolução (com 64 filtros de dimensões 3×3 pixels e *stride* igual a um), duas camadas de *batch normalization* e ativação *PReLU*. Após a sequência de blocos residuais, o sinal é propagado por mais uma camada convolucional (com 64 filtros de dimensões 3×3 pixels e *stride* igual a um), com aplicação de *batch normalization*. Depois, ocorre uma adição deste sinal de saída com o sinal utilizado como entrada dos blocos residuais, devido à uma conexão de atalho.

A seguir, o sinal é propagado por dois blocos de *Up-Sampling*, responsáveis pelo aumento de resolução da imagem com dimensões de LR para as dimensões de HR . Cada bloco é composto por uma camada convolucional, uma camada de *PixelShuffle* e ativação *PReLU*.

A camada de *PixelShuffle* é responsável por rearranjar todos os elementos matriciais presentes nos diferentes canais (filtros) da saída da camada convolucional em uma única imagem, obtendo a resolução da imagem HR . Esta operação utiliza como parâmetro o fator de aumento de resolução desejada (r), onde o número de canais deve corresponder a r^2 . Por exemplo, para uma rede de SR construída para o aumento de $2X$ ($r = 2$), são necessários 4 canais para compor a imagem em alta resolução. A Figura 11 apresenta um exemplo desta operação. [40] [43]

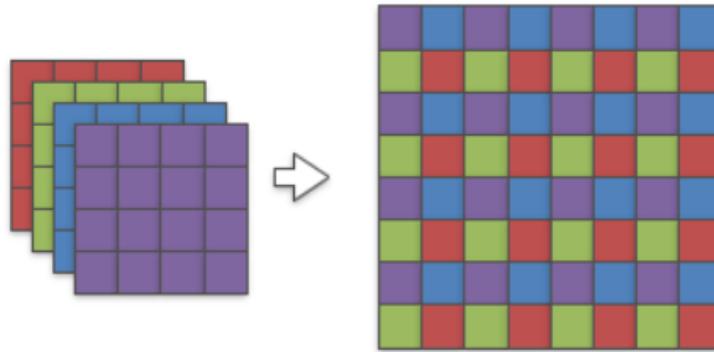


Figura 11: Demonstração da operação de *PixelShuffle* para $r = 2$. Os quatro canais com dimensão de LR compõe os pixels da imagem de dimensões HR. Retirado de [40].

Após o *Up-Sampling*, o sinal passa pela camada convolucional final (com 3 canais de dimensões 3×3 pixels e *stride* igual a um), originando a imagem HR como saída.

3.5.3.2 Arquitetura EDSR

A arquitetura *Enhanced Deep Super Resolution (EDSR)* [42] foi desenvolvida com base na *SRResNet*, visando aperfeiçoar seu desempenho.

A primeira modificação diz respeito à remoção das camadas de *batch normalization* de toda a rede, o que aumentou substancialmente o desempenho da rede. Além disso, tal eliminação também diminuiu aproximadamente 40% do uso de memória da GPU (em comparação a *SRResNet*), o que possibilita aumentar o tamanho/profundidade do modelo, utilizando o mesmo recurso computacional. A Figura 12 apresenta uma comparação entre os blocos residuais presentes na arquitetura *SRResNet* e na arquitetura *EDSR*.

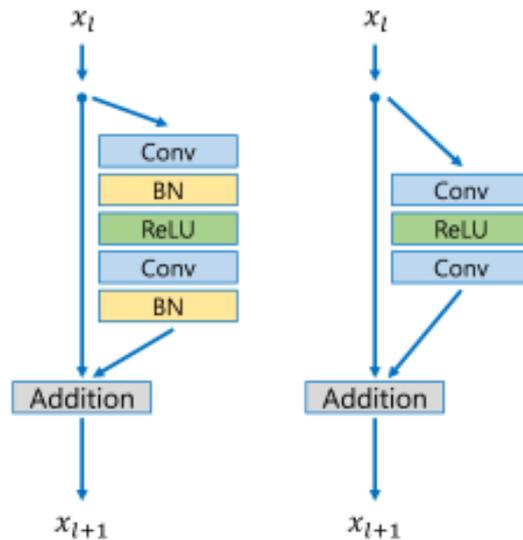


Figura 12: Comparação entre a composição dos blocos residuais na arquitetura SRResNet (esquerda) e na EDSR (direita). Adaptado de [42].

Outra alteração diz respeito ao aumento do número de parâmetros da rede. Os autores alegam que aumentar o número de filtros da rede, ou invés da profundidade da rede, maximiza a capacidade de desempenho do modelo. A arquitetura EDSR também remove qualquer camada de ativação ReLU de fora dos blocos residuais. A função de perda utilizada também é alterada para a função L1 (*Absolute Error*). A Figura 13 apresenta a arquitetura EDSR.

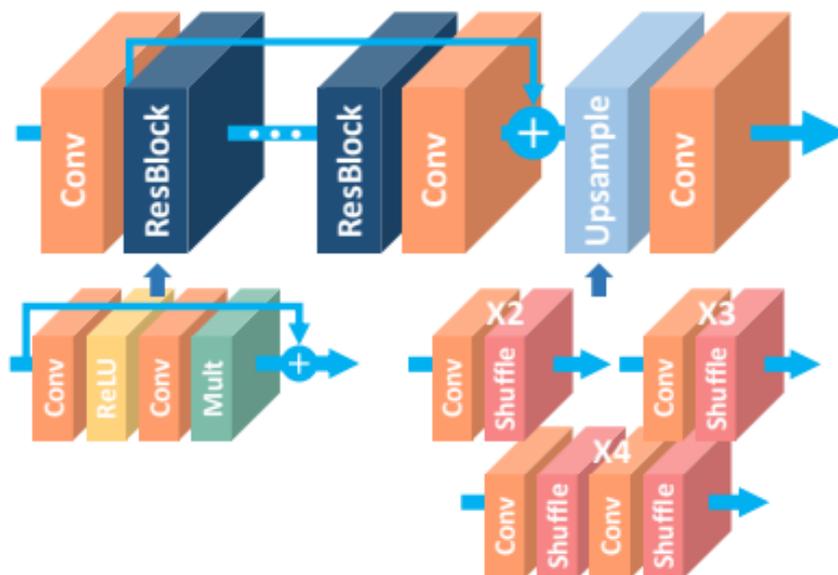


Figura 13: Esquema da arquitetura EDSR. Retirado de [42].

Considerando as modificações citadas, a rede possui como entrada uma camada convolucional, seguida dos blocos residuais, compostos apenas por 2 camadas convolucionais, uma ativação ReLU e uma camada de multiplicação pelo fator de *Residual Scaling*. Após os blocos, o sinal é propagado para uma camada convolucional. Depois, ocorre uma operação de adição dessa saída com o sinal que antecede os blocos residuais. A seguir, o sinal passa por um único bloco de *Up-Sampling*, composto por uma camada convolucional e uma camada de *PixelShuffle*. Por fim, o sinal é alimentado em uma camada convolucional.

A Tabela 1 apresenta uma comparação entre algumas especificações que resumem as mudanças executadas para elaboração da arquitetura EDSR.

Tabela 1: Comparação entre as especificações da *SRResNet* e *EDSR*. Adaptado de [42].

Opções	<i>SRResNet</i>	<i>EDSR</i>
Nº Blocos residuais	16	32
Nº Filtros	64	256
Uso de <i>BN</i>	Sim	Não
Função de erro	L2	L1

A Figura 14 apresenta um gráfico para avaliação comparativa de diversos modelos de SR, feita por Wang et al. [34]. A arquitetura EDSR (em verde, na direita) apresenta um ótimo desempenho, permanecendo entre os 5 melhores modelos avaliados.

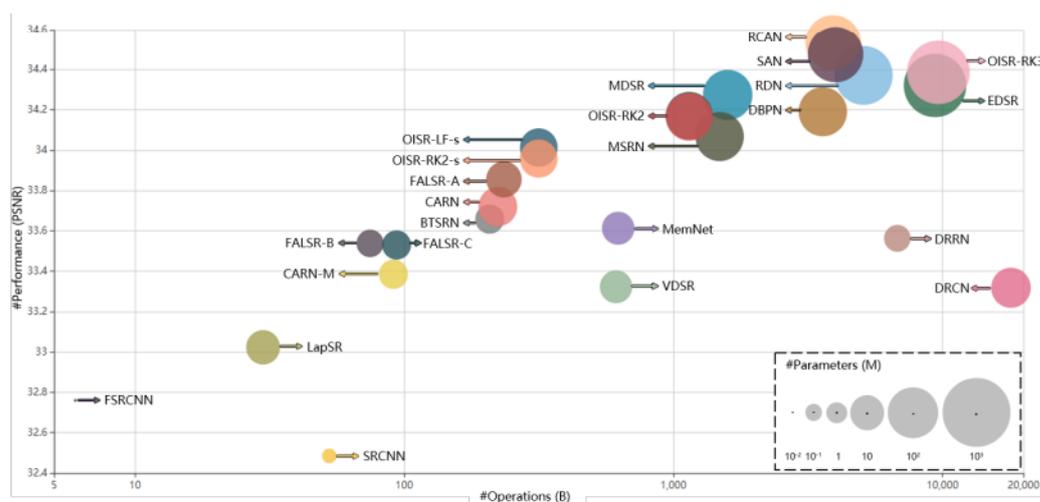


Figura 14: Gráfico Desempenho versus número de operações, para avaliação comparativa de diversos modelos de SR. Retirado de [34].

3.5.4 Conjunto de dados

O conjunto de dados (*dataset*) é a base para qualquer treinamento de modelo de DL. Os conjuntos devem ser robustos, em questão de quantidade de dados, e bem adequados ao propósito da tarefa. No caso da *SISR*, os conjuntos de dados devem conter pares de imagens correspondentes em baixa resolução e alta resolução.

Existe uma grande variedade de conjuntos de dados concebidos para o treinamento de modelos de *SISR*. Entre os mais utilizados na literatura, têm-se os conjuntos DIV2K, BSDS500, T91, Set5, Set14, entre outros. Um importante detalhe nesses conjuntos de dados é que as imagens de baixa resolução são digitalmente geradas a partir da imagem de alta resolução. Isto significa que as imagens de alta resolução são deterioradas digitalmente para simular um modelo de degradação. Isso é realizado por meio da aplicação de *down-sampling*, baseado em alguma técnica de interpolação, aplicação de filtros passa-baixa e adição de ruído. Imagens que possuem esta característica são denominadas de imagens sintéticas (*synthetic data*). [33]

Como é possível presumir, utilizar imagens sintéticas para o treinamento de um modelo que irá lidar com degradações reais não é ideal. O atual estágio da pesquisa em SR já reconhece que modelos treinados com um conjunto de imagens sintéticas não são adequados para lidar com um conjunto de imagens realistas (*real-world images*), assim como também evidencia que o treinamento utilizando um conjunto de imagens reais apresenta um desempenho muito inferior aos modelos treinados com imagens sintéticas. Em outras palavras, o desempenho dos modelos baseados em imagens sintéticas é superestimado. [33]

Um grande empenho na área de SR tem sido empregado para o desenvolvimento de métodos focados em imagens realísticas, abordagem conhecida como *Real-World Single Image Super Resolution (RSISR)*. Para isso, alguns conjuntos de dados foram elaborados, tais como RealSR, DRealSR, City100, SR-RAW, entre outros. [33]

Direcionando estas concepções para o campo da microtomografia, os conjuntos de imagens tomográficas também podem ser sintéticos ou realísticos.

Os conjuntos sintéticos podem ser gerados de duas formas. A primeira diz respeito à variação do número de projeções utilizados para a reconstrução das imagens tomográficas. Ou seja, as imagens de alta resolução são imagens

reconstruídas com maior número de projeções, e imagens de baixa resolução são reconstruídas com menor número de projeções. O outro método é realizar capturas de maior resolução, que serão as imagens de alta resolução, e gerar imagens de baixa resolução por meio da degradação digital. É importante ressaltar que simular os efeitos de degradação que ocorrem em uma aquisição tomográfica de baixa resolução é uma tarefa extremamente difícil.

Os conjuntos realísticos, por outro lado, se baseiam na aquisição em maior e menor resoluções das mesmas amostras para utilizar como imagens de alta e baixa resolução, respectivamente. A criação deste conjunto apresenta alguns desafios técnicos, devido à necessidade de registrar os pares de imagens adquiridos de forma adequada.

Seguindo os atuais entendimentos do campo de SISR discutidos, este trabalho propôs a criação de um conjunto de dados realístico específico para imagens de aglomerados de minério de ferro.

3.5.5 Métricas de Avaliação de Qualidade

A avaliação da qualidade (*quality assessment*) de uma imagem pode ser realizada diretamente pela percepção humana. Por outro lado, a avaliação humana pode produzir resultados com pouca confiabilidade, devido à subjetividade da tarefa, além de ser uma tarefa trabalhosa e que não pode ser automatizada. Sendo assim, métricas objetivas para avaliar a qualidade de uma imagem são fundamentais. [33]

Algumas métricas foram desenvolvidas para a avaliação de qualidade, por exemplo PSNR, SSIM, IFC, LPIPS, NIQE, PIQE e NRQM. Entre essas opções, deve se dar destaque ao *PSNR*.

A *Peak Signal-to-Noise Ratio* (PSNR) é uma das métricas mais utilizadas para avaliação de qualidade em reconstrução de imagens. Em SR, essa definição equivale à razão entre o valor máximo de pixel e o erro quadrático médio (MSE) entre as imagens. Seja uma imagem de referência (*groundtruth*) I e uma versão \hat{I} (imagem de baixa resolução ou imagem reconstruída por SR), o PSNR entre I e \hat{I} é:

$$PSNR = 10 \cdot \log_{10} \left(\frac{L^2}{\frac{1}{N} \sum_{i=1}^N (I(i) - \hat{I}(i))^2} \right) \quad (11)$$

Na Equação (11), L representa o valor máximo do pixel (255 para imagens 8-bit) e N indica o número de pixels.

Apesar da métrica PSNR ser amplamente empregada em SR, seu uso é severamente criticado. Diversos trabalhos demonstram como os valores de PSNR são inconsistentes para fundamentar um aperfeiçoamento de qualidade visual [33][44][45][46][47]. Por exemplo, a Figura 15 apresenta uma desvantagem do PSNR, baseado no cálculo do MSE. Imagens com distorções que afetam acentuadamente a qualidade visual possuem o mesmo valor de MSE.

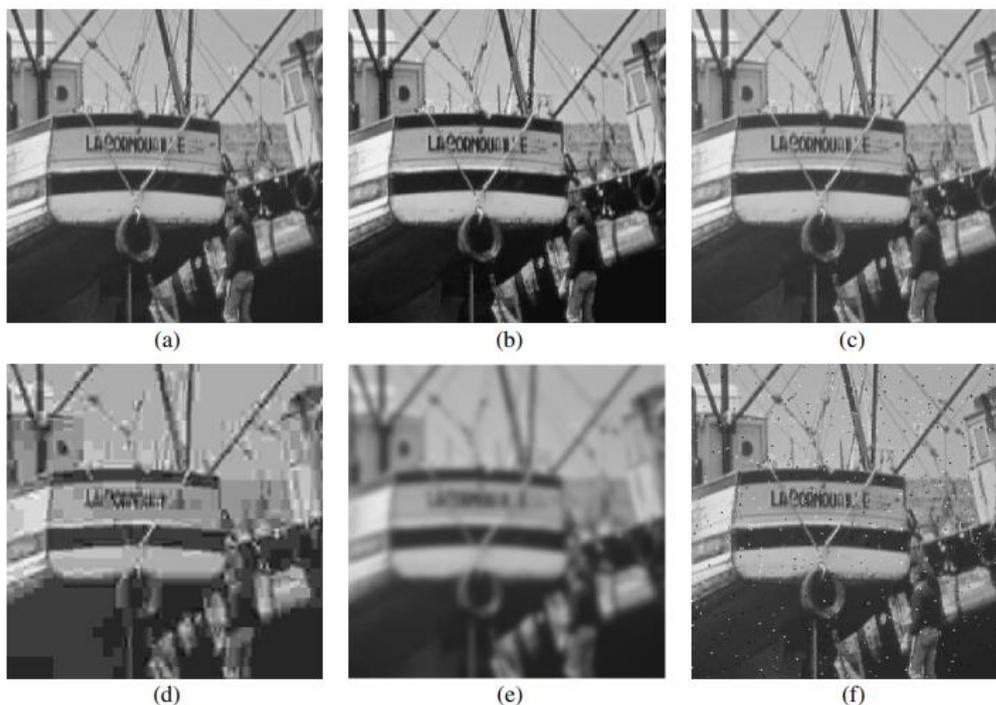


Figura 15: Comparação da qualidade visual de uma imagem de barco com diferentes tipos de distorções. Todas as imagens possuem $MSE = 210$. Retirado de [44].

A escolha de métricas para a avaliação de qualidade ainda é um tópico em forte discussão no campo de SR. Em geral, é extremamente complicado propor uma função que quantifique o aumento da qualidade de uma imagem de forma consistente, devido à natureza complexa e subjetiva da percepção de qualidade. Cada métrica possui desvantagens, sendo necessário encontrar a mais adequada para determinada tarefa.

3.5.6 Super Resolução em MicroCT

Os métodos de SR baseados em DL representam um recurso atrativo para lidar com as limitações inerentes à caracterização por microCT. A possibilidade de aumentar digitalmente a resolução de uma imagem capturada em menores resoluções permite reduzir o tempo de aquisição e aumentar o volume de amostra a ser adquirido.

Procurando esses benefícios, a aplicação da SISR na área da microCT já tem sido explorada.

Hendriksen et al. [48] desenvolveram um sistema “*on-the-fly*” para melhorar a resolução de imagens tomográficas (de qualquer objeto), utilizando a SR baseada em DL. Os modelos foram treinados utilizando apenas 1 ou 9 fatias das imagens adquiridas em alta e baixa resolução, com base nas redes MS-D e U-Net. Os experimentos foram conduzidos utilizando uma imagem simulada de um cilindro de espuma. Os resultados foram avaliados pelas métricas PSNR e SSIM. Os autores concluíram que os resultados demonstram uma significativa melhora na resolução das imagens tomográficas, além da atenuação de ruídos.

Bizhani et al. [49] aplicaram diferentes CNNs para três tarefas de restauração de imagens de rochas: *denoising*, *deblurring* e SR. Para a SR, os autores alegam que dentre vários modelos, as redes EDSR, WDSR e DFCAN se mostraram superiores nos experimentos. Um conjunto de dados sintético com cerca de 27000 imagens de 256x256 pixels foi empregado. Os resultados foram avaliados por meio de PSNR, SSIM e *multi-scale* SSIM, demonstrando aumento na qualidade visual das imagens.

Roslin et al. [50] recorreram ao uso de modelos de SR para um conjunto de imagens de granodiorito, empregando uma rede U-Net 3D. Os resultados foram avaliados por meio das métricas PSNR e SSIM, além da abordagem de segmentação de poros e grãos, que revela que os modelos de SR provocaram aperfeiçoamento na análise quantitativa das segmentações.

4 Metodologia

A obtenção de um modelo de Aprendizado Profundo com bom desempenho depende, principalmente, do cuidadoso preparo de um conjunto de dados. No caso de um modelo de Super Resolução, este conjunto de dados deve apresentar pares de imagens perfeitamente correspondentes em baixa e alta resolução, além de ser suficientemente vasto e diversificado para o correto aprendizado da rede neural. Esta exigência é garantida por meio de diferentes etapas de pré-processamento das imagens tomográficas adquiridas.

A escolha de um algoritmo de Super Resolução, seus respectivos parâmetros de treinamento, e as configurações dos equipamentos também são elementos relevantes para o método desenvolvido.

Além disso, a escolha de métodos para avaliar o desempenho dos modelos de SR treinados e seus impactos é indispensável para a discussão deste estudo.

A Figura 16 apresenta um esquema que sintetiza as etapas cruciais para execução e reprodutibilidade deste trabalho, que serão descritas nesta seção.

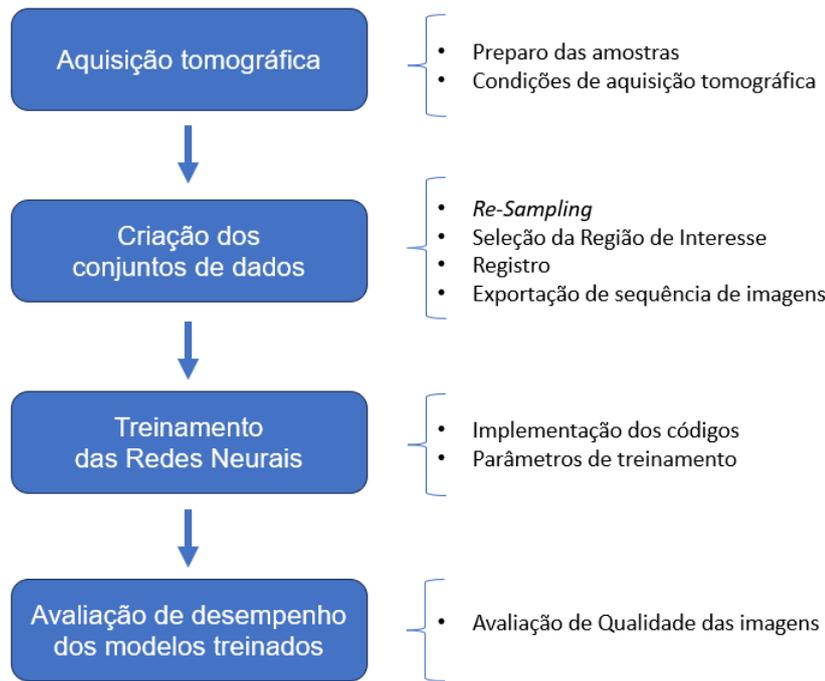


Figura 16: Esquema das etapas presentes na metodologia desenvolvida.

4.1 Preparo das Amostras

As amostras de briquetes de minério de ferro possuem um volume que, juntamente com a alta concentração de ferro em sua composição, dificulta imensamente a aquisição de imagens tomográficas com boas resoluções e menor quantidade de ruídos e artefatos.

Por isso, 4 amostras de briquetes foram enviadas pela empresa Vale S.A. já devidamente cortadas, possibilitando a obtenção de melhores resoluções durante a captura dessas amostras no microtomógrafo. Para a obtenção desses cortes, as amostras foram primeiramente cortadas em seu maior eixo, resultando em duas faces com espessuras próximas a 7 mm. Estas duas faces foram, então, cortadas em fatias menores de até 7 mm, em orientação longitudinal (L) em uma das faces e em orientação transversal (T) na outra. Assim, cada face foi cortada em 3 ou 4 amostras. A nomenclatura de cada fatia é composta pelo nome original da amostra, seguido da orientação do corte (L ou T) e da letra atribuída à ordem da fatia. A Figura 17 ilustra as fatias da amostra F-33 C e suas possíveis nomenclaturas.

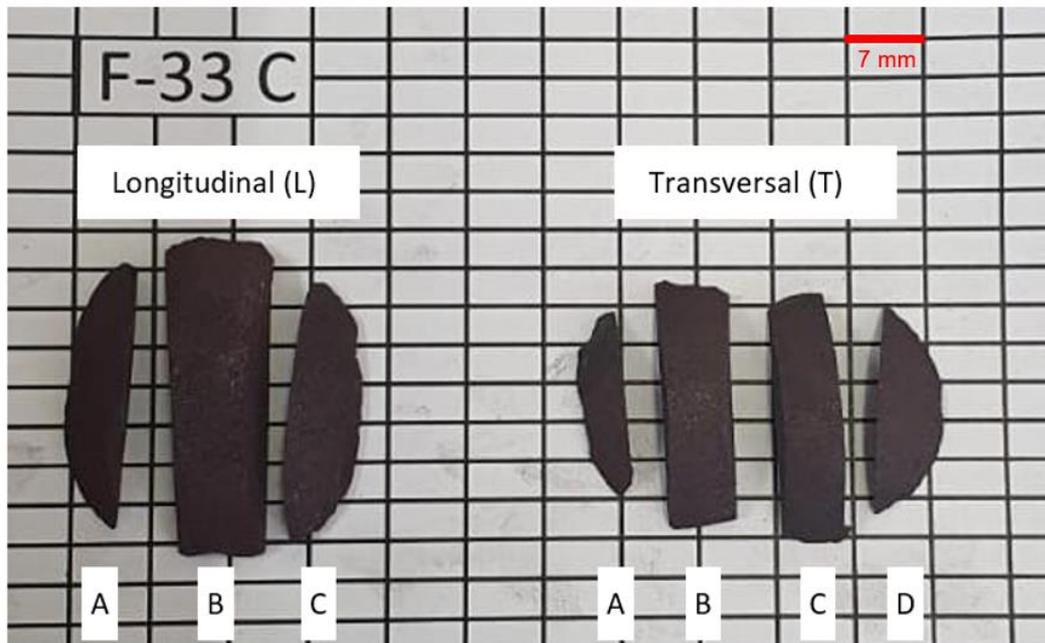


Figura 17: Recortes da amostra de briquete F-33 C. Ao lado esquerdo, uma das faces foi cortada longitudinalmente em 3 fatias; ao lado direito, a outra face foi cortada transversalmente em 4 fatias.

Para que as amostras fossem colocadas no porta amostras do equipamento de microtomografia, elas foram fixadas, com cola, em uma pequena barra de alumínio. A Figura 18 apresenta a amostra fixada na barra, inserida no porta-amostra.



Figura 18: Uma amostra com a pequena barra fixada em seu maior eixo.

4.2 Captura em MicroCT

Para a realização das capturas tomográficas, o equipamento utilizado foi o microtomógrafo de modelo Xradia 510 Versa, da fabricante ZEISS (Figura 19). As características gerais do equipamento estão disponíveis na Tabela 2. Os parâmetros de captura foram definidos no software Scout-and-Scan Control System (versão 16.0.11592.42189).



Figura 19: Equipamento da ZEISS, modelo Xradia 510 versa, utilizado para captura tomográfica das amostras de briquete.

Tabela 2: Características principais do modelo Zeiss Xradia 510 Versa.

Xradia 510 Versa	
Resolução espacial	0,7 μm
Voltagem	30-160 kV
Potência Máxima	10 W
Lentes Objetivas	0,4X, 4X, 20X e 40X
Tamanho máx. da amostra	300 mm
Capacidade	15 kg

Para o treinamento de diferentes modelos de super resolução, diversas resoluções de captura foram previamente contempladas. Foram definidas as

resoluções de captura em: 2 μm , 4 μm , 6 μm e 16 μm . Essas quatro resoluções de captura têm como objetivo o treinamento de:

- a) Um modelo de SR capaz de aumentar a resolução de imagens capturadas em 16 μm para 6 μm (modelo SR-I). A captura em 16 μm foi escolhida pois é uma análise mais veloz e a única capaz de adquirir todo o volume das amostras. Além disso, a diferença de resoluções é notável, o que torna este modelo um caso mais desafiador para a aplicação de Super Resolução.
- b) Um modelo de SR capaz de aumentar a resolução de imagens capturadas em 6 μm para 2 μm (modelo SR-II).
- c) Um modelo de SR capaz de aumentar a resolução de imagens capturadas em 4 μm para 2 μm (modelo SR-III). A captura em 4 μm será proveitosa para embasar uma comparação adicional em relação ao desempenho do modelo de SR-II (de 6 μm para 2 μm).

Como cada amostra varia de volume e formato, as condições de captura foram relativamente distintas. Cada amostra foi colocada no microtomógrafo (Figura 20) e centralizada em relação ao feixe de raios-x e detector.

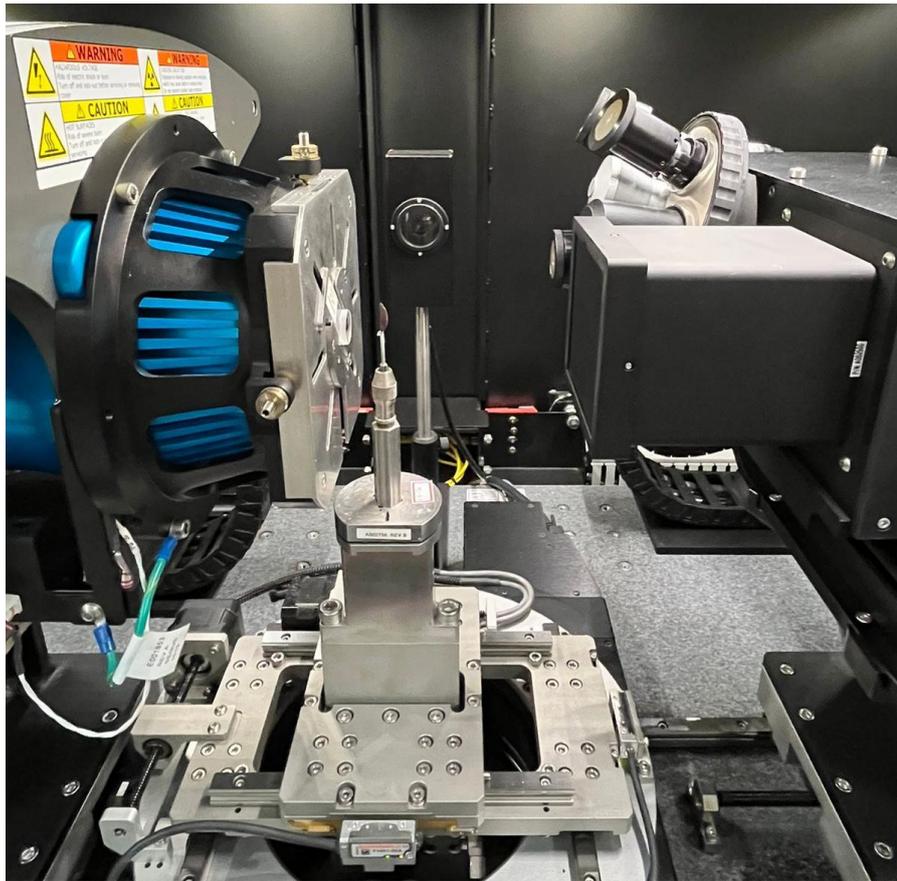


Figura 20: Interior do equipamento de microCT. Observa-se a fonte de raios-x à esquerda, a amostra a ser analisada, e do detector à direita.

Para obter as resoluções desejadas, deve-se escolher a lente objetiva apropriada e alterar a distância entre a fonte e amostra, ou da distância entre amostra e detector. A lente de 0,4X foi utilizada para capturas de resoluções piores que $5\ \mu\text{m}$. Para resoluções melhores que $5\ \mu\text{m}$, utilizou-se a lente de 4X.

O tempo de exposição, voltagem da fonte e potência da fonte são parâmetros que irão impactar o contraste da aquisição tomográfica, assim como a relação entre sinal e ruído. A potência foi sempre utilizada no valor máximo do equipamento (10W). Entretanto, o tempo de exposição e a voltagem da fonte foram escolhidas em meio à análise da intensidade de transmissão do feixe de raios-x. Em geral, procurou-se uma transmissão de intensidade próxima a 9000, com uma boa relação sinal-ruído.

O valor de bin utilizado foi sempre 2, uma vez que a utilização de bin igual a 1 implica em uma piora da relação sinal-ruído, que precisaria ser compensada com um aumento do tempo de captura. O número de projeções utilizado foi sempre 1601. Além disso, utilizou-se o filtro HE-1 para atenuação do efeito de Beam Hardening em todas as aquisições.

As variadas condições de captura afetam o tempo de aquisição necessário para cada amostra. Todas essas informações estão resumidas na Tabela 3, lembrando que o valor de bin, número de projeções, potência da fonte e filtro utilizados foram os mesmos para todas as aquisições.

Tabela 3: Resumo das condições de captura para cada amostra.

Amostra	Resolução	Lente	Exposição	Voltagem	Tempo de análise (h)
F33C_TA	16 um	0,4x	1,0 s	140kV	1,08
	06 um	0,4x	1,0 s	140kV	1,08
	02 um	4x	3,0 s	140kV	2,03
F12_TD	16 um	0,4x	1,0 s	140kV	1,08
	06 um	0,4x	1,1 s	140kV	1,13
	02 um	4x	3,3 s	140kV	2,17
F14_TD	16 um	0,4x	1,0 s	120kV	1,08
	06 um	0,4x	1,0 s	120kV	1,08
	04 um	4x	1,0 s	120kV	1,08
	02 um	4x	2,2 s	120kV	1,67
F34B_TA	16 um	0,4x	1,0 s	120kV	1,08
	06 um	0,4x	1,0 s	120kV	1,08
	04 um	4x	1,0 s	120kV	1,08
	02 um	4x	3,20 s	120kV	2,17

4.3 Reconstrução das Imagens Tomográficas

A etapa de reconstrução das imagens tomográficas foi feita no software Scout-and-Scan Control System Reconstructor, da ZEISS.

No processo de reconstrução, os parâmetros de *center-shift* e de correção de Beam Hardening são fatores muito importantes para obter imagens adequadas.

A falta de alinhamento entre o eixo de rotação da amostra e o centro do detector afeta gravemente a qualidade da imagem reconstruída. O parâmetro de *center-shift* permite corrigir este deslocamento e garantir a qualidade da reconstrução.

O fenômeno de Beam Hardening ocorre quando raios-X de menor energia são mais atenuados que raios-X de alta energia, gerando artefatos na captura. A correção de Beam Hardening minimiza esses artefatos indesejados.

4.4 Criação do Conjunto de Dados

Um conjunto de dados apropriado para o treinamento e validação de um modelo de Super Resolução segue as mesmas características de qualquer conjunto de dados para algoritmos baseados em aprendizado supervisionado: Deve conter os elementos de entrada (*inputs*) e os alvos (*targets*) correspondentes, essenciais para o cálculo de erro que possibilita a etapa de aprendizado.

No caso da SISR, os elementos de entrada são imagens capturadas em menor resolução, enquanto os alvos são as imagens capturadas em maior resolução. É importante ressaltar que cada par de imagem de baixa e alta resolução correspondem a mesma região da amostra.

Sendo assim, para criar um conjunto de dados deste tipo, que garanta pares correspondentes de imagens tomográficas em baixa e alta resolução, algumas etapas de processamento digital de imagens são essenciais e serão detalhadas a seguir. Tais etapas foram executadas nos softwares ORS Dragonfly (Versão 2020.2 para Windows) [51], para as manipulações de imagens 3D, e FIJI [52], para manipulações das sequências de imagens 2D.

Deve-se notar que os processos detalhados adiante foram efetuados para cada um dos três conjuntos de dados. Ou seja, considerando os modelos propostos, as etapas a seguir foram feitas para os pares de imagens em 16 e 6 μm , 6 e 2 μm , e 4 e 2 μm .

4.4.1 Re-Sampling

Para possibilitar a manipulação e processamento simultâneo de um par de imagens de pior e melhor resolução, ambas imagens devem possuir o mesmo número de pixels. Essa equivalência entre as dimensões dos pares de imagens é alcançada por um processo chamado de *re-sampling*.

Isso pode ser realizado pelo aumento do número de pixels da imagem de baixa resolução (*up-sampling*) ou pela diminuição do número de pixels da imagem de alta resolução (*down-sampling*), por meio de algoritmos clássicos de interpolação.

No software Dragonfly, é possível fornecer o valor do tamanho de voxel (resolução da captura) ao abrir uma imagem. Ao fornecer este valor, qualquer operação feita entre as imagens levará em conta a diferença entre resoluções como fator de *re-sample*.

Como os algoritmos de Super Resolução utilizados neste projeto funcionam apenas para aumentos por fatores de 2X, 3X e 4X, existe uma notável particularidade no modelo de 16 μm para 6 μm , cujo fator de aumento é um número não-inteiro entre 2X e 3X. Para a criação deste conjunto de dados, a imagem de 16 μm é tratada como uma imagem de 18 μm . Ou seja, é feito um down-sample na imagem de 16 μm , para que atinja as dimensões equivalentes a 18 μm , possibilitando um aumento de fator inteiro de 3X.

4.4.2 Recorte da Região de interesse

Na Figura 21 é possível visualizar que duas capturas de resoluções distintas possuem volumes discrepantes. Portanto, é primordial selecionar uma Região de Interesse (*RoI*), que exista nas duas imagens.

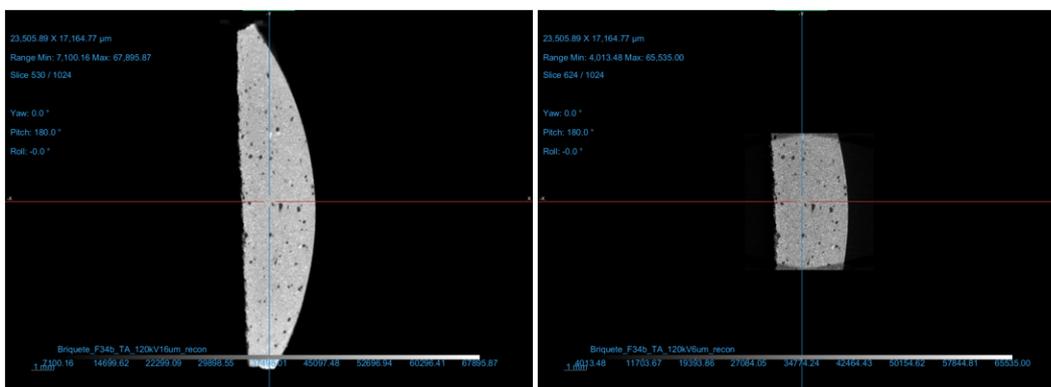


Figura 21: Visualização da diferença de volumes, na mesma escala, entre a captura em 16 μm (esquerda) e 6 μm (direita) da amostra F33C-TA.

Para isso, seleciona-se uma região em formato cilíndrico, a partir da imagem de melhor resolução (imagem de menor volume). Com essa região selecionada, efetua-se um recorte (*crop*) da mesma no par de imagens. O resultado dessa etapa é o par de volumes cilíndricos de uma mesma região da amostra (Figura 22).

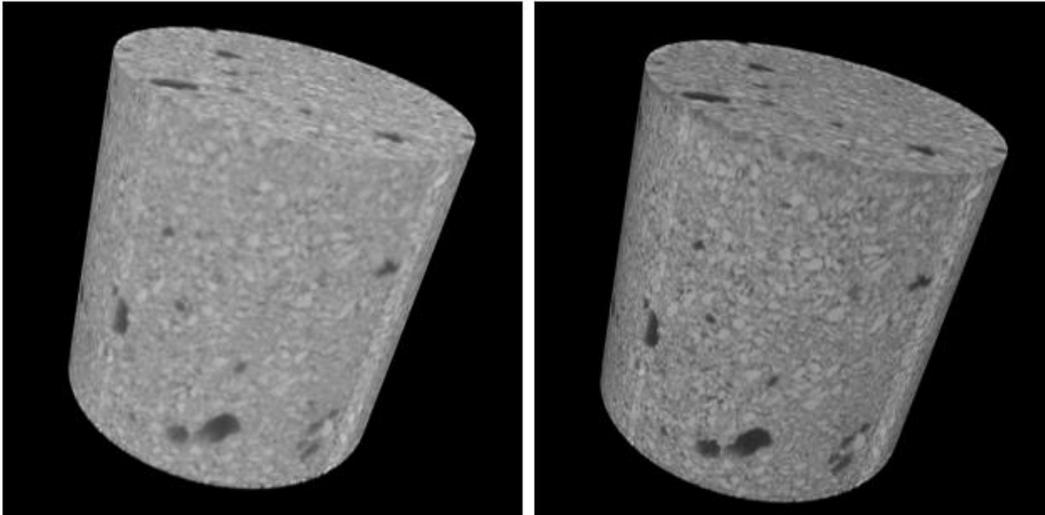


Figura 22: Visualização dos dois volumes cilíndricos, 6 μm (esquerda) e 2 μm (direita), gerados a partir da seleção de uma Região de Interesse.

4.4.3 Registro

Apesar da etapa de recorte das *RoIs* ter como objetivo gerar pares de alta e baixa resolução que apresentem a mesma região de uma amostra, é muito provável que esses pares não estejam perfeitamente alinhados, devido aos deslocamentos e distorções nas orientações dos eixos x , y e z . Este desalinhamento pode ser observado na Figura 23. Portanto, a aplicação de um processamento mais sofisticado, a fim de garantir que as regiões estejam representando o mesmo volume da amostra, é necessária. Esta etapa é conhecida como registro.

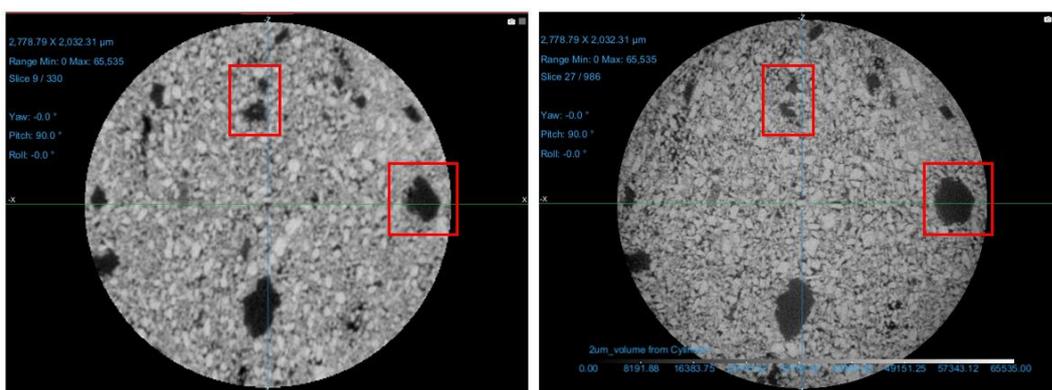


Figura 23: Fatias de uma *RoI* em 6 μm (esquerda) e 2 μm (direita). É possível observar que diversas estruturas pertencentes ao briquete, como os poros destacados pelos retângulos vermelhos, não são idênticos, revelando que as fatias não estão alinhadas.

O processo de registro foi feito no DragonFly, utilizando o tipo de registro elástico (não-rígido). Especificamente, foi utilizada a transformada “*Affine*”

Transform”, com otimizador “*standard gradient descent*” e interpolação “*B-Spline*”. A métrica de similaridade selecionada foi “*Mattes mutual information*”, sem emprego de penalidades.

A etapa de registro gera uma nova imagem, com as correções necessárias para ser fielmente correspondente com a imagem de maior resolução (Figura 24).

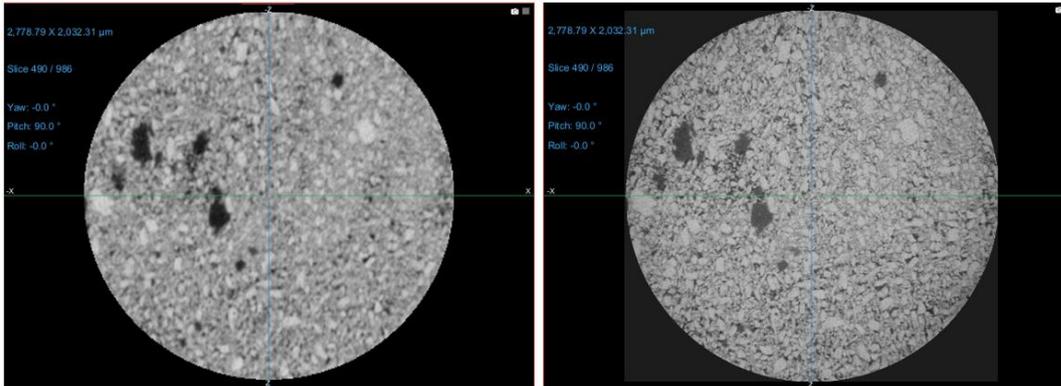


Figura 24: Fatias dos pares de imagens devidamente registradas. É possível observar que após o registro, as mesmas estruturas do briquete estão presentes nas fatias.

É necessário realizar um *down-sample* no eixo z da imagem gerada pelo processo de registro. Isso é feito para que ambas as imagens possuam a mesma quantidade de fatias/projeções. Além disso, o contraste da imagem gerada pelo registro é diferente do contraste original, devido à uma peculiaridade do DragonFly. Por isso, após o registro também é necessário reajustar o contraste da imagem registrada.

4.4.4 Exportar sequência de imagens

Após todos os processamentos que requerem manipulação de imagens tridimensionais serem executados no DragonFly, as imagens são exportadas para o software FIJI. É por meio deste programa em que se realiza um último ajuste antes de salvar a sequência de imagens 2D que compõe as imagens 3D.

Para este ajuste, é necessário reiterar que CNNs se beneficiam de imagens com características diferentes, já que assim se evita *overfitting* e se garante a generalização do modelo. Numa sequência de imagens tomográficas, uma imagem 2D acaba sendo muito similar à fatia seguinte (Figura 25). Portanto, torna-se mais efetivo reduzir a quantidade de imagens da sequência, para que o conjunto de dados possua imagens mais diversas.

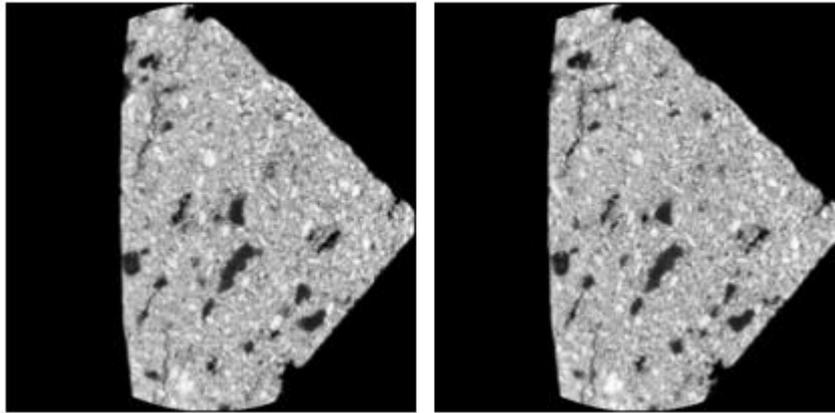


Figura 25: Fatia número 1 da sequência (direita) e a fatia seguinte (número 2). As imagens são muito semelhantes.

Logo, é feita uma redução de um fator de 5 vezes no total de imagens do Stack, de forma que imagens subsequentes possuam características diferentes, como na Figura 26.

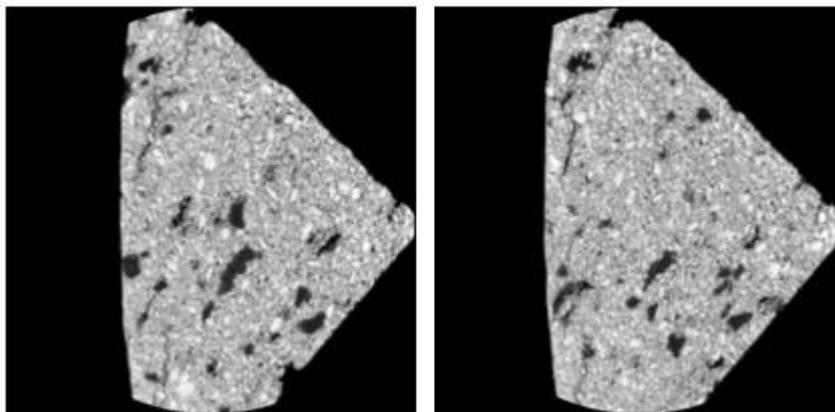


Figura 26: Fatia número 1 da sequência (direita) e a fatia seguinte (número 2), após a redução por fator de 5. As imagens já possuem características distintas.

Finalmente, as imagens são salvas como uma sequência de imagens bidimensionais, no formato *.tiff*.

4.4.5 Divisão do conjunto de dados

Por fim, os conjuntos de dados foram divididos em três subconjuntos: conjunto de treinamento, validação e teste. Essa divisão foi feita com os seguintes pesos: 80% para conjunto de treinamento, 10% para validação, e 10% para teste. É importante ressaltar que os conjuntos de validação e teste possuem funções diferentes. O conjunto de validação é utilizado para avaliar o treinamento no final de cada época de aprendizado, enquanto o conjunto de testes é empregado para avaliar o desempenho do modelo após o treinamento.

A Tabela 4 resume o tamanho dos conjuntos criados para cada um dos modelos.

Tabela 4: Tamanho dos conjuntos de dados para os diferentes modelos.

Modelo	Número de Amostras	Subdivisão	Número de Imagens
SR-I	4	Treinamento	210
		Validação	26
		Teste	26
		Conjunto Total	262
SR-II	4	Treinamento	203
		Validação	25
		Teste	25
		Conjunto Total	253
SR-III	2	Treinamento	149
		Validação	18
		Teste	18
		Conjunto Total	185

4.5 Pipeline de dados

Normalmente, é necessário que o conjunto de dados passe por determinadas operações antes que seja alimentado em uma rede neural. Por exemplo, todos os inputs de uma rede neural devem possuir as mesmas dimensões. Entretanto, as imagens tomográficas possuem variação em seus tamanhos, a depender do volume da amostra. Assim, é necessário garantir que todas as imagens tenham a mesma dimensão requerida pela rede neural. Este conjunto de operações pode ser entendido como pipeline de dados (*data pipeline*).

O pipeline de dados deste projeto foi construído principalmente para resolver o problema da dimensão de entrada fixa requerida pela rede. Para isso, o pipeline realiza uma etapa de *patching*: ele recebe uma imagem do conjunto de dados e gera os ladrilhos (*patches*) dessa imagem, de acordo com o tamanho especificado.

Para executar o *patching*, define-se o tamanho do ladrilho, que será o mesmo input da rede (por exemplo, 128x128, 256x256 etc.). Caso o valor do tamanho das imagens (do conjunto de dados) não seja múltiplo do valor de tamanho do ladrilho, a etapa de *patching* realiza um *padding* automático na

imagem. Neste caso, o *padding* é a adição de linhas e colunas de valor zero (pixels pretos) nas extremidades da imagem, a fim de que o tamanho da imagem se torne múltiplo do valor de tamanho do ladrilho. A Figura 27 ilustra o processo de patching.

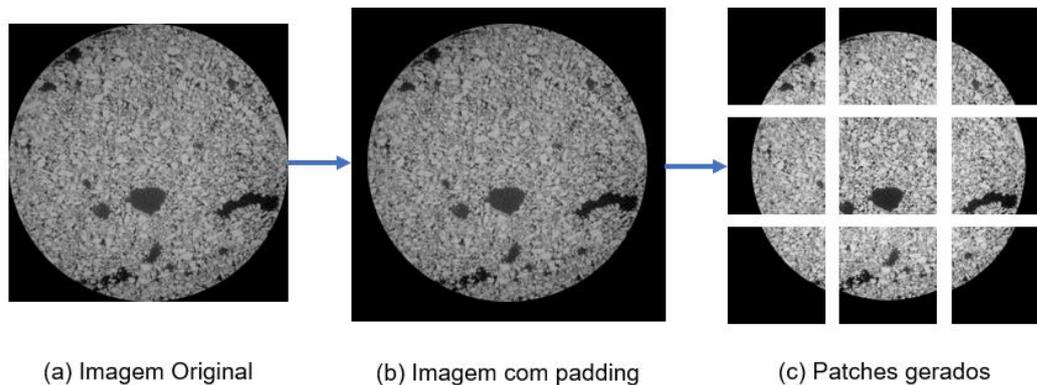


Figura 27: Exemplo da operação de patching. A imagem original (a) possui 964x964 pixels de tamanho. O padding automático expande a imagem para uma dimensão de 1152x1152 pixels. Assim, o pipeline gera 9 patches de 384x384 pixels.

Assim, todas as imagens do conjunto de dados, independentemente do tamanho, passam pelo pipeline de dados. Ou seja, o treinamento da rede é feito com os ladrilhos gerados pelo pipeline. É importante ressaltar que este pipeline não é uma etapa adicional da criação de conjunto de dados. Isso porque o pipeline é executado na etapa de treinamento da rede neural. A geração dos patches é feita durante a primeira época de treinamento e armazenada em cache, de forma a acelerar e otimizar o treinamento.

O pipeline de dados foi implementado em linguagem Python, utilizando as funções do módulo `tf.data` [53], do framework TensorFlow [54], na versão 2.7.

4.6 Treinamento da Super Resolução

Os treinamentos dos modelos em arquitetura EDSR foram feitos a partir da implementação feita por Martin Krasser baseada em TensorFlow 2.x, disponível como código aberto na plataforma GitHub [55].

A rede EDSR foi escolhida para este trabalho por seu notável desempenho, por ter sido bem revisada na literatura, e ser uma implementação consideravelmente simples, comparado a outros modelos.

Além da adição do pipeline de dados, as outras mudanças da implementação original giram em torno do tipo de imagem de entrada para a rede

e a normalização desses dados. Originalmente, a arquitetura da rede possuía como entrada imagens de três canais (RGB). Essa característica foi modificada a fim de permitir o processamento de imagens do tipo *16bit* (tons de cinza). A normalização foi alterada para se adequar à essa mudança de tipo da imagem, mantendo os dados de entrada no intervalo [0,1].

Conforme mencionado na seção 4.4.5, os conjuntos de dados são separados em subconjuntos de treinamento, validação e teste, na proporção 80%/10%/10%. Cada modelo possui um número distinto de imagens e patches no conjunto de dados. Como a rede é treinada com os patches, é mais apropriado expressar sua quantidade, expostos na Tabela 5.

Tabela 5: Resumo dos conjuntos de treinamento, validação e teste para os modelos propostos.

Modelo	Subdivisão	Número de Imagens	Número de Patches
SR-I	Treinamento	210	840
	Validação	26	104
	Teste	26	104
SR-II	Treinamento	203	1827
	Validação	25	225
	Teste	25	225
SR-III	Treinamento	149	2384
	Validação	18	288
	Teste	18	288

Os hiper parâmetros da rede foram mantidos iguais para todos os modelos treinados. O número de blocos residuais (*Residual blocks depth*) utilizados foi de 32, com 256 filtros. O otimizador empregado foi o Adam. Um importante parâmetro do otimizador é a taxa de aprendizado (*learning rate*), que controla o passo em que os pesos serão atualizados. A implementação utiliza, por padrão, um *learning rate* com taxa de decaimento (*Piecewise Constant Decay*, com *boundaries* = [200000] e *values* = [1e-4, 5e-5]). A função de perda usada foi L1, e a métrica observada durante a validação foi o PSNR.

O número de batch foi de 4 imagens. As redes foram treinadas entre aproximadamente 26 e 28 horas, apresentando diferentes números de épocas/iterações completadas.

Para qualquer outro parâmetro não descrito, manteve-se a opção padrão da implementação de Krasser [55].

Os treinamentos foram conduzidos em um computador com a seguinte configuração de hardware: 01 processador AMD Ryzen Threadripper PRO 3955WX, 512 GB de memória RAM e 01 GPU Nvidia RTX A6000 48GB.

4.7 Etapa de Teste

A etapa de teste dos modelos possibilita qualificar o desempenho dos modelos treinados sobre imagens do conjunto de teste, ou seja, imagens que os modelos nunca tiveram acesso.

4.7.1 Análise Visual

Inicialmente, o comportamento dos modelos de SR foi avaliado por meio da própria percepção humana, recorrendo-se a comparações entre as imagens de baixa resolução, alta resolução e super resolvidas.

4.7.2 PSNR

A métrica de PSNR foi utilizada para quantificar a qualidade visual entre baixa e alta resolução (fazendo uso de *up-sampling* da baixa resolução por interpolação bicúbica), e entre a alta resolução e imagem super resolvida. Os valores de PSNR apresentados são uma média de todos os valores calculados para o conjunto de teste.

4.7.3 Segmentação de poros

Como o PSNR não é capaz de expressar corretamente a qualidade da reconstrução/recuperação de detalhes gerados pelo modelo de SR, este trabalho propõe realizar uma segmentação de regiões porosas, a fim de quantificar o impacto da SR em uma tarefa clássica de análise de imagens digitais.

A segmentação foi realizada pelo método de limiarização. A escolha destes limiares é uma etapa subjetiva. A sensibilidade dos resultados em função da escolha de limiares também foi discutida. A partir de análises quantitativas sobre as imagens binárias resultantes do processo de segmentação, obtiveram-se medidas de número de poros, área total de poros e fração em área dos poros.

4.7.4 Precision, Recall e F1-Score

Apesar dos resultados quantitativos, obtidos na etapa de segmentação de poros, possibilitarem avaliar se a SR gera impactos positivos, não é possível determinar como os pixels de poros da imagem super resolvida têm relação com

a imagem HR. Por isso, outra forma de investigar o impacto da SR foi realizada por meio das métricas *Precision*, *Recall* e *F1-Score*, amplamente utilizadas para avaliação de desempenho em sistemas de classificação. Os cálculos dessas métricas foram feitos com as imagens binárias geradas na etapa de segmentação.

Para a formulação das métricas, é necessário definir alguns conceitos. Os verdadeiros positivos (VP) representam a contagem dos pixels que existem tanto na imagem de referência (HR) quanto na imagem a ser comparada (super resolvida ou LR). Os VPs equivalem ao resultado da operação AND entre as imagens de referência e a ser comparada. Os falsos positivos (FP) representam os pixels que existem apenas na imagem a ser comparada. E os falsos negativos (FN) representam os pixels que existem apenas na imagem de referência.

O *precision* é uma métrica sensível à quantidade de falsos-positivos presentes na imagem a ser comparada. Ou seja, revela que fração de pixels da imagem a ser comparada realmente existe na imagem de referência. Um *precision* baixo demonstra que uma boa quantidade dos pixels da imagem inferida não estão presentes na imagem de referência. É formulada como:

$$Precision = \frac{VP}{VP + FP} = \frac{\textit{imagem AND}}{\textit{imagem a ser comparada}} \quad (12)$$

O *recall* é uma métrica sensível à quantidade de falsos-negativos presentes na imagem a ser comparada. Ou seja, revela que fração de pixels da imagem de referência estão presentes na imagem inferida. Um *recall* baixo demonstra que poucos pixels presentes na imagem de referência fazem parte da imagem inferida. É formulada como:

$$Recall = \frac{VP}{VP + FN} = \frac{\textit{imagem AND}}{\textit{imagem HR}} \quad (13)$$

O *F1-Score* compõe uma média harmônica das métricas *precision* e *recall*. É formulada como:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (14)$$

4.7.5 Segmentação de Volume 3D

Além das análises em imagens bidimensionais, a segmentação de poros em uma amostra tridimensional também foi conduzida. Os valores de fração de volume dos poros para cada uma das resoluções foram averiguados.

5

Resultados e discussões

5.1 Resultados dos treinamentos

5.1.1 Modelo SR-I: 16 para 06 (3X)

O modelo SR-I foi treinado por aproximadamente 26 horas, alcançando 900 épocas (189000 iterações).

5.1.1.1 Análise Visual

A Figura 28 apresenta uma visualização das versões de LR, super resolvida e HR de uma imagem presente no conjunto de testes. Comparando as imagens, é possível observar que a imagem super resolvida se aproxima da versão HR, em termos de detalhamento e nitidez de estruturas como partículas ricas em ferro (regiões claras) e poros (regiões escuras). O contraste gerado pelo modelo de SR também se aproximou do contraste original da imagem HR.

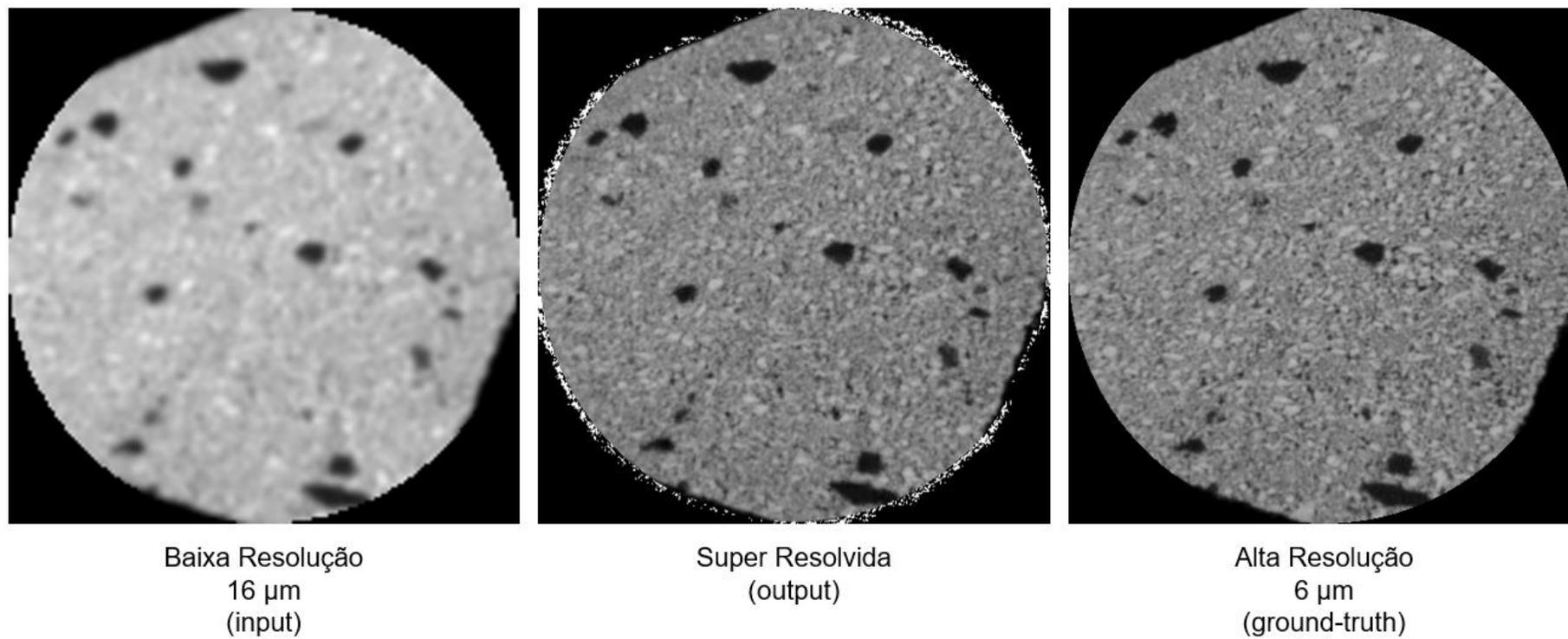


Figura 28: Visualização de uma imagem de baixa resolução (esquerda), imagem super resolvida pelo modelo SR-I (centro), e a imagem de alta resolução (direita). Corte da amostra F34B-TA.

Deve-se notar que a imagem super resolvida possui alguns artefatos na região exterior à fronteira da amostra. O modelo de SR gerou uma auréola de pixels claros na imediação da superfície da amostra. Uma das possíveis explicações para esse comportamento está associado com o Efeito de Volume Parcial (*Partial Volume Effect*) presente nas bordas das amostras, fenômeno inerente às aquisições tomográficas. A região próxima à fronteira da amostra possui uma transição gradual entre os tons mais claros do briquete e os pixels pretos do fundo. Esse efeito é discrepante entre imagens LR e HR, como apresenta a Figura 29. A rede neural provavelmente teve dificuldade em lidar com tal discrepância para um conjunto completo de treinamento. Esse artefato foi removido pela eliminação dos pixels situados na região exterior ao contorno da amostra em baixa-resolução, que foi sobreposto à imagem SR.

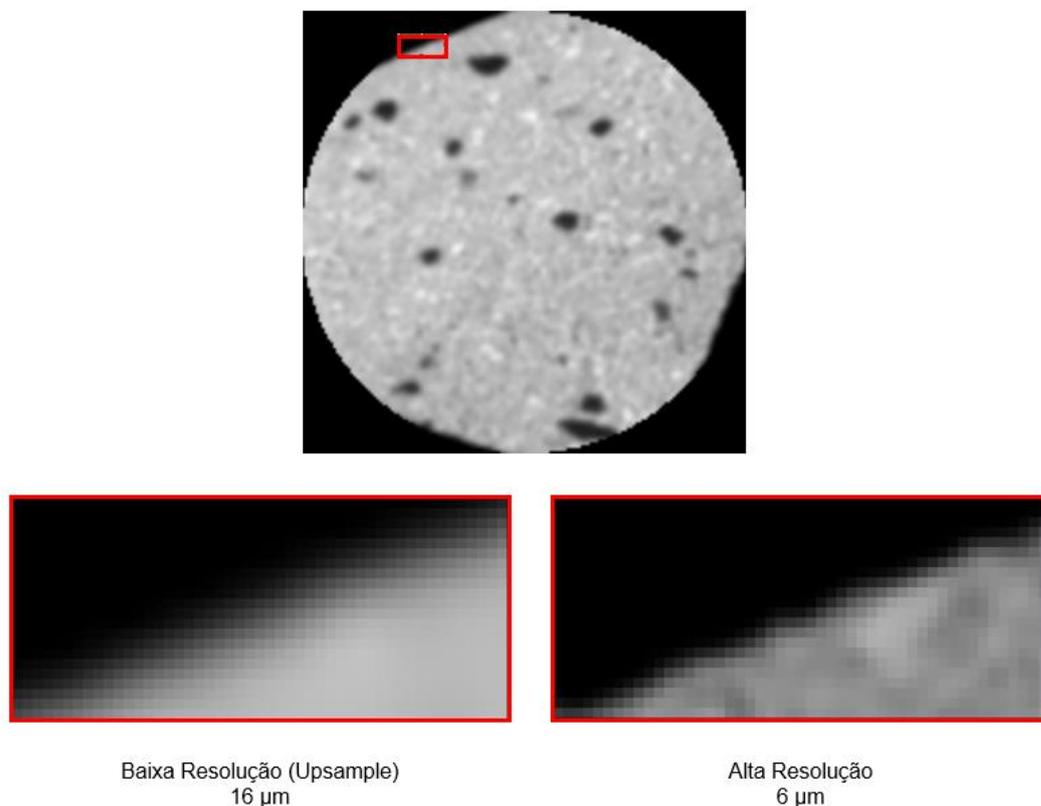


Figura 29: Exemplo do efeito halo presente em regiões de fronteira de uma imagem LR e HR, para a região de interesse indicada.

A Figura 30 expõe o esquema comparativo de resoluções para uma região ampliada, indicada em vermelho.

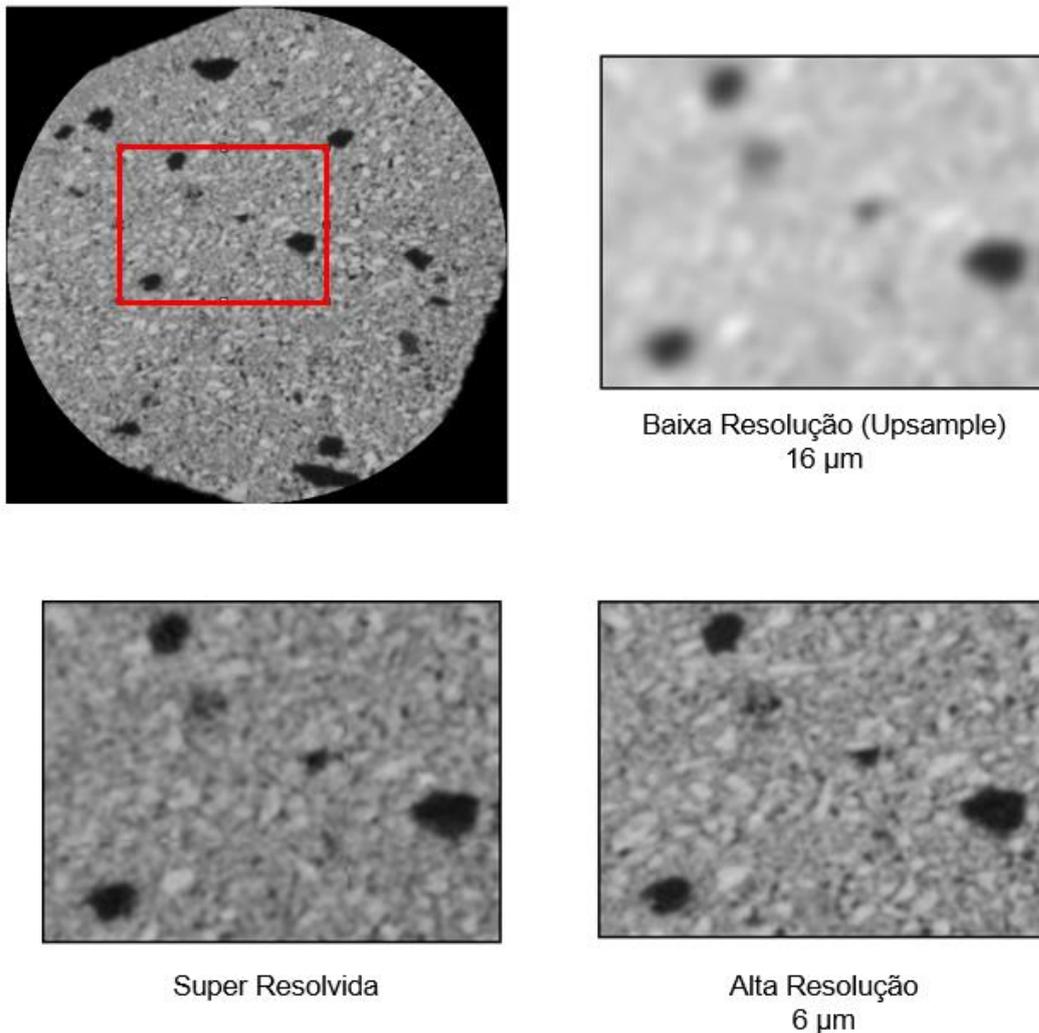


Figura 30: Visualização das versões da *ROI* em baixa resolução, super resolvida e alta resolução

5.1.1.2 PSNR

O conjunto de testes, contendo 26 imagens, foi utilizado para gerar o valor médio dos PSNRs. A média para PSNR entre as imagens super resolvidas e as imagens HR foi de 21,51 dB, enquanto a média para PSNR entre as imagens LR e imagens HR foi de 16,16 dB.

O aumento da métrica evidencia a melhora da qualidade visual das imagens super resolvidas, em comparação às imagens LR.

5.1.1.3 Segmentação de Poros

A segmentação de poros proposta foi baseada na imagem presente na Figura 28. A Figura 31 apresenta a visualização do processo de segmentação. Primeiro, foi realizada a segmentação por limiarização na imagem de LR (Figura

31a), de acordo com a faixa tonal do histograma (Figura 31b). Depois, realizou-se a segmentação na imagem de alta resolução (Figura 31c), de acordo com a faixa tonal do histograma (Figura 31d). Por último, a segmentação da imagem super resolvida (Figura 31e) foi executada com a faixa tonal presente na Figura 31f.

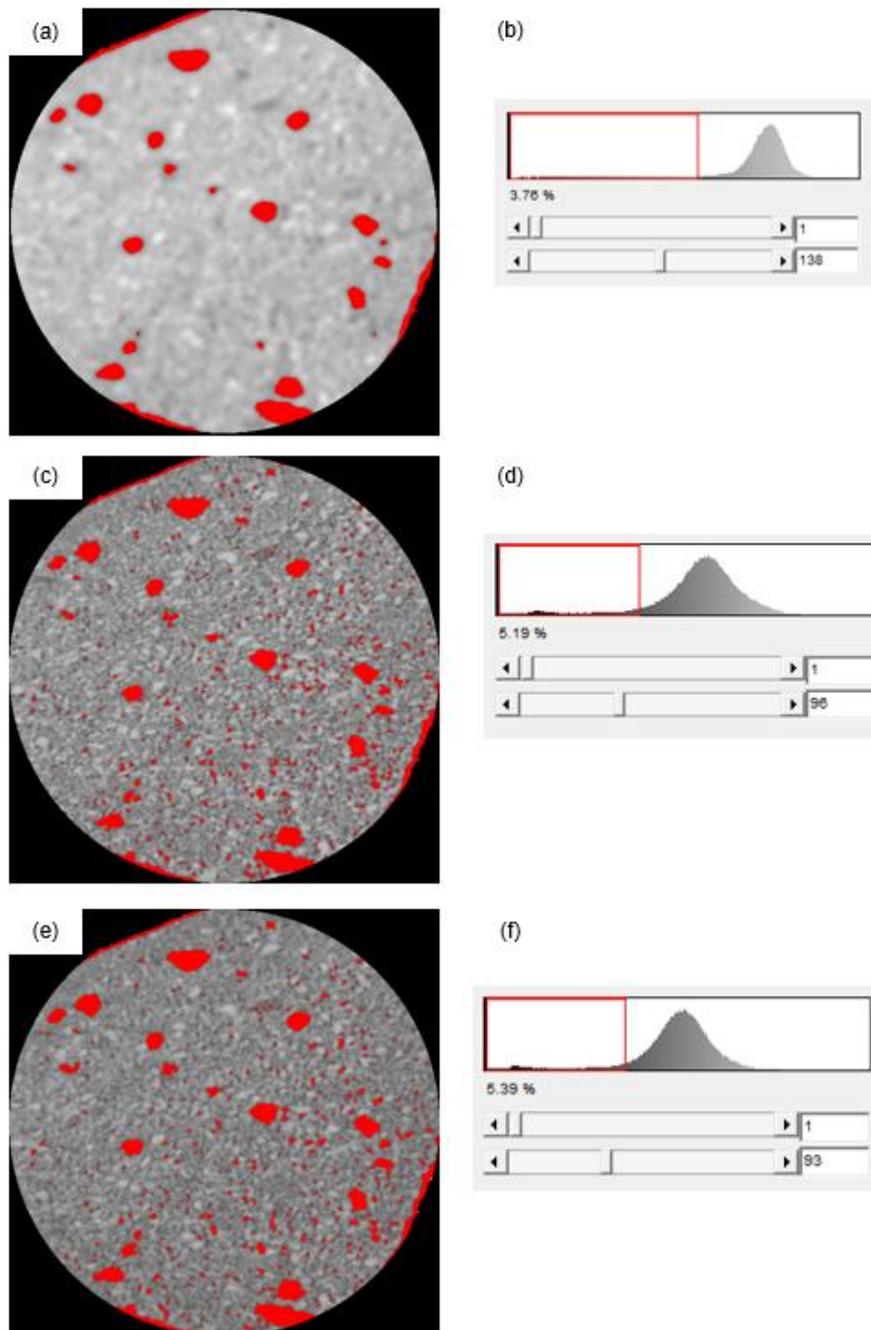


Figura 31: Processo de segmentação das regiões porosas. (a) e (b) se referem a imagem de LR, com limiar de 0-138. (c) e (d) se referem a imagem de HR, com limiar de 1-96. (e) e (f) se referem a imagem super resolvida, com limiar de 1-93.

As imagens binárias resultantes do processo de segmentação são apresentadas na Figura 32. Algumas regiões da borda da amostra foram incluídas

na segmentação, sendo facilmente eliminadas em uma filtragem de valores destoantes de *feret*, em uma etapa de pós-processamento. O *feret*, neste contexto, representa a maior distância entre dois pontos do contorno de um objeto. Os objetos indesejados, por serem extremamente alongados, possuem valores de *feret* muito superiores aos outros objetos.

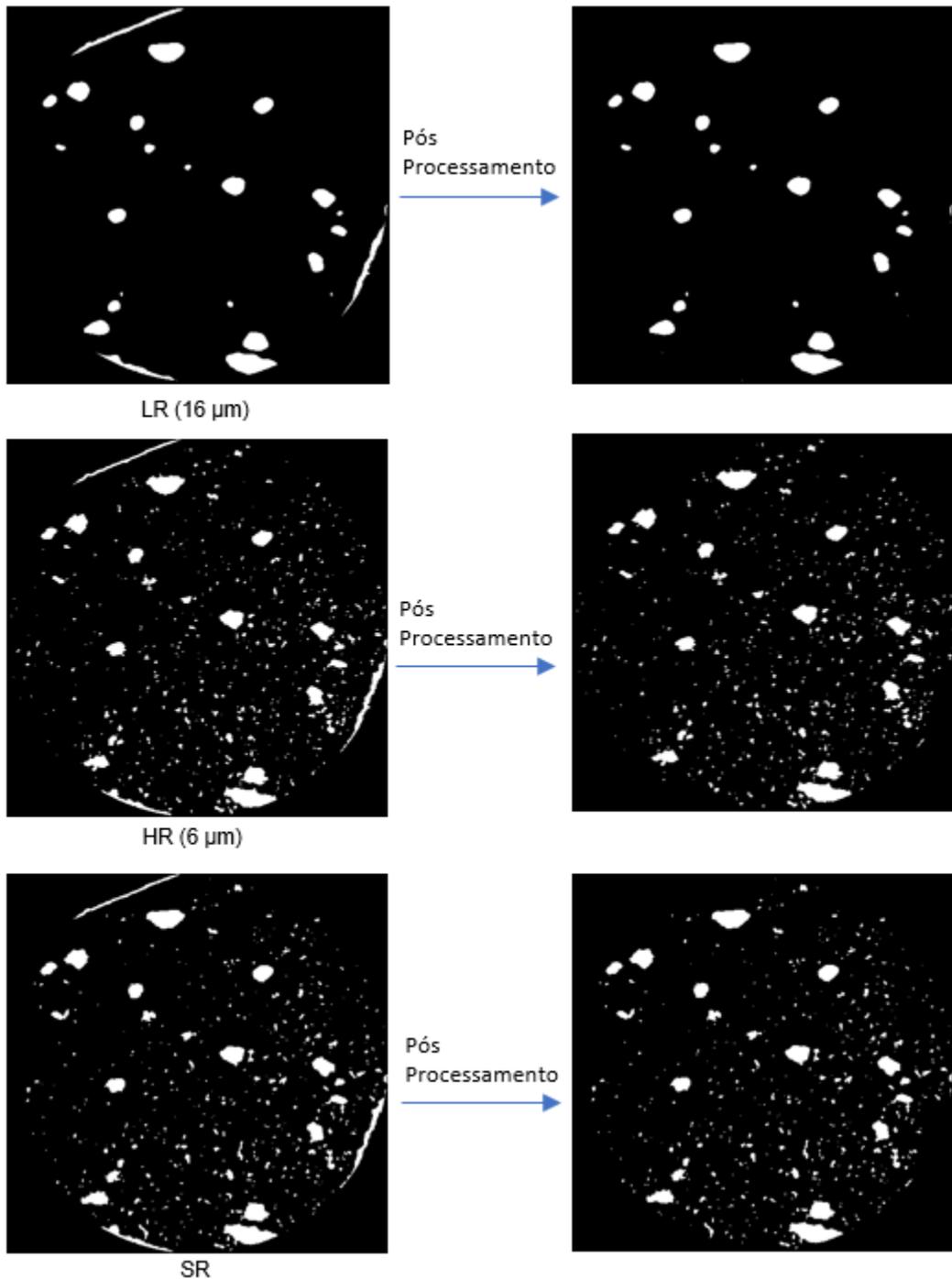


Figura 32: Visualização das imagens binárias obtidas para segmentação de poros nas versões LR, HR e SR.

É possível observar que a segmentação da LR detecta apenas poros de maiores diâmetros. A imagem HR possui uma enorme quantidade de poros pequenos. Além disso, os poros de maiores diâmetros em HR possuem silhuetas muito detalhadas, em comparação aos mesmos poros em LR. Neste caminho, pode-se notar que a segmentação da imagem super resolvida identificou uma grande quantidade de poros menores, como na imagem HR. Entretanto, já não é possível afirmar que esses poros são exatamente os mesmos nas duas imagens.

A Tabela 6 apresenta os valores obtidos, em uma análise quantitativa da imagem binária, da área total de poros e sua respectiva fração de área, assim como a quantidade de poros identificados para as diferentes resoluções da imagem.

Tabela 6: Análises quantitativas feitas a partir da segmentação de poros.

Modelo	Resolução	Área Total de Poros (pixels)	Fração em área	Contagem de poros
SR-I	Baixa (16 μm)	6366	3,8%	26
	Super Resolvida	10016	6,0%	467
	Alta Resolução (06 μm)	9565	5,8%	535

Pode-se notar uma significativa variação da fração em área dos poros entre a LR e HR de 2,0 pontos percentuais, isto é, um aumento efetivo de 50,2%. A contagem de poros saiu de 26 poros, na LR, para 535 poros, na HR.

Analisando os dados da imagem super resolvida, constata-se que a SR superestimou levemente a fração em área dos poros, apresentando um valor de 0,2 ponto percentual maior que da imagem HR, isto é, um erro de 4,7%. É importante ressaltar que a segmentação por limiar é uma técnica baseada em uma avaliação subjetiva, pois não existe nenhuma métrica objetiva para definir qual faixa tonal é mais adequada para segmentar os poros. Portanto, estas pequenas variações percentuais na área total de poros incluem erros associados, e não são evidências adequadas do impacto da SR em análises quantitativas das imagens de aglomerado. Por isso, é pertinente entender como a mudança do limiar pode afetar essas análises. A Tabela 7 apresenta algumas experiências, diminuindo ou aumentando o limiar superior com passo de 2 unidades.

Tabela 7: Valores de fração em área para diferentes faixas tonais.

Resolução	Faixa tonal de segmentação	Área Total de Poros (pixels)	Fração em área	Contagem de poros
Alta Resolução (06 μm)	1-92	8389	5,1%	425
	1-94	8960	5,4%	482
	1-96 (referência)	9565	5,8%	535
	1-98	10301	6,2%	598
	1-100	11104	6,7%	647

Por meio dos dados obtidos na Tabela 7, vemos que pequenas mudanças no limiar superior produzem variações perceptíveis na fração em área de poros. Considerando, então, o erro associado a essa escolha de limiares, a diferença de 0,2 ponto percentual observada entre a fração em área de poros das imagens HR e SR é sensível à escolha dos limiares, podendo diminuir ou aumentar consideravelmente.

Outro dado extremamente relevante na Tabela 6 diz respeito ao número de poros contabilizados. Na imagem super resolvida, foram identificados 467 poros, um aumento de 17,96 vezes em relação à contagem de poros da imagem LR. Esse valor também representa aproximadamente 87% dos 535 poros quantificados na imagem HR. Essa característica demonstra que apesar da SR apresentar uma área total de poros pouco maior do que a HR, ela é capaz de simular bem a quantidade de poros presentes no briquete.

Apesar destes resultados quantitativos exporem que a imagem super resolvida se aproxima da imagem HR, eles não garantem que os pixels de poros da imagem super resolvida realmente têm relação com os pixels da imagem HR. Para ilustrar esse questionamento, a Figura 33 apresenta a imagem binária super resolvida com as demarcações de poros das imagens de baixa resolução, em azul, e alta resolução, em vermelho. Por meio desta comparação, é possível notar que muitos dos poros pequenos da imagem super resolvida não existem na imagem de alta resolução, assim como muitos poros pequenos da imagem de alta resolução não existem na imagem super resolvida. Ainda assim, a imagem super resolvida consegue identificar boa parte dos poros de menor diâmetro que de fato estão presentes na imagem de alta resolução.

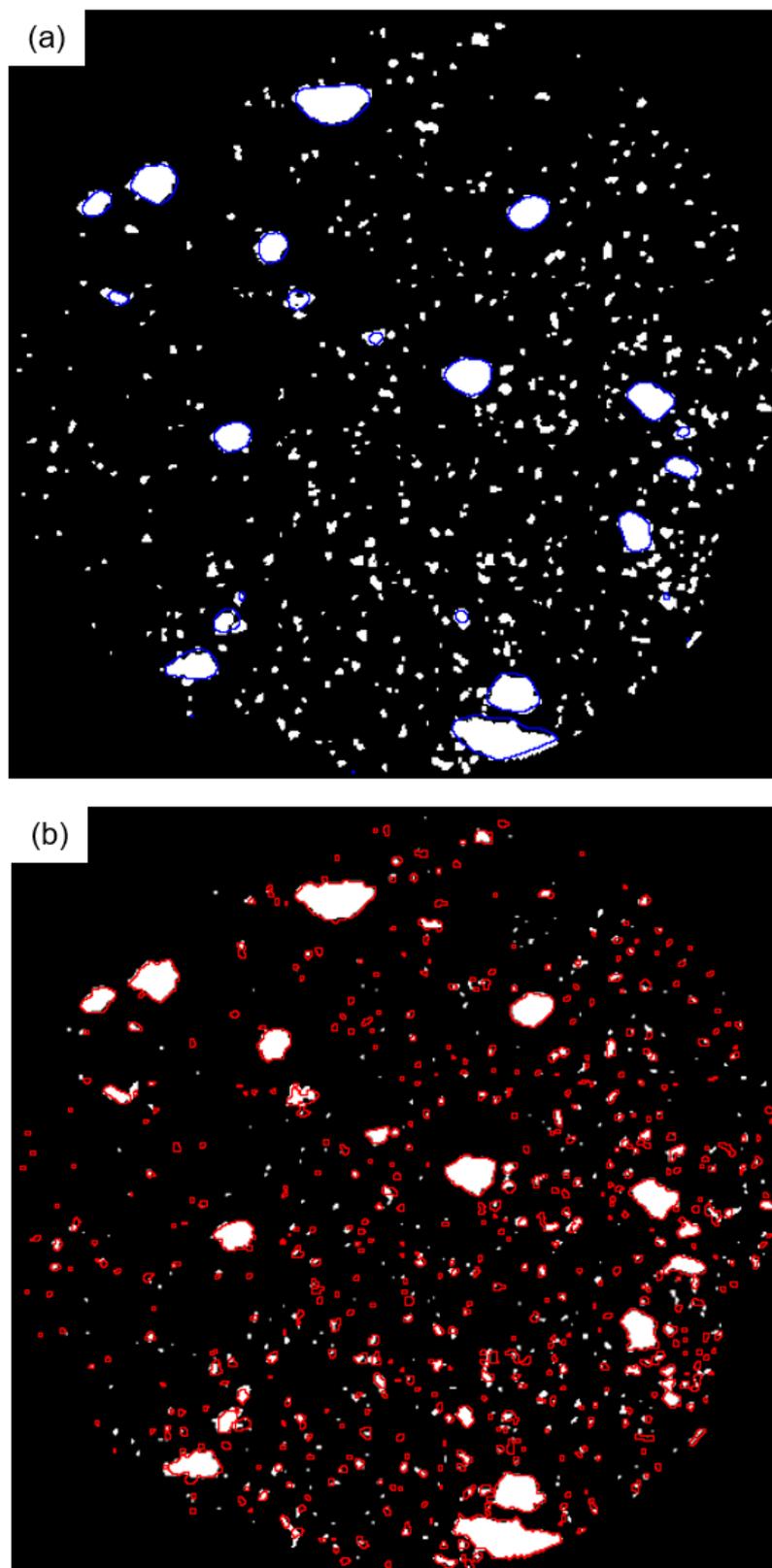


Figura 33: Comparação de poros segmentados. As imagens se referem à binária super resolvida com a seleção, em azul, de poros da imagem de baixa resolução (a) e com a seleção, em vermelho, de poros da imagem de alta resolução (b).

5.1.1.4 Precision, Recall e F1-Score

Neste sentido, é extremamente relevante compreender como os poros gerados pela SR se relacionam com os poros presentes na imagem HR. Uma maneira eficiente de aferir essa proximidade é aplicar métricas de avaliação de modelos de classificação, tais como *precision* e *recall*. Como descrito na metodologia, essas métricas foram aplicadas para o par de imagens LR e HR, e para os pares SR e HR. A Figura 34 apresenta as imagens resultantes da operação *AND* entre as imagens LR e HR, e SR e HR.

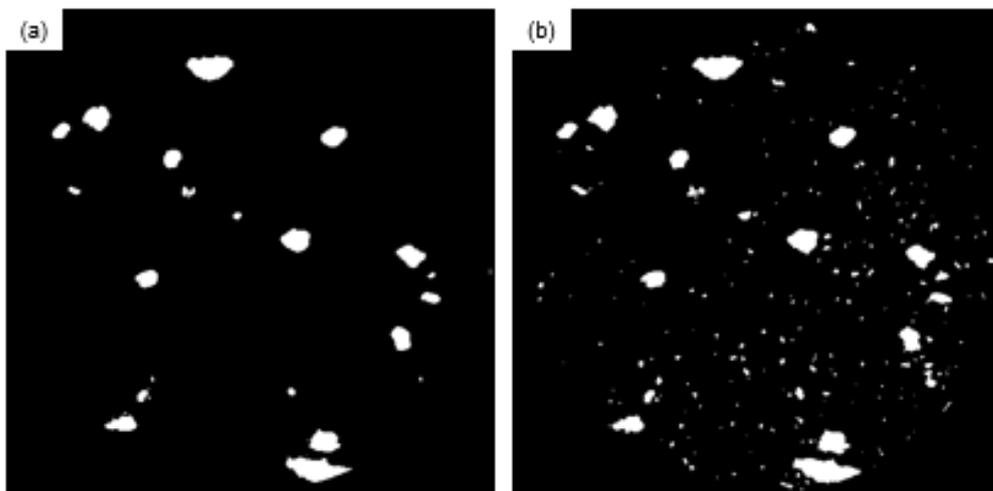


Figura 34: Resultados da operação *AND* entre LR e HR (a), e SR e HR (b).

A partir dessas imagens *AND*, foi possível medir os valores presentes na Tabela 8.

Tabela 8: Medidas de *Precision* e *Recall*.

	LR e HR	SR e HR
Precision	85,45%	71,17%
Recall	54,04%	74,27%
F1-Score	66,21%	72,69%

A redução de 14,28 pontos percentuais no *precision* para o modelo SR exprime que a imagem SR possui mais pixels falsos-positivos, em relação a imagem HR, do que a imagem LR. Já o aumento de 20,23 pontos percentuais no *recall* indica que a imagem SR possui menos pixels falsos-negativos, em relação a imagem HR, do que a imagem LR. Este aumento expressivo no *recall* é coerente com o aumento da fração em área de poros observada nas imagens SR. O

aumento efetivo de 6,48 pontos percentuais no F1-Score é uma evidência de que a melhora de resolução da imagem LR, realizada pelo modelo de SR, estabelece impactos positivos para a análise digital de imagens.

5.1.1.5 Segmentação de Volume 3D

Após as análises envolvendo uma imagem bidimensional, é válido atestar o impacto da SR em uma amostra tridimensional. Seguindo os métodos anteriores, a segmentação de poros foi efetuada para todas as camadas de uma amostra. A escolha de limiares foi feita a partir dos valores encontrados na análise da camada bidimensional, com um ajuste a partir da observação de diferentes camadas. Portanto, os limiares definidos para LR foram de 1-135, para HR foram de 0-96, e para SR foram de 1-93. A Figura 35 apresenta uma visualização dos poros segmentados para cada resolução.

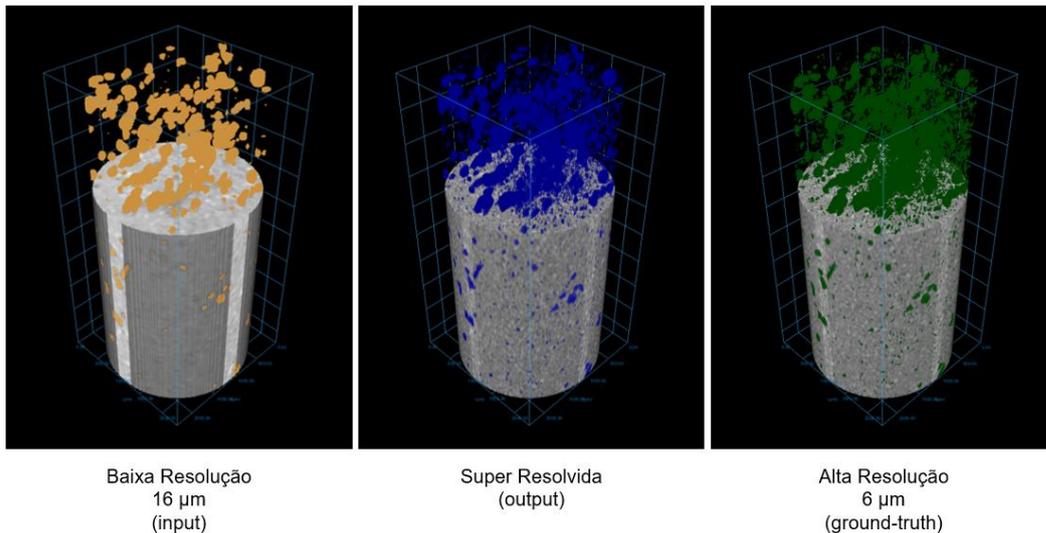


Figura 35: Visualização dos poros segmentados em LR (esquerda), super resolução (centro) e HR (direita).

A Tabela 9 expõe as medidas de porosidade encontradas.

Tabela 9: Valores de porosidade das amostras 3D.

Modelo	Resolução	Fração em Volume
SR-I	Baixa (16 μm)	2,9%
	Super Resolvida	4,6%
	Alta Resolução (06 μm)	4,5%

Seguindo os resultados da segmentação bidimensional, a super resolução superestima em apenas 0,1 ponto percentual, isto é, 2,2% a fração em área dos poros presentes na captura de alta resolução.

5.1.2 Modelo SR-II: 06 para 02 (3X)

O modelo SR-II foi treinado por aproximadamente 26 horas, alcançando 427 épocas (194712 iterações).

5.1.2.1 Análise Visual

A Figura 36 apresenta a comparação entre versões LR, super resolvida e HR de uma imagem do conjunto de testes. Assim como no modelo SR-I, é possível constatar que a imagem super resolvida possui um detalhamento e nitidez de estruturas que se aproxima da versão HR. É possível observar que o modelo SR-II também gera os artefatos ao redor da amostra, mas em menor grau.

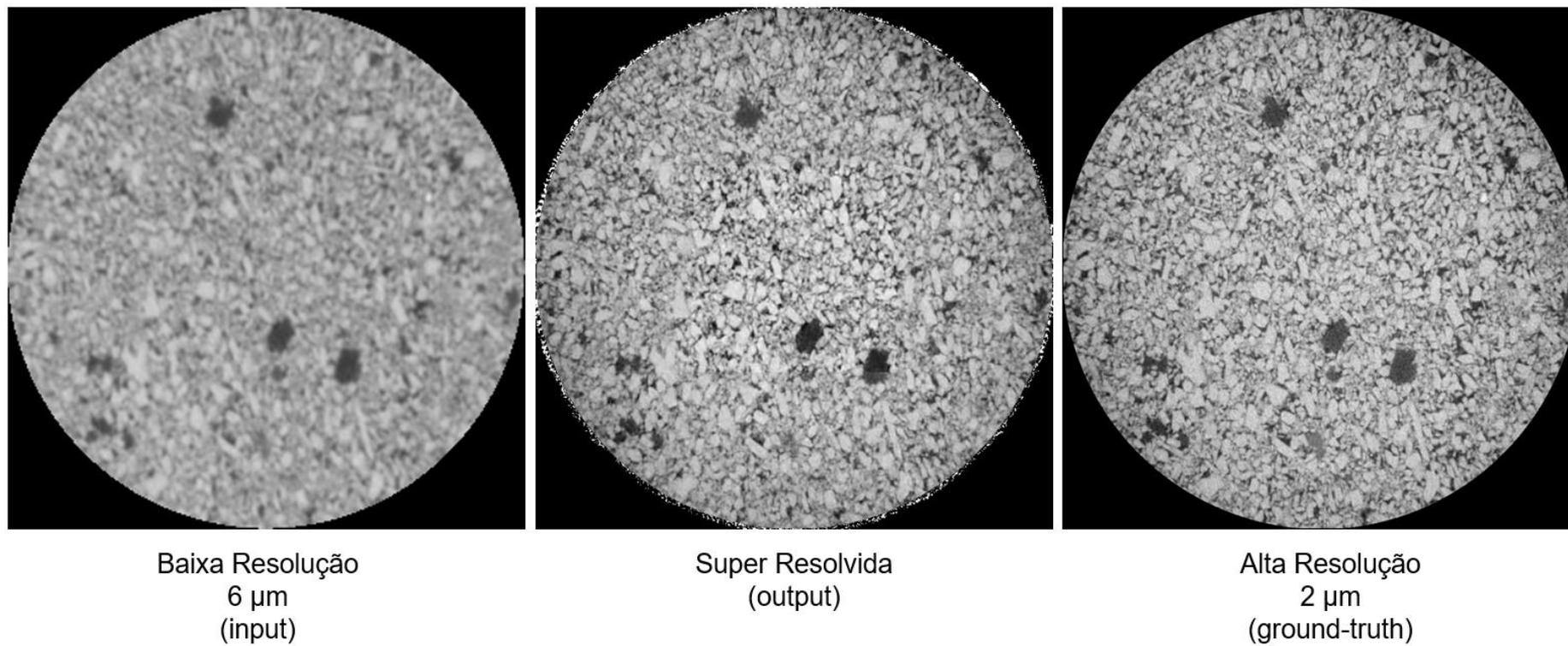


Figura 36: Comparação das imagens LR (esquerda), super resolvida pelo modelo SR-II (centro), e HR (direita). Corte da amostra F34B-TA.

A Figura 37 apresenta outro comparativo para uma região ampliada da imagem.

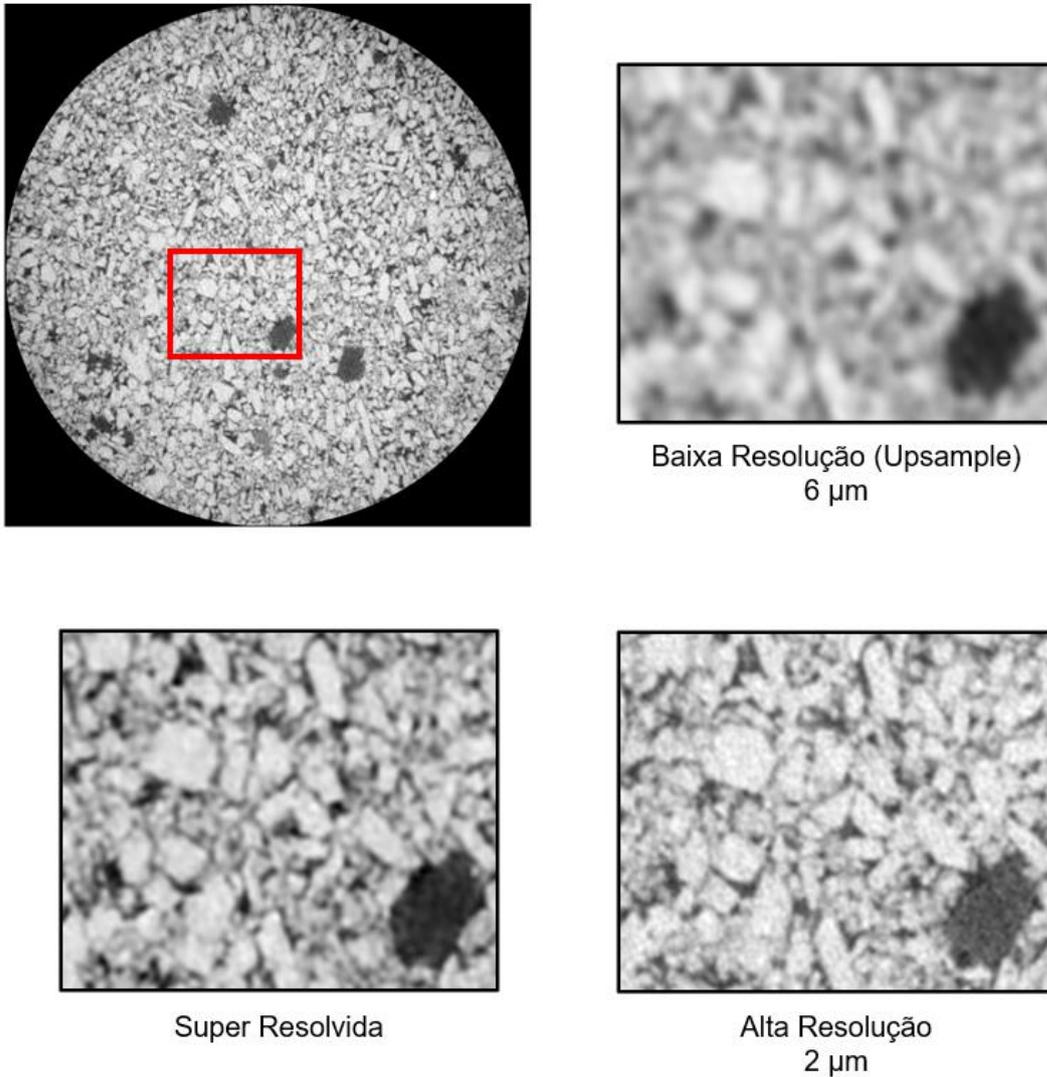


Figura 37: Visualização das versões da *ROI* em baixa resolução, super resolvida e alta resolução.

Os resultados do modelo SR-II também podem ser comparados às imagens capturadas em resolução de 4 μm. A Figura 38 apresenta uma comparação entre as imagens super resolvidas, 4 μm e 2 μm.

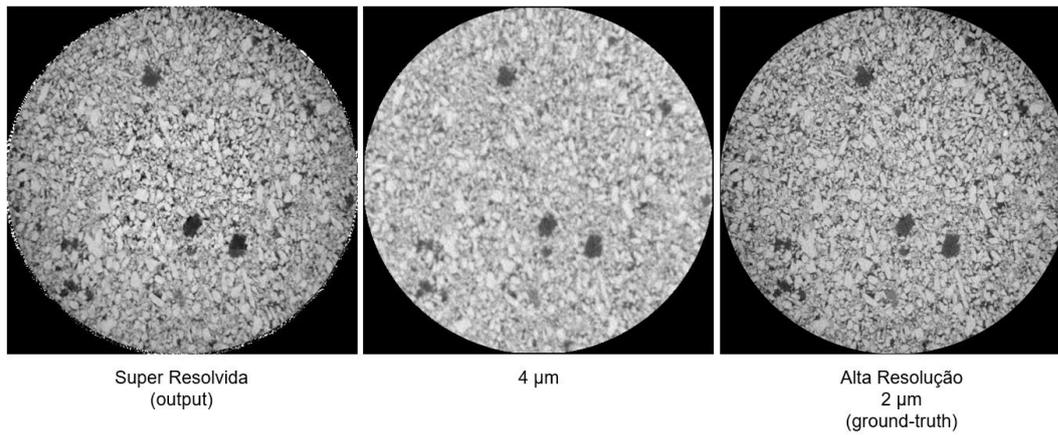


Figura 38: Comparação das imagens super resolvida (esquerda), 4 μm (centro), e HR (direita).

A Figura 39 também expõe um esquema comparativo, entre essas resoluções, para uma região ampliada da imagem. Nota-se que a captura de 4 μm expressa uma melhora quase insignificante de nitidez em relação à imagem de 6 μm . Entretanto, a imagem de 4 μm ainda está distante, em questão de qualidade, da captura de 2 μm . Neste ponto, a imagem super resolvida possui nitidez e contrastes entre partículas e poros mais próximos à captura de 2 μm .

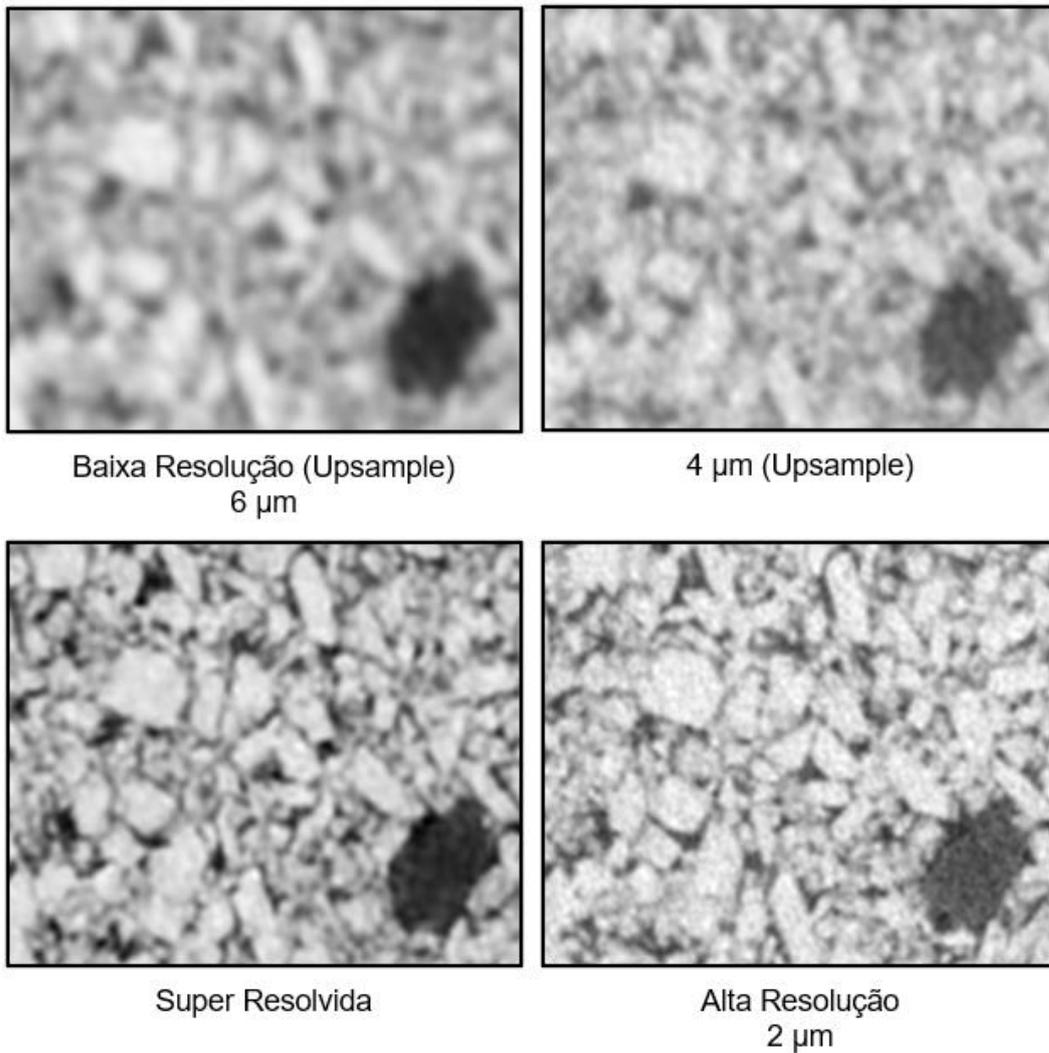


Figura 39: Visualização das versões da Rol super resolvida, em 4 μm e em 2 μm .

5.1.2.2 PSNR

O valor médio dos PSNRs foi calculado para o conjunto de testes contendo 25 imagens. A média para PSNR entre as imagens super resolvidas e as imagens HR foi de 20,36 dB, enquanto a média para PSNR entre as imagens LR e imagens HR foi de 20,12 dB.

Apesar do valor de PSNR das imagens super resolvidas ser maior que o valor de PSNR das imagens LR, essa diferença não evidencia objetivamente a melhora da qualidade visual das imagens super resolvidas, em comparação às imagens LR.

5.1.2.3 Segmentação de Poros

A Figura 40 apresenta a pré-visualização do processo de segmentação por limiarização.

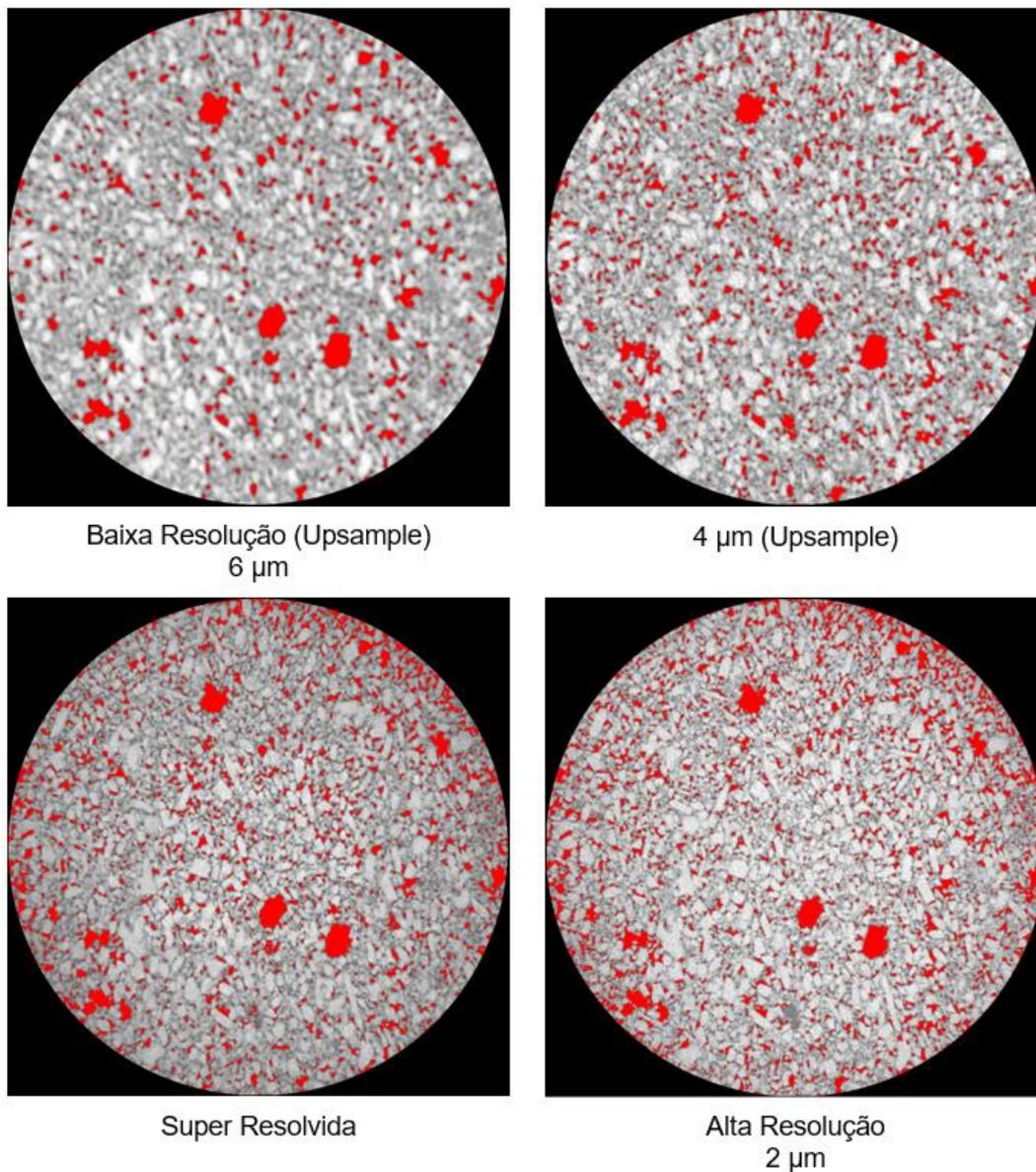


Figura 40: Processo de segmentação das regiões porosas. A segmentação da LR foi feita na faixa tonal de 1-118; de 4 μm foi feita entre 1-115; da super resolvida foi feita entre 1-100; e da HR foi feita entre 1-101.

As imagens binárias resultantes do processo de segmentação são apresentadas na Figura 41. Neste caso, não foi necessário nenhum pós-processamento para eliminar partículas incorretas.

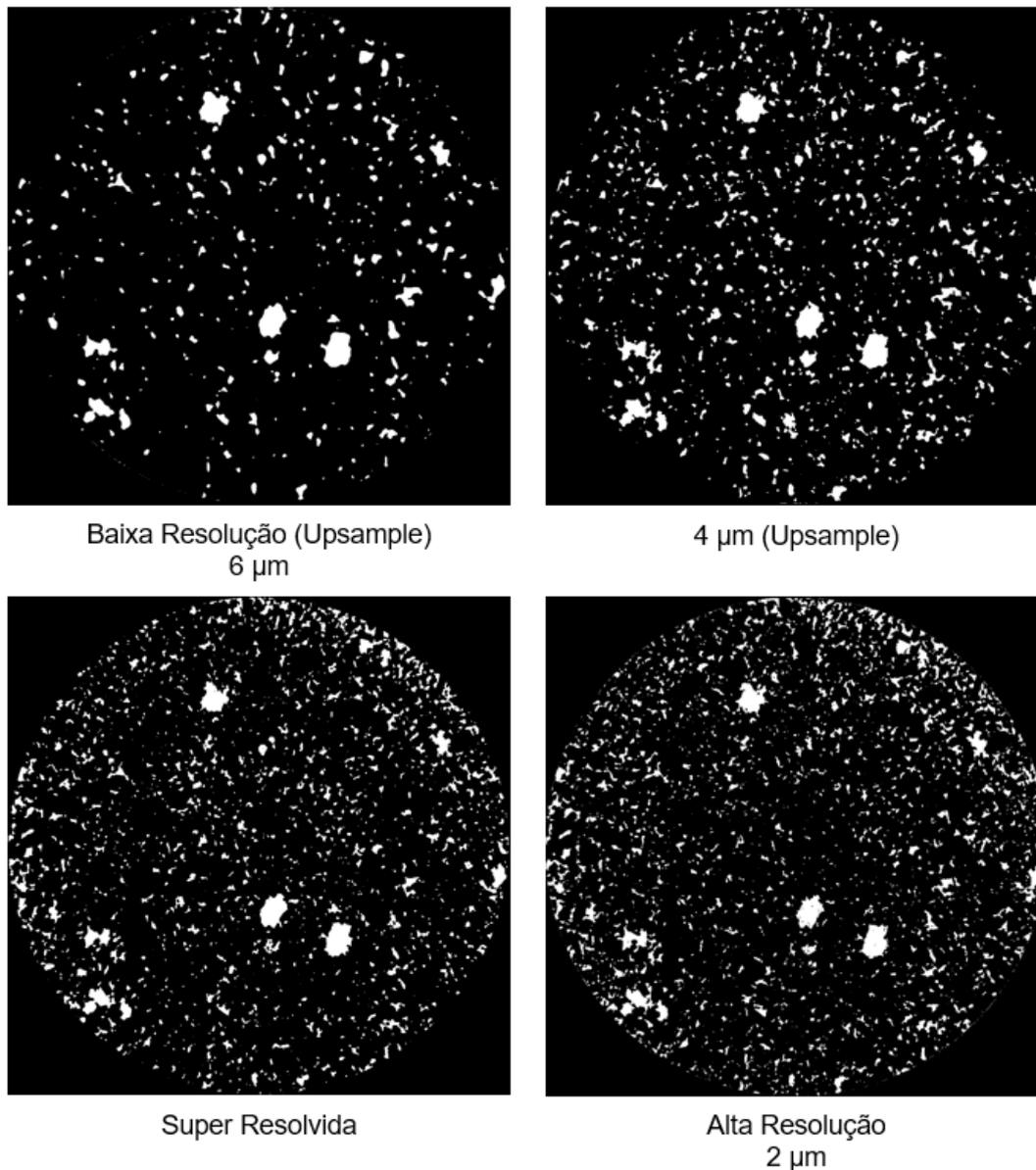


Figura 41: Visualização das imagens binárias obtidas para segmentação de poros nas versões LR, 4 μm, SR e HR.

Ainda que o aumento de resolução de 6 μm para 2 μm seja menos abrupto do que aumento de 16 μm para 6 μm, em questão de nitidez e detalhes da estrutura dos briquetes, as diferenças entre as segmentações continuam significantes.

A Tabela 10 apresenta os valores obtidos para área total de poros, a respectiva fração de área e quantidade de poros identificados para as diferentes resoluções da imagem.

Tabela 10: Análises quantitativas feitas a partir da segmentação de poros.

Modelo	Resolução	Área Total de Poros (pixels)	Fração em área	Contagem de poros
SR-II	Baixa (6 μm)	40417	5,6%	467
	4 μm	60222	8,3%	1005
	Super Resolvida	69081	9,5%	1732
	Alta Resolução (2 μm)	75695	10,3%	2218

O modelo SR-II expõe uma grande variação entre a fração de área entre a segmentação LR e HR, em uma diferença de 4,74 pontos percentuais, isto é, 87,3%. Em contraste com o modelo SR-I, a fração em área do modelo SR-II não ultrapassou o valor encontrado para a imagem HR. A fração em área da imagem SR se mostra maior que da imagem de 4 μm , se aproximando por 0,82 ponto percentual (0,96%) de diferença da imagem HR. Além disso, a SR também é capaz de reconhecer mais poros que a imagem de 4 μm . A contagem de poros da imagem super resolvida também foi a que mais se aproximou da contagem da imagem HR.

5.1.2.4 Precision, Recall e F1-Score

A Figura 42 apresenta as imagens resultantes da operação *AND* entre as imagens LR e HR, 4 μm e HR, e SR e HR.

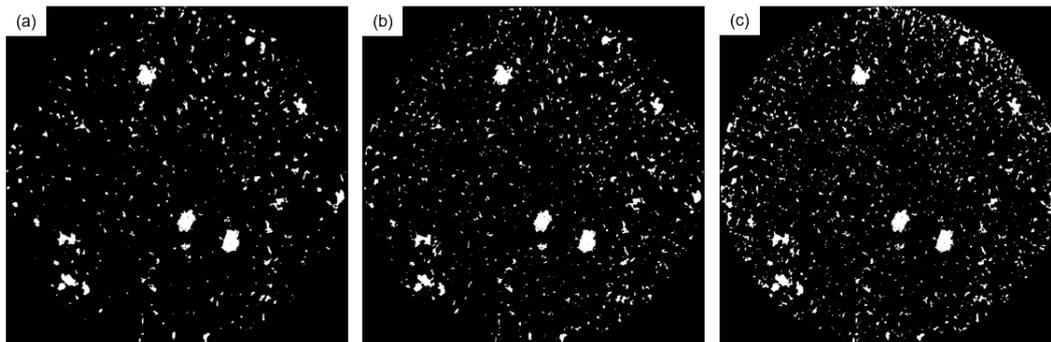


Figura 42: Resultados da operação AND entre LR e HR (a), 4 μm e HR (b), e SR e HR (c).

A Tabela 11 apresenta as métricas obtidas.

Tabela 11: Medidas de Precision e Recall.

	LR e HR	4 μm e HR	SR e HR
Precision	76,27%	59,83%	70,08%
Recall	40,72%	47,60%	63,95%
F1-Score	53,10%	53,02%	66,87%

Apesar do precision alto, a imagem LR possui recall muito baixo, o que indica alta quantidade de pixels falsos negativos, ou seja, pixels que existem na imagem HR, mas não se encontram na imagem LR. A SR apresentou recall de 23,23 pontos percentuais maior que da imagem LR. A queda de 6,19 pontos percentuais no precision da SR, em comparação com LR, indica que a imagem super resolvida possui mais falsos positivos, um comportamento esperado.

Em suma, o modelo SR-II apresentou excelentes impactos na análise de poros segmentados, exibindo, inclusive, métricas superiores às capturas em 4 μm .

5.1.2.5 Segmentação de Volume 3D

A segmentação da LR foi feita com limiar 1-114, a segmentação da imagem HR foi feita com limiar 1-102, e a SR foi feita com limiar 1-100. A Figura 44 apresenta a visualização dos poros segmentados para cada resolução.

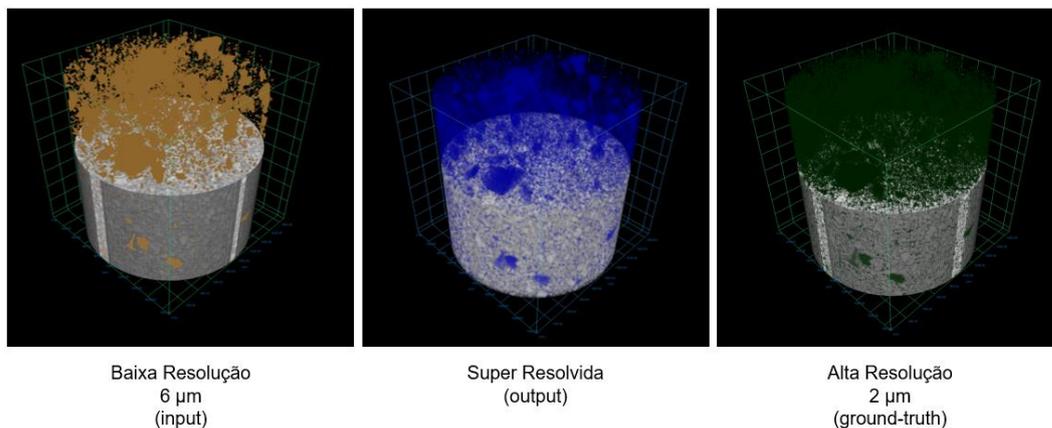


Figura 43: Visualização dos poros segmentados em LR (esquerda), super resolução (centro) e HR (direita).

A Tabela 12 expõe as medidas de porosidade encontradas.

Tabela 12: Valores de porosidade das amostras 3D.

Modelo	Resolução	Fração em Volume
SR-I	Baixa (6 μm)	4,4%
	Super Resolvida	9,0%
	Alta Resolução (2 μm)	8,2%

Para a análise de todo o volume, a fração em área da imagem super resolvida superestimou a fração encontrada na imagem HR por 0,8 ponto percentual (9,8%). Esses resultados ainda são coerentes, tendo em vista a sensibilidade da escolha de limiar.

5.1.3 Modelo SR-III: 04 para 02 (2X)

O modelo SR-III foi treinado por aproximadamente 28 horas, alcançando 390 épocas (232440 iterações).

5.1.3.1 Análise Visual

A Figura 44 apresenta a comparação entre versões LR, super resolvida e HR de uma imagem do conjunto de testes. Novamente, imagem super resolvida demonstra possuir um detalhamento e nitidez de estruturas maior que o presente na versão LR, se aproximando da versão HR. O modelo SR-II ainda gera, mas com menor intensidade, os artefatos ao redor da amostra.

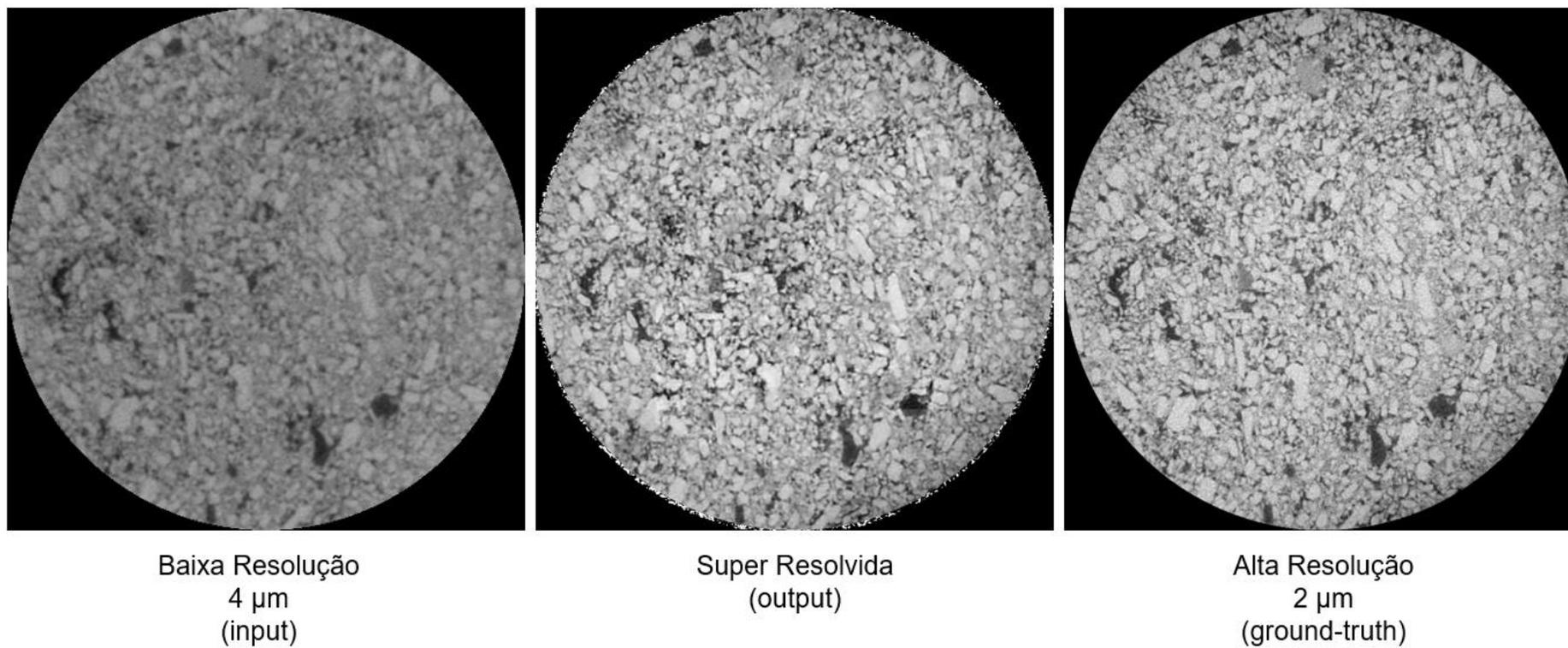


Figura 44: Comparação das imagens LR (esquerda), super resolvida pelo modelo SR-III (centro), e HR (direita). Corte da amostra F34B-TA.

A Figura 45 apresenta outro comparativo para uma região ampliada da imagem.

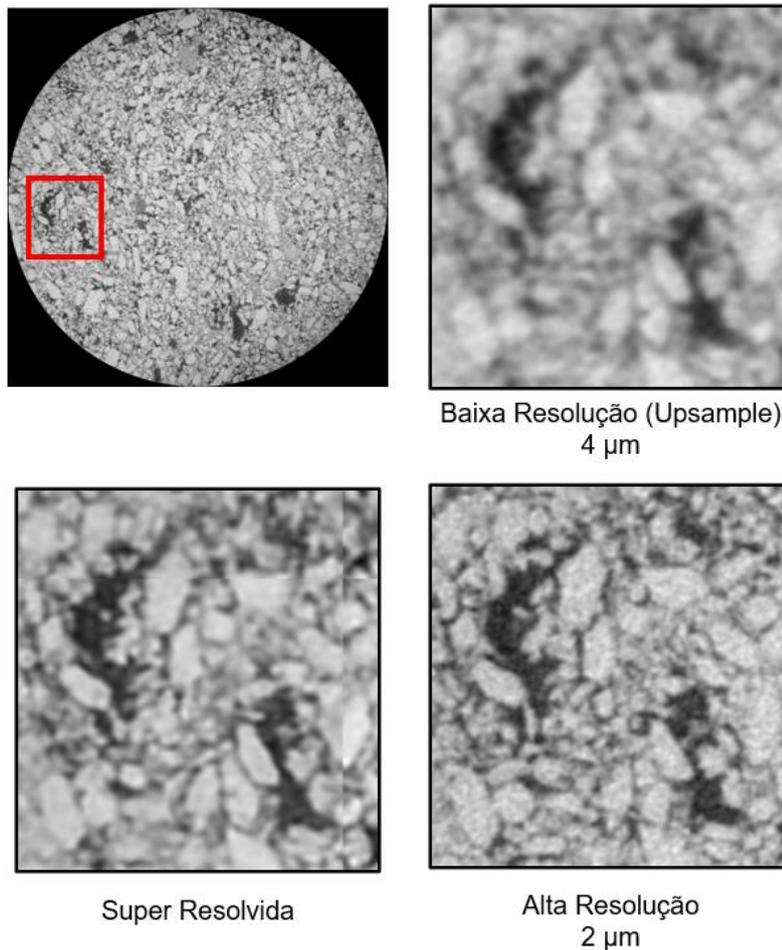


Figura 45: Visualização das versões da *RoI* em baixa resolução (esquerda), super resolvida (centro) e alta resolução (direita).

5.1.3.2 PSNR

O valor médio dos PSNRs foi calculado para o conjunto de testes contendo 18 imagens. A média para PSNR entre as imagens super resolvidas e as imagens HR foi de 19,28 dB, enquanto a média para PSNR entre as imagens LR e imagens HR foi de 18,59 dB.

Novamente, apesar de ocorrer um aumento da métrica, a diferença entre os valores não demonstrou nenhum padrão em relação aos modelos anteriores. Sendo assim, o PSNR não evidencia objetivamente a melhora da qualidade visual observada das imagens super resolvidas, em comparação às imagens LR.

5.1.3.3 Segmentação de Poros

A Figura 46 apresenta a pré-visualização do processo de segmentação por limiarização, de acordo com as faixas tonais escolhidas. A visualização da segmentação da imagem HR conta com as anotações, em cor amarela, dos poros segmentados da imagem LR, enquanto a visualização da segmentação da imagem super resolvida conta com as anotações, em verde dos poros segmentados da imagem HR.

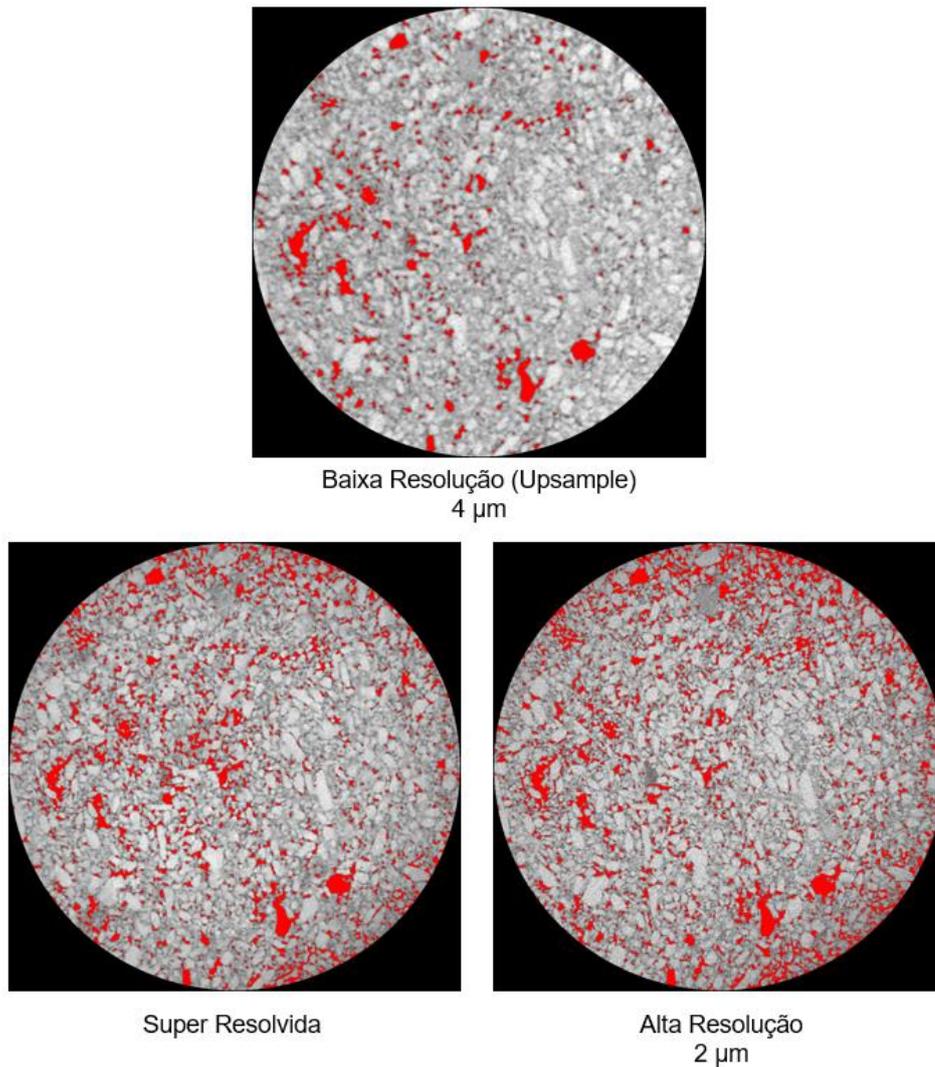


Figura 46: Processo de segmentação das regiões porosas. A segmentação da LR foi feita na faixa tonal de 1-114; da super resolvida foi feita entre 1-104; e da HR foi feita entre 1-100.

As imagens binárias resultantes do processo de segmentação são apresentadas na Figura 47.

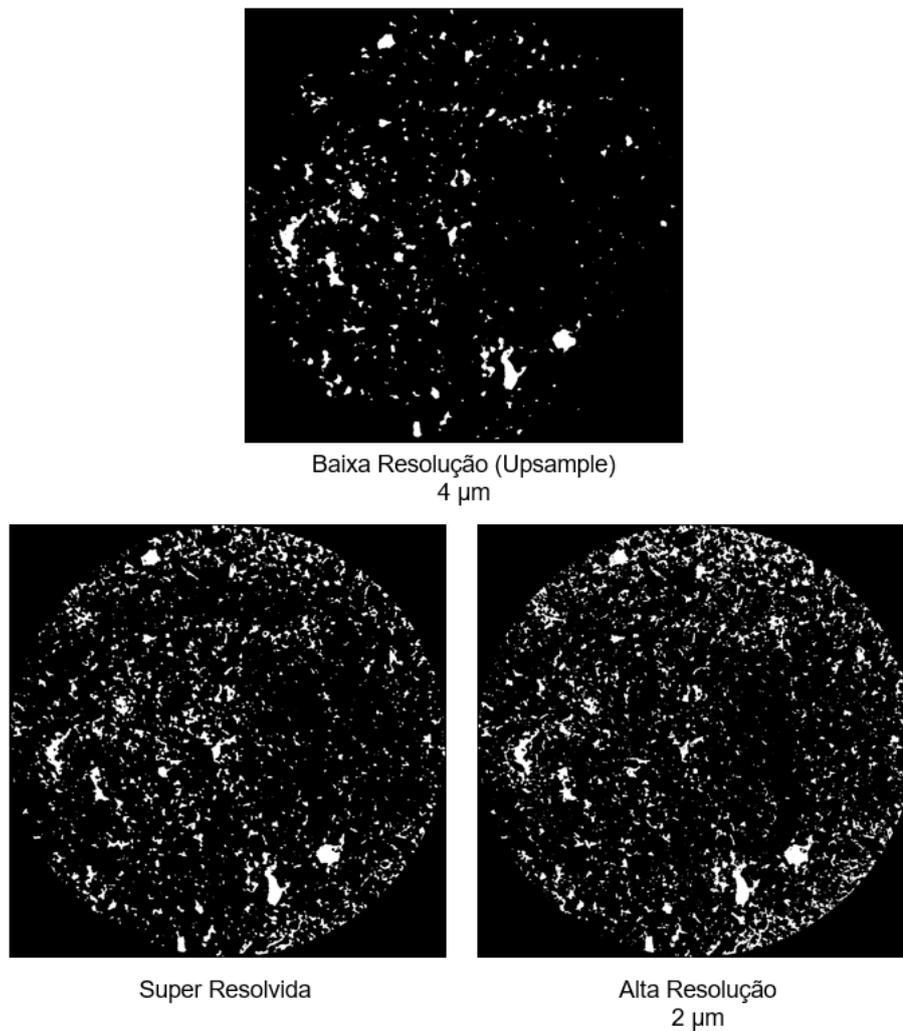


Figura 47: Visualização das imagens binárias obtidas para segmentação de poros nas versões LR, HR e SR.

Seguindo o mesmo comportamento dos outros modelos, a imagem binária da LR exibe uma quantidade de poros muito inferior aos resultados da imagem super resolvida e da HR, que por sua vez reconhece poros de menores diâmetros.

A Tabela 13 apresenta os valores obtidos para área total de poros, a respectiva fração de área e quantidade de poros identificados para as diferentes resoluções da imagem.

Tabela 13: Análises quantitativas feitas a partir da segmentação de poros.

Modelo	Resolução	Área Total de Poros (pixels)	Fração em área	Contagem de poros
SR-III	Baixa (4 μm)	30996	4,4%	484
	Super Resolvida	66067	9,3%	1600
	Alta Resolução (2 μm)	78734	11,1%	2379

A variação da fração em área dos poros entre a LR e HR foi de 6,7 pontos percentuais (154% de aumento em termos de área total), enquanto a contagem de poros foi de 484 poros para LR, e 2379 para HR.

Quanto a imagem super resolvida, constata-se que a SR apresentou apenas 1,8 ponto percentual (19,2% em área total) de diferença para a fração em área aferida para a imagem HR. A quantidade de poros contabilizados na imagem super resolvida também destoa muito à quantidade identificada na imagem LR, se aproximando bem da quantidade presente na HR.

5.1.3.4 Precision, Recall e F1-Score

A Figura 48 apresenta as imagens resultantes da operação *AND* entre as imagens LR e HR, e SR e HR.

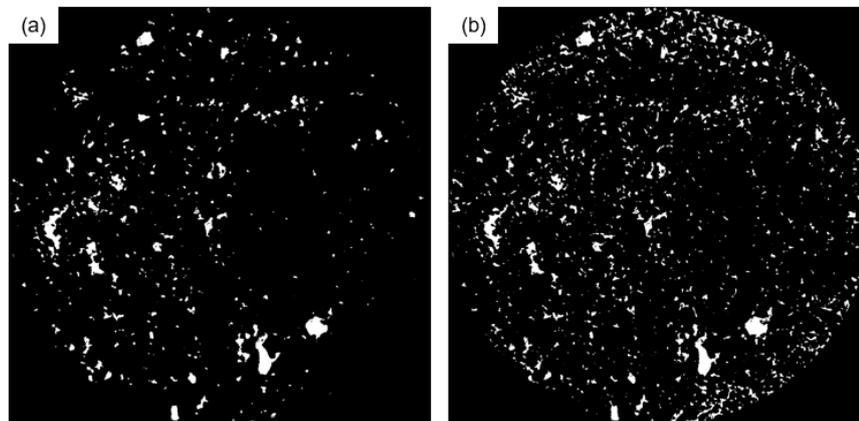


Figura 48: Resultados da operação *AND* entre LR e HR (a), e SR e HR (b).

A Tabela 14 apresenta as métricas obtidas.

Tabela 14: Medidas de Precision e Recall.

	LR e HR	SR e HR
Precision	85,93%	69,66%
Recall	33,83%	58,45%
F1-Score	48,55%	63,57%

Assim como nos outros modelos, apesar do precision alto, a imagem LR possui recall muito baixo, o que indica alta quantidade de pixels falsos negativos. A SR apresentou recall 24,62 pontos percentuais maior que da imagem LR, demonstrando novamente que é capaz de recuperar informações da estrutura

porosa do briquete. A queda de 16,27 pontos percentuais no precision da SR, em comparação com LR, indica que o modelo SR-III possui maior tendência a gerar pixels falsos positivos.

Em suma, o modelo SR-III também apresentou impactos positivos na análise de poros segmentados, mesmo gerando mais falsos negativos.

5.1.3.5 Segmentação de Volume 3D

A segmentação da LR foi feita com limiar 1-108, a segmentação da imagem HR foi feita com limiar 1-100, e a SR foi feita com limiar 1-100. A Figura 49 apresenta a visualização dos poros segmentados para cada resolução.

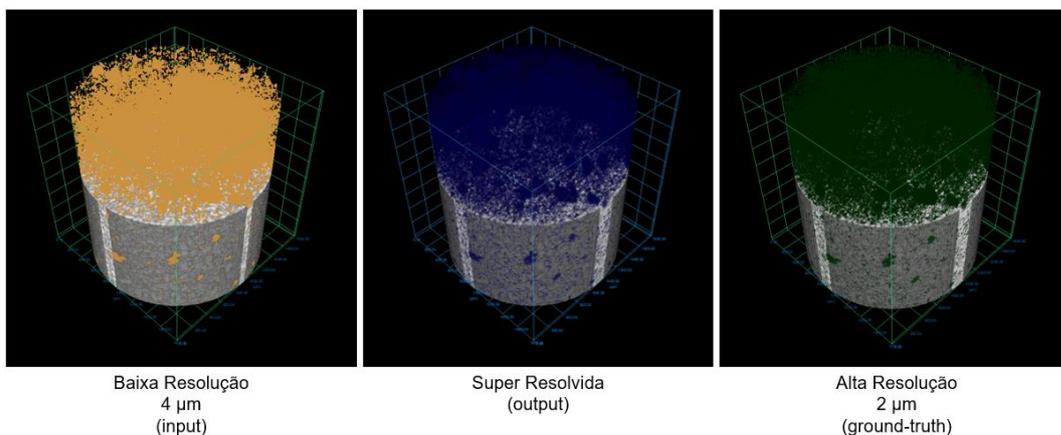


Figura 49: Visualização dos poros segmentados em LR (esquerda), super resolução (centro) e HR (direita).

A Tabela 15 expõe as medidas de porosidade encontradas.

Tabela 15: Valores de porosidade das amostras 3D.

Modelo	Resolução	Fração em Volume
SR-III	Baixa (4 µm)	6,3%
	Super Resolvida	12,5%
	Alta Resolução (2 µm)	10,0%

O aumento na taxa de falsos-positivos pode ser observado na segmentação tridimensional, na qual a super resolução superestimou a quantidade realista de poros em 2,5 pontos percentuais (25% de erro). Ainda assim, a super resolução apresenta uma fração em área mais próxima da HR.

5.2 Considerações finais

Os três modelos treinados obtiveram desempenhos satisfatórios, que evidenciam a recuperação efetiva de detalhes e redução de ruídos durante o processo de aumento de resolução por SR. Indo além, outra análise essencial é estudar o benefício de cada modelo quanto à economia efetiva de tempo.

Para isso, é necessário investigar o tempo de análise em relação ao volume adquirido durante uma captura tomográfica. A Tabela 16 expõe esses dados para a amostra F34B-TA. A categoria “proporção de volume” se refere a quantas vezes um volume é maior que o menor volume (captura de 2 μm).

Tabela 16: Dados de tempo de análise e volume adquirido para a amostra F34B-TA.

Amostra	Resolução	Tempo de captura (h)	Volume adquirido (mm^3)	Proporção de volume	Número de fatias
F34B_TA	16 μm	1,08	103,00	16,6	1018
	06 μm	1,08	58,66	9,5	1018
	04 μm	1,08	33,89	5,5	1002
	02 μm	2,17	6,19	1	998

A Tabela 17 apresenta o tempo levado para gerar as imagens super resolvidas, o número de camadas processadas e o tempo médio de processamento por camada.

Tabela 17: Dados de tempo para aplicação da SR na amostra F34B-TA.

Modelo	Tempo de processamento	Fatias Processadas	Tempo de Processamento por Fatia (min/fatia)
SR-I	1,3 min	289	0,00450
SR-II	29,1 min	313	0,09297
SR-III	21,7 min	466	0,04657

A partir dos dados referentes ao modelo SR-I, pode-se estimar que a captura do volume completo da amostra em resolução de 6 μm levaria, no mínimo, 2,16 horas (referente a 2 capturas). O mesmo volume em resolução de 16 μm levou apenas 1,08 hora. O processamento das 1018 camadas presentes da imagem de 16 μm leva, em média, 4,6 minutos (0,08 horas). Portanto, o uso do modelo SR-I é capaz de gerar uma estimativa adequada da imagem de 6 μm , com uma redução

de 1 hora (incluindo o tempo de processamento). Efetivamente, o uso do modelo SR-I poupa 1,08 hora de uso do microtomógrafo.

Em relação ao modelo SR-II, a captura do volume em 6 μm levou 1,08 horas, enquanto a captura do mesmo volume em 2 μm levaria, no mínimo, 21,7 horas (referente a 10 capturas). O processamento das 1018 camadas da imagem em 6 μm leva, em média, 94,6 minutos (1,58 horas). Assim, o modelo SR-II é capaz de gerar uma estimativa adequada da imagem de 2 μm , com redução de 19 horas (incluindo o tempo de processamento). Efetivamente, o emprego do modelo SR-II poupa 20,62 horas de uso do tomógrafo. Além disso, é possível comparar ao tempo mínimo de 2,16 horas (referente a 2 capturas) necessárias para a captura do volume em resolução de 4 μm . Como os resultados apontaram que a imagem super resolvida está mais próxima da captura em 2 μm do que a captura de 4 μm está, o modelo SR-II efetivamente poupa 1,08 hora de uso da fonte de Raios-X, em relação à captura em 4 μm .

Já para o modelo SR-III, a captura do volume em 4 μm levou 1,08 horas, enquanto a captura do mesmo volume em 2 μm levaria, no mínimo, 6,48 horas. O processamento das 1002 camadas da imagem em 4 μm leva, em média, 46,7 minutos (0,78 horas). Assim, o modelo SR-III é capaz de gerar uma estimativa adequada da imagem de 2 μm , com redução de 5,7 horas (incluindo o tempo de processamento). O Modelo SR-III efetivamente poupa 5,4 horas de uso da fonte de Raios-X.

Para facilitar a visualização de todas essas comparações, a Tabela 18 apresenta um resumo dos dados discutidos. Na tabela, são dados os tempos de aquisição para as versões LR e HR de cada um dos modelos. A "Tempo SR" se refere à soma do tempo de aquisição da imagem LR e seu tempo de processamento pelo respectivo modelo de SR. Também são apresentados a economia de horas de tomografia, devido ao uso da SR, em percentual.

Tabela 18: Resumo das comparações de tempo.

Modelo	Tempo LR (h)	Tempo HR (h)	Tempo SR (Tempo LR + Tempo Processamento) (h)	Economia de Tempo
SR-I	1,08	2,16	1,16	46%
SR-II	1,08	21,7	2,66	88%
SR-III	1,08	6,48	1,86	71%

6 Conclusão

A metodologia desenvolvida para este trabalho, envolvendo as etapas de aquisição de imagens em diferentes resoluções, processamento digital de imagens e treinamento de modelos, foi capaz de gerar modelos de SR capazes de recuperar informações de alta frequência (presentes nas imagens HR) a partir de uma imagem LR. O desempenho dos modelos foi investigado com sucesso por meio de métricas e análises quantitativas que indicam um impacto positivo da SR em análises digitais de imagem.

Em geral, os três modelos treinados foram capazes de criar pixels de regiões porosas que estão perfeitamente correlacionados aos poros da imagem HR. Este comportamento foi verificado pelo aumento do *recall* (diminuição de falsos negativos) para imagens super resolvidas. Por outro lado, os modelos também tendem, em menor grau, a gerar pixels falsos-positivos, evidenciado pela diminuição do *precision* (aumento de falsos positivos). A combinação destes comportamentos expõe que a SR é extremamente eficaz em estimar uma imagem HR, ao passo que as imagens super resolvidas podem gerar, de forma branda, algumas distorções nas estruturas dos aglomerados.

Com os resultados obtidos, também é possível concluir que o PSNR não é uma métrica adequada para avaliar modelos de SR, uma vez que os aumentos de PSNR foram pouco significativos para os modelos SR-II e SR-III, em contraste com as outras métricas obtidas.

Tendo em vista a eficácia da SR quanto ao aumento de resolução, a técnica também alcança os objetivos desejados em relação à atenuação das limitações das capturas tomográficas. A aplicação de SR traz, efetivamente, economia ao tempo de captura, enquanto possibilita o aumento de resolução de amostras capturas com maior volume (pior resolução).

Devido ao potencial da SR constatado neste projeto, sua metodologia pode ser aperfeiçoada, a começar pela captura tomográfica de mais amostras de aglomerado. Quanto ao treinamento das redes, pode-se realizar uma investigação mais profunda a partir do ajuste fino (*fine-tuning*) dos parâmetros da rede. Tendo

em vista também o rápido avanço do campo de DL, outras arquiteturas de redes neurais também poderão ser comparadas, assim como novas métricas de avaliação de imagem podem ser empregadas.

Enfim, os resultados discutidos indicam um impacto positivo da aplicação de SR em microCT. O desenvolvimento desta técnica pode ser um fator influente para a popularização da técnica na indústria, uma vez que o uso da SR possibilita redução no tempo de captura e a utilização de microtomógrafos com menor poder de resolução.

7

Referências bibliográficas

- [1] GEERDES, M.; TOXOPEUS, H.; VLIET, C. van der. **Modern Blast Furnace Ironmaking**: an introduction. 4. ed. [S.l.], IOS Press BV, 2020.
- [2] LU, L. **Iron Ore**: Mineralogy, Processing and Environmental Sustainability. 2. ed. [S.l.], Woodhead Publishing, 2022.
- [3] J. M. Mourão. Estudo Prospectivo do Setor Siderúrgico, Brasília, CGEE/ABM, 2008.
- [4] LUZ, A. B.; SAMPAIO, J.A.; FRANÇA, S. C. A. **Tratamento de Minérios**. 5. ed., Rio de Janeiro: CETEM/MCT, 2010.
- [5] FERNÁNDEZ-GONZÁLEZ, D.; PIÑUELA-NOVAL, J.; VERDEJA, L. F. 'Iron Ore Agglomeration Technologies'. *Iron Ores and Iron Oxide Materials*. **InTech**. 2018. doi:10.5772/intechopen.72546.
- [6] AUGUSTO, K. S.; PACIORNIK, S. Porosity characterization of iron ore pellets by x-ray microtomography, **Materials Research**, v. 21, n. 2, 2018, doi: 10.1590/1980-5373-mr-2017-0621.
- [7] BEZERRA, E. T. V.; AUGUSTO, K. S.; PACIORNIK, S. Discrimination of pores and cracks in iron ore pellets using deep learning neural networks. **REM-International Engineering Journal**, v. 73, p. 197-203, 2020.
- [8] AUGUSTO, K. S. **Microtomografia Computadorizada de Raios X Aplicada à Caracterização de Porosidade em Pelotas de Minério de Ferro**. Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Química e de Materiais, 2016.
- [9] EISBERG, R.; RESNICK, R. **Física Quântica**. 3 ed. New York, John Wiley & Sons Inc., 1994.
- [10] POPOV, O. et al. Quantitative Microstructural Analysis and X-ray Computed Tomography of Ores and Rocks—Comparison of Results, **Minerals**, v. 10, n. 2, p. 129, 2020.
- [11] MANTOVANI, I. F. **Microtomografia e nanotomografia de raios X aplicada à caracterização multiescalar de sistemas porosos carbonáticos**. Tese de Doutorado. Florianópolis, 2013.
- [12] KUBAT, M. **An Introduction to Machine Learning**. 2. ed. [S.l.] Springer Nature, 2017.

- [13] LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning, **Nature**, v. 521, n. 7553, p. 436–444, 2015. DOI: 10.1038/nature14539.
- [14] SHRESTHA, A.; MAHMOOD, A. Review of deep learning algorithms and architectures, **IEEE Access**, v. 7, p. 53040–53065, 2019. DOI: 10.1109/ACCESS.2019.2912200.
- [15] KARIRI, E. et al. Exploring the Advancements and Future Research Directions of Artificial Neural Networks: A Text Mining Approach, **Appl. Sci.**, v. 13, p. 3186, 2023. DOI: <https://doi.org/10.3390/>
- [16] MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity, **The bulletin of mathematical biophysics**, v. 5, p. 115-133, 1943.
- [17] SOARES, P.; SILVA, J. P. Aplicação de Redes Neurais Artificiais em Conjunto com o Método Vetorial da Propagação de Feixes na Análise de um Acoplador Direcional Baseado em Fibra Ótica. **Revista Brasileira de Computação Aplicada**, 2011. DOI: 10.5335/rbca.2013.1803
- [18] HAYKIN, S. **Redes Neurais: Princípios e prática**. 2. ed. São Paulo: Bookman, 2008.
- [19] WANG, Y. et al. The influence of the activation function in a convolution neural network model of facial expression recognition, **Applied Sciences**, v. 10, n. 5, 2020. DOI: 10.3390/app10051897.
- [20] APICELLA, A. et al. A survey on modern trainable activation functions, **Neural Networks**, v. 138, p. 14-32, 2021.
- [21] DUBEY, S. R.; SINGH, S. K.; CHAUDHURI, B. B. Activation functions in deep learning: A comprehensive survey and benchmark, **Neurocomputing**, v. 503, p. 92-108, 2022.
- [22] HE, K.; ZHANG, X.; REN, S.; SUN, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: **Proceedings of the IEEE international conference on computer vision**, p. 1026–1034, 2015.
- [23] ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain, **Psychological Review**, v. 65, n. 6, p. 386–408, 1958. DOI: 10.1037/h0042519.
- [24] FIORIN, D. V.; MARTINS, F. R.; SCHUCH, N. J. et al. Aplicações de redes neurais e previsões de disponibilidade de recursos energéticos solares, **Rev. Bras. Ensino Fís.**, v. 33, n.1, p. 1309, 2011.
- [25] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**, [S.I.], MIT Press, 2016.

- [26] SVENSSON, T. **Semantic Segmentation of Iron Pellets with Neural Networks**. Dissertação (Mestrado em Engenharia Espacial), Lulea University of Technology, 2019.
- [27] HE, K., ZHANG, X., REN, S. “Deep Residual Learning for Image Recognition”. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**, p. 770-778, 2016.
- [28] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**, p. 1-9, 2015.
- [29] BENGIO, Y.; SIMARD, P.; FRASCONI, P. Learning long-term dependencies with gradient descent is difficult. **IEEE transactions on neural networks**, v. 5, n. 2, p. 157-166, 1994.
- [30] PASCANU, R., MIKOLOV, T., BENGIO, Y. On the difficulty of training Recurrent Neural Networks, In: **International conference on machine learning**. Pmlr, 2013. p. 1310-1318.
- [31] IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: **International conference on machine learning**. pmlr, 2015. p. 448-456.
- [32] OOI, Y. K.; IBRAHIM, H. Deep Learning Algorithms for Single Image Super-Resolution: A Systematic Review, **Electronics**, v. 10(7), p. 867, 2021.
- [33] CHEN, H. et al. Real-world single image super-resolution: A brief review, **Information Fusion**, v. 79, p. 124-145, 2022.
- [34] WANG, Z.; CHEN, J.; HOI, S. C. H. et al. Deep Learning for Image Super-resolution: A Survey. **IEEE transactions on pattern analysis and machine intelligence**, v. 43, n. 10, p. 3365-3387, 2020.
- [35] SHAO, WZ; ELAD, M. Simple, accurate, and robust nonparametric blind super-resolution. In: Image and Graphics: **8th International Conference, ICIG 2015**, Springer International Publishing, 2015. p. 333-348.
- [36] LI, X. et al. Deep learning methods in real-time image super-resolution: a survey, **Journal of Real-Time Image Processing**, v. 17, p. 1885–1909, 2020.
- [37] DONG, C. et al. Learning a deep convolutional network for image super-resolution. In: **Computer Vision–ECCV 2014: 13th European Conference**. Springer International Publishing, 2014. p. 184-199.

- [38] KIM, J.; LEE, J.K.; LEE, K.M. Accurate image super-resolution using very deep convolutional networks. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**, 2016, p. 1646–1654.
- [39] KIM, J.; LEE, J.K.; LEE, K.M. Deeply-Recursive Convolutional Network for Image Super-Resolution. In: **Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition** 2016, p. 1637–1645.
- [40] SHI, W. et al. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. In: **Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition**, 2016; p. 1874–1883.
- [41] LEDIG, C. et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. **Proceedings of the IEEE conference on computer vision and pattern recognition**, 2017, p. 4681-4690.
- [42] LIM, B. et al. Enhanced Deep Residual Networks for Single Image Super-Resolution. In: **Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops**, 2017, p. 1132–1140.
- [43] SHI, W. et al. Is the deconvolution layer the same as a convolutional layer?. **arXiv preprint arXiv:1609.07009**, 2016.
- [44] WANG, Z. et al. Image quality assessment: from error visibility to structural similarity. **IEEE transactions on image processing**, v. 13, n. 4, p. 600-612, 2004.
- [45] WANG, Z.; BOVIK, A. C. Mean squared error: Love it or leave it? A new look at signal fidelity measures. **IEEE signal processing magazine**, v. 26, n. 1, p. 98-117, 2009.
- [46] REIBMAN, A. R.; SCHAPER, T. Subjective performance evaluation of super-resolution image enhancement. **Second Int. Wkshp on Video Proc. and Qual. Metrics (VPQM'06)**, 2006.
- [47] REIBMAN, A. R.; BELL, R. M.; GRAY, S. Quality assessment for super-resolution image enhancement. In: **2006 International Conference on Image Processing**. p. 2017-2020. 2006.
- [48] HENDRIKSEN, A. A. et al. On-the-fly machine learning for improving image resolution in tomography. **Applied Sciences**, v. 9, n. 12, p. 2445, 2019.
- [49] BIZHANI, M.; ARDAKANI, O. H.; LITTLE, E. Reconstructing high fidelity digital rock images using deep convolutional neural networks. **Scientific reports**, v. 12, n. 1, p. 1-14, 2022.
- [50] ROSLIN, A. et al. Processing of micro-CT images of granodiorite rock samples using convolutional neural networks (CNN), Part I: Super-resolution

enhancement using a 3D CNN. **Minerals Engineering**, v. 188, p. 107748, 2022.

- [51] Dragonfly 2020.2 [Computer software]. Object Research Systems (ORS) Inc, Montreal, Canada, 2020; software disponível em <<http://www.theobjects.com/dragonfly>>.
- [52] Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. **BMC Bioinformatics**, 18(1).
- [53] **Module: tf.data**. Tensorflow. Disponível em <https://www.tensorflow.org/api_docs/python/tf/data>. Acesso em: 14/06/2023.
- [54] ABADI, M. et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. **arXiv preprint arXiv:1603.04467**, 2016. Software disponível em <[tensorflow.org](https://www.tensorflow.org)>.
- [55] Krasser, M. **Single Image Super-Resolution with EDSR, WDSR and SRGAN**. Disponível em < <https://github.com/krasserm/super-resolution>>. Acesso em: 14/06/2023.