

Ada Cristina França da Silva

# **Uma Rede Neural Convolutacional para Segmentação de Imagens de Minério de Ferro**

Brasil

2022, v-0.0.1

Ada Cristina França da Silva

## Uma Rede Neural Convolutacional para Segmentação de Imagens de Minério de Ferro

Defesa apresentada ao Programa de Pós-Graduação em Engenharia de Computação e Sistemas da Universidade Estadual do Maranhão como parte das exigências para obtenção do título de Mestre em Engenharia de Computação e Sistemas.



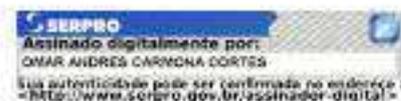
---

Dr. RENATO PORFÍRIO ISHII, UFMS  
Examinador(a) Externo à IES



---

Dr. FÁBIO MANOEL FRANÇA LOBATO, PECS/UEMA  
Examinador(a) Interno



---

Dr. OMAR ANDRES CARMONA CORTES, PECS/UEMA  
Presidente(a)

Brasil  
2022, v-0.0.1

Silva, Ada Cristina França da.

Uma rede neural convolucional para segmentação de imagens de minério de ferro / Ada Cristina França da Silva. – São Luís, 2022. 65

Dissertação (Programa de Pós-Graduação em Engenharia de Computação e Sistemas) - Universidade Estadual do Maranhão, 2022.

Orientador: Prof. Dr. Omar Andres Carmona Cortes.

1.Minério de ferro. 2.Segmentação. 3.Redes neurais. 4.CNN. 5.Visão computacional. I.Título.

CDU: 004.9:622

*“A matemática mostra o que é. É a linguagem das relações invisíveis entre as coisas. Mas para usar e aplicar essa linguagem, devemos ser capazes de apreciar plenamente, sentir, apreender o invisível, o inconsciente..*  
*(Lovelace, Ada)*

# Agradecimentos

Agradeço primeiramente aos meus pais, o meu querido pai Adauto Vaz da Silva pelo apoio aos meus estudos desde nova. A minha mãe que trabalhou bastante para garantir que eu tivesse a melhor educação possível. Eles sempre serão meus amores.

Ao meu noivo, João Pedro Augusto Costa, pelo apoio emocional, mentoria e ajuda nos momentos difíceis.

Aos meus queridos amigos de empresa, que trouxeram momentos leves nas viagens para Carajás-PA e de muito trabalho para a implementação deste projeto.

Ao meu orientador, Omar Carmona Cortês, sem o qual não seria possível a conclusão desse trabalho, e que em nesses momentos de pandemia e isolamento conseguiu fornecer conselhos para a finalização deste trabalho.

# Resumo

Esse trabalho apresenta um mecanismo de busca que possui como objetivo encontrar uma Rede Neural Convolutiva (CNN) eficiente com base na arquitetura de uma rede chamada de UNet. O objetivo principal é segmentar regiões de minério fino em uma imagem para facilitar o processo de produção de minério de ferro em uma usina. As imagens foram coletadas a partir de instrumentação localizada em uma usina de minério de ferro em uma etapa de peneiramento. Foi possível obter um *dataset* composto de 688 imagens e suas segmentações. Os resultados da CNN otimizada mostram que a arquitetura encontrada alcançou um valor de 80% de Intersect Over Union (IoU) em comparação as outras redes analisadas: UNet sem otimização e DeepLabV3+, que alcançaram 75% e 78%, respectivamente. De modo geral, os resultados encontrados foram considerados promissores e a solução da CNN otimizada foi capaz de detectar as regiões de minério fino, com a utilização em um ambiente industrial.

**Palavras-chaves:** Minério de Ferro. Segmentação. Redes Neurais. CNN. Visão Computacional

# Abstract

This work presents a mechanism used to search a Convolutional Neural Network (CNN) based on the architecture of a network called UNet. The main aim is to segment fine iron ore regions in an image to facilitate the iron ore production process in a plant. The images were collected from an iron ore plant in a screening step. It was possible to obtain a *dataset* composed of 688 images and their label segmentation. The results of the optimized architecture shows that the found architecture of CNN achieved value of 80% of Intersect Over Union(IoU) and the other comparison networks: UNet without optimization and DeepLabV3+, which achieved 75% and 78%, respectively. In general, the results found were considered promising and the optimized CNN solution was able to detect the fine ore regions.

**Key-words:** Iron Ore. Segmentation. Neural Network. CNN. Computer Vision.

# Lista de ilustrações

Figura 1 – Fluxograma de Usina de Beneficiamento . . . . .	23
Figura 2 – Peneira Vibratória . . . . .	24
Figura 3 – Peneira . . . . .	25
Figura 4 – Processo de estratificação . . . . .	26
Figura 5 – Minério de Ferro: Finos . . . . .	26
Figura 6 – Minério de Ferro: Superfinos . . . . .	27
Figura 7 – Minério de Ferro: Granulado . . . . .	27
Figura 8 – Minério de Ferro: Pelotas . . . . .	28
Figura 9 – Minério de Ferro - Peneira . . . . .	28
Figura 10 – Convolução tradicional . . . . .	30
Figura 11 – <i>Depth Wise Convolution</i> aplicada em uma imagem de dimensões $12 \times 12 \times 3$ com um kernel de tamanho $5 \times 5$ . . . . .	31
Figura 12 – Representação de uma convolução com Dilatação. (a) Dilatação = 1, (b) Dilatação = 2, (c) Dilatação = 3 . . . . .	32
Figura 13 – Operações de <i>Pooling</i> . . . . .	32
Figura 14 – Arquitetura tradicional da UNet . . . . .	34
Figura 15 – Dimensões da NAS . . . . .	35
Figura 16 – Base de dados do minério de ferro após o pré-processamento . . . . .	41
Figura 17 – Base de dados com Regiões de Finos segmentadas. Segmentação realizada por especialista. . . . .	42
Figura 18 – Etapas da Metodologia . . . . .	43
Figura 19 – Treinamento da <i>UNet</i> . . . . .	48
Figura 20 – <i>Loss - DeepLabV3</i> . . . . .	49
Figura 21 – Célula Otimizada - NASUNet . . . . .	51
Figura 22 – Célula Otimizada - <i>UNet</i> Tradicional . . . . .	52
Figura 23 – Célula Otimizada - <i>Deep Lab V3</i> . . . . .	52
Figura 24 – Célula Otimizada - <i>NASUNet</i> . . . . .	53
Figura 25 – Imagens reais (Coluna da esquerda); Imagem segmentada por modelo neural (Coluna do meio); Imagens segmentadas pelo especialista (Coluna da direita) . . . . .	54
Figura 26 – Imagens de minério de ferro granulado . . . . .	55
Figura 27 – Segmentações ruins . . . . .	55
Figura 28 – Aplicação - usina Carajas . . . . .	64

# Lista de tabelas

Tabela 1 – Comparação de trabalhos relacionados . . . . .	19
Tabela 2 – Configuração do Cromossomo . . . . .	44
Tabela 3 – Parâmetros . . . . .	46
Tabela 4 – Configuração do Cromossomo . . . . .	47
Tabela 5 – <i>NASUNet</i> - Configuração . . . . .	50
Tabela 6 – Resultados . . . . .	52
Tabela 7 – Kruskal . . . . .	53

# Lista de abreviaturas e siglas

CIS	<i>Coarse Image Segmentation</i>
CNN	<i>Convolutional Neural Network</i>
CSV	<i>Comma-separated values</i>
FC-DenseNet	<i>Fully Convolutional DenseNet</i>
FIS	<i>Fine Image Segmentation</i>
GA	<i>Genetic Algorithm</i>
GCN	<i>Graph Convolutional Network</i>
IBRAM	Instituto Brasileiro de Mineração
IoU	<i>Intersect Over Union</i>
mIoU	<i>Mean Intersect Over Union</i>
MRL	<i>Manufacturing Readiness Levels</i>
MSE	<i>Mean Squared Error</i>
NAS	<i>Neural Architecture Search</i>
OS	<i>Over-Segmentation</i>
PIB	Produto Interno Bruto
PSPNet	<i>Pyramid scene parsing network</i>
RNA	Redes Neurais Artificiais
ROI	<i>Region of interest</i>
SA	<i>Segmentation Precision</i>
TRL	<i>Technology Readiness Level</i>
US	<i>Under-Segmentation</i>
VGG19	<i>Very Deep Convolution Neural Network</i>

# Lista de símbolos

\* Convolução

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>14</b>
<b>1.1</b>	<b>Justificativa</b>	<b>15</b>
<b>1.2</b>	<b>Objetivos</b>	<b>16</b>
1.2.1	Objetivo Geral	16
1.2.2	Objetivos Específicos	16
<b>1.3</b>	<b>Trabalhos correlatos</b>	<b>17</b>
1.3.1	Segmentação de Minério de Ferro	17
1.3.2	Otimização e Busca Neural	19
<b>1.4</b>	<b>Estrutura</b>	<b>20</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>22</b>
<b>2.1</b>	<b>Usina</b>	<b>22</b>
2.1.1	Peneiramento	23
2.1.2	Minério de Ferro	26
<b>2.2</b>	<b>Redes Neurais Artificiais</b>	<b>29</b>
<b>2.3</b>	<b>Redes Neurais Convolucionais</b>	<b>29</b>
<b>2.4</b>	<b>A Rede Neural UNet</b>	<b>33</b>
<b>2.5</b>	<b>Busca de Arquitetura Neural</b>	<b>34</b>
<b>2.6</b>	<b>Meta-heurísticas</b>	<b>36</b>
2.6.1	Algoritmos Genéticos	36
<b>2.7</b>	<b>Considerações Finais</b>	<b>37</b>
<b>3</b>	<b>MECANISMO DE BUSCA NEURAL</b>	<b>39</b>
<b>3.1</b>	<b>Aquisição das Imagens</b>	<b>39</b>
3.1.1	Pré-Processamento	40
3.1.2	Geração de Ground Truth	41
<b>3.2</b>	<b>Criação do Mecanismo de Busca Neural</b>	<b>42</b>
<b>3.3</b>	<b>Considerações Finais</b>	<b>44</b>
<b>4</b>	<b>EXPERIMENTOS COMPUTACIONAIS</b>	<b>45</b>
<b>4.1</b>	<b>Hardwares e Softwares aplicados</b>	<b>45</b>
<b>4.2</b>	<b>Configurações</b>	<b>46</b>
<b>4.3</b>	<b>Resultados</b>	<b>47</b>
4.3.1	Teste não paramétrico	53
4.3.2	Discussão	54
4.3.3	Nível de Maturidade Tecnológica	55

<b>5</b>	<b>CONCLUSÃO</b> . . . . .	<b>57</b>
<b>5.1</b>	<b>Trabalhos Futuros</b> . . . . .	<b>58</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>59</b>
	<b>APÊNDICES</b>	<b>63</b>
	<b>APÊNDICE A – SOLUÇÃO IMPLEMENTADA</b> . . . . .	<b>64</b>
	<b>APÊNDICE B – PRODUÇÕES CIENTÍFICAS</b> . . . . .	<b>65</b>

# 1 Introdução

A cadeia de produção do minério de ferro é inicializada pelo processo de extração do minério. O processo é iniciado com a utilização de diversos equipamentos que possuem como objetivo lavar áreas em que há abundância do mineral, por fim, o material é transportado até a usina. A etapa da usina é definida como beneficiamento do minério, que possui como objetivo aumentar o valor do produto e adequá-lo aos padrões das empresas de siderurgia. O beneficiamento é composto por etapas físicas e químicas que dependem da qualidade do minério da lavra adquirida e qual o resultado desejado, como exemplo de etapas desse processo, pode-se destacar a britagem e o peneiramento. A britagem e o peneiramento são etapas fundamentais para redução da granulometria do material através do uso sequencial de britadores e peneiras.

O processo na usina possui várias limitações para aumento de produção e variáveis que precisam ser analisadas pelos times de operação em tempo real. As principais decisões são baseadas na qualidade do material, quantidade de material proveniente da lavra, disponibilidade e manutenção dos equipamentos, além de paradas operacionais, um exemplo é quando a umidade do material é identificada e as taxas de vazão dos equipamentos precisam ser ajustadas. A complexidade para analisar todas as variáveis e a necessidade de padronizar as decisões traz grandes oportunidades de melhoria com uso de técnicas da área de otimização, aprendizado de máquina e automação.

Na etapa do peneiramento, a análise granulométrica do material é uma técnica utilizada para obter melhor compreensão sobre os ajustes que implicam em aumento ou diminuição da produção de cada equipamento. Neste contexto, a granulometria é fundamental, pois ajuda os integrantes da operação a entender se o material está sendo peneirado de forma eficiente, de forma que não tenha a granulometria incorreta em 20% do material.

Analisar o minério em uma peneira pode ser considerado uma atividade humana com alto nível de complexidade, pois não há padronização no resultado e o operador precisa se encontrar fisicamente próximo dos equipamentos para as análises. A área de visão computacional pode padronizar os resultados de granulometria com mais confiabilidade, agilizar as análises com a automatização do processo, e trazer mais segurança operacional. Nesse contexto, a proposta desse trabalho é criar uma arquitetura eficiente de rede neural convolucional para o problema de segmentação de minério de ferro e comparar os resultados com redes neurais de estado da arte para área de visão computacional. A métrica de avaliação das redes neurais é a Intersect Over Union (IoU), pois é uma boa métrica independente da variação de escala das segmentações, sendo que a CNN base para a busca

da arquitetura é a UNet. Além disso, a arquitetura encontrada será comparada com a UNet padrão e com a DeepLabV3+, considerando que a DeepLabV3+ representa o estado da arte em segmentação de imagens (AHMED et al., 2020) (MINGWEI et al., 2020).

## 1.1 Justificativa

A área de redes convolucionais vem apresentando avanços consideráveis desde o ano de 2010, contando com exemplos concretos de utilização em diversas atividades econômicas. Nesse âmbito, a indústria possui um cenário ideal para a utilização e agilização de novas tecnologias. No setor da mineração, em especial a automatização de uma inspeção de minério de ferro com tecnologia de visão computacional, pode garantir ganhos de aumento de produtividade, segurança dos trabalhadores e padronização nos processos.

Com a automatização de um cálculo de granulometria, os times operacionais podem ter essa variável de forma mais rápida e confiável. Por ser útil na otimização de estratégias nos processos da usina, por exemplo, a decisão de aumento ou diminuição do fluxo de minério, essa variável pode trazer ganho de produtividade no processo de beneficiamento. A granulometria pode ser utilizada para modelos de otimização onde visam garantir um equilíbrio entre capacidade de um equipamento e a sua eficiência (SALES, 2020).

No contexto da segurança dos trabalhadores, a automatização do cálculo de granulometria pode trazer mais segurança para o processo industrial de beneficiamento de minério, pois o processo de análise granulométrica em muitas usinas atualmente é realizado por inspeção visual em campo, próxima de um equipamento de peneiramento. O deslocamento e a aproximação de uma pessoa nesses equipamentos industriais pode levar a acidentes de trabalho, como exemplo, a projeção de minério das peneiras.

Por fim, sem um processo automático de cálculo de granulometria, a inspeção visual não garante uma padronização. Essa inspeção baseada na percepção de um operador, pode ser alterada por diferentes percepções de operadores, que podem variar por problemas na percepção em cores, até estresses causados por fadiga física ou fadiga visual. A automatização e a criação de um método para cálculo de granulometria garante uma padronização e simplificação do processo.

O uso de visão computacional com o intuito de identificar a granulometria de minério de ferro em ambientes industriais é um desafio para métodos tradicionais. Esses desafios podem consistir desde problemas na iluminação; a tendência das partículas de apresentar um centro mais iluminado e bordas com maior sombreamento, o que pode causar super segmentação, como no caso do algoritmo watershed; ruídos do plano de fundo da região de interesse e ruídos pelos movimentos das partículas. Além disso, as partículas sobrepostas de minério de ferro são um fator que pode influenciar negativamente a qualidade da segmentação (DUAN et al., 2020).

Nesse cenário, as novas arquiteturas de redes neurais convolucionais garantem um processo mais preciso e com velocidades de classificação em tempo real. Essas arquiteturas são ideais para lidar com os desafios de visão computacional. A busca pela arquitetura ideal requer um esforço por parte de um pesquisador, pois é necessário encontrar os parâmetros e hiper parâmetros de uma rede neural, considerando que existem muitas possibilidades de combinação, o que representa um grande esforço computacional. Por isso, esse trabalho aplica um mecanismo de busca para encontrar uma arquitetura de rede neural convolucional eficiente para o processo de segmentação de imagens de minério de ferro.

## 1.2 Objetivos

Destaca-se nesta seção o objetivo geral e os objetivos específicos a serem alcançados neste trabalho.

### 1.2.1 Objetivo Geral

O objetivo principal deste trabalho é aplicar uma técnica de busca neural para identificar uma arquitetura de rede neural artificial profunda para segmentação de regiões de interesse em imagens de minério de ferro, utilizando como base um conjunto de imagens capturadas em um ambiente industrial.

### 1.2.2 Objetivos Específicos

Para alcançar o objetivo geral desse trabalho, foi necessário realizar os seguintes objetivos específicos:

- Definir um processo de aquisição adequado para as imagens em um ambiente industrial;
- Gerar uma base com segmentações manuais realizadas por um especialista na área de mineração;
- Estudar e implementar técnicas de Busca e Otimização Neural;
- Estudar e implementar técnicas de segmentação de imagens;
- Analisar as vantagens do método proposto para a segmentação de minério fino.

## 1.3 Trabalhos correlatos

Neste capítulo são apresentados trabalhos disponíveis na literatura com foco em segmentação de imagens, segmentação de minério de ferro, métodos para identificar a granulometria de grãos, além de trabalhos relacionados com a busca e otimização de redes neurais. Cada trabalho apresentado nesta seção considerou uma etapa diferente da mineração para a coleta e criação da base de dados, desde etapas de laboratório para o estudo desse minério, em etapas transporte desse material pela usina com correias transportadoras, nos processos iniciais de extração do minério de ferro das minas e em ambientes simulados.

Em relação ao mecanismo para realização da segmentação, foi selecionada a rede *UNet*. Esta arquitetura foi utilizada como *backbone*, rede base para o modelo de rede neural encontrado neste trabalho. Na literatura podem ser encontrado diversos casos de uso da *UNet*, na área médica (SIDDIQUE et al., 2020), segmentação de áreas urbanas em imagens de satélite (MCGLINCHY et al., 2019), detecção de navios (KARKI; KULKARNI, 2021) e segmentação de estradas (HOU et al., 2021).

### 1.3.1 Segmentação de Minério de Ferro

O uso de redes neurais convolucionais aplicadas a problemas de segmentação vem substituindo soluções de processamento de imagens tradicionais. No setor da mineração, o algoritmo de *Watershed* (ROERDINK; MEIJSTER, 2003) é o algoritmo clássico para soluções de separação da granulometria do minério de ferro. Com as novas arquiteturas das redes neurais profundas, a segmentação de imagens com rede convolucionais vem sendo analisada para problemas de segmentação de minério de ferro com resultados superiores aos métodos tradicionais e serão apresentados nesta seção.

A distribuição do minério de ferro é um indicador importante para tomada de decisão em tempo real, essa análise granulométrica pode ser feita em algumas etapas como as correias transportadoras. No artigo de Liu et al. (2020), métodos tradicionais de extração de características de imagem foram usados para pré-processamento das imagens e as redes de segmentação *UNet* e *ResUNet* foram utilizadas somente para identificar o contorno dos grãos de minério. O resultado pode ser comparado com métodos tradicionais usando três métricas: *Segmentation Precision(SA)*, *Over-Segmentation(OS)* e *Under-Segmentation(US)*, nos experimentos é possível identificar que a técnica proposta teve uma média de SA de 0,94 e o *Watershed* de 0,64, além disso, os valores de US do algoritmo proposto alcançaram um valor de 0,10, enquanto o *Watershed* um resultado de 0,34.

Em Svensson (2019), uma base de dados com 180 imagens microscópicas de minério de ferro fino foi experimento para a segmentação de modelos das seguintes redes neurais convolucionais: *PSPNet*, *FC-DenseNet*, *DeepLabV3+* e *GCN*. Os experimentos foram

aplicados para identificar a melhor arquitetura, o tamanho ideal da base de dados de treino e o impacto das estratégias de aumento de dados (HERNANDEZ-GARCIA, 2020). Com as técnicas de aumento da base dados, a acurácia média dos modelos aumentou em 3,14% e a média de *IoU* melhorou em 6,84%.

Chen et al. (2021) propôs um método de segmentação e rastreamento de rochas em vídeos, utilizando duas redes de convolução: Mask R-CNN e Deep Sort. O modelo de Mask R-CNN foi melhorado com a utilização de uma função de *loss* baseada no *ArcFace Loss* para incrementar o reconhecimento de características. Com uma base de dados de 180 imagens, estratégias de aumento de dados foram aplicadas para aumentar a base para 900 imagens. Os experimentos identificaram um resultado de 84,27% de *F1 Score* para o modelo melhorado e 79,99% para o modelo padrão, um aumento de 4,28%.

Com o objetivo de identificar e calcular a granulometria de pelotas de minério de ferro, Duan et al. (2019) aprimorou a rede neural de segmentação *UNet* para um modelo mais leve denominado de *lightweight UNet*, com a proposta de ter menos parâmetros e a introdução de camadas de normalização para reduzir o tempo computacional. Os resultados dos experimentos mostram que o modelo só necessita de 1,5s a 3s para processar uma imagem de 256 x 256, um tempo computacional aceitável para o uso em tempo real. Além disso, através da métrica *DICE* foi observado que a *UNet* proposta teve resultados muitos superiores a *UNet* tradicional (0,8597 x 0,6035).

Já no trabalho de (OLIVIER et al., 2020), com a utilização de uma base de dados com 718 imagens capturadas em um ambiente para simular uma correia transportadora, uma *VGG16* foi implementada com o propósito de segmentar os grãos de minério de ferro. Com a métrica de erro, *MSE*, os experimentos com a base de testes obtiveram uma média de erros de -0,012 e um desvio padrão de 0,107.

Em pilhas de minério, Thurley (2013) utilizou do algoritmo *Watershed* para segmentar fragmentos de minério de ferro. Regiões com minério de ferro muito fino, foram caracterizadas e segmentadas como regiões de finos. A comparação dos resultados foi feita com a separação de *frames* com grãos maiores (grupo B e C) e *frames* com maiores regiões de finos (grupo A). Como resultado, foi possível identificar uma distribuição maior de finos no grupo A.

O trabalho de Ma et al. (2020) utilizou uma CNN para a segmentação de imagens de minério de ferro diverso e de métodos tradicionais para minérios com granulometria maior, em correias transportadoras. A solução proposta possui duas etapas, uma primeira onde é feito a detecção de qual o tipo de material: correia transportadora vazia, material mais diverso ou material com granulometria maior. Dependendo dessa primeira detecção, dois algoritmos podem ser utilizados, *FIS* (*Fine Image Segmentation*) para um material mais diverso, ou *CIS* (*Coarse Image Segmentation*) para uma material com granulometria maior. A comparação dos resultados foi feita com a distribuição da área identificada pelos

algoritmos e por trabalhos manuais, os experimentos identificaram um erro máximo de 5,61% em relação à distribuição de área manual.

No conjunto das redes neurais convolucionais existem diversas arquiteturas de segmentação de imagens com a finalidade semelhante ao trabalho de [Liu et al. \(2020\)](#). Para uma melhor visualização dos modelos apresentados, a Tabela 7 ilustra algumas redes e suas respectivas finalidades:

Tabela 1 – Comparação de trabalhos relacionados

Trabalho	Finalidade	Modelos	Número de Imagens
<a href="#">Svensson (2019)</a>	Segmentação Minério Super Fino	<i>BiSeNet</i> , <i>DeepLabv3+</i> , <i>FC-DenseNet56</i> , <i>GCN</i> e <i>PSPNet</i>	180 Imagens
<a href="#">Liu et al. (2020)</a>	Segmentação de Minério de Ferro	<i>ResNet</i> e <i>UNet</i>	39 Imagens
<a href="#">Chen et al. (2021)</a>	Segmentação e Rastreamento de Rochas em Vídeo	<i>Mask-R CNN</i> e <i>Deep Sort</i>	180 Imagens
<a href="#">Duan et al. (2019)</a>	Segmentação de Pelotas de minério de Ferro	LightWeight UNet	500 Imagens
<a href="#">Olivier et al. (2020)</a>	Segmentação de Minério Ferro	<i>VGG16</i>	781 Imagens
<a href="#">Thurley (2013)</a>	Segmentação de Minério Ferro em pilhas	<i>Watershed</i>	-
<a href="#">Ma et al. (2020)</a>	Segmentação de Minério Ferro	<i>CNN</i>	2880 Imagens

### 1.3.2 Otimização e Busca Neural

Uma das etapas desse trabalho consiste em otimizar uma rede de segmentação para identificar minério fino. Com base nesta premissa, alguns trabalhos de otimização de redes neurais foram analisados de forma a contribuir na construção da proposta. O trabalho de [Weng et al. \(2019\)](#) propõe uma arquitetura de busca com base na arquitetura da rede UNet, buscando a melhor composição de células que formam grafos que representam operações primitivas como convoluções e operações de *pooling*. Como resultado, o modelo de *NASUNet* obteve melhores resultados de *IoU* se comparado com a *UNet* tradicional e o modelo de *FC-DenseNet*.

O trabalho de [Domingos et al. \(2022\)](#) utilizou um algoritmo genético para evoluir redes neurais convolucionais, neste artigo a rede neural *Deep-Emotive* foi otimizada com base em um espaço de busca de 5 parâmetros: Camadas de convolução, Filtros, Tamanho do Filtro, Tipo de Subamostragem e Neurônios da camada densa. Os experimentos foram realizados em duas bases de dados distintas para classificação de emoções: *FER-2013*, no qual a solução otimizada teve 1,68% de aumento de acurácia, e a *Cohn-Kanade AU* com aumento de 0,68%

Nenhuma das propostas apresentadas analisa o material em tempo real em uma etapa específica de peneiramento na usina, onde as condições da imagem são prejudicadas por vibração, alterações na iluminação e o ambiente é inerente a partículas de minério de ferro. Este trabalho se diferencia dos trabalhos apresentados, pois a base de dados foi adquirida em uma etapa diferente, na etapa do peneiramento. Além disso, o propósito deste experimento é segmentar regiões do minério de ferro com uma granulometria menor, também denominado de minério fino ou *Sinter Feed*.

## 1.4 Estrutura

O Capítulo 2 apresenta uma introdução sobre a etapa da cadeia de produção do minério de ferro que é o foco deste trabalho: a usina. Esse capítulo apresenta um resumo para melhor entendimento do trabalho, dos principais processos para beneficiamento do minério e conceitos básicos sobre peneiramento. Adicionalmente, há uma subseção de resumo sobre as características do minério de ferro, assim como os tipos de minérios que serão analisados nas imagens. Além disso, o capítulo apresenta uma visão geral sobre a fundamentação teórica dos métodos de aprendizado de máquina e otimização. Com ênfase em redes neurais convolucionais, assim como definições básicas para o entendimento da técnica e como esta pode ser utilizada em problemas de segmentação.

O Capítulo 3 apresenta os materiais e métodos utilizados para obter os resultados dos experimentos. Esse capítulo detalha o processo de aquisição e pré-processamento da base de dados para treinamento dos modelos que serão testados. Por fim, apresenta o detalhamento do processo de otimização empregado para a busca de um arquitetura eficiente para a segmentação das imagens de minério de ferro.

O Capítulo 4 detalha as métricas utilizadas para comparação dos resultados e os parâmetros selecionados para os modelos de segmentação e a meta-heurística de otimização. Adicionalmente são discutidos os resultados encontrados após o treinamento e teste dos três modelos: UNet tradicional, DeepLabV3 e NASUNet. Neste capítulo são apresentados alguns estudo de casos do modelo na base de teste.

O Capítulo 5 apresenta as conclusões que podem ser inferidas dos resultados obtidos, as produções científicas publicadas através dessa pesquisa, assim como possibilidades de

trabalhos futuros.

## 2 Fundamentação Teórica

Este capítulo apresenta a fundamentação teórica utilizada neste trabalho, a qual é importante para a compreensão do contexto e propósito da segmentação de minério de ferro, assim como o entendimento das técnicas computacionais adotadas para o alcance dos objetivos apresentados. A seguir são apresentados o entendimento sobre o ambiente industrial de tratamento do minério e informações sobre os tipos de minério de ferro. Apresentam-se os conceitos sobre redes neurais, técnicas de pesquisa neural e as métricas de avaliação dos resultados.

### 2.1 Usina

A usina é o local onde acontece a etapa de beneficiamento do minério de ferro. Nessa etapa, um conjunto de operações e equipamentos separam o material de alto valor econômico, do material que não possui utilidade, também chamado de ganga. Além de eliminar as impurezas, o beneficiamento possui etapas para garantir que o grão do minério esteja no tamanho ideal para o produto desejado. Cada processo de beneficiamento é específico para cada depósito do mineral, uma vez que possuem seu composto mineralógico diferente. Em algumas reservas, o mineral pode não ser encontrado em seu estado natural com a presença de mais um mineral, para isso, a configuração do tratamento é essencial.

Apesar das diferenças entre os beneficiamentos, há alguns padrões que são seguidos na configuração de uma usina. A etapa inicial do beneficiamento é o processo de caracterização do material, no qual estudos são feitos em laboratório para identificar os elementos que constituem as rochas das lavras de minério. Informações de textura, granulometria e densidade são importantes para entender como será o comportamento do minério na usina.

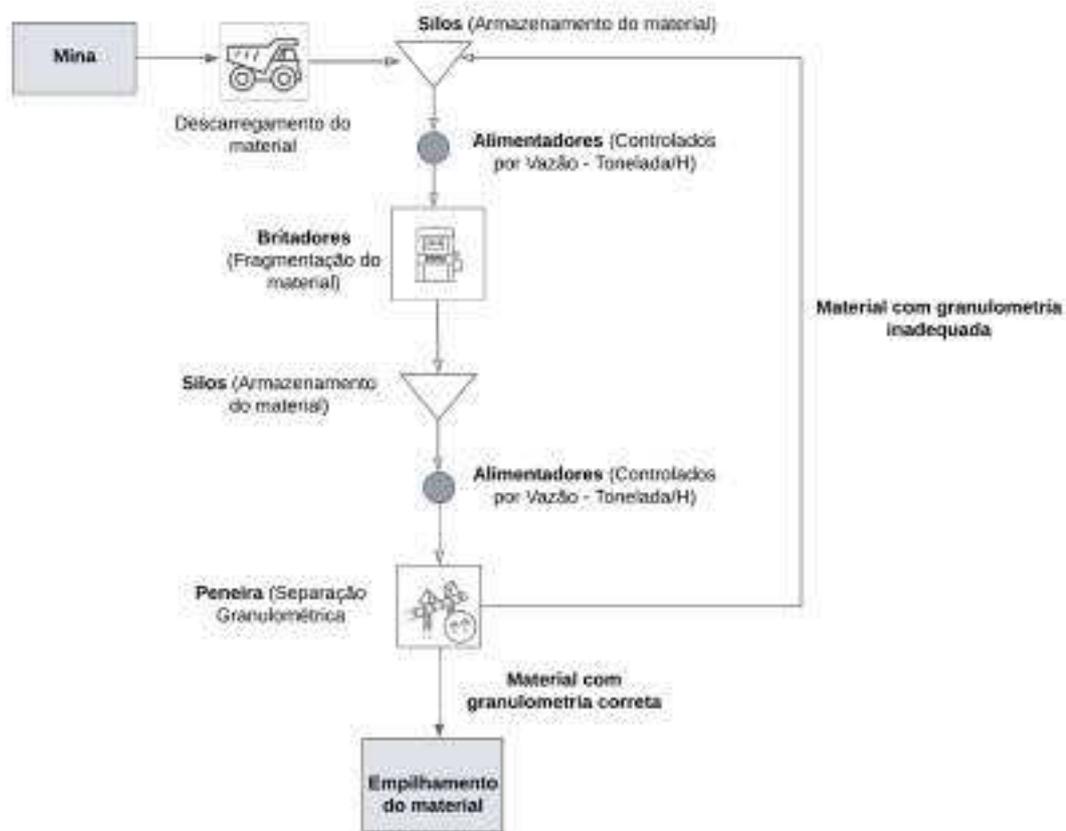
A composição das operações da usina é construída de acordo com as características do material a ser tratado para garantir o processamento adequado e o maior aproveitamento do mineral, garantindo um valor comercial.

Assim há duas operações básicas que serão descritas em detalhes: fragmentação e separação, que podem ser feitas por etapas de peneiramento e britagem, sendo que essas duas operações serão contextualizadas nas subseções seguintes.

O Fluxograma 1 é um exemplo de uma planta de usina de beneficiamento de minério de ferro, na qual é possível identificar 4 elementos principais no processo de Fragmentação e separação granulométrica: Silos, Alimentadores, Britadores e a Peneira. De acordo com algumas variáveis, como a qualidade do material no processo, os alimentadores (que graficamente podem ser visto antes dos equipamentos) são alterados para diminuir

ou aumentar a vazão de minério de ferro que alimenta cada equipamento. O material com granulometria inadequada retorna para o processo de fragmentação com os britadores, pois se a quantidade desse material for superior a 20% de minério fino, os britadores perdem sua eficiência e necessitam de mais manutenções, por consequência a usina perde sua eficiência. Esse é o fator principal que torna a análise granulométrica essencial na etapa do peneiramento.

Figura 1 – Fluxograma de Usina de Beneficiamento



Fonte: Acervo do autor

### 2.1.1 Peneiramento

Existem várias formas para a separação granulométrica de minério, desde separações físicas até separações químicas. O peneiramento a seco é um dos modelos mais utilizados para a separação, visto que possui ganhos ambientais e de eficiência, esse processo consiste

em separar o material em 2 ou mais grupos, sendo a peneira uns dos equipamentos escolhidos para a separação granulométrica. A peneira vibratória consiste em um equipamento com um ou mais telas com superfícies abertas para a queda de acordo com o tamanho do material, sendo que a vibração do equipamento permite aumentar o contato do material com a tela. A Figura 2 é um exemplo de uma peneira vibratória.

Figura 2 – Peneira Vibratória



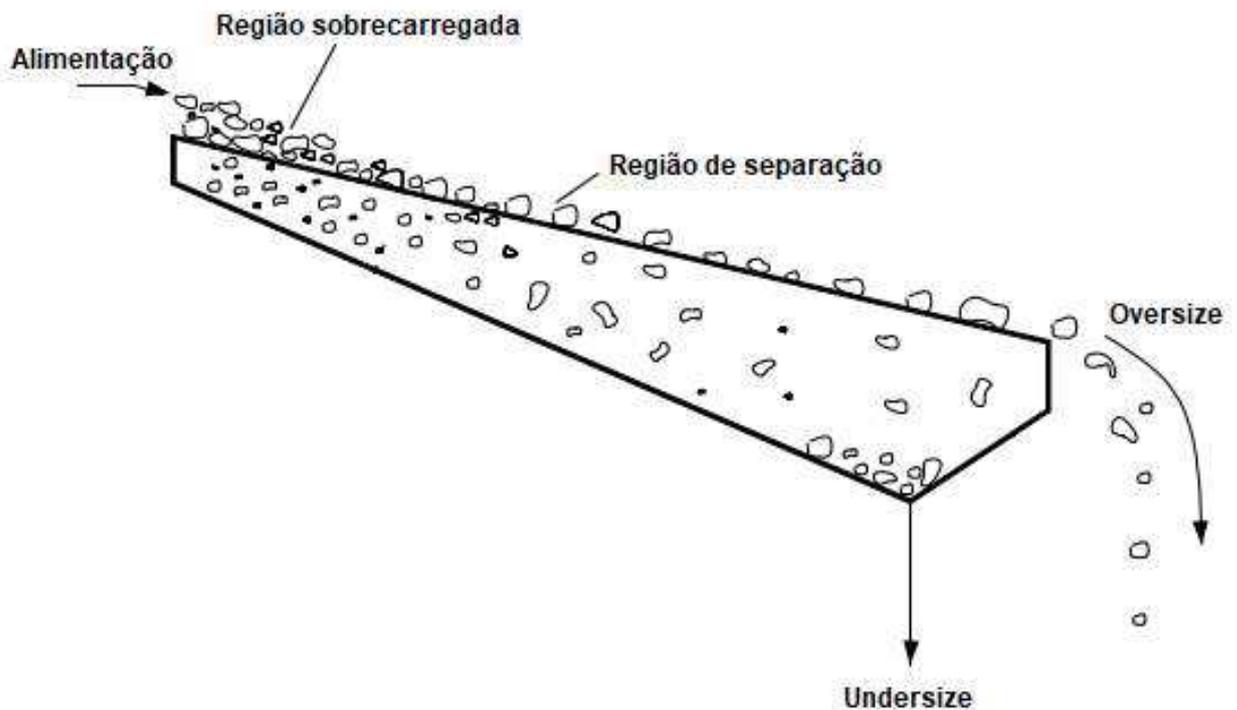
Fonte: (VALE, 2013)

Os produtos gerados pelo peneiramento são chamados de *Oversize* (Material com granulometria maior) e *Undersize* (Material com granulometria menor) como mostrado na Figura 3. A alimentação das peneiras é controlada de acordo com a eficiência do peneiramento e de variáveis no processo da usina (quantidade de material disponível, qualidade do material e disponibilidade dos equipamentos)

A eficiência pode ser definida como a quantidade de partículas mais finas que passam pela abertura da tela do peneiramento, pela quantidade de alimentação total da peneira e a quantidade de finos presente na alimentação; conforme a Equação 2.1, na qual  $E$  é a eficiência;  $P$  é o passante de fino em tonelada/hora (t/h),  $\alpha$  é a quantidade de fino na alimentação também em tonelada/hora(t/h) e  $A$  é a alimentação em tonelada/hora (t/h).

$$E = \frac{P}{(aA)} X 100 \quad (2.1)$$

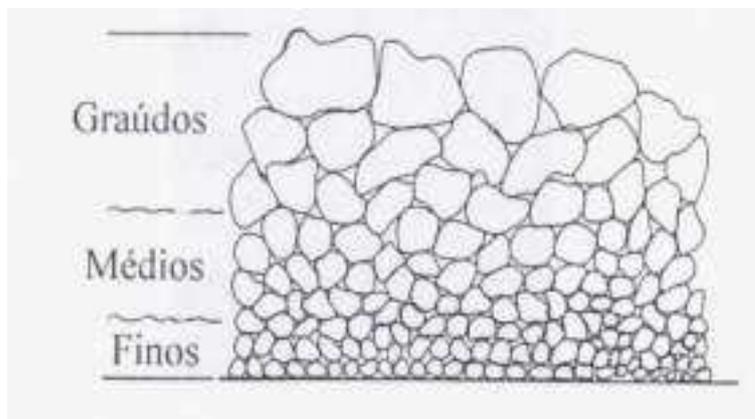
Figura 3 – Peneira



Fonte: (VALADAO; CORREA, 2007)

No peneiramento há dois conceitos importantes que caracterizam o processo, a estratificação, causada pelo movimento de vibração das peneiras, no qual as partículas maiores se agrupam nas camadas superiores da peneira e as partículas menores se localizam mais próximo da tela de separação, conforme a Figura 4; e a probabilidade de separação, que se caracteriza pela probabilidade das partículas finas passarem pelas aberturas da tela da peneira.

Figura 4 – Processo de estratificação



Fonte: (CORREIA, 2010)

### 2.1.2 Minério de Ferro

O minério de ferro é encontrado na natureza no formato de rocha e misturado a outros elementos. Esse minério é um dos mais abundantes da crosta terrestre, no qual representa cerca de 5% da massa total. O minério brasileiro apresenta variações, com condições diferentes de constituintes mineralógicos, tamanho e morfologia dos poros, porosidade, além de forma e superfície das partículas. Conforme essas características são classificadas nas lavras de minério, os processos da usina são definidos para atender os requisitos do produto final para as siderurgias. Os produtos de minério podem ser distribuídos em não aglomerados e aglomerados. A seguir são mostrados alguns exemplos de produtos:

- Finos (*Sinter Feed*): são partículas com tamanho entre 0,15 mm e 6,3 mm. Esse produto é utilizado em processo de aglomeração (sinterização ou pelletização) e pode ser visto na Figura 5;

Figura 5 – Minério de Ferro: Finos



Fonte: (VALADAO; CORREA, 2007)

- Superfinos (*Pellet Feed*): são partículas menores que 0,15mm. Esse produto é utilizado para a aglomeração por pelotização, sendo visto na Figura 6;

Figura 6 – Minério de Ferro: Superfinos



Fonte: (VALADAO; CORREA, 2007)

- Granulado (*Lump Ore*): são partículas de tamanho entre 31,7 mm e 6,3mm de forma irregular. Esse produto é obtido somente com processo de regulação de granulometria, podendo ser visto na Figura 7;

Figura 7 – Minério de Ferro: Granulado



Fonte: (VALADAO; CORREA, 2007)

- Pelotas (Pellets): minério de ferro de forma esférica, no qual a granulometria possui o diâmetro aproximado de 12mm, podendo ser visto na Figura 8;

Figura 8 – Minério de Ferro: Pelotas



Fonte: (VALADAO; CORREA, 2007)

A Figura 9 apresenta imagens das amostras de minério que possuem partículas de superfino, fino e granulado. Essas imagens serão utilizadas para treinamento e teste das técnicas de visão computacional desenvolvidas.

Figura 9 – Minério de Ferro - Peneira



Fonte: Acervo do autor

Nesta seção foram apresentados os conceitos fundamentais de uma usina, explicando o cenário principal da produção de minério de ferro através da apresentação da estrutura de uma usina, dos conceitos de peneiramento e das variações de granulometria mineral. A próxima seção apresenta os conceitos de Redes Neurais Artificiais, técnica utilizada no desenvolvimento deste trabalho para segmentação das regiões de finos em imagens de minério de ferro.

## 2.2 Redes Neurais Artificiais

Uma Rede Neural Artificial (RNA) pode ser descrita como um grafo computacional composto por unidades elementares de grande complexidade, e é obtida conectando-as de maneiras específicas (AGGARWAL, 2018). A unidade básica de uma RNA de processamento é denominada de neurônio artificial, no qual, para um determinado conjunto de entradas é fornecido uma saída, conforme a expressão da Equação 2.2:

$$f(x) = \sigma \left( \sum_{i=1}^N x_i w_i + b \right) \quad (2.2)$$

na qual  $x_i$  é um vetor de entrada e  $w_i$  é um vetor de pesos associado a entrada  $i$ ,  $b$  é o termo *bias* e  $\sigma$  é a função de ativação. Neste modelo, ocorre uma iteração entre os pesos das arestas e posteriormente a aplicação de uma função de ativação, cujo valor de *threshold* define se o neurônio será ativado.

As diversas possibilidades de combinações das unidades simples, torna possível a construção de um modelo capaz de representar funções mais complexas. Redes Neurais com mais camadas ocultas são chamadas de Redes Neurais Profundas. As arquiteturas de redes neurais profundas melhoram a representação de um problema e podem aprender qualquer função matemática se dispuserem dos dados necessários, assim como podem ser mais sucessíveis ao *overfitting* com arquiteturas mais complexas. Atualmente, o treinamento desses modelos se tornou viável, considerando o aumento dos recursos computacionais e da disponibilidade desses recursos.

Redes Neurais Profundas em geral, são utilizadas em problemas que dispõem de uma grande quantidade de dados para serem analisados. Dados não estruturados, como imagens, sons e outros tipos de sinais, necessitam de uma análise mais complexa. Por concepção, as redes neurais convolucionais são projetadas para funcionar com entradas em formato de matrizes, que geralmente representam imagens, que por suas características possuem fortes dependências espaciais em regiões locais (BAZGIR et al., 2020).

## 2.3 Redes Neurais Convolucionais

O tipo de Rede Neural Profunda mais utilizada para classificação, segmentação ou instanciação de imagens é chamada de rede neurais convolucional (CNN) (RAWAT et al., 2017). Essas redes são projetadas para funcionar com entradas estruturadas em grade, no qual há uma forte dependência espacial entre regiões locais, como uma imagem, em que os pixels mais próximos representam alguma forma específica.

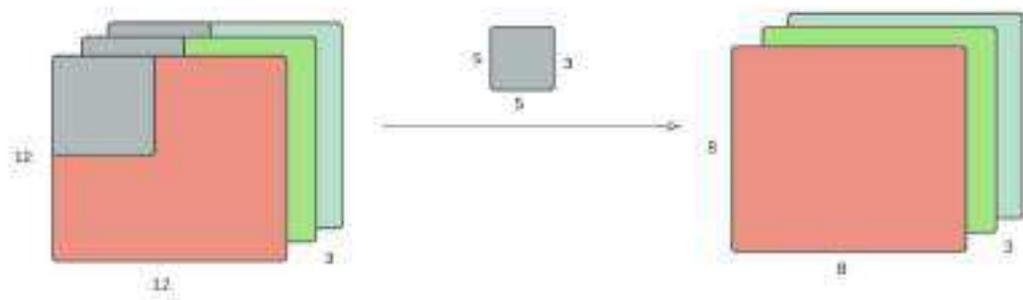
As CNNs usam as primeiras camadas para identificar formas primitivas, como linhas retas e curvas. Em camadas mais profundas, formas mais complexas podem ser

identificadas. Este tipo de estrutura de entrada de rede espera uma entrada em três dimensões ( $H \times W \times P$ ), representando a altura ( $H$ ), largura ( $W$ ) e profundidade ( $P$ ) das imagens que estão sendo inseridas. A estrutura de uma CNN pode ser concebida pelo enfileiramento de diferentes camadas: **Camada de Convolução**, **Camada de Pooling** e **Camada Totalmente Conectada**:

- **Camada de Convolução:** A operação de convolução é responsável por extrair características dos dados de entrada, esse processo de extração consiste na aplicação de filtros que deslizam em todo o conjunto de informações. O treinamento da rede é responsável por identificar os pesos desses filtros. A proposta das camadas de convolução é de diminuir a quantidade de parâmetros a serem aprendidos. Em uma rede neural tradicional, um conjunto de dados com dimensões de  $32 \times 32 \times 3$ , necessitaria de  $32 \times 32 \times 3$  parâmetros para cada neurônio. Em resumo, somente com dois neurônios seriam necessários 6.144 parâmetros. A proposta da CNN é tornar o aprendizado de imagens viável e mais eficiente (ALBAWI et al., 2017).

Cada filtro atua em toda a profundidade do mapa de características, ou seja para uma imagem de 3 canais de  $32 \times 32 \times 3$ , um filtro de  $5 \times 5$  precisa ter dimensão  $5 \times 5 \times 3$ . Geralmente, este kernel possui valores parametrizados de acordo com o tamanho da entrada, um exemplo são os valores de  $3 \times 3$  ou  $16 \times 16$ . Ao passar pelas camadas de convolução, conforme mostrado na Figura 10, o produto entre os valores de entrada e cada kernel gera uma sobreposição que permite a conectividade entre as camadas.

Figura 10 – Convolução tradicional



Fonte: Acervo do autor

Uma CNN possui opções de hiperparâmetros que permitem providenciar várias oportunidades para diminuir a quantidade de parâmetros aprendíveis. Esses hiperparâmetros definem as operações de **Padding**, quantidade de pixels adicionados para a extremidade da imagem; **Stride**, quantidade de pixels que serão passados em cada operação do filtro

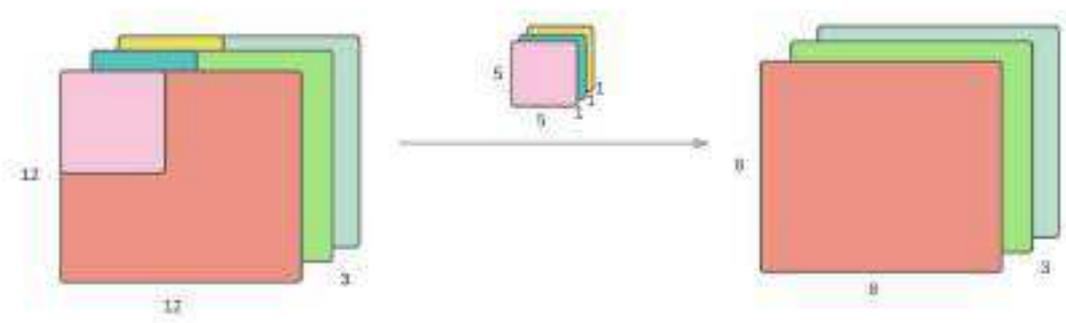
de convolução; **Profundidade de saída**, que representa a quantidade de filtros e por resultado, o tamanho da profundidade de saída. A dimensão de saída de uma imagem quadrada pode ser representada pela Equação 2.3, em que  $F$  representa o tamanho do Filtro,  $S$  é o *stride* selecionado e  $P$ , o valor de *padding*.

$$D = \frac{W - F + 2P}{S} + 1 \quad (2.3)$$

Além do método tradicional de convolução, é importante descrever dois outros métodos de convolução avaliados nesse trabalho, *depth wise convolution* (CHOLLET, 2017) e convoluções com dilatação.

A *depth wise convolution* lida com a dimensão de profundidade (número de canais de uma imagem). Cada canal de imagem possui uma interpretação diferente que pode ser analisada de forma isolada, o processo de cálculo consiste em duas etapas: uma etapa de multiplicação de cada filtro de convolução por cada canal de imagem e uma etapa chamada de *pointwise convolution* que possui filtros com profundidade igual a um, com a finalidade de aumentar a profundidade da imagem depois do processo de *depth wise*. A Figura 11 é uma representação da convolução de *depth wise*. Esse tipo de convolução auxilia na diminuição do *overfitting* do modelo, uma vez que há menos parâmetros para serem aprendidos.

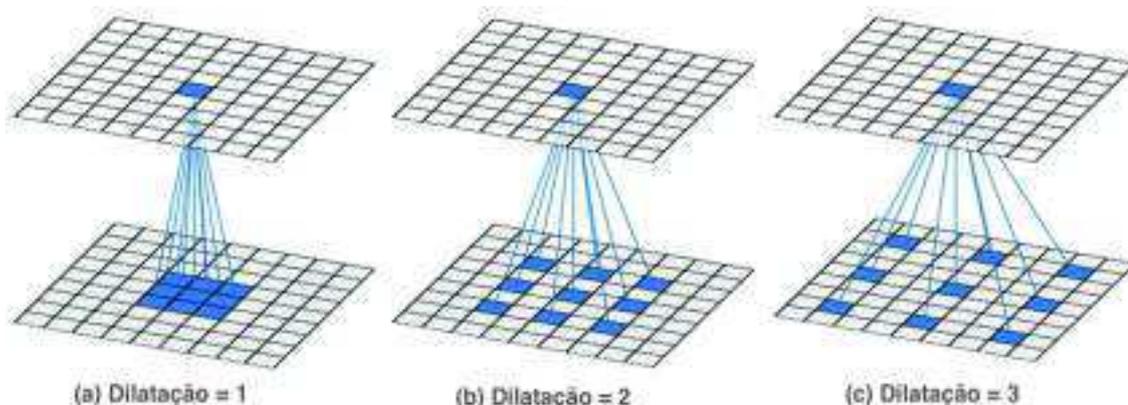
Figura 11 – *Depth Wise Convolution* aplicada em uma imagem de dimensões  $12 \times 12 \times 3$  com um kernel de tamanho  $5 \times 5$



Fonte: Acervo do autor

A dilatação na convolução é uma técnica para a expansão do kernel com a inserção de valores nulos, o objetivo é expandir a cobertura de informação no processo de convolução, a Figura 12 demonstra a técnica de dilatação.

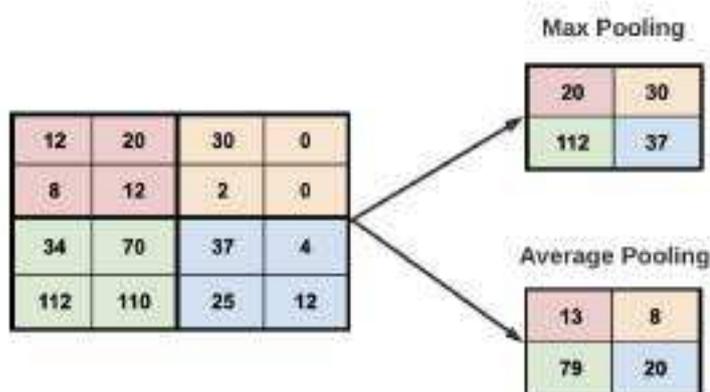
Figura 12 – Representação de uma convolução com Dilatação. (a) Dilatação = 1, (b) Dilatação = 2, (c) Dilatação = 3



Fonte: (CUI et al., 2019)

• **Pooling**: esta camada é responsável pela redução de dimensionalidade do vetor de características, esse tipo de operação sumariza as principais informações, transformando o vetor de entrada em um vetor de características condensadas. Geralmente, esse tipo de operação é posterior ao processo de convolução, sendo as vantagens principais a redução de parâmetros ao longo do modelo e a diminuição do *overfitting*. Um exemplo de redução de dimensionalidade é a entrada de um mapa com uma dimensão de  $24 \times 24$ , que pode ser resumido, após o agrupamento, como  $12 \times 12$ , sem perder informações relevantes. As técnicas principais de *pooling* podem ser o **average pooling**, que consiste em calcular a média de um grupo de pixels e o **max pooling**, que identifica o pixel com maior valor dentro de um grupo, conforme exemplo na Figura 13.

Figura 13 – Operações de Pooling



Fonte: Acervo do autor

- **Camada Totalmente Conectada:** esta é a camada final de uma rede convolucional. A finalidade é conectar os neurônios das camadas anteriores com a camada de saída, e criar o processo de classificação com base na extração de características das camadas de convolução e *pooling*.

## 2.4 A Rede Neural *UNet*

Como previamente mencionado, as CNNs são adequadas para trabalhar com imagens, sendo que a parte totalmente conectada funciona como classificador ou regressor. Porém, as vezes é necessário segmentar partes de interesse na imagem. No processo de segmentação, são necessárias informações sobre a localização e a classificação de cada pixel, e para esse objetivo foram introduzidas as redes neurais totalmente convolucionais (*FCN - Fully Connected Networks*).

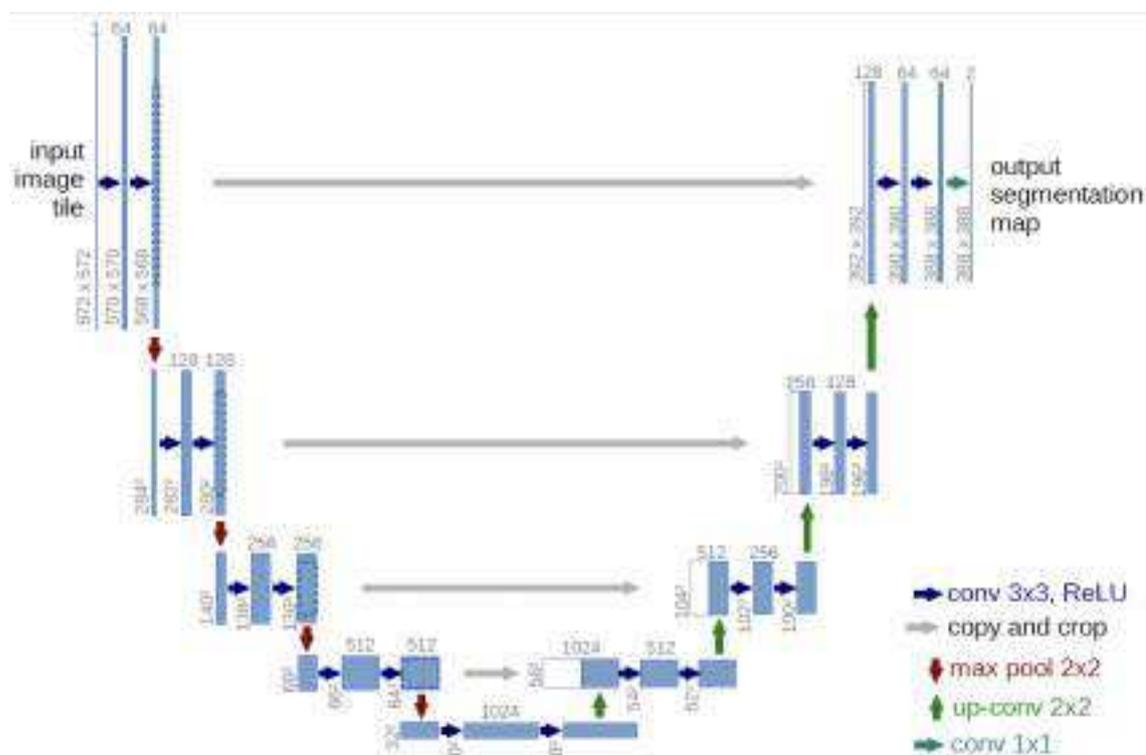
A rede *UNet* é uma FCN proposta por Long et al. (2015) com o objetivo de produzir imagens médicas segmentadas. Embora algumas pesquisas que utilizam essa rede de segmentação estejam associadas à área médica, alguns artigos mostram que a capacidade dessa rede neural é vasta, como mencionado na Seção 1.3, há uma diversidade de possibilidades de utilização da *UNet*. Portanto, esse modelo pode ser utilizado para qualquer tipo de tarefa de segmentação em que o local desejado e as classes previstas precisem ser associadas aos pixels de cada imagem.

A rede é composta por duas partes: um caminho de contração (*downsampling*) e um caminho de expansão (*upsampling*). A ideia é que no lugar de operadores de *pooling*, existam operadores de *upsampling* em caminhos expansivos, além disso, há uma conexão direta entre o caminho de *downsampling* e o *upsampling*, essa conexão permite o acesso de informações com resoluções iguais que facilitam a localização dos pixels. Existem muitas técnicas para aumentar a resolução da imagem, como a interpolação cúbica e a convolução transposta (DUMOULIN; VISIN, 2016). Na segmentação da *UNet*, a convolução transposta é usada em caminhos expansivos.

A Figura 14 apresenta a imagem do artigo original que representa uma estrutura em forma de “U”. Nessa arquitetura, pode-se observar que cada processo é composto por duas camadas de convolução, que aumentam a profundidade do mapa de características. Essas operações de convolução são distribuídas junto com camadas de *pooling*, que diminuem o tamanho da imagem. No caminho para *upsampling*, existem sucessivos processos de

convoluções intercambiadas com técnicas de convolução transposta.

Figura 14 – Arquitetura tradicional da UNet



Fonte: (LONG et al., 2015)

## 2.5 Busca de Arquitetura Neural

O design de modelos de redes neurais profundas é crucial para a performance final de um modelo. Métodos de pesquisas, também chamados de *Neural Architecture Search (NAS)* vem sendo utilizados para de forma automática, construir redes neurais eficientes com recursos computacionais limitados e pouca intervenção de um pesquisador, esse conceito pode ser considerado um subnível das técnicas de AutoML (HE et al., 2019) ou até mesmo ser definido como um macroprocesso de melhoria de hiper parâmetros. De acordo com Elsken et al. (2019) esses métodos de pesquisa podem ser definidos em três dimensões: Espaço de busca, Estratégia de busca e Estratégia de avaliação, conforme o fluxo da Figura 15.

Figura 15 – Dimensões da NAS



Fonte: Acervo do autor

O Espaço de busca refere-se em como a arquitetura pode ser representada. De forma específica, esse espaço pode ser determinado por operações pré-definidas, *templates* de arquiteturas, métodos de conexão das redes, número de camadas de convolução, etc. A qualidade desse espaço é importante para o desempenho computacional, pois quanto maior o domínio, maior o custo para encontrar o melhor resultado. Há algumas abordagens de espaço para *NAS*, como as seguintes:

- Espaço de Busca Global: Esse espaço de busca abrange todos os componentes necessários de uma rede neural, isso significa que a estratégia de busca necessita otimizar um grande espaço;
- Espaço de Busca Modular: Diferente do espaço de busca global, o espaço de busca modular considera a arquitetura de rede neural como módulos diferentes, podendo o projetista definir uma solução no espaço de busca como a configuração de uma camada da rede. Dessa forma, o espaço de busca é reduzido;
- Espaço de Busca Discreto: Somente operações básicas são incluídas para serem modificadas na estrutura da rede neural, como exemplo, pode-se citar a busca de parâmetros como o *kernel* das convoluções.

A Estratégia de busca é o passo que define como o espaço de soluções será explorado. A técnica escolhida deve selecionar uma metodologia que seja capaz de identificar bons candidatos de arquiteturas e, ao mesmo tempo, ser eficiente para convergir mais rapidamente. Há diversas estratégias de buscas, algoritmos evolutivos, aprendizagem por reforço, estratégias baseadas no gradiente, *random search* e *grid search*.

O objetivo da busca por uma arquitetura é obter um modelo que apresente performance satisfatória ao avaliar dados novos. A estratégia de performance refere-se ao processo de avaliação das arquiteturas geradas pela estratégia de busca. Um exemplo simples dessa avaliação pode ser o resultado da métrica aplicada aos dados de validação no processo de treinamento do modelo, entretanto, o alto custo computacional para essas avaliações fizeram outras estratégias de performance surgirem.

Segundo [Elsken et al. \(2019\)](#), a performance pode ser estimada utilizando medidas com 'fidelidades mais baixas', que incluem treinamentos menores, treinamentos em uma amostra dos dados, imagens com resolução menor, ou até menos filtros e menos células na arquitetura. Outra forma de avaliação é a extrapolação da curva ([DOMHAN et al., 2014](#)), na qual o aprendizado inicial é 'extrapolado' para identificar se o processo de parametrização irá gerar um modelo com desempenho insatisfatório já nas primeiras épocas, essa extrapolação permite descartar mais rapidamente uma execução ruim. Outra abordagem inclui um processo de *transfer learning*, em que novas arquiteturas inicializam seus pesos com bases em arquiteturas que já foram treinadas.

## 2.6 Meta-heurísticas

Uma heurística é uma estratégia utilizada com o objetivo de encontrar soluções para um problema em situações onde é admissível identificar soluções não ótimas, mas em que é necessário garantir que serão encontradas soluções possíveis e adequadas. Essa estratégia abrange algoritmos que não são dependentes de um problema específico e que guiam a construção de algoritmos de otimização ([TALBI, 2020](#)), sendo utilizados em cenários onde é muito custoso computacionalmente testar todas as possibilidades de soluções.

Um problema de otimização pode ser definido pela Equação 2.4, onde se busca por uma solução ótima  $x$  dentro de um espaço de soluções  $X$ , maximizando ou minimizando uma função objetivo  $f(x)$ , enquanto há um conjunto de restrições  $(r_1, r_2, \dots, r_n)$ .

$$f(x) \in X | \{r_1, r_2, \dots, r_n\} \quad (2.4)$$

A seguir é apresentada a meta-heurística Algoritmo Genético (AG), utilizada neste trabalho para otimizar a arquitetura neural para identificação de minério de ferro fino.

### 2.6.1 Algoritmos Genéticos

Algoritmo Genético (AG) é um modelo matemático baseado no mecanismo de seleção natural, em que existe o princípio da seleção dos mais aptos durante o processo de reprodução, apresentado no trabalho de [JOHN \(1992\)](#). Por definição, um AG consiste em alguns elementos: população de cromossomos, tipo de seleção, fitness, processo de crossover e o processo de mutação de genes.

A população de cromossomos representa um conjunto de possíveis soluções para o problema. A representação padrão dos indivíduos consiste em um vetor de  $n$  valores binários, no qual o valor 0 poderia indicar a não inclusão de uma variável, e o valor 1 a inclusão na solução candidata. As representações com valores reais são adequadas quando

a representação binária correspondente necessita de muitos bits para caracterizar um indivíduo ou quando o problema a ser selecionado exige muitas variáveis reais.

A seleção de novos indivíduos para uma nova população se dá pela verificação da função *fitness* de cada cromossomo, esta função indica o quão apto um indivíduo candidato está para resolver o problema em questão. O método de seleção, geralmente o método da roleta ou torneio é utilizado para encontrar candidatos, em que há uma probabilidade maior de selecionar candidatos baseado nos melhores valores de *fitness*.

Os operadores de *crossover* e mutação são aplicados em uma população para gerar novos indivíduos. O *crossover* é responsável pela "geração de filhos", em que cromossomos são combinados para gerar indivíduos com características semelhantes aos pais. A mutação é aplicada sobre os cromossomos na tentativa de aumentar a variabilidade da população e possivelmente sair de ótimos locais.

A geração inicial de uma população pode ser realizada de diversas formas, a mais comum é a criação de cromossomos a partir de uma distribuição aleatória. O algoritmo 1 mostra a estrutura de um Algoritmo Genético canônico.

---

**Algoritmo 1** Estrutura Básica AG

---

```
1: iteracoes ← 0
2: inicializar();
3: avaliar();
4: enquanto não critério_parada() faça
5:   selecionar();
6:   recombinar();
7:   mutar();
8:   avaliar();
9:   iteracoes ← iteracoes + 1;
10: fim enquanto
```

---

Como critérios de parada do algoritmo genético, são utilizados decisões que levam em consideração a aptidão dos indivíduos, limitação de gerações, erro mínimo de acordo com um valor especificado pelo projetista e tempo de execução. Enquanto o critério de parada não for satisfeito, a avaliação e os operadores genéticos são aplicados na próxima geração.

## 2.7 Considerações Finais

Esse capítulo apresentou os conceitos para o entendimento do modelo utilizado no processo de segmentação do minério de ferro, dando ênfase em redes neurais convolucionais. Além disso, apresenta o processo de busca de uma arquitetura neural, definindo tipos de

estratégias e espaços de busca mais comuns na literatura. Por fim, foi descrita a abordagem de busca utilizada: o Algoritmo Genético.

O próximo capítulo mostrou a metodologia do trabalho, definindo em detalhes todos os passos necessários para a utilização da arquitetura identificada pelo processo de busca de arquitetura no problema de segmentação de minério fino.

## 3 Mecanismo de Busca Neural

Neste capítulo, apresenta-se a metodologia adotada na busca de uma rede neural convolucional ótima. De início, é apresentado o processo de aquisição de imagens para os experimentos e o pré-processamento para a geração da base de dados. Em seguida, é descrito todo o processo de identificação de arquitetura ideal para a segmentação de finos em imagens.

### 3.1 Aquisição das Imagens

Uma base de dados foi utilizada neste estudo. Esta base foi adquirida em uma usina em Carajás - PA, Brasil. A aquisição das imagens ocorreu em um período de 1 ano, seguindo as diferentes mudanças no comportamento do minério por causa do período chuvoso da região. Por causa do ambiente de vibração constante, várias câmeras foram testadas para ter o modelo ideal da captura dos *frames* de imagens, assim como a posição ideal para instalação do equipamento, uma vez que não seria possível instalar próximo da peneira em constante vibração. Além disso, com o material em constante deslocamento, foi preciso otimizar a velocidade do obturador da câmera para evitar a trepidação da imagem. Por isso, os seguintes protocolos foram estabelecidos:

- Câmera com Zoom de  $20x$ ;
- Imagens devem ser limitadas para a região do minério de ferro;
- Imagens devem ser capturadas da região final do equipamento de peneiramento;
- Iluminação na lateral do equipamento de peneiramento;
- Distância de 5 metros entre a câmera e o equipamento de peneiramento;
- Instalação da câmera em ângulo frontal ao equipamento de peneiramento;
- Resolução de  $1280 \times 720$ ;
- Alta velocidade no obturador da câmera  $[1/45500, 2s]$ .

De forma a gerar imagens para os experimentos, vários *frames* foram selecionados para a realização de um pré-processamento e criação da base de dados final.

### 3.1.1 Pré-Processamento

A etapa de pré-processamento foi aplicada com o objetivo de melhorar os resultados a serem obtidos no processo de segmentação, reduzir a quantidade de ruídos e definir qual o melhor formato de entrada para os modelos.

As imagens capturadas pela câmera instalada foram recortadas para identificar somente a região final do equipamento de peneiramento, o que gerou uma imagem de resolução de 1024 x 256. De forma a gerar uma imagem quadrada para os modelos, cada frame capturado pela imagem foi recortado e gerou 4 *regions of interest* - ROI de 256 x 256. Depois do processo de recorte, as imagens são convertidas para a escala de cinza, com o intuito de gerar um mapa de características menor.

De todas as imagens resultantes, um processo de *data augmentation* foi aplicado para garantir uma maior diversidade dos dados. Duas técnicas de aumento de dados foram utilizadas:

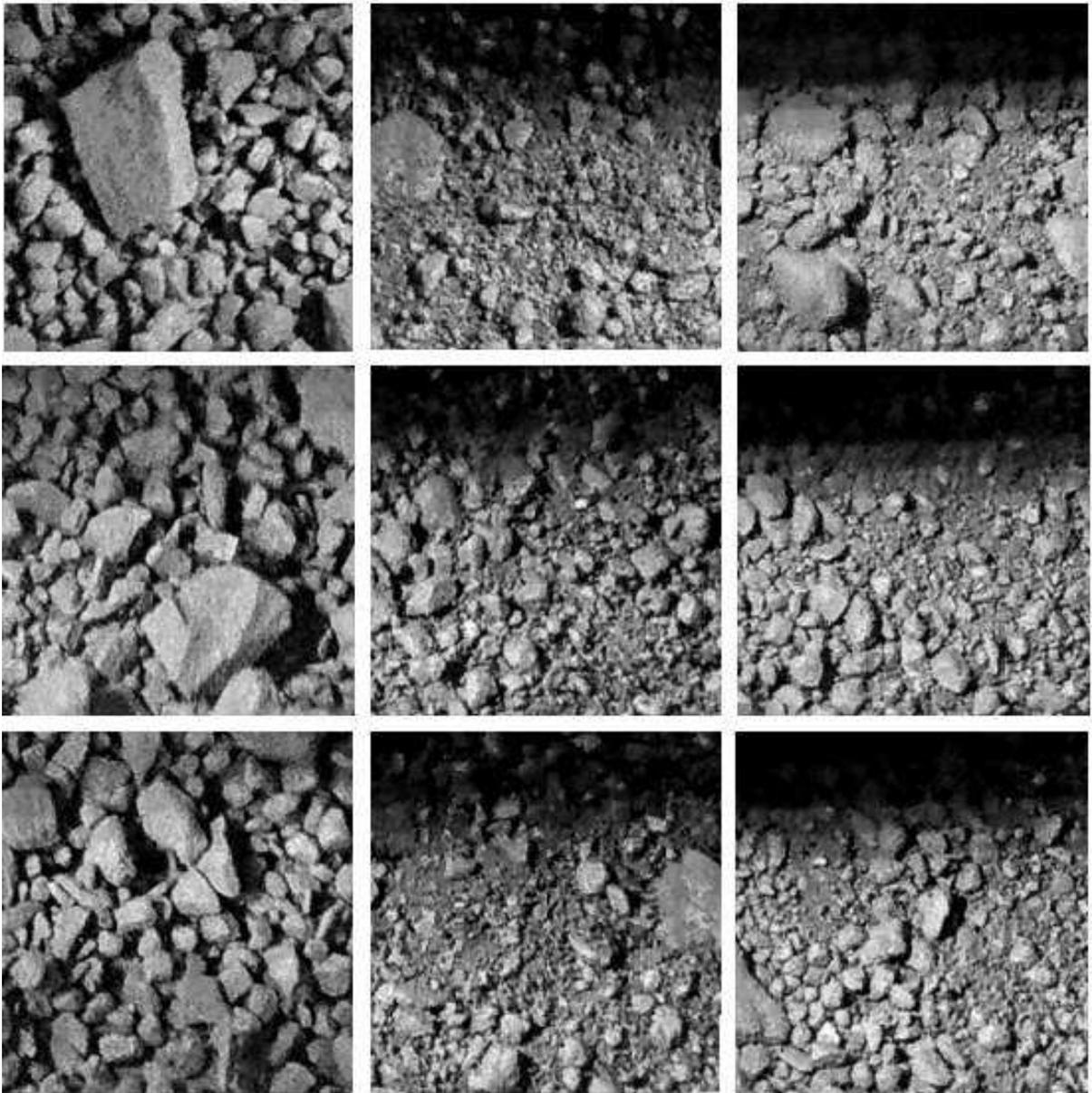
- Translação das imagens em ângulos gerados de forma aleatória;
- Filtro de brilho com intensidade aleatória.

Para facilitar esse trabalho, a biblioteca de aumento de dados *imgaug* (JUNG et al., 2020) foi utilizada para criar as funções de translação e aumento da intensidade de brilho.

No final, uma base de 688 imagens com dimensões de 256 x 256 foi gerada para os experimentos.

Um exemplo das imagens captadas pode ser visto na Figura 16.

Figura 16 – Base de dados do minério de ferro após o pré-processamento



Fonte: Acervo do autor

### 3.1.2 Geração de Ground Truth

Para cada imagem da base de dados, foi aplicado um processo manual de segmentação da região de finos, seguindo as orientações de um especialista da área de mineração. Uma região de finos pode ser definida como uma área que possui maior proporção de finos, mas que pode conter alguns grãos de minério com uma granulometria maior. Alguns exemplos da segmentação feita pelo especialista podem ser visualizados na Figura 17, no qual a região em cinza foi classificada como região de fino. Os resultados obtidos pelo processo de segmentação foram armazenados em um arquivo de extensão '.CSV', contendo os pontos  $x$ ,  $y$  de um polígono de cada região de finos, ou seja, cada imagem pode possuir

$n$  polígonos.

Figura 17 – Base de dados com Regiões de Finos segmentadas. Segmentação realizada por especialista.



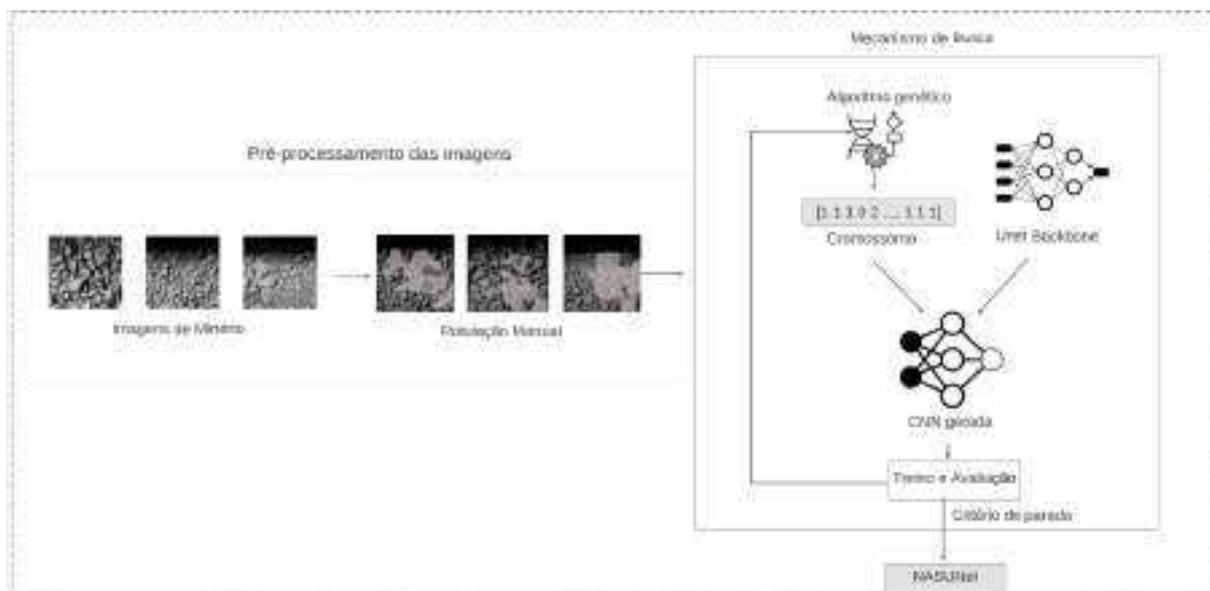
Fonte: Acervo do autor

Para identificar a região de finos, somente uma classe é mapeada para a segmentação com a rede neural. Uma matriz de  $256 \times 256$  é preenchida com valores de **0 (zero)**, representando a ausência de fino, e se um respectivo pixel em uma imagem de minério de ferro se caracteriza por ser um fino, o valor de **1 (um)** é utilizado. Para facilitar o processo de teste e diminuir o tempo de processamento, todas as máscaras geradas foram salvas em formato *png*.

## 3.2 Criação do Mecanismo de Busca Neural

No processo de identificação de uma arquitetura otimizada para a segmentação de finos em imagens, a rede de segmentação *UNet* foi empregada como uma rede base, chamada de *backbone*, para a busca de uma arquitetura otimizada. Baseado no trabalho de [Weng et al. \(2019\)](#), um processo de busca foi realizado para otimizar a arquitetura de forma modular com a utilização de um algoritmo genético (informações sobre a configuração podem ser vistas na Seção 4.2), no qual uma célula foi otimizada e possui diferentes operações identificadas pelo algoritmo de NAS. O método proposto é composto pelas seguintes etapas: (1) Aquisição das imagens de minério de ferro; (2) Rotulação manual e geração de região de fino; (3) Geração de arquitetura de *NASUNet* (Indivíduo); (4) Treino e Avaliação de Modelo; (5) Critério de Parada baseado no tempo de busca. O fluxo simplificado do trabalho pode ser visto na Figura 18.

Figura 18 – Etapas da Metodologia



Fonte: Acervo do autor

A representação de todo o espaço de busca foi resumido para um espaço discreto, onde cada cromossomo possui 15 genes para representar as configurações da arquitetura. O 4 primeiros genes indicam o tipo de convolução utilizada em cada neurônio da célula buscada, as opções são uma convolução, duas convoluções seguidas ou uma convolução com dilatação. Os genes de 5 até 10 indicam a conexão dos neurônios entre si, de forma a otimizar o espaço de busca, cada neurônio só pode se conectar com os neurônios seguintes, e caso um neurônio não possua nenhuma conexão, ele é descartado da célula. O tamanho do *kernel* pode variar entre  $2 \times 2$  até  $5 \times 5$  e é representado pelo gene 11. Os genes de 12 até 15, representam a quantidade de *dropout* de cada célula, podendo variar entre {"0", "0,2", "0,5"} A configuração do cromossomo pode ser vista na Tabela 2. A utilização da biblioteca *pymoo* permitiu configurar o domínio de cada gene de um intervalo a partir de um valor mínimo e um valor máximo, assim como a geração de valores no domínio discreto.

Tabela 2 – Configuração do Cromossomo

Gene	Opções	Configuração
1	{Conv2D, 2 Conv2D, Conv2D com dilatação}	Convolução do primeiro neurônio
2	{Conv2D, 2 Conv2D, Conv2D com dilatação}	Convolução do segundo neurônio
3	{Conv2D, 2 Conv2D, Conv2D com dilatação}	Convolução do terceiro neurônio
4	{Conv2D, 2 Conv2D, Conv2D com dilatação}	Convolução do quarto neurônio
5	{Sim, Não}	Conexão de neurônio {primeiro, segundo}
6	{Sim, Não}	Conexão de neurônio {primeiro, terceiro}
7	{Sim, Não}	Conexão de neurônio {primeiro, quarto}
8	{Sim, Não}	Conexão de neurônio {segundo, terceiro}
9	{Sim, Não}	Conexão de neurônio {segundo, quarto}
10	{Sim, Não}	Conexão de neurônio {terceiro, quarto}
11	{(2,2), (3, 3), (4, 4), (5, 5)}	Tamanho do Kernel
12	{'0', '0, 2', '0,5'}	Dropout da primeira célula
13	{'0', '0, 2', '0,5'}	Dropout da segunda célula
14	{'0', '0, 2', '0,5'}	Dropout da terceira célula
15	{'0', '0, 2', '0,5'}	Dropout da quarta célula

### 3.3 Considerações Finais

Neste capítulo, foram apresentados os métodos necessários para a construção da base de dados, como: Aquisição das imagens, Pré-processamento e Geração de segmentação de cada imagem coletada. Por fim, abordou-se o método proposto para a identificação de uma arquitetura otimizada para a segmentação de minério fino.

## 4 Experimentos Computacionais

A avaliação do processo de treinamento foi realizada com a utilização de uma métrica de similaridade chamada de *Intersect Over Union (IoU)* (REZATOFIGHI et al., 2019). O *IoU* é definido pela razão entre a interseção e a união de duas regiões  $X$  e  $Y$ , conforme equação 4.1:

$$IoU = \frac{X \cap Y}{X \cup Y} \quad (4.1)$$

O foco na área das regiões permite que essa métrica seja invariante em termo de escalas, por isso, bastante utilizada em segmentações e detecções de objetos (BEERS, 2021).

Para o processo de treinamento, a função de perda (*loss*) selecionada foi baseada no cálculo de *IoU*, no qual o objetivo é a minimização do valor negativo dessa métrica, conforme a equação 4.2.

$$L_{IoU} = -IoU \quad (4.2)$$

Para avaliação da eficiência da arquitetura neural localizada pelo processo de otimização, seus resultados foram comparados a outras duas arquiteturas que representam o estado da arte em termos de redes neurais convolucionais aplicadas a tarefa de segmentação:

1. *UNet* padrão (WENG et al., 2019)
2. *DeepLabV3* (YURTKULU et al., 2019)

Na etapa de treinamento dos modelos de segmentação, a base de imagens foi dividida de acordo com a seguinte configuração: 70% treino, 20% teste e 10% validação.

### 4.1 Hardwares e Softwares aplicados

Os experimentos realizados neste trabalho contaram com o suporte de ferramentas e bibliotecas. Para o treinamento e teste dos modelos, a plataforma do *Colab Pro* (INC., 2017) foi utilizada, com possibilidades de acesso em uma *GPU* do tipo T4 e P100 se estiver disponível, além de um limite de memória RAM de 25GB.

A implementação das redes neurais foram realizadas com a linguagem *Python* e com a utilização do *framework Tensorflow* (ABADI et al., 2015).

De forma a agilizar a criação do mecanismo de busca, o *framework Pymoo* (Blank; Deb, 2020) do *Python* foi utilizado para otimização da busca neural.

## 4.2 Configurações

Todas as redes utilizadas neste experimento foram treinadas utilizando os parâmetros apresentados na Tabela 3. O valor de *batch* adotado foi escolhido com base na limitação de memória para treinamento das redes, o otimizador *Adam* foi utilizado por sua rápida convergência, e o parâmetro de *Learning Rate* é alterado após 5 épocas quando não há aumento nos valores de validação, conforme o range de valores da Tabela 3. O critério de parada padrão é de finalização do treinamento após 100 épocas, adicionalmente, o processo de treino é finalizado caso o valor de perda da validação não diminua por 10 épocas consecutivas, condição esta sendo adicionada de forma a evitar *overfitting* (YING, 2019).

Tabela 3 – Parâmetros

Parâmetro	Valor
<i>Batch</i>	5
<i>Otimizador</i>	Adam
<i>Learning Rate</i>	[1,0e-04; 1,0e-05]

A configuração utilizada pela rede *UNet* tradicional consistiu em uma *VGG19* (SIMONYAN; ZISSERMAN, 2014) para as camadas de *downsampling*, *dropout* de 0,5 depois de cada célula das operações de *upsampling*, além disso, foi aplicado um *transfer-learning* da *VGG19* com os pesos do pré-treino da *image-net* (DENG et al., 2009).

Se tratando da rede neural *DeepLabV3* (YURTKULU et al., 2019), a principal característica consiste na utilização de várias camadas de convolução com dilatação (CHEN et al., 2017). A configuração padrão da rede utilizou uma *ResNet50* (HE et al., 2015) como estrutura inicial, e *transfer-learning* dos pesos da *image-net*.

Em relação ao Algoritmo Genético (JOHN, 1992), a função objetivo foi definida como a maximização do resultado de *Intersect Over Union* dos dados de teste. As seguintes configurações adotadas para o mecanismo de busca desenvolvido neste trabalho podem ser vistas na tabela 4, a taxa de *crossover* de 80 % foi utilizada para construir uma diversidade na população, permitindo que o algoritmo genético faça uma boa exploração no espaço de busca. O fator de mutação utilizado foi o valor adotado como o padrão pela biblioteca *pymoo*, essa taxa é calculada utilizando a função 4.3, em que  $n$  é a quantidade de variáveis do problema. Um valor mais baixo na taxa de mutação garante um refinamento do espaço de busca e uma diminuição na perda de informações no processo de mutação.

$$\frac{1}{n} \quad (4.3)$$

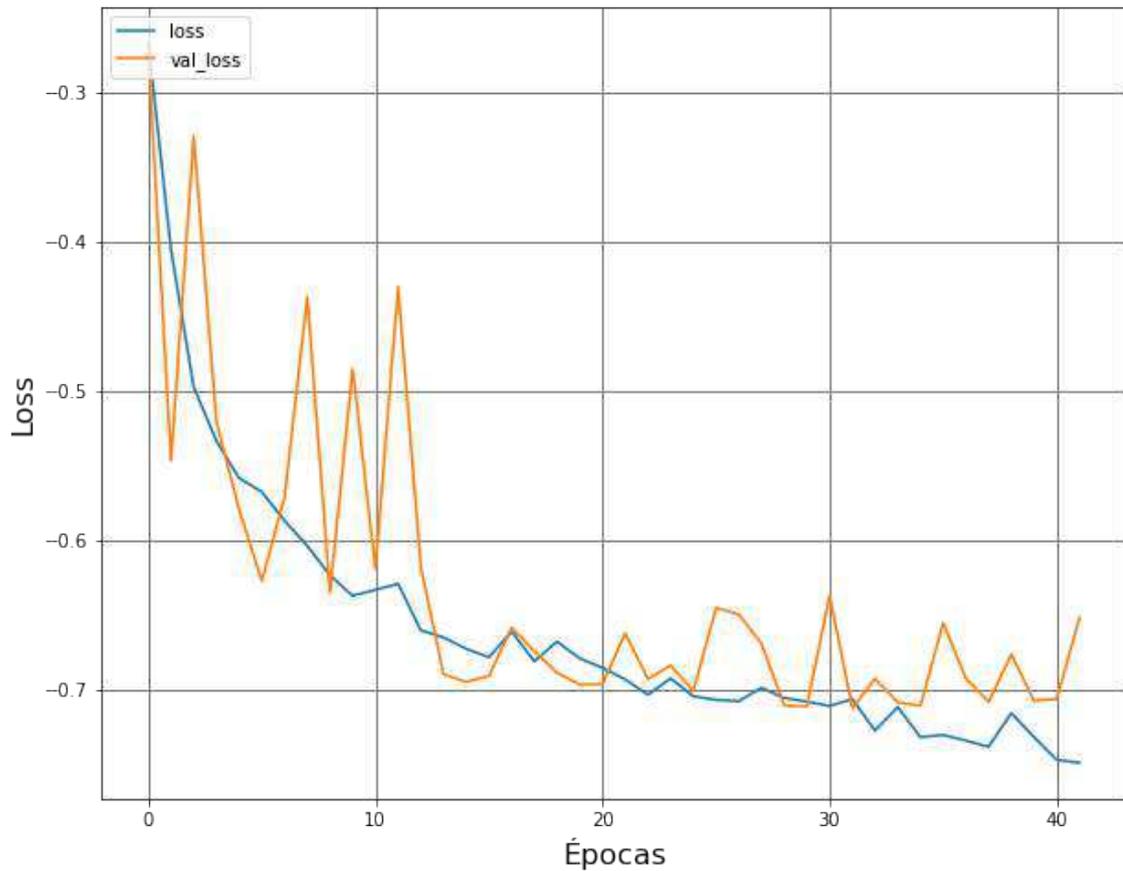
Tabela 4 – Configuração do Cromossomo

	<b>Configuração</b>
Taxa de <i>Crossover</i>	80%
Função de Avaliação	$\sum_{i=1}^{\text{size}(teste)} IoU$
Tamanho da População	5 indivíduos
Seleção	Seleção por torneio
Taxa de Mutação	6%%
Codificação Cromossomos	Real
Tempo de Execução	15 horas

### 4.3 Resultados

A Figura 19 apresenta as informações relacionadas ao treinamento da *UNet* tradicional. No gráfico, a série em azul representa o erro em relação aos dados de treinamento, enquanto a série representada em cor laranja apresenta a variação do erro do modelo em relação ao conjunto de dados de validação. A etapa de treinamento durou 40 épocas, sendo interrompida devido ao critério de parada que finaliza o processo após 10 épocas consecutivas sem ocorrer diminuição do erro de validação.

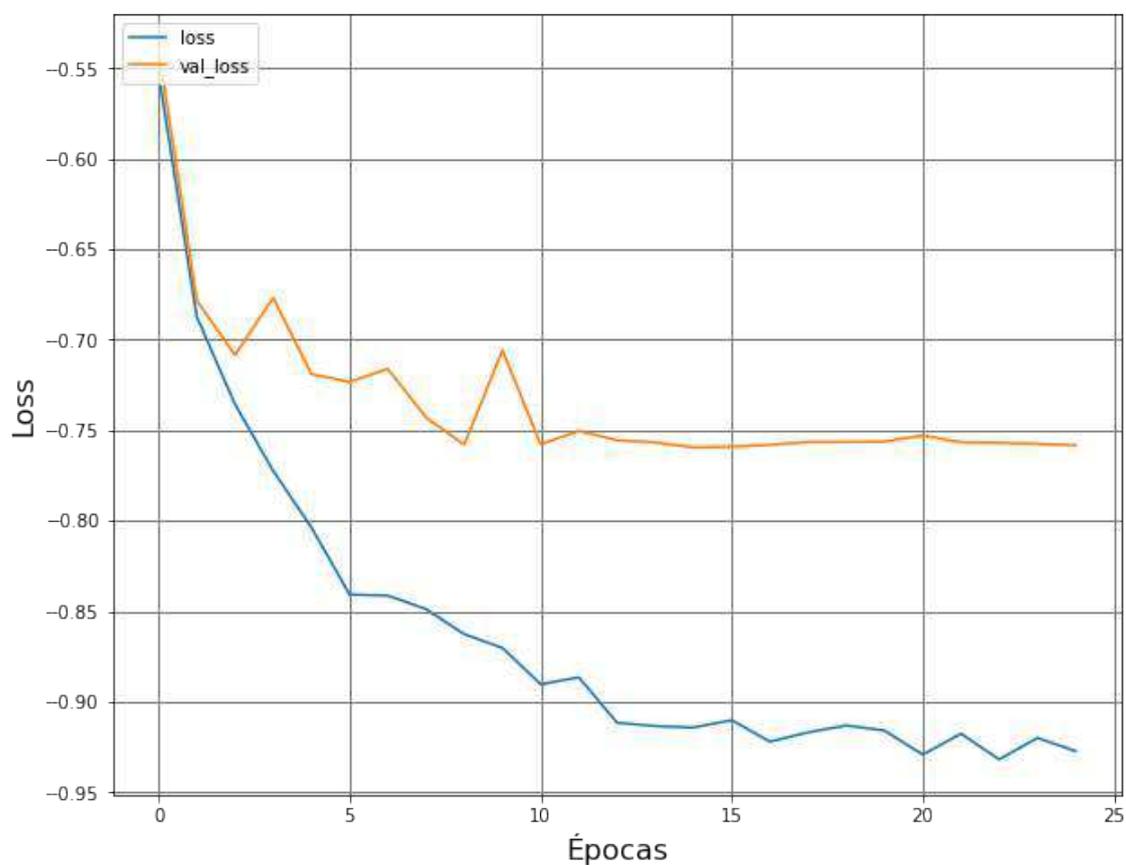
Figura 19 – Treinamento da UNet



Fonte: Acervo do autor

A Figura 20 apresenta as informações relacionadas ao treinamento da rede *DeepLabV3*. No gráfico, a série em azul representa o erro em relação aos dados de treinamento, enquanto a série representada em cor laranja apresenta a variação do erro do modelo em relação ao conjunto de dados de validação. A etapa de treinamento durou somente 25 épocas, seguindo o critério já apresentado de interrupção do processo de treinamento após 10 épocas consecutivas sem ocorrer diminuição do erro de validação.

Figura 20 – Loss - DeepLabV3



Fonte: Acervo do autor

A busca neural baseada no *backbone* da rede da *UNet* identificou que para os dados apresentados, o melhor conjunto de características, que representa o melhor indivíduo foi:

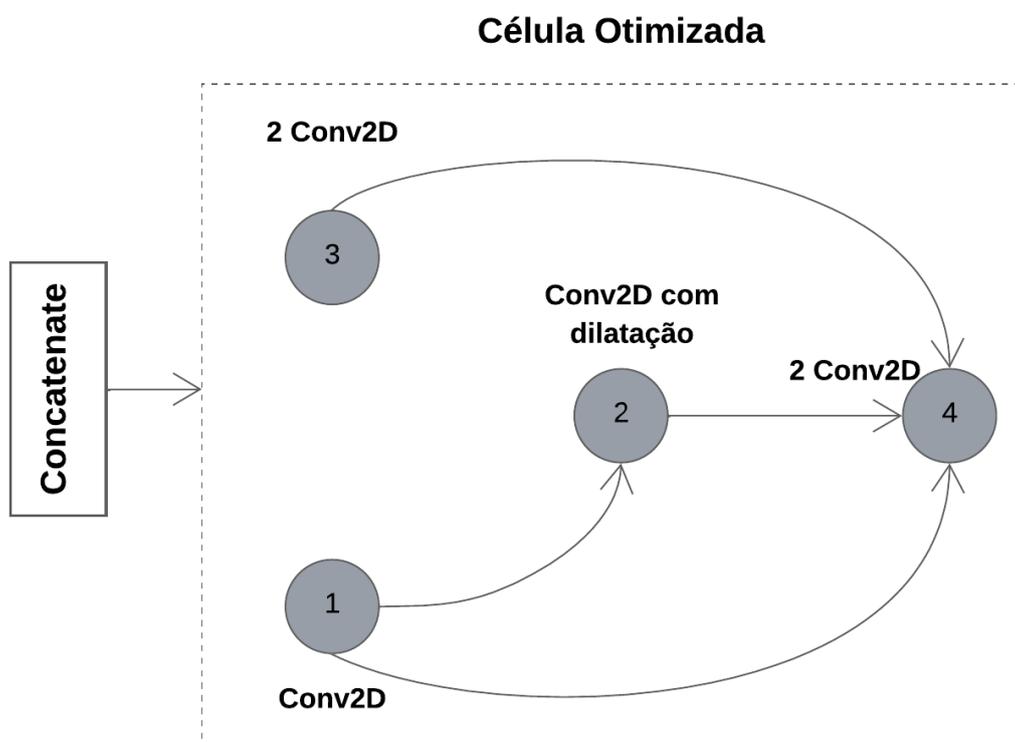
- [0 2 1 1 1 0 1 0 1 1 3 2 0 1 0]

A Tabela 5 apresenta a configuração conforme cada gene do indivíduo encontrado. É possível identificar que menos operações de *dropout* e um tamanho de *kernel* maior representaram um melhor resultado. A representação gráfica da estrutura da célula pode ser vista na Figura 21.

Tabela 5 – NASUNet - Configuração

<b>Valor</b>	<b>Configuração</b>	<b>Resultado</b>
0	Convolução do primeiro neurônio	<b>Conv2D</b>
2	Convolução do segundo neurônio	<b>Conv2D com dilatação</b>
1	Convolução do terceiro neurônio	<b>2 camadas de Conv2D</b>
1	Convolução do quarto neurônio	<b>2 camadas de Conv2D</b>
1	Conexão do primeiro neurônio com o segundo	<b>Sim</b>
0	Conexão do primeiro neurônio com o terceiro	<b>Não</b>
1	Conexão do primeiro neurônio com o quarto	<b>Sim</b>
0	Conexão do segundo neurônio com o terceiro	<b>Não</b>
1	Conexão do segundo neurônio com o quarto	<b>Sim</b>
1	Conexão do terceiro neurônio com o quarto	<b>Sim</b>
3	Tamanho do Kernel	<b>4x4</b>
2	Dropout primeira célula	<b>0,5</b>
0	Dropout segunda célula	<b>0</b>
1	Dropout terceira célula	<b>0,25</b>
0	Dropout terceira célula	<b>0</b>

Figura 21 – Célula Otimizada - NASUNet



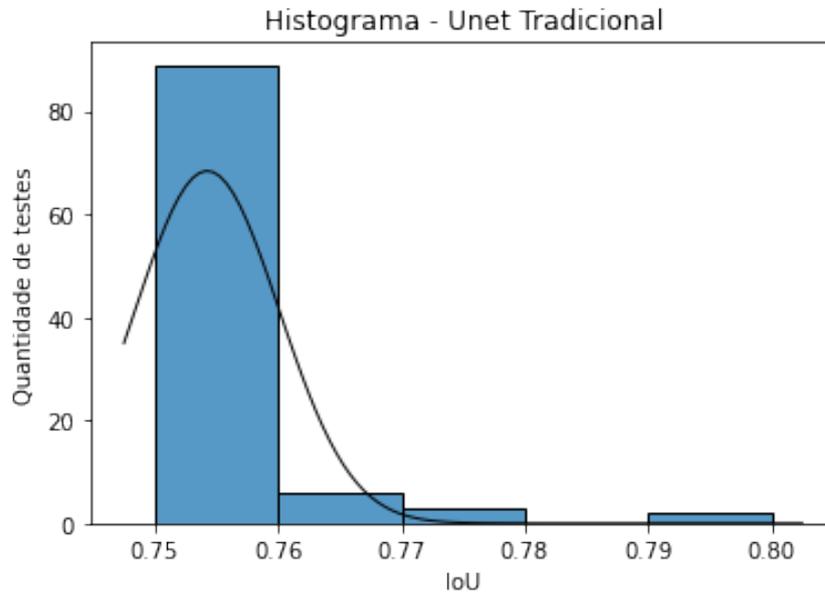
Fonte: Acervo do autor

Os resultados dos dados de teste apresentaram uma diferença de aproximadamente 2% entre o *IoU* obtido com a arquitetura de busca neural comparado ao resultado obtido pela arquitetura *Deep Lab V3*. Se comparado com arquitetura tradicional da *UNet*, o aumento é de aproximadamente 5%. A Tabela 6 apresenta os resultados da média de 100 execuções de teste para cada uma das 3 redes. A representação dos valores de *IoU* dos modelos treinados pode ser analisada através dos histogramas da Figura 22, que apresenta a distribuição dos resultados de teste do treinamento de 100 vezes do modelo da *UNet* tradicional, no qual é possível verificar uma concentração dos valores de entre 0,75 e 0,76, a Figura 23 que apresenta os resultados para a rede *DeepLabV3*, em que é possível verificar uma distribuição normal, com a média em 0,7804, e a Figura 24 com a distribuição dos resultados NASUNet, em que é possível verificar uma cauda longa para a esquerda, com uma concentração entre valores de 0,81 e 0,82.

Tabela 6 – Resultados

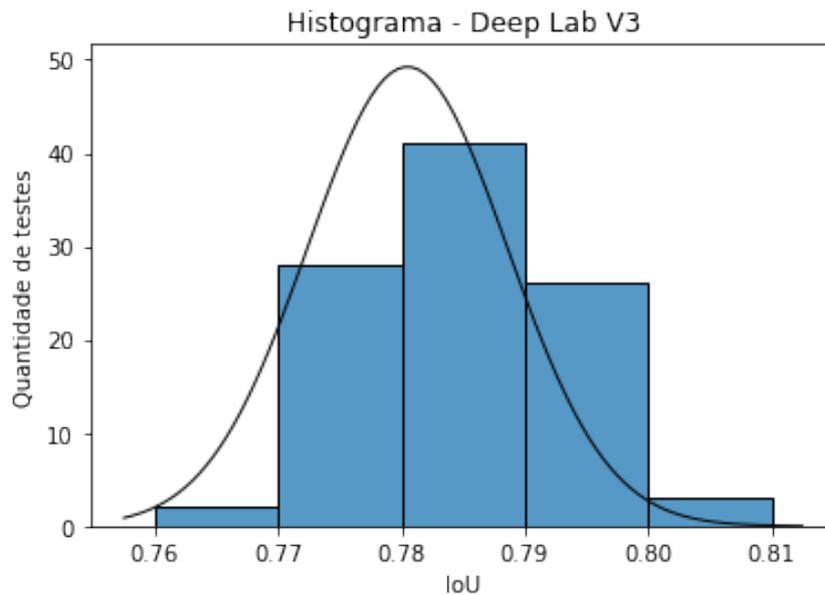
Modelo	<i>IoU</i> - média	Desvio padrão
<i>UNet</i>	0,7542	0.005
<i>Deep Lab V3</i>	0,7804	0.008
<i>NASUNet</i>	<b>0,8034</b>	0.016

Figura 22 – Célula Otimizada - *UNet* Tradicional

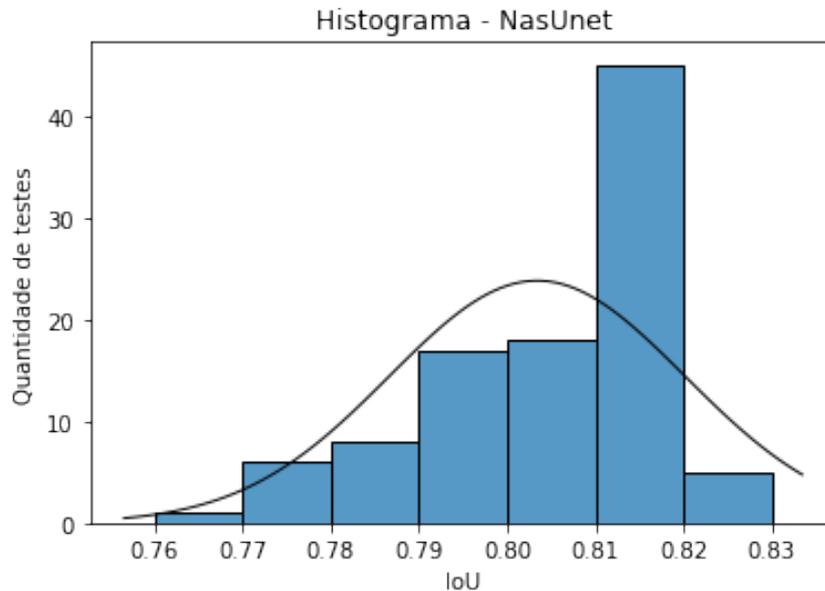


Fonte: Acervo do autor

Figura 23 – Célula Otimizada - *Deep Lab V3*



Fonte: Acervo do autor

Figura 24 – Célula Otimizada - *NASUNet*

Fonte: Acervo do autor

### 4.3.1 Teste não paramétrico

De forma a comparar os resultados estatisticamente, um teste não paramétrico foi aplicado nas distribuições, esse teste é utilizado quando não conseguimos garantir que há uma distribuição normal nos dados, que podem ser visto nos histogramas da Seção 4.3. O teste de Kruskal-Wallis (KRUSKAL; WALLIS, 1952) foi o tipo de teste não paramétrico selecionado, esse teste analisa uma hipótese nula em que todos os grupos de dados possuem o mesmo valor de mediana.

De forma a obter comprovação estatística do valor de *IoU* obtido pelo modelo de *NASUNet*, o teste de Kruskal foi aplicado para avaliar a significância estatística do aumento de 2% e 5% em relação aos modelos de *UNet* e *DeepLabV3*, respectivamente. Ao utilizar o teste foi possível obter os seguintes resultados:

Tabela 7 – Kruskal

Comparação	p-value
<i>UNet</i> x <i>NASUNet</i>	8.70e-34
<i>DeepLabV3</i> x <i>NASUNet</i>	5.060e-20

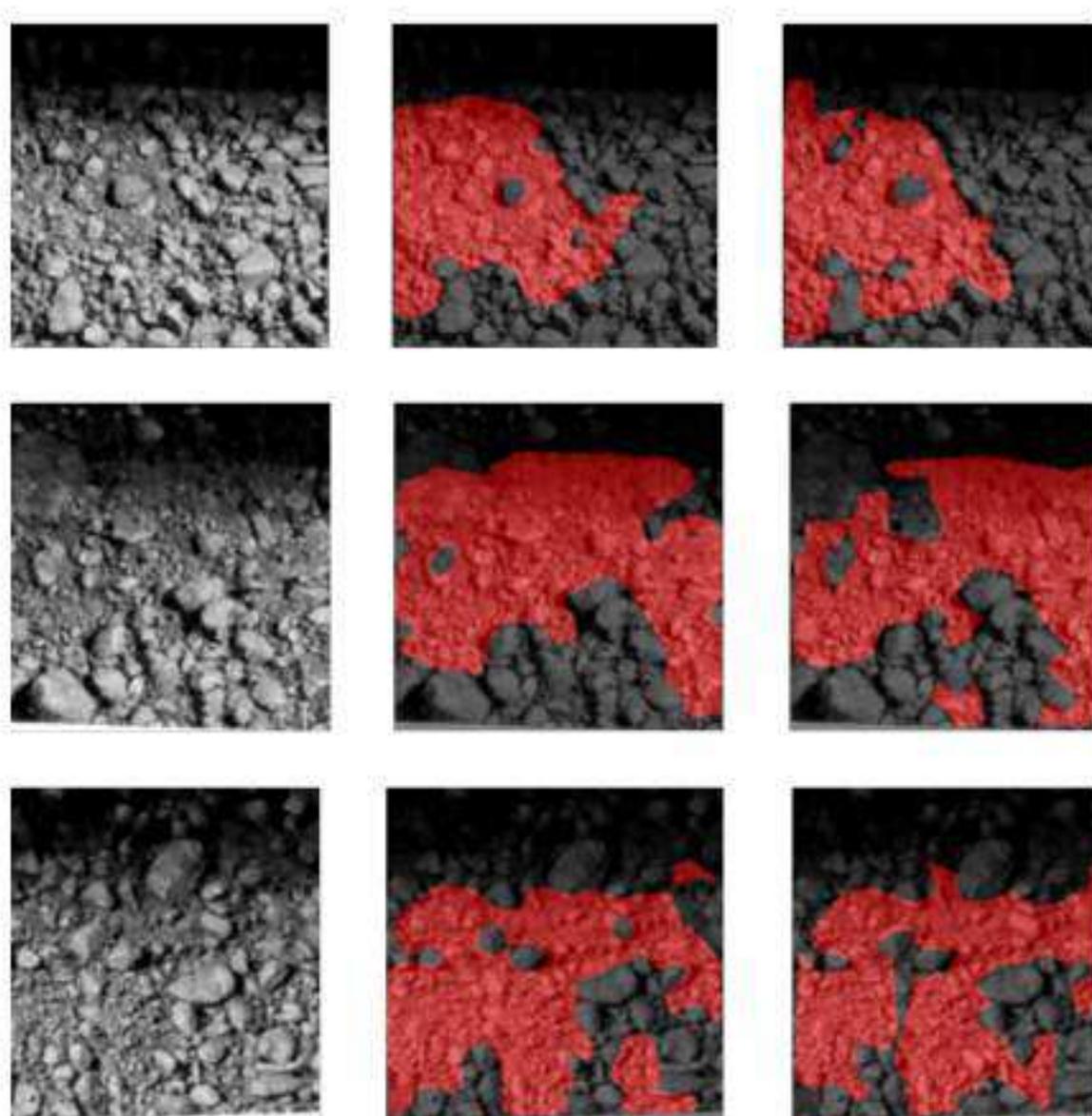
Com base no valor de *p-value* menor que 0,05, é possível ter confiança de que as médias dos resultados não são as mesmas para os três modelos testados, com um nível de significância de 5%.

### 4.3.2 Discussão

Nesta seção são abordados alguns exemplos de imagens de segmentação encontrados com base nas técnicas aplicadas.

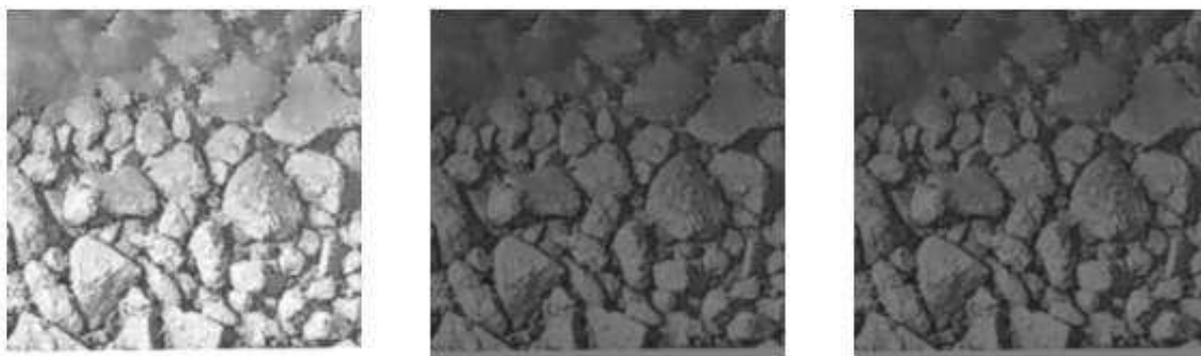
A Figura 25 apresenta casos em que a segmentação identificou corretamente a área de interesse, diferenciando áreas de minério granulado de regiões de fino.

Figura 25 – Imagens reais (Coluna da esquerda); Imagem segmentada por modelo neural (Coluna do meio); Imagens segmentadas pelo especialista (Coluna da direita)



A Figura 26 apresenta um caso em que não há região de finos na imagem, somente material granulado. Nesses casos, o modelo de segmentação não teve problemas para classificar somente como não fino toda a imagem de granulado.

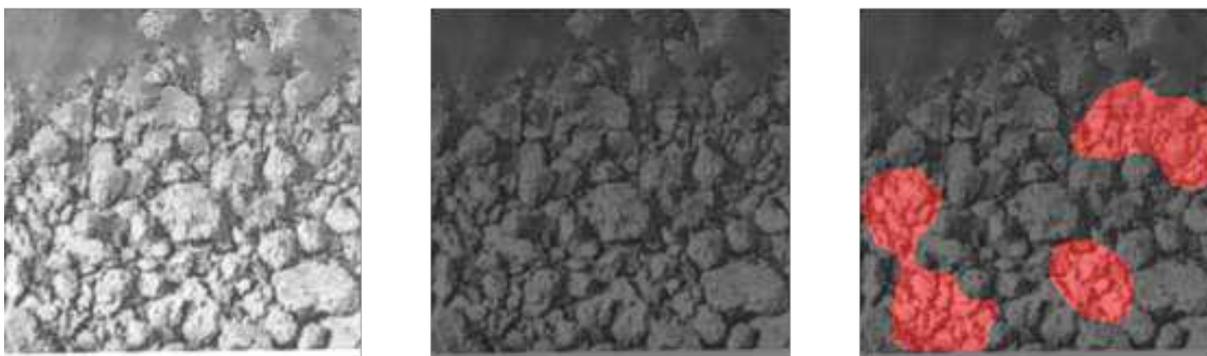
Figura 26 – Imagens de minério de ferro granulado



Fonte: Acervo do autor

Em alguns casos, devido à trepidação proveniente do peneiramento, alguns *frames* possuem a imagem ruidosa. Nesses cenários, a solução acaba erroneamente identificando finos em área de granulado, como o exemplo da Figura 27.

Figura 27 – Segmentações ruins



Fonte: Acervo do autor

De modo geral, a solução foi capaz de detectar regiões de finos quando o minério possui granulometria menor. Os falsos positivos encontrados geralmente representavam regiões bem pequenas de segmentação, nesse sentido, ao analisar a limitação de 20% de área da imagem, conforme explicado na Seção 2.1, esses falsos positivos não geram um alarme para diminuir a produção da Usina. Os falsos negativos acabam gerando um impacto maior em problemas para a solução, visto que a não identificação acaba ocasionando um fluxo de finos maior do que esperado em etapas posteriores da Usina.

### 4.3.3 Nível de Maturidade Tecnológica

O modelo de segmentação de finos se encontra em um ambiente operacional e sendo utilizada com entrada para tomada de decisões em um ambiente industrial do peneiramento. De acordo com o *Technology Readiness Level (TRL)* (MANKINS, 1995),

que avalia a maturidade de projetos de pesquisa, desenvolvimento e inovação, essa solução se encontra em um nível de TRL/MRL 7, pois o protótipo está demonstrado e avaliado em ambiente operacional.

## 5 Conclusão

Neste trabalho foi apresentado um mecanismo de busca desenvolvido para encontrar uma rede neural convolucional eficiente para segmentação de minério de ferro fino. A arquitetura da *UNet* foi utilizada como base para a construção de uma estrutura otimizada para o problema, chamada de *NASUNet*.

A arquitetura encontrada foi testada em uma base de dados coletada durante vários períodos do dia e em estações diferentes. Dos vídeos coletados, 688 imagens foram segmentadas para construção do *dataset* utilizado nas etapas de treino e teste dos modelos.

O modelo de *NASUNet* foi treinado utilizando o *dataset* construído, e obteve como resultado o valor de 80% de *intersect over union*. Os resultados obtidos podem ser considerados promissores, levando em consideração que as imagens foram coletadas em um ambiente real de produção, portanto:

- O ruído das imagens capturadas, devido ao cenário de um equipamento em constante vibração;
- A variação do formato do material, devido aos diferentes períodos do ano em que as imagens foram coletadas;
- O aspecto similar de cor em todo o material.

De forma a alcançar o objetivo principal definido neste trabalho, foram realizadas as seguintes contribuições para atingir os objetivos específicos:

- Aquisição das imagens de minério de ferro em um ambiente industrial;
- Criação de uma base de dados pré-processada com segmentações manuais realizadas por um especialista na área de mineração;
- Utilização de um mecanismo de otimização para realização de busca da arquitetura otimizada para segmentação de minério;
- Modelagem do conjunto de parâmetros utilizados pelo mecanismo de busca para encontrar um modelo otimizado;
- Avaliação de efetividade do método de segmentação para as regiões de fino do minério de ferro.

A segmentação de finos por visão computacional é uma necessidade de várias usinas de minério, e esse fluxo de trabalho pode auxiliar e ser aplicado no setor de mineração com o intuito de identificar a granulometria dos materiais. Adicionalmente, é importante destacar que a solução implementada e analisada neste trabalho foi avaliada e se encontra em produção em um dos equipamentos de peneiramento em uma indústria na região de Carajás, um diagrama resumido da aplicação pode ser visto no Apêndice A. Por fim, os resultados deste trabalho geraram uma publicação em uma conferência e uma publicação em um periódico, mais detalhes sobre as produções científicas podem ser encontrados no Apêndice B.

## 5.1 Trabalhos Futuros

Mesmo com os resultados considerados satisfatórios para a contribuição do trabalho e metodologia criada, ainda são necessárias algumas melhorias como trabalhos futuros:

- Ampliar a base de dados com amostras de minério de ferro de outras áreas de lavra, pois apesar de a aquisição das imagens ter sido realizada durante um período de um ano, o tipo de minério dessa pesquisa foi limitado para a região de Carajás;
- A segmentação proposta por esse trabalho, foi a identificação de minério fino, um trabalho mais específico para identificar e segmentar também o minério de ferro granulado permite ter um melhor entendimento do material proveniente da lavra;
- Por fim, sugere-se analisar o desempenho dos modelos em plataformas embarcadas, visto que a maioria das soluções industriais precisam ser descentralizadas e não possuem infraestrutura robusta, e a utilização de embarcados facilita o processo de implementação de uma aplicação.

# Referências

- ABADI, M. et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>. Citado na página 45.
- AGGARWAL, C. *Neural Networks and Deep Learning: A Textbook*. [S.l.: s.n.], 2018. ISBN 978-3-319-94462-3. Citado na página 29.
- AHMED, I. et al. Comparison of deep-learning-based segmentation models: Using top view person images. *IEEE Access*, v. 8, p. 136361–136373, 2020. Citado na página 15.
- ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. In: *2017 International Conference on Engineering and Technology (ICET)*. [S.l.: s.n.], 2017. p. 1–6. Citado na página 30.
- BAZGIR, O. et al. Representation of features as images with neighborhood dependencies for compatibility with convolutional neural networks. *Nature Communications*, v. 11, 09 2020. Citado na página 29.
- BEERS, F. van. Capsule networks with intersection over union loss for binary image segmentation. In: . [S.l.: s.n.], 2021. Citado na página 45.
- Blank, J.; Deb, K. pymoo: Multi-objective optimization in python. *IEEE Access*, v. 8, p. 89497–89509, 2020. Citado na página 46.
- CHEN, L.-C. et al. *Rethinking Atrous Convolution for Semantic Image Segmentation*. 2017. Citado na página 46.
- CHEN, M. et al. Rock particle motion information detection based on video instance segmentation. *Sensors (Basel, Switzerland)*, v. 21, 2021. Citado 2 vezes nas páginas 18 e 19.
- CHOLLET, F. *Xception: Deep Learning with Depthwise Separable Convolutions*. 2017. Citado na página 31.
- CORREIA, J. C. G. Classificação e peneiramento. *Tratamento de Minérios*, p. 257–298, 2010. Citado na página 26.
- CUI, X. et al. Multiscale spatial-spectral convolutional network with image-based framework for hyperspectral imagery classification. *Remote Sensing*, v. 11, p. 2220, 09 2019. Citado na página 32.
- DENG, J. et al. Imagenet: A large-scale hierarchical image database. In: IEEE. *2009 IEEE conference on computer vision and pattern recognition*. [S.l.], 2009. p. 248–255. Citado na página 46.
- DOMHAN, T.; SPRINGENBERG, J. T.; HUTTER, F. Extrapolating learning curves of deep neural networks. In: . [S.l.: s.n.], 2014. Citado na página 36.

- DOMINGOS, D.; CORTES, O. C.; LOBATO, F. Evoluindo redes neurais convolucionais na detecção de emoções usando micro ags. In: . [S.l.: s.n.], 2022. Citado na página 20.
- DUAN, J. et al. Detection and segmentation of iron ore green pellets in images using lightweight u-net deep learning network. *Neural Computing and Applications*, v. 32, p. 5775–5790, 2019. Citado 2 vezes nas páginas 18 e 19.
- DUAN, J. et al. Detection and segmentation of iron ore green pellets in images using lightweight u-net deep learning network. *Neural Computing and Applications*, v. 32, 05 2020. Citado na página 15.
- DUMOULIN, V.; VISIN, F. A guide to convolution arithmetic for deep learning. *ArXiv*, abs/1603.07285, 2016. Citado na página 33.
- ELSKEN, T.; METZEN, J. H.; HUTTER, F. *Neural Architecture Search: A Survey*. 2019. Citado 2 vezes nas páginas 34 e 36.
- HE, K. et al. *Deep Residual Learning for Image Recognition*. 2015. Citado na página 46.
- HE, X.; ZHAO, K.; CHU, X. Automl: A survey of the state-of-the-art. *CoRR*, abs/1908.00709, 2019. Disponível em: <<http://arxiv.org/abs/1908.00709>>. Citado na página 34.
- HERNANDEZ-GARCIA, A. Data augmentation and image understanding. *arXiv preprint arXiv:2012.14185*, 2020. Citado na página 18.
- HOU, Y. et al. C-unet: Complement unet for remote sensing road extraction. *Sensors (Basel, Switzerland)*, v. 21, 2021. Citado na página 17.
- INC., A. *Google Colab*. 2017. <<https://colab.research.google.com/>>. Citado na página 45.
- JOHN, H. H. Adaptation in natural and artificial systems. *MIT Press (Cambridge, MA)*, 1992. Citado 2 vezes nas páginas 36 e 46.
- JUNG, A. B. et al. *imgaug*. 2020. <<https://github.com/aleju/imgaug>>. Online; accessed 01-Feb-2020. Citado na página 40.
- KARKI, S.; KULKARNI, S. Ship detection and segmentation using unet. In: *2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*. [S.l.: s.n.], 2021. p. 1–7. Citado na página 17.
- KRUSKAL, W. H.; WALLIS, W. A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, v. 47, p. 583–621, 1952. Citado na página 53.
- LIU, X. et al. Ore image segmentation method using u-net and res\_unet convolutional networks. *RSC Advances*, v. 10, p. 9396–9406, 2020. Citado 2 vezes nas páginas 17 e 19.
- LONG, J.; SHELHAMER, E.; DARRELL, T. *Fully Convolutional Networks for Semantic Segmentation*. 2015. Citado 2 vezes nas páginas 33 e 34.
- Ma, X. et al. A New Belt Ore Image Segmentation Method Based on the Convolutional Neural Network and the Image-Processing Technology. *Minerals*, v. 10, n. 12, p. 1115, dez. 2020. Citado 2 vezes nas páginas 18 e 19.

- MANKINS, J. C. Technology readiness levels. a white paper. *Office of Space Access and Technology*. NASA, 1995. Disponível em: <[http://www.artemisinnovation.com/images/TRL\\_White\\_Paper\\_2004-Edited.pdf](http://www.artemisinnovation.com/images/TRL_White_Paper_2004-Edited.pdf)>. Citado na página 55.
- MCGLINCHY, J. et al. Application of unet fully convolutional neural network to impervious surface segmentation in urban environment from high resolution satellite imagery. In: . [S.l.: s.n.], 2019. p. 3915–3918. Citado na página 17.
- MINGWEI, Z. et al. Nas-hris: Automatic design and architecture search of neural network for semantic segmentation in remote sensing images. *Sensors*, v. 20, p. 5292, 09 2020. Citado na página 15.
- OLIVIER, L. E.; MARITZ, M. G.; CRAIG, I. K. Estimating ore particle size distribution using a deep convolutional neural network. *IFAC-PapersOnLine*, v. 53, p. 12038–12043, 2020. Citado 2 vezes nas páginas 18 e 19.
- RAWAT et al. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, v. 29, p. 1–98, 06 2017. Citado na página 29.
- REZATOFIGHI, H. et al. *Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression*. 2019. Citado na página 45.
- ROERDINK, J.; MELJSTER, A. The watershed transform: Definitions, algorithms and parallelization strategies. *Fundam Inf*, v. 41, 10 2003. Citado na página 17.
- SALES, C. G. de. *Peneiramento Industrial: Metodologia para dimensionamento e simulação*. Tese (Doutorado) — Universidade Federal de Minas Gerais - UFMG, 2020. Citado na página 15.
- SIDDIQUE, N. et al. U-net and its variants for medical image segmentation: theory and applications. 11 2020. Citado na página 17.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv 1409.1556*, 09 2014. Citado na página 46.
- SVENSSON, T. *Semantic Segmentation of Iron Ore Pellets with Neural Networks (Dissertation)*. Tese (Doutorado) — Luleå University of Technology, 2019. Disponível em: <<http://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-74352>>. Citado 2 vezes nas páginas 17 e 19.
- TALBI, E.-G. Machine learning into metaheuristics: A survey and taxonomy of data-driven metaheuristics. In: . [S.l.: s.n.], 2020. Citado na página 36.
- THURLEY, M. J. Automated image segmentation and analysis of rock piles in an open-pit mine. *2013 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, p. 1–8, 2013. Citado 2 vezes nas páginas 18 e 19.
- VALADAO, G.; CORREA, A. *Introdução ao Tratamento de Minérios*. [S.l.: s.n.], 2007. ISBN 978-8570414786. Citado 4 vezes nas páginas 25, 26, 27 e 28.
- VALE. *Peneiramento Sem água*. 2013. Disponível em: <[https://issuu.com/jorgesa10/docs/peneiramento\\_sem\\_agua](https://issuu.com/jorgesa10/docs/peneiramento_sem_agua)>. Citado na página 24.

WENG, Y. et al. Nas-unet: Neural architecture search for medical image segmentation. *IEEE Access*, v. 7, p. 44247–44257, 2019. Citado 3 vezes nas páginas 19, 42 e 45.

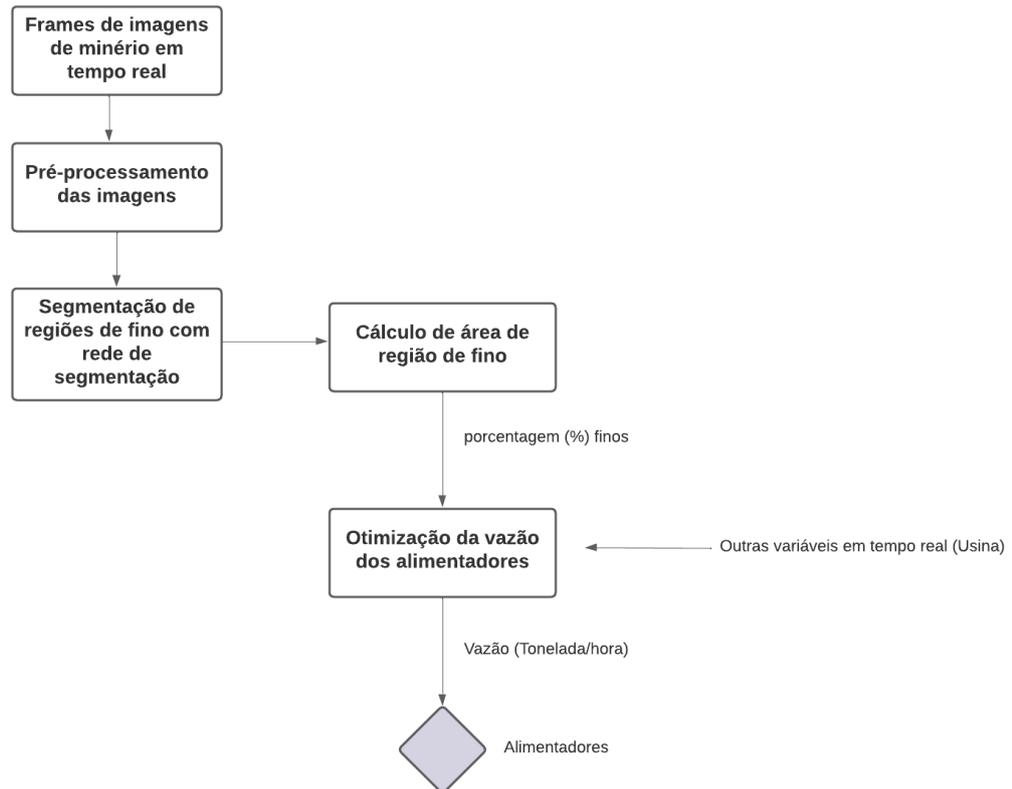
YING, X. An overview of overfitting and its solutions. *Journal of Physics: Conference Series*, v. 1168, p. 022022, 02 2019. Citado na página 46.

YURTKULU, S. C.; ŞAHIN, Y. H.; UNAL, G. Semantic segmentation with extended deeplabv3 architecture. In: IEEE. *2019 27th Signal Processing and Communications Applications Conference (SIU)*. [S.l.], 2019. p. 1–4. Citado 2 vezes nas páginas 45 e 46.

# Apêndices

# APÊNDICE A – Solução Implementada

Figura 28 – Aplicação - usina Carajas



Fonte: Acervo do autor

## APÊNDICE B – Produções Científicas

Artigo científico aprovado na conferência ISDA (22nd International Conference on Intelligent Systems Design and Applications) com o título de GANASUNet: An Efficient Convolutional Neural Architecture for Segmenting Iron Ore Images.

da Silva, Ada Cristina França ; CORTES, O. A. C. . Segmentation of Fine-Grained Iron Ore Using Deep Learning and IoT. In: Vikram Bali; Vishal Bhatnagar; Deepti Aggarwal; Shivani Bali; Mario José Diván. (Org.). Cyber-Physical, IoT, and Autonomous Systems in Industry 4.0. 1ed.: CRC Press, 2021, v. , p. 1-