

**Thalita Dias Pinheiro Caldas**

**Desenvolvimento de uma metodologia para  
caracterização de fases no pellet feed utilizando  
microscopia digital e aprendizagem profunda**

**Tese de Doutorado**

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Engenharia de Materiais e de Processos Químicos e Metalúrgicos do Departamento de Engenharia Química e de Materiais.

Orientador: Prof. Sidnei Paciornik  
Coorientadora: Dra. Karen Soares Augusto

Rio de Janeiro  
Agosto de 2023

**Thalita Dias Pinheiro Caldas**

**Desenvolvimento de uma metodologia para  
caracterização de fases no pellet feed utilizando  
microscopia digital e aprendizagem profunda**

Tese apresentada como requisito parcial para  
obtenção do grau de Doutor pelo Programa de Pós-  
graduação em Engenharia de Materiais e de  
Processos Químicos e Metalúrgicos 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. Richard Bryan Magalhães Santos**

Departamento de Engenharia Química e de Materiais – PUC-Rio

**Dr. Otávio da Fonseca Martins Gomes**

Centro de Tecnologia Mineral – CETEM

**Dr. Vinícius de Moraes Oliveira**

Centro de Tecnologia de Ferrosos – Vale S.A.

**Prof. Leonardo Evangelista Lagoeiro**

Departamento de Geologia – UFPR

Rio de Janeiro, 29 de agosto de 2023

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

### **Thalita Dias Pinheiro Caldas**

Thalita Dias Pinheiro Caldas graduou-se em Licenciatura plena em Química pela Universidade Estadual do Piauí (2016). Possui mestrado em Engenharia de Materiais e Processos Químicos e Metalúrgicos pela PUC-Rio (2019), com ênfase em Ciência dos Materiais, Microscopia Digital e Processamento e Análise Digital de Imagens.

### Ficha Catalográfica

Caldas, Thalita Dias Pinheiro

Desenvolvimento de uma metodologia para caracterização de fases no pellet feed utilizando microscopia digital e aprendizagem profunda / Thalita Dias Pinheiro Caldas; orientador: Sidnei Paciornik; coorientadora: Karen Soares Augusto. – 2023.

117 f.: il. color.; 30 cm

Tese (doutorado)–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. Minério de ferro. 3. Caracterização microestrutural. 4. Microscopia digital. 5. Análise de imagens. 6. Aprendizagem Profunda. 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**

A Deus, por me dar a vida e dedicar a mim seu infinito amor e misericórdia, por me proteger e ajudar a sempre superar as maiores dificuldades. Aos meus pais, por todos os ensinamentos e valores. Por todo o amor e todos os sacrifícios que fizeram por mim. Aos meus irmãos João Neto e Sávio Dias por me apoiarem em cada decisão e me incentivarem a não desistir dos meus objetivos.

Ao meu namorado Allan Cytryn por todo amor e incentivo nesta caminhada, tê-lo ao meu lado deixou o percurso muito mais leve e agradável.

À minha família do Rio de Janeiro representada por Dazinha Moraes e Matheus Brandão, eu tenho eterna gratidão por todo cuidado, amor e proteção.

À minhas amigas Adrielle Morália, Láine Machado, Lana Teles, Joseane Rodrigues, Bárbara Cortez, Dnane Almeida, Carla Bandeira e Débora Arrais, por todo o apoio, cumplicidade e confiança. A todos os meus avós, tios e primos pelo incentivo.

Ao meu orientador Sidnei Paciornik, por quem tenho imensa admiração e carinho, sou grata por todo o conhecimento compartilhado, e toda a atenção dedicada não só no desenvolvimento desta tese, mas desde o início do mestrado quando me acolheu no grupo de pesquisa.

À minha querida coorientadora Dra. Karen Augusto que foi parte indispensável para o desenvolvimento deste trabalho, e para minha formação profissional. Sempre paciente ao orientar e atenta a cada etapa no desenvolvimento da metodologia do trabalho. Sou muito grata por todo o apoio, incentivo e confiança. Me inspiro muito em todo o seu potencial como pesquisadora e orientadora.

Ao Dr. Julio Iglesias por ter contribuído de forma essencial para o conteúdo deste trabalho. Ao Dr. Marcos Henrique por sempre se mostrar disposto a ajudar com excelentes contribuições.

Aos amigos da PUC, Igor Nogueira, Laura Menezes, Isabel Almeida, Francisco José, Camila Gomes e Khriissy Medeiros por todo o companheirismo durante o doutorado. Em especial ao Richard Bryan por ter sempre ideias brilhantes e me ajudar tanto nas automatizações das rotinas desenvolvidas neste trabalho. E ao Bernardo Amaral pela excelente contribuição na criação das imagens sintéticas utilizadas neste trabalho.

À Vale pela bolsa concedida, disponibilização das amostras e principalmente pela oportunidade de contribuir com o importante projeto que deu origem à esta tese de doutorado. Em especial ao Alei Domingues por toda a assistência prestada durante todo o desenvolvimento do trabalho e importantes contribuições sobre o conhecimento do material utilizado.

Ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), pela contribuição financeira que possibilitou a dedicação integral ao programa de pós-graduação, e à PUC-Rio pela bolsa de isenção de mensalidades.

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

## Resumo

Caldas, Thalita Dias Pinheiro; Paciornik, Sidnei; Augusto, Karen Soares. **Desenvolvimento de uma metodologia para caracterização de fases no pellet feed utilizando microscopia digital e aprendizagem profunda.** Rio de Janeiro, 2023. 117p. Tese de Doutorado – Departamento de Engenharia Química e de Materiais, Pontifícia Universidade Católica do Rio de Janeiro.

O minério de ferro é encontrado na natureza como agregado de minerais, dentre os principais minerais presentes em sua composição estão: hematita, magnetita, goethita e quartzo. Dada a importância do minério de ferro para a indústria, há um crescente interesse por sua caracterização com o objetivo de avaliar a qualidade do material. Com o avanço de pesquisas na área de análise de imagens e microscopia, rotinas de caracterização foram desenvolvidas utilizando ferramentas de Microscopia Digital e Processamento e Análise Digital de Imagens capazes de automatizar grande parte do processo. Porém esbarrava-se em algumas dificuldades, como por exemplo identificar e classificar as diferentes texturas das partículas de hematita, as diferentes formas de seus cristais ou discriminar quartzo e resina em imagens de microscopia ótica de luz refletida. Desta forma, a partir da necessidade de se construir sistemas capazes de aprender e se adaptar a possíveis variações das imagens deste material, surgiu a possibilidade de estudar a utilização de ferramentas de *Deep Learning* para esta função. Este trabalho propõe o desenvolvimento de uma nova metodologia de caracterização mineral baseada em *Deep Learning* utilizando o algoritmo *Mask R-CNN*. Através do qual é possível realizar segmentação de instâncias, ou seja, desenvolver sistemas capazes de identificar, classificar e segmentar objetos nas imagens. Neste trabalho, foram desenvolvidos dois modelos: Modelo 1 que realiza segmentação de instâncias para as classes compacta, porosa, martita e goethita em imagens obtidas em Campo Claro e o Modelo 2 que utiliza imagens adquiridas em Luz Polarizada Circularmente para segmentar as classes monocristalina, policristalina e martita. Para o Modelo 1 foi obtido F1-score em torno de 80% e para o Modelo 2 em torno de 90%. A partir da segmentação das classes foi possível extrair atributos importantes de cada partícula, como distribuição de quantidade, medidas de forma, tamanho e fração de área. Os resultados obtidos foram muito promissores e indicam que a metodologia desenvolvida pode ser viável para tal caracterização.

## Palavras - chave

Minério de ferro; Caracterização Microestrutural; Microscopia Digital; Análise de Imagens; Aprendizagem Profunda; Segmentação de Instâncias.

## Abstract

Caldas, Thalita Dias Pinheiro; Paciornik, Sidnei; Augusto, Karen Soares. **Development of a methodology for phase characterization in pellet feed using digital microscopy and deep learning**. Rio de Janeiro, 2023. 117p. Tese de Doutorado – Departamento de Engenharia Química e de Materiais, Pontifícia Universidade Católica do Rio de Janeiro.

Iron ore is found in nature as an aggregate of minerals. Among the main minerals in its composition are hematite, magnetite, goethite, and quartz. Given the importance of iron ore for the industry, there is a growing interest in its characterization to assess the material's quality. With the advancement of image analysis and microscopy research, characterization routines were developed using Digital Microscopy and Digital Image Processing and Analysis tools capable of automating a large part of the process. However, it encountered some difficulties, such as identifying and classifying the different textures of hematite particles, the different shapes of its crystals, or discriminating between quartz and resin in optical microscopy images of reflected light. Therefore, from the need to build systems capable of learning and adapting to possible variations of the images of this material, the possibility of studying the use of Deep Learning tools for this function arose. This work proposes developing a new mineral characterization methodology based on Deep Learning using the *Mask* R-CNN algorithm. Through this, it is possible to perform instance segmentation, that is, to develop systems capable of identifying, classifying, and segmenting objects in images. In this work, two models were developed: Model 1 performs segmentation of instances for the compact, porous, martite, and goethite classes in images obtained in Bright Field, and Model 2 uses images acquired in Circularly Polarized Light to segment the classes monocrystalline, polycrystalline and martite. For Model 1, F1-score was obtained around 80%, and for Model 2, around 90%. From the class segmentation, it was possible to extract important attributes of each particle, such as quantity distribution, shape measurements, size, and area fraction. The obtained results were very promising and indicated that the developed methodology could be viable for such characterization.

## Keywords

Iron ore; Microstructural Characterization; Digital Microscopy; Image Analysis; Deep Learning; Instance Segmentation.

# Sumário

1 Introdução	17
2 Objetivos	20
3 Revisão Bibliográfica	21
3.1. Minério de Ferro	21
3.2. Caracterização Microestrutural e Microscopia Ótica	26
3.3. Conceitos Básicos de Inteligência Artificial	32
3.4. Redes Neurais Convolucionais	35
3.5. <i>Mask R-CNN</i>	40
3.6 Métricas de Validação	42
4 Materiais e Métodos	44
4.1. Aquisição de Imagens	44
4.2. Descrição das Classes	44
4.3. Descrição da Hierarquia dos Modelos	48
4.4. Construção do Banco de Dados	51
4.4.1. Rotulagem Imagens BF	51
4.4.2. Rotulagem Imagens CPOL	54
4.5. Descrição do Banco de Dados	56
4.5.1. Criação de Imagens Sintéticas	57
4.6. Descrição das etapas de treinamento	59
5 Resultados e Discussão	61
5.1. Resultados Modelo 1	61
5.1.1. Desenvolvimento do Modelo	61



5.1.2. Pós-processamento	82
5.1.3. Extração de Atributos	86
5.2. Resultados Modelo 2	91
5.2.1. Desenvolvimento do Modelo	91
5.2.2. Pós-processamento	102
5.2.3. Extração de Atributos	104
5.3. Utilização do Sistema Desenvolvido	109
6 Conclusões	111
7 Referências Bibliográficas	113

## Lista de figuras

Figura 1 - Produtos obtidos do beneficiamento do minério de ferro classificados de acordo com sua granulometria. Adaptado de [15].	22
Figura 2 - Representação simplificada das etapas de formação das pelotas. Adaptado de [16].	22
Figura 3 - Imagem de amostra de minério de ferro contendo as principais partículas delineadas e rotuladas com sua respectiva classificação. Em A) imagem obtida por Campo Claro e em B) imagem obtida por Luz Polarizada Circularmente.	25
Figura 4 - Imagem de uma amostra de minério de ferro obtida em modo BF no MO.	27
Figura 5 - Imagem de uma área de amostra de minério de ferro obtida em a) modo BF b) modo CPOL no MO.	28
Figura 6 - Representação simplificada da formação de mosaicos. a) ladrilhos em captura, as setas vermelhas indicam a direção de aquisição à medida que a platina se movimenta b) concatenação dos ladrilhos após a captura completa.	29
Figura 7 - Demonstração da tentativa de segmentação de partículas de hematita compacta (mais brilhantes na imagem). Em A) a imagem original e em B) a tentativa de segmentar o limiar mais alto (próximo a 255 no histograma).	30
Figura 8 - Demonstração simplificada da estrutura de um neurônio biológico. [37]	33
Figura 9 - Demonstração esquemática simplificada de um neurônio artificial. Adaptado de [37]	34
Figura 10 - Representação de uma RNA simples com uma camada de entrada, duas ocultas e uma de saída. Adaptado de [38]	34
Figura 11 – Representação de Classificação de imagens com RNA	35
Figura 12 – Representação esquemática das camadas básicas de uma CNN.	35
Figura 13 - Exemplo ilustrativo de uma operação de convolução entre uma imagem I e um kernel K gerando um mapa de características $I*K$ . [38]	36
Figura 14 - Representação gráfica da operação de <i>max pooling</i> em uma imagem 4x4 utilizando um kernel 2x2.	37
Figura 15 - Exemplificação dos tipos de segmentação semântica e de instâncias.	40

Figura 16 - Representação esquemática de uma matriz de confusão.	42
Figura 17 - Exemplos e descrição das características das classes de partículas minerais de interesse para o Modelo 1 (imagens BF).	46
Figura 18 - Exemplos e descrição das características das classes de partículas minerais para o Modelo 2 (imagens CPOL).	47
Figura 19 - Exemplo de imagem de entrada utilizada para A) Modelo 1A que tem as partículas de Hematita Compacta como classe de interesse e B) para o Modelo 1B que tem as partículas Porosas e Martitas como classes de interesse.	48
Figura 20 - Exemplo de imagem de entrada utilizada para o Modelo 2 que tem as partículas Monocristalinas, Policristalinas e Martitas como classes de interesse.	49
Figura 21 - Fluxograma representativo da hierarquia dos modelos desenvolvidos.	50
Figura 22 - Interface do VIA com uma imagem rotulada e com as respectivas classes atribuídas pelo operador.	52
Figura 23 - Representação de um dos possíveis erros da rotulagem automática. Em A) a rotulagem incorreta e em B) visualização das partículas, que são separadas e foram incluídas em um mesmo rótulo.	53
Figura 24 – Processo de rotulagem realizado nas imagens CPOL com base na rotulagem das imagens BF. A) rotulagem transferida da imagem BF para a CPOL com as respectivas associações de classes realizadas, e Martita (Modelo1) em preto – mostrada com a seta vermelha. B) imagem CPOL corretamente rotulada para o Modelo 2.	55
Figura 25 - Exemplo de imagem sintética criada e completamente rotulada.	58
Figura 26 - Exemplo de imagem de inferência obtida a partir do Modelo 1B.	60
Figura 27 - Exemplos de FN da classe Compacta. Em A está apresentado o <i>ground truth</i> referente à inferência em B e em C está apresentado o <i>ground truth</i> referente à inferência em D.	63
Figura 28 - Matriz de Confusão que relaciona a predição do modelo com o <i>ground truth</i> do Treinamento 1.	64
Figura 29 - Exemplos de FP da classe Goethita. A coluna da esquerda apresenta o <i>ground truth</i> e a da direita apresenta a inferência.	65
Figura 30 - Exemplo do pré-processamento realizado para a retirada da Goethita. Em A está apresentado um exemplo de imagem original e em B a imagem pré-processada com todos os pixels com tonalidade de Goethita colocados em preto.	66
Figura 31 - Matriz de confusão obtida para a validação do Treinamento 2.	67

Figura 32 - Matriz de Confusão referente à validação do Treinamento 4.	69
Figura 33 - Exemplo de imagem inferida pelo modelo obtido a partir do Treinamento 4, contendo uma partícula de martita não identificada	70
Figura 34 - Exemplo de imagem inferida pelo modelo obtido a partir do Treinamento 5.	71
Figura 35 – Imagem de validação inferida pelo modelo obtido a partir do Treinamento 5. Comparação por meio da sobreposição da inferência sobre a rotulagem.	72
Figura 36 - Gráfico comparativo das métricas gerais obtidas para as validações dos treinamentos 5, 6 e 7.	73
Figura 37 - Exemplos de partículas das classes Porosa e Martita que causaram dúvidas durante a rotulagem das imagens.	74
Figura 38 - Imagem de validação inferida pelo modelo obtido a partir do Treinamento 8. Comparação por meio da sobreposição da inferência sobre a rotulagem.	75
Figura 39 - Gráfico comparativo das métricas gerais obtidas para as validações dos treinamentos 5, 6, 7, 8, 9 e 10.	77
Figura 40 - Imagem inferida pelo modelo obtido a partir do Treinamento 12 para observação da quantidade de acertos referentes a classe de interesse do modelo: Compacta.	78
Figura 41 - Exemplo de inferência realizada pelo modelo obtido a partir do Treinamento 12 em uma imagem com problema de foco em (a), e o resultado da imagem inferida sobreposta ao <i>ground truth</i> em (b).	79
Figura 42 - Imagem inferida pelo modelo obtido a partir do Treinamento 13 que considerou as classes Goethita e background.	80
Figura 43 - Exemplificação das etapas do Modelo 1. A) a imagem original BF, B) a imagem inferida pelo modelo 1B identificando a classe Compacta (em vermelho), C) a imagem com as Compactas identificadas em preto que é <i>input</i> para Modelo 1B D) imagem inferida pelo Modelo 1B (Porosas em verde e Martitas em vermelho), E) imagem binária da classe Goethita obtida a partir da limiarização na imagem original.	81
Figura 44 – Processamento realizado após a inferência do Modelo 1A. A) Exemplo de imagem original utilizada para inferência do modelo 1A, B) imagem RGB obtida a partir da inferência do Modelo 1A, em vermelho a classe compacta, C) imagem em 8bits, D) imagem binária com as partículas da classe compacta segmentadas.	82
Figura 45 - Processamento realizado após a inferência do Modelo 1B. A) Exemplo de imagem original utilizada para inferência do modelo 1B, B) Imagem RGB obtida a partir da inferência do modelo 1B, C) Imagem	

em tons de cinza obtida a partir do canal R, D) Imagem em tons de cinza obtida a partir do canal G.	83
Figura 46 - Demonstração da segmentação da magnetita presente numa partícula de Martita.	83
Figura 47 – Comparação entre A) imagem original e B) imagem binária referente a segmentação da Goethita com base em (A).	84
Figura 48 - Comparação entre A) imagem original rotulada e B) imagem contendo todas as classes de interesse do Modelo 1 identificadas por cores diferentes conforme a legenda na imagem.	85
Figura 49 - Distribuição de Fração de Área para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme a legenda dos gráficos.	87
Figura 50 - Distribuição de AR para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme a legenda dos gráficos.	88
Figura 51 - Distribuição de <i>Feret</i> Máximo para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme a legenda dos gráficos.	89
Figura 52 - Distribuição de Circularidade para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme as legendas dos gráficos.	90
Figura 53 - Distribuição de <i>Roundness</i> para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme as legendas dos gráficos.	91
Figura 54 - Representação do processamento realizado no banco de dados para o Modelo 2. Em A) a imagem original BF utilizada no Modelo 1 e em B) a imagem CPOL utilizada para o Modelo 2 com as Martitas identificadas no Modelo 1 em preto.	92
Figura 55 - Exemplos de partículas danificadas após a limiarização da Goethita. A) e C) antes e B) e D) depois.	95
Figura 56 - Exemplo de imagem utilizada no Treinamento 3, após a limiarização da Goethita e a retirada do fundo com objetos espúrios.	96
Figura 57 - Exemplos de partículas da classe Policristalina (rotulagem em amarelo) que foram inferidas pelo modelo como Martita (máscara de segmentação em verde).	98
Figura 58 - Exemplo do pós-processamento utilizado para reduzir os ruídos de buracos nas partículas e objetos muito pequenos que são resquícios da limiarização da goethita: A) Antes do pós-processamento e B) imagem pós processada.	99

- Figura 59 - Comparação de partículas rotuladas. Antes, A) e C), e depois, B) e D), de gerar a nova rotulagem 100
- Figura 60 - Gráficos comparativos das métricas obtidas a partir das validações dos Treinamentos 1,2,3,4 e 5. As métricas individuais por classe em A) *Precision* B) *Recall* C) *F1-score* e em D) as métricas gerais. 101
- Figura 61 – Exemplificação do processamento realizado no Modelo 2. A) imagem original utilizada como input do modelo (com martitas do Modelo 1 em preto), B) imagem processada com limiarização da Goethita, retirada do fundo para redução de ruído e exclusão de objetos muito pequenos, C) input do modelo contendo as rotulagens correspondentes, D) Imagem de inferência do Modelo 2 com todas as classes de interesse identificadas corretamente conforme mostrado na legenda em cores. 102
- Figura 62 – Exemplo de pós processamento realizado após a inferência do Modelo 2 na imagem A) original, em B) Imagem RGB obtida a partir da inferência do Modelo 2; em vermelho estão as partículas da classe Monocristalina, em verde as da classe Martita e em azul as da classe Policristalina. C) Imagem em tons de cinza obtida a partir do canal R (Monocristalinas), D) Imagem em tons de cinza obtida a partir do canal G (Martitas), E) Imagem em tons de cinza obtida a partir do canal B (Policristalinas). 103
- Figura 63 - Comparação entre A) imagem original rotulada e B) imagem contendo todas as classes de interesse do Modelo 2 identificadas por cores diferentes conforme a legenda na imagem. 104
- Figura 64 - Distribuição de Fração de Área para as classes Monocristalina, Policristalina e Martita, representadas pelas cores conforme a legenda dos gráficos. 105
- Figura 65 - Distribuição de AR para as classes Monocristalina, Policristalina e Martita, representadas pelas cores conforme a legenda dos gráficos. 106
- Figura 66 - Distribuição de *Feret* Máximo para as classes Monocristalina, Policristalina e Martita, representadas pelas cores conforme a legenda dos gráficos. 107
- Figura 67 - Distribuição de Circularidade para as classes Monocristalina, Policristalina e Martita, representadas pelas cores conforme a legenda dos gráficos. 108
- Figura 68 - Distribuição de *Roundness* para as classes Monocristalina, Policristalina e Martita, representadas pelas cores conforme a legenda dos gráficos. 109

## Lista de tabelas

Tabela 1 - Descrição das métricas usadas para validar os resultados no presente trabalho. Adaptado de [42]	43
Tabela 2 - Quantidades de imagens e partículas dos <i>datasets</i> utilizados no Modelo1.	56
Tabela 3 - Quantidades de imagens e partículas de cada classe utilizadas no Modelo 2.	56
Tabela 4 - Descrição dos <i>datasets</i> utilizados para o Modelo1 incluindo as imagens sintéticas geradas.	58
Tabela 5 - Descrição das principais características dos treinamentos realizados para o desenvolvimento do Modelo 1.	61
Tabela 6 - Métricas gerais e individuais calculadas com base nos resultados obtidos para o Treinamento 1.	62
Tabela 7 – Relação das métricas gerais e individuais obtidas para o Treinamento 2.	67
Tabela 8 - Métricas gerais e individuais calculadas a partir da validação do modelo obtido do Treinamento 4.	69
Tabela 9 - Métricas gerais e individuais calculadas a partir da validação do modelo obtido do Treinamento 5.	71
Tabela 10 - Métricas gerais e individuais calculadas a partir da validação do modelo obtido do Treinamento 9.	76
Tabela 11 - Métricas gerais e individuais calculadas a partir da validação do modelo obtido do Treinamento 10.	76
Tabela 12 - Comparação entre métricas de validação dos modelos obtidos nos Treinamentos 11 e 12.	78
Tabela 13 - Descrição das principais características dos treinamentos realizados para o desenvolvimento do Modelo 2.	93
Tabela 14 - Métricas gerais e individuais obtidas a partir da validação do modelo obtido no Treinamento 1.	94
Tabela 15 - Métricas gerais e individuais obtidas a partir da validação do modelo obtido no Treinamento 2.	95
Tabela 16 - Comparação entre as métricas gerais obtidas para as validações dos Treinamentos 1,2 e 3 do Modelo 2.	96
Tabela 17 - Métricas gerais e individuais obtidas a partir da validação do modelo obtido no Treinamento 3.	97

Tabela 18 - Métricas gerais e individuais obtidas a partir da validação do Treinamento 5.



# 1

## Introdução

A produção do aço é muito importante em escala mundial dada a larga utilidade deste material em diversos setores da indústria. O aço é uma liga formada por ferro (Fe) e carbono (C), que apresenta um conjunto de propriedades interessantes, dentre elas boa resistência e versatilidade. O minério de ferro é a principal matéria-prima utilizada no processo siderúrgico, pois dele é extraído o ferro metálico para a produção do aço. Para que esteja adequado para a utilização nos fornos de redução, o minério é submetido a processos de beneficiamento que obtêm como produto o granulado e finos de minério (*sinter feed* e *pellet feed*). O granulado possui granulometria adequada para alimentar os fornos, já os finos de minério passam por processos de aglomeração para, junto com o granulado, alimentarem os fornos de redução.[1][2]

O Brasil tem a segunda maior produção de minério de ferro do mundo. O minério é encontrado na natureza como um agregado de minerais e o principal mineral de ferro presente no minério estudado nesta tese, oriundo do Quadrilátero Ferrífero (MG) e Carajás (PA), é a hematita ( $\text{Fe}_2\text{O}_3$ ), estando presentes também magnetita ( $\text{Fe}_3\text{O}_4$ ), goethita ( $\text{Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ ), e quartzo.[3] Dada a importância desse material para a indústria e o crescente interesse em avaliar sua qualidade, diversas pesquisas visando sua caracterização vêm sendo desenvolvidas. [3][4] Estas pesquisas buscam em geral avaliar a microestrutura do minério, a qual interfere no desempenho do material no interior dos fornos de redução. A presença de partículas porosas no minério, por exemplo, melhora a redutibilidade do material enquanto a presença de partículas compactas confere mais resistência ao material.

Para observar a microestrutura do minério de ferro, os pesquisadores têm utilizado ferramentas de microscopia, principalmente a Microscopia Ótica de Luz Refletida, visto que a maior parte das fases minerais deste material podem ser discriminadas visualmente por meio desta técnica. Caracterizar o minério de ferro não é uma tarefa simples, visto que sua microestrutura contém partículas de diferentes minerais, os quais apresentam características morfológicas e texturais variadas. A hematita, por exemplo, pode estar na forma de partículas compactas ou porosas. As compactas podem ser classificadas em monocristalinas e policristalinas. E as hematitas porosas são classificadas em microcristalinas e martitas. [3]

Um dos problemas da caracterização está na escolha do modo de operação do Microscópio Ótico para adquirir as imagens. Em imagens obtidas no Campo Claro, é possível observar a maioria das partículas minerais, mas não é possível diferenciar, por exemplo hematitas compactas monocristalinas de hematitas compactas policristalinas. Para que seja possível realizar esta observação, utiliza-se

luz circularmente polarizada, tornando possível a visualização dos grãos internos às partículas policristalinas. [3] Tendo isso em vista, este trabalho utilizou pares de imagens obtidas no modo Campo Claro e no modo de Luz Polarizada Circular, para ter uma visualização de todas as partículas de interesse.

Uma outra dificuldade é a automação do processo de caracterização. Durante muito tempo, técnicos especializados nesse tipo de material, observavam amostras em um Microscópio Ótico e de forma manual e subjetiva realizavam sua caracterização. Porém o ambiente industrial requer rapidez e qualidade nas análises e esses processos ficam comprometidos quando realizados somente por um operador humano. Pensando na diversidade das partículas minerais presentes no minério de ferro e na dificuldade de sua caracterização, o desenvolvimento de metodologias automatizadas passou a ser considerado. O objetivo é identificar as diferentes partículas nas imagens, atribuir cada uma delas a classes específicas e extrair medidas úteis. Essas tentativas de caracterização automática se davam tradicionalmente através de ferramentas de Microscopia Digital e Processamento e Análise Digital de Imagens.

A Microscopia Digital permite a automação de várias funções do microscópio ótico por meio do controle de *softwares*. [5] Processamento e Análise Digital de Imagens é uma área de estudo que envolve um conjunto de técnicas que permitem, realizar uma série de modificações e análises nas imagens, muitas vezes de forma automática. [6] Por meio destas técnicas tornou-se possível desenvolver rotinas capazes de analisar as partículas presentes nas imagens de minério de ferro realizando: identificação, classificação e medidas úteis para o controle de qualidade do material. Porém a utilização destes métodos apresentava uma série de limitações, como por exemplo a dificuldade de adaptação do sistema a diferentes condições de aquisição de imagem ou a grande semelhança textural entre classes diferentes de partículas.

Estas limitações estão diretamente ligadas à dificuldade de realizar uma boa segmentação das partículas. O grande problema da segmentação, principalmente dos tipos de hematita nas imagens, está na similaridade de tonalidade, que dificulta a segmentação por limiarização. Pensando em problemas relacionados com os métodos existentes para a classificação destas partículas minerais, surge a possibilidade da utilização de novas ferramentas de visão computacional baseadas em Inteligência Artificial, uma área de estudo que visa o desenvolvimento de máquinas capazes de imitar o raciocínio humano para realizar tarefas complexas. O uso destas ferramentas poderia melhorar, dentre outras coisas, a capacidade do sistema de classificar partículas de forma mais independente da diferença de tonalidade entre as partículas. A possibilidade de criação destes sistemas se dá através da Aprendizagem de Máquina (*Machine Learning*).

*Machine Learning* é uma subárea de Inteligência Artificial através da qual as máquinas são capazes de aprender a desempenhar determinada função originalmente realizada por seres humanos, como reconhecer diferentes objetos em uma imagem. Dentro da área de *Machine Learning* existem redes neurais artificiais capazes de aprender a desempenhar funções muito complexas a partir de um grande

conjunto de dados e ainda se adaptar a possíveis variações, esta subárea é denominada *Deep Learning*. Variados tipos de rede *Deep Learning* têm sido aplicados na resolução de problemas relacionados à classificação de imagens, em especial as Redes Neurais Convolucionais, que são capazes de extrair informações das próprias imagens conhecidas e desta forma, aprender a classificar novas imagens. [7][8]

O presente trabalho propõe, portanto, o desenvolvimento de uma metodologia inovadora para a caracterização de amostras de finos de minério de ferro, que englobe as seguintes partículas minerais: hematita compacta (monocristalina e policristalina), hematita porosa (martita) e goethita. Desta forma, foi utilizado um algoritmo baseado em *Deep Learning* chamado *Mask R-CNN* [9], que é o estado da arte para segmentação de instâncias. Por meio desta aplicação foi possível treinar modelos capazes de identificar as partículas de cada classe na imagem, classificá-las e realizar a segmentação. A partir da segmentação bem sucedida foi possível extrair medidas de interesse das partículas, tais como fração de área e parâmetros de tamanho e forma.

O desenvolvimento desta metodologia se deu a partir de uma parceria entre o Grupo de Análise de Imagens e Microscopia Digital da PUC-Rio e o Centro de Tecnologia de Ferrosos da empresa Vale. Para possibilitar a compreensão do conteúdo, este documento está organizado nos seguintes capítulos. O primeiro compreende a presente introdução, o segundo capítulo expõe os principais objetivos da pesquisa. O terceiro capítulo traz uma fundamentação teórica para o entendimento do problema que se busca solucionar, mostrando trabalhos relevantes para tal. No quarto capítulo é apresentada a metodologia na qual o trabalho se norteia, no quinto estão apresentados os resultados obtidos nesta pesquisa. No capítulo seis, as conclusões do trabalho e as perspectivas futuras, e por fim as referências bibliográficas utilizadas.

## 2 Objetivos

O presente trabalho tem como principal objetivo desenvolver uma metodologia de caracterização automática de minério de ferro por microscopia ótica aplicando conhecimentos de visão computacional e *Deep Learning*. Tendo em vista este objetivo principal, os específicos são:

- Obter um conjunto de treinamento representativo de todas as classes de partículas de interesse (hematita, martita e goethita) presentes no minério de ferro, utilizando imagens de MO;
- Desenvolver uma hierarquia de classes de acordo com as características observadas em cada classe de partícula, correlacionando os pares de imagem adquiridas nos modos Campo Claro e Luz Polarizada Circularmente
- Treinar e validar um sistema que se torne capaz de identificar, classificar e segmentar as partículas de interesse utilizando um algoritmo de segmentação de instâncias;
- A partir das imagens de inferência obtidas pelo modelo treinado, desenvolver uma metodologia de pós processamento para obter imagens segmentadas por classe de partícula;
- Extrair medidas quantitativas de todas as classes de partículas identificadas.

### 3

## Revisão Bibliográfica

Este capítulo apresenta um embasamento teórico para o desenvolvimento do presente trabalho. Aborda as principais publicações relacionadas a esta área de estudo, iniciando pelo material analisado nesta pesquisa, o minério de ferro.

### 3.1. Minério de Ferro

Desde a antiguidade, o homem utilizava metais para a fabricação de armas e ferramentas de trabalho. Há aproximadamente 6 mil anos surgiu a Metalurgia que se trata de um conjunto de processos físicos e químicos capaz de extrair metais a partir de minerais. Os minerais são substâncias sólidas, inorgânicas, cristalinas, com composição química e propriedades físicas bem definidas, resultantes de processos geológicos sem intervenção humana. [10][11] Neste período e durante um longo tempo que se seguiu, o bronze foi a liga metálica mais utilizada para a fabricação de utensílios – Idade do Bronze – até aproximadamente 1000 a.C., quando se dá início à Idade do Ferro. Esse processo de utilização do ferro como matéria-prima para produzir diferentes utensílios evoluiu com o tempo e então surgiu o conceito de Siderurgia.

A Siderurgia é um ramo da Metalurgia que possibilita a obtenção de produtos do aço através do minério de ferro e parte do princípio da reação de uma fonte de carbono com sopros de ar para extrair ferro a partir do minério. O minério de ferro é encontrado na natureza na forma de agregados minerais que contém um ou mais minerais economicamente e/ou tecnologicamente viáveis para a extração.[11] O aço é um material primordial para diferentes setores de produção, como transportes, setor elétrico, utensílios do cotidiano, embalagens etc. Sua produção (em usina integrada) se dá em algumas etapas: redução, refino e transformação mecânica. A primeira, trata exatamente da redução do minério de ferro para a obtenção do ferro gusa. [12][13]

Os fornos de redução são alimentados com granulados e aglomerados de minério de ferro, que se diferenciam por suas granulometrias e são obtidos a partir do tratamento do minério. O tratamento ou beneficiamento se dá por uma série de processos que buscam obter um produto adequado química, física e metalurgicamente para os processos aos quais serão submetidos posteriormente. Esse processo inicia geralmente com uma operação de cominuição do minério bruto com o objetivo de fragmentar o sólido por um conjunto de ações mecânicas, como por exemplo britagem e moagem. Depois da fragmentação, o minério passa por uma etapa de classificação que é realizada de acordo com os tamanhos de partícula. A Figura 1 apresenta os diferentes produtos do tratamento do minério de ferro. [11][14]



Figura 1 - Produtos obtidos do beneficiamento do minério de ferro classificados de acordo com sua granulometria. Adaptado de [15].

O granulado é utilizado diretamente nos fornos de redução por ter uma granulometria adequada (31,7 – 6,3 mm) para tal aplicação. Já o *sinter feed* (6,3 – 0,15 mm) e o *pellet feed* (< 0,15 mm) necessitam passar por processos de aglomeração devido às suas granulometrias muito finas. Estes são submetidos, respectivamente, aos processos de sinterização, o que dá origem ao sinter, e pelletização, que gera a pelota. Desta forma, o forno de redução é alimentado com uma mistura de granulado, sinter e pelota como matéria-prima e coque como combustível. Além de agentes redutores para produzir o metal quente que esteja adequado para o processo siderúrgico. Este trabalho analisará amostras de *pellet feed*, um dos tipos de finos de minério que não passaram por processos de aglomeração. [10] Para que o *pellet feed* esteja adequado para a pelletização o material precisa atender algumas especificações, como possuir teor total de Fe superior a 64%, por exemplo. [16]

A formação das pelotas é realizada por uma série de fenômenos que envolvem um concentrado de *pellet feed* e água, conforme mostrado na Figura 2. Em resumo, esta formação se dá predominantemente por forças capilares, e esta depende da distribuição de tamanho das partículas de *pellet feed* e do quanto estas partículas são hidrofílicas. Tendo isso em vista, é importante conhecer determinadas propriedades do *pellet feed* para a formação das pelotas são elas: distribuição de tamanho de partículas, morfologia, hidrofiliidade e tipo de minério utilizado.

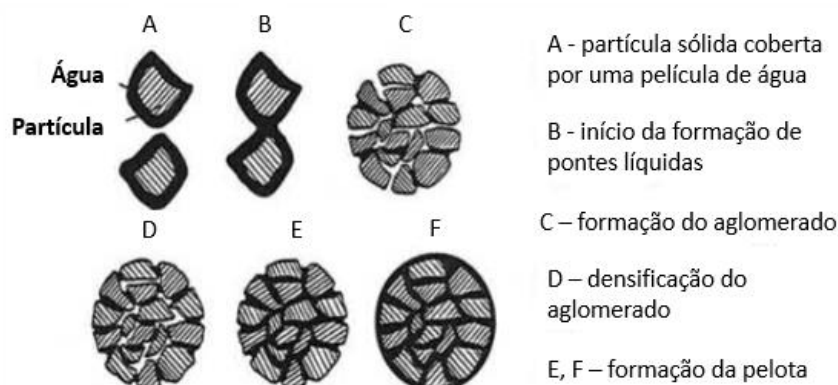


Figura 2 - Representação simplificada das etapas de formação das pelotas. Adaptado de [16].

Conforme mencionado anteriormente o tipo de minério está diretamente ligado ao desempenho do *pellet feed* no processo de pelletização. Além disso, a

qualidade dos minérios também está relacionada a características químicas (composição), físicas (granulometria) e metalúrgicas (desempenho no processo siderúrgico)[17][18]. O minério de ferro brasileiro é destaque na indústria mundial devido aos altos teores de ferro presentes em sua composição, cerca de 62%, enquanto a média do restante do mundo está em torno de 51,6%. No caso do minério de ferro, o ferro metálico é obtido principalmente a partir dos óxidos. [12][19] Os principais óxidos de ferro discutidos no presente trabalho serão a hematita ( $\text{Fe}_2\text{O}_3$ ), com teor de ferro de 69,9% e a magnetita ( $\text{Fe}_3\text{O}_4$ ), que possui teor de ferro aproximado a 72,4%. Além dos óxidos, existe a goethita ( $\text{Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ ), que é um óxido de ferro hidratado. Estes minerais são compostos basicamente por Fe, O e/ou OH, e podem ser discriminados, por seus conteúdos teóricos de cada elemento e estruturas cristalinas. Estes serão tratados como principais deste estudo pois o minério brasileiro estudado neste trabalho (oriundo das minas do Quadrilátero Ferrífero - MG e de Carajás - PA) apresenta, em geral em sua composição hematita, magnetita, goethita e alguns minerais de ganga como o quartzo. [3][16]

Para seguir nesta discussão é importante ressaltar que as nomenclaturas de partículas minerais apresentadas aqui são baseadas em classificações realizadas pelo Centro de Tecnologia de Ferrosos (CTF) na VALE e serão explicadas no decorrer do trabalho. O principal óxido de ferro presente nas amostras de minério de ferro estudadas neste trabalho é a hematita, o qual apresenta altos teores de ferro e ocorre em diferentes tipos de rocha. É um material altamente anisotrópico e apresenta birrefletância quando é observado em um microscópio ótico. [1] Existem algumas formas nas quais a hematita pode ser encontrada: partículas compactas (monocristalinas e policristalinas) ou porosas, estas classes estão relacionadas com as condições de cristalização do material e são importantes para o processamento do mesmo. [5]

A hematita compacta é um mineral de alto grau que contém excelentes características físicas e químicas, como alto teor de ferro, baixo teor de sílica e uma textura uniforme. É importante mencionar que nesta discussão o termo “textura” se refere à distribuição espacial dos diferentes tipos de minerais que podem estar contidos numa mesma partícula de minério ou rocha. [4] Estas características da hematita possibilitam que este mineral seja utilizado na obtenção do ferro via redução direta.[20] O conhecimento sobre a presença de partículas de hematita compacta monocristalina e policristalina no minério é um fator importante para o processo de cominuição do minério (britagem e moagem), ou seja, quando este minério é submetido a redução de granulometria, que é feita a partir da quebra destas partículas.

As partículas monocristalinas são compostas por um só cristal sólido e contínuo, por isso são mais duras e difíceis de quebrar do que as que apresentam descontinuidades (policristalinas). Quanto mais facilitada for a cominuição, mais eficiente é o processo e menos energia é gasta, e isso é esperado pela indústria. Estas observações são possíveis em imagens obtidas com polarização circular exatamente pela propriedade de birrefletância da hematita. Além desta classificação, os cristais internos nas partículas policristalinas podem ser classificados por sua forma em granular, lamelar ou lobular. Estas informações são

utilizadas como parâmetros de classificação por técnicos treinados nos laboratórios da Vale desde 1990. Porém, este trabalho não tratará da classificação dos grãos, apenas das partículas.

As partículas de hematita porosa podem ser microcristalinas ou martitas. Entende-se que a hematita se forma a partir da oxidação da magnetita em ambientes de superfície. Então martita é na verdade de um termo textural utilizado para se referir a hematita com hábito de magnetita. A magnetita começa a oxidar, e num processo de martitização a partícula referida como martita é formada. Nas amostras estudadas neste trabalho, não estão presentes partículas de hematita microcristalina, apenas martita. [21]

A morfologia das partículas influencia no desempenho siderúrgico do minério. Pois a presença de partículas porosas melhora a redutibilidade do minério, ou seja, ao ser submetido a redução em um forno, o minério é reduzido com mais facilidade pois os gases permeiam melhor o material. Quando as partículas compactas estão presentes conferem uma maior resistência mecânica. [21] A Figura 3 apresenta imagens obtidas em Campo Claro (Figura 3A), na qual é possível discriminar os minerais mencionados anteriormente. Na imagem, as partículas estão delineadas e rotuladas com sua respectiva classificação. Na Figura 3B é possível observar o mesmo campo da imagem obtida a partir do modo de polarização circular, no qual é possível observar, principalmente, os grãos internos nas partículas de hematita policristalina.



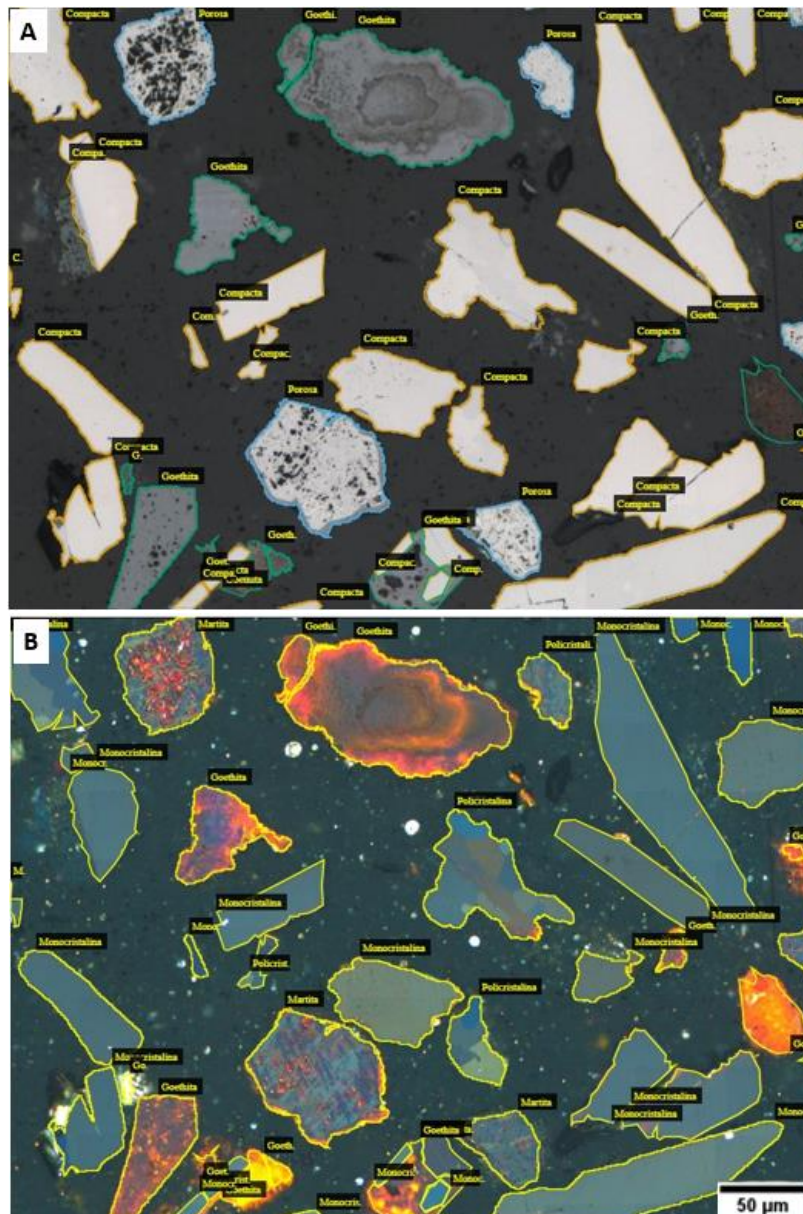


Figura 3 - Imagem de amostra de minério de ferro contendo as principais partículas delimitadas e rotuladas com sua respectiva classificação. Em A) imagem obtida por Campo Claro e em B) imagem obtida por Luz Polarizada Circularmente.

A porosidade, portanto, contribui para a redutibilidade do minério de ferro. Quanto mais poroso, mais facilmente o minério é reduzido. Consequentemente, menos energia é consumida para realizar a redução. Em contrapartida, as partículas compactas aumentam a resistência dos aglomerados. Todos os conceitos abordados, o conhecimento sobre as características dos minerais presentes no minério de ferro e sua importância para todo o processo siderúrgico, conferiram o interesse pela caracterização microestrutural desse material. Corroborando com estas afirmações, Donskoi et al. [22] citou que uma simples caracterização envolvendo a composição química ou composição mineral do minério de ferro não é suficiente para entender o comportamento do material. Pois minérios que possuem, por exemplo, mineralogias análogas, se comportam de forma diferente em um mesmo processamento, portanto

para obter uma análise mais completa é preciso considerar também as informações texturais.

Desta forma, este trabalho propõe uma caracterização de todos estes tipos de partícula citados, que foram divididos em classes minerais (hematita compacta, hematita porosa, magnetita, goethita) e classes texturais (monocristalinas, policristalinas e martita). Estas classes serão descritas com mais detalhe e exemplificadas na seção de Materiais e Métodos. As próximas seções discorrerão a respeito das técnicas comumente utilizadas para tal caracterização e sua relevância, apresentando trabalhos anteriormente publicados que abordaram este tema.

### 3.2. Caracterização Microestrutural e Microscopia Ótica

Os minerais de ferro possuem características particulares que tornam amostras de minério de ferro distintas entre si, seja dentre outros motivos, pela forma como foram processadas ou localidade da qual foram extraídas. Devido a larga utilização deste material na indústria, principalmente para a produção de aço, a busca pela melhoria da qualidade tem sido cada vez maior. Para obter informações suficientes ao controle de qualidade é necessário extrair uma série de características das amostras a serem analisadas, como por exemplo: minerais presentes, distribuição de porosidade, tamanho médio dos grãos e morfologia etc.[14][23]

O Microscópio Ótico (MO) é um instrumento provido de técnicas capazes de fornecer excelentes visualizações para posterior extração de informações desejadas para a referida caracterização. Além de ser uma técnica rápida e de baixo custo quando comparada por exemplo com a Microscopia Eletrônica de Varredura (MEV). Tipicamente, um MO de boa qualidade é capaz de atingir uma resolução espacial de aproximadamente 0,2  $\mu\text{m}$ . O sistema presente no MO é composto por um feixe de luz visível que incide sobre a amostra e um conjunto de lentes que ampliam sua visualização. Esse sistema pode trabalhar em dois modos baseados na interação da luz visível com a amostra: o modo de luz transmitida e de luz refletida. O primeiro é utilizado quando se tem necessidade de caracterizar materiais transparentes, e o segundo quando se pretende olhar para amostras opacas. No caso do minério de ferro, trabalha-se com a luz refletida, pois este modo promove uma melhor visualização das fases minerais presentes neste material. [24]

Para a formação das imagens por meio do MO é necessário preparar as amostras de forma que a superfície sobre a qual a luz incidirá esteja o mais plana possível para que seja refletida de forma especular e as imagens sejam capturadas por meio da câmera CCD (*Charge-Coupled Device*). [24] Por ser um instrumento muito versátil, o MO apresenta diversas técnicas para aquisição de imagens, e algumas das mais utilizadas em estudos relacionados à ciência dos materiais são: Campo Claro (*Bright Field* - BF), Campo Escuro (*Dark Field* - DF) e luz polarizada de forma linear (LPOL) ou de forma circular (CPOL). O modo mais amplamente utilizado é o modo BF, muito aplicado na visualização da microestrutura de metais e cerâmicas, por exemplo. Neste modo de operação, um cone de luz é focalizado na amostra e o mecanismo de contraste na imagem é dado pela absorção ou espalhamento da luz. [25] A Figura 4 exibe uma imagem de uma amostra de minério de ferro obtida em modo BF no MO. É possível perceber um bom contraste entre os diferentes tipos de minerais presentes, uma vez que cada um reflete a luz de forma específica. Em modo BF, a maioria dos objetos aparece clara sobre um fundo escuro.

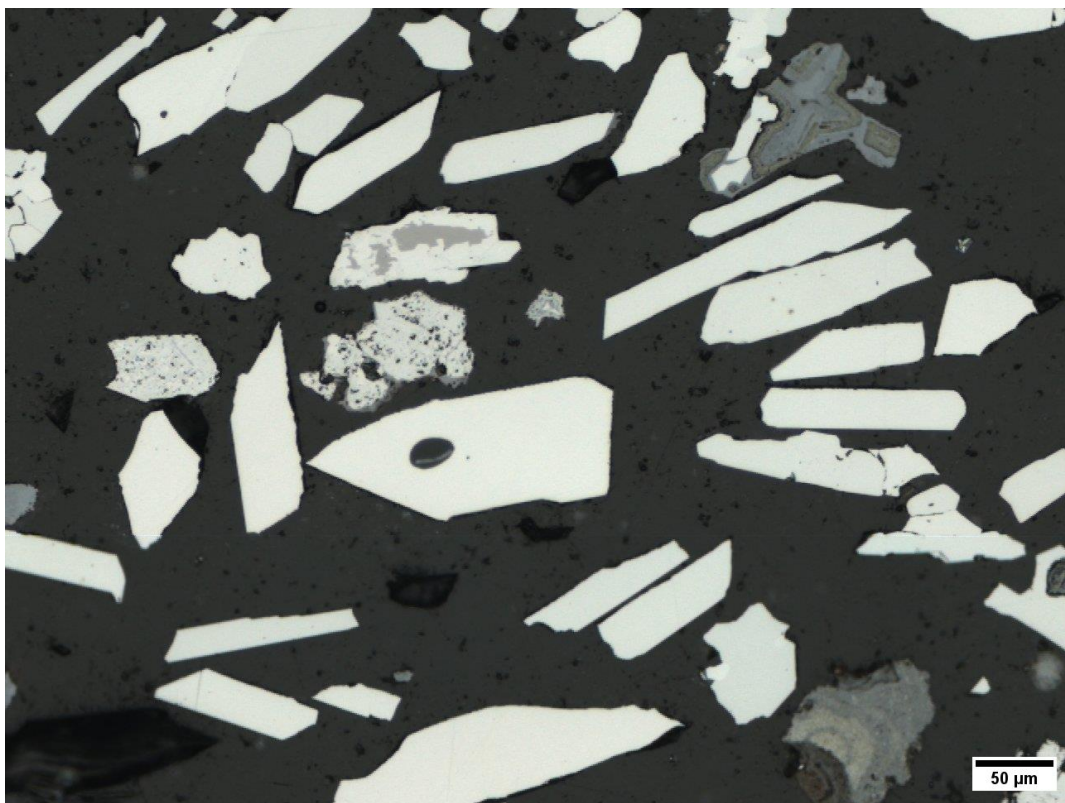


Figura 4 - Imagem de uma amostra de minério de ferro obtida em modo BF no MO.

Para o modo de operação de luz polarizada, são introduzidos dois elementos de polarização (polarizadores) no sistema ótico do microscópio. Estes polarizadores fazem com que o campo elétrico da luz, que naturalmente vibra em todos os planos, vibre em um só plano ou plano de polarização. A luz também pode ser polarizada de forma circular, o que ocorre na utilização do modo CPOL. Este modo de polarização dispõe de uma placa de quarto de onda, que alinha a onda LPOL a  $45^\circ$  em relação ao seu plano inicial de polarização. Essa onda é dividida em dois componentes com uma diferença de fase de  $90^\circ$ , e a combinação delas resulta em uma onda polarizada circularmente onde a trajetória do campo elétrico é helicoidal. [3] Os modos de luz polarizada são amplamente utilizados para o estudo de materiais anisotrópicos, tornando possível realizar discriminação entre grãos e a maneira como estão dispostos nas partículas do material. [25]

Ao observar partículas de hematita em imagens adquiridas no modo CPOL, por exemplo, é possível discriminar seus cristais e saber que se trata de uma partícula monocristalina ou policristalina, o que no modo BF não é tão visível. A Figura 5 exibe imagens de uma mesma área de uma amostra de minério de ferro obtidas em modos BF (Figura 5a) e CPOL (Figura 5b). É possível perceber, como principal diferença entre as imagens, os vários cristais de hematita bem discriminados na partícula maior da imagem obtida em modo CPOL.

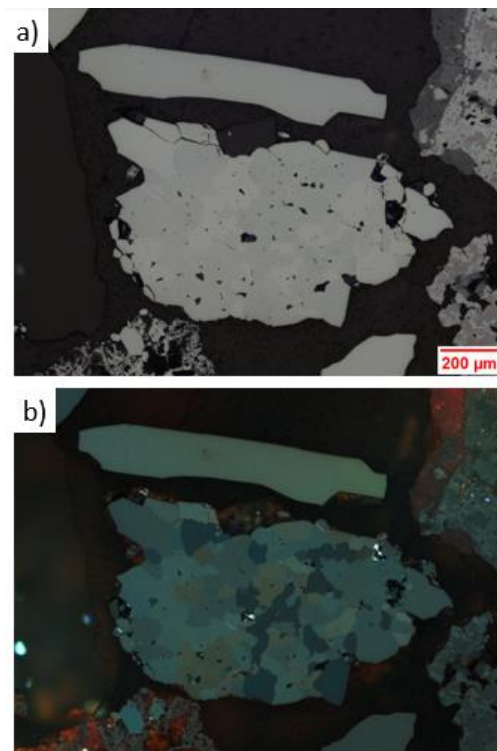


Figura 5 - Imagem de uma área de amostra de minério de ferro obtida em a) modo BF b) modo CPOL no MO.

Todos estes modos de operação apresentados podem ser combinados tornando mais completa a análise de diversos materiais, incluindo o minério de ferro. Para que estas análises sejam cada vez mais rápidas e eficientes, é possível utilizar um sistema no qual o MO é acoplado ao computador. Esse sistema controlado por softwares é chamado de Microscopia Digital (MD), que une a aquisição, automação do microscópio e análise de imagens. A MD também possibilita a redução de equívocos associados ao operador, uma vez que a visão humana é muitas vezes suscetível a erros. A utilização de MD torna a caracterização microestrutural do minério de ferro bem mais robusta. [5]

A MD permite automatizar, controlando pelo computador, a troca de lentes para capturar imagens em diferentes aumentos [5], a movimentação da platina [26], correção de fundo [27], dentre outras funções importantes para a criação de rotinas que operem cada vez mais de forma independente de um operador humano. Na aquisição de imagens de minério de ferro, por exemplo, geralmente se utiliza a MD para cobrir áreas maiores das amostras, obtendo assim uma melhor representatividade. Essa possibilidade se dá a partir da formação de mosaicos. Estes podem ser obtidos a partir do controle e motorização da platina onde está localizada a amostra. Para realizar a aquisição, o operador precisa informar ao software a área da amostra a ser capturada. A Figura 6 exibe uma representação esquemática da direção de formação de mosaicos no MO. A platina se desloca no plano x-y à medida que são capturados pequenos “ladrilhos” da área total informada pelo operador. Esses ladrilhos na verdade são capturados com um certo grau de sobreposição que não está apresentado na imagem abaixo. Depois disso, eles são



concatenados (Figura 6b) formando uma imagem representativa da área de interesse (mosaico).

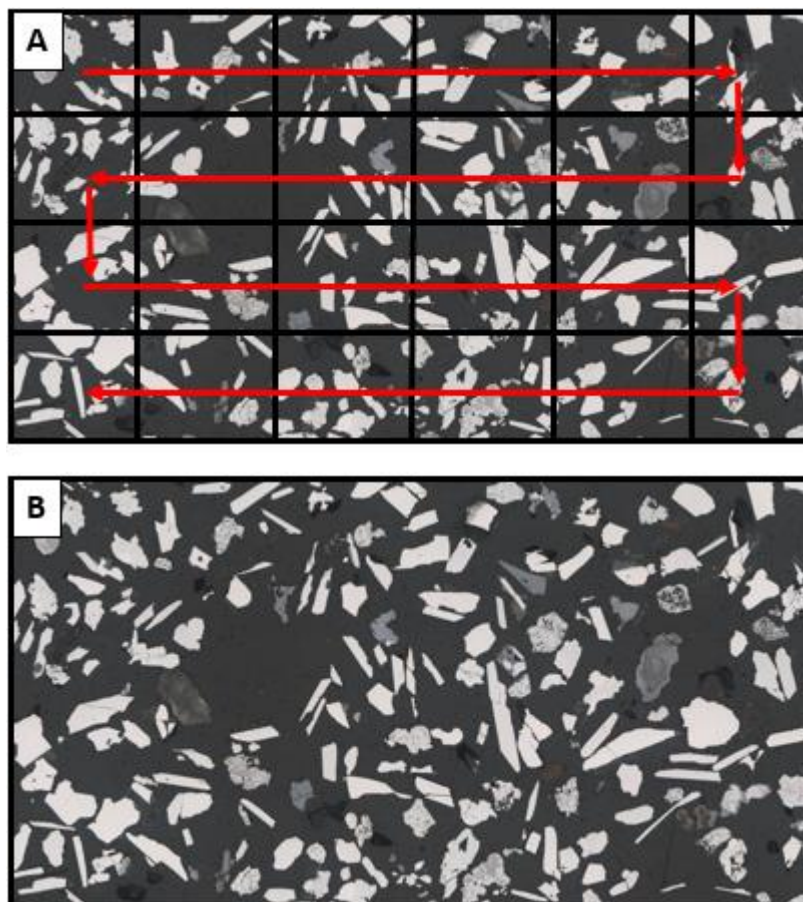


Figura 6 - Representação simplificada da formação de mosaicos. a) ladrilhos em captura, as setas vermelhas indicam a direção de aquisição à medida que a platina se movimenta b) concatenação dos ladrilhos após a captura completa.

A aquisição de mosaicos pode ser realizada tanto da forma descrita acima (*fullframe*) quanto utilizando *subframes* adquirindo subáreas de ladrilhos centralizadas. Esta segunda abordagem aproveita uma iluminação mais homogênea, o que é crítico em imagens CPOL. Sua desvantagem é o aumento do tempo de aquisição já que cada *subframe* cobre uma área muito menor. A partir de todas as ferramentas de MD descritas aliadas a ferramentas de Processamento e Análise Digital de Imagens (PADI), é possível desenvolver rotinas que automatizam grande parte da análise mineral.

Durante muito tempo a caracterização microestrutural de amostras de minério de ferro foi realizada por um operador humano que observava as imagens obtidas no MO e realizava algumas medições de forma subjetiva e suscetível a erros. A partir desta dificuldade, e tendo em vista a alta demanda do ambiente industrial, diversos estudos foram realizados buscando desenvolver metodologias automatizadas e cada vez menos dependentes do operador humano. [28] Dentre as diversas pesquisas desenvolvidas ao longo do tempo estão a elaboração de metodologias automáticas utilizando MD e PADI para a identificação e quantificação de fases do sinter em amostras de minério de ferro [27], classificação

de partículas minerais no minério de ferro combinando imagens BF e CPOL [29], classificação de tipos de hematita no minério de ferro [3], reconhecimento de grãos de hematita policristalina por análise de imagens [4], dentre outros.

Para o caso da caracterização das partículas minerais, o desenvolvimento dessas metodologias automáticas não é nada trivial. Alguns minerais presentes nesse tipo de amostra possuem tonalidades semelhantes, dificultando a segmentação automática e se torna problemático fixar um limiar capaz de distinguir esses objetos. A Figura 7 apresenta uma demonstração de tentativa de segmentação de partículas de hematita compacta por limiarização. As partículas de hematita compacta são as que possuem tonalidade mais brilhante e aspecto mais uniforme observados na Figura 7a. Estas características deveriam ser suficientes para que a segmentação destas partículas, por limiar, fosse bem sucedida. Mas o problema está no fato de que outros tipos de hematita, como as porosas observadas na imagem (indicadas pelas setas vermelhas na Figura 7a), também possuem essas características em parte da partícula. Portanto, ao tentar segmentar as compactas (Figura 7b), parte das partículas porosas também são selecionadas no mesmo limiar.

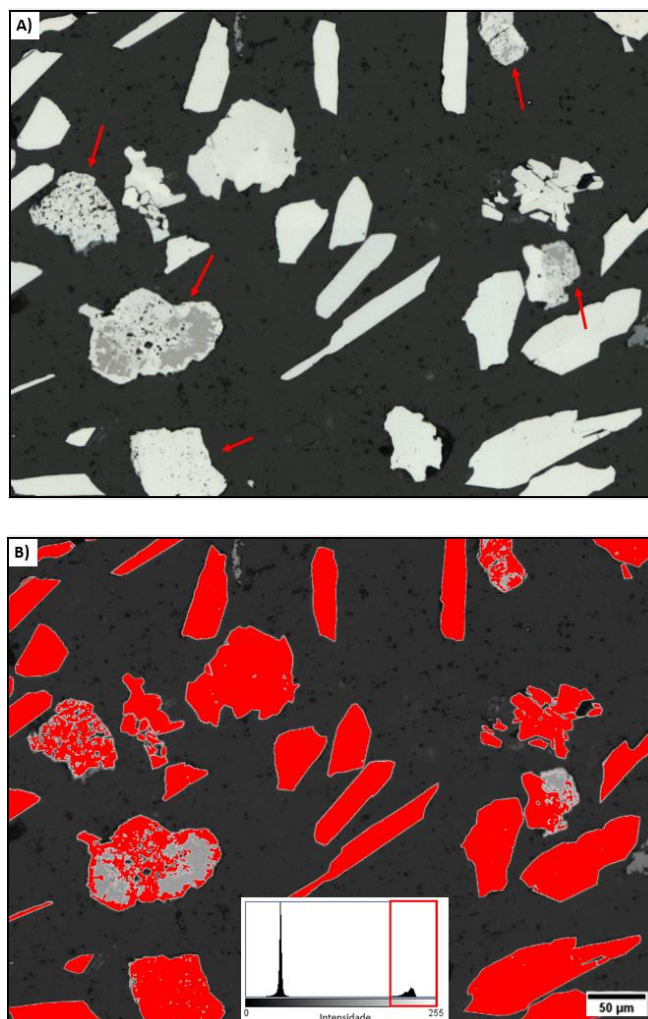


Figura 7 - Demonstração da tentativa de segmentação de partículas de hematita compacta (mais brilhantes na imagem). Em A) a imagem original e em B) a tentativa de segmentar o limiar mais alto (próximo a 255 no histograma).

Além disso, quando se deseja separar hematitas monocristalinas e policristalinas em imagens CPOL, também encontram-se problemas. Visando solucionar ambos os problemas, Iglesias et al.[29], desenvolveram uma metodologia automática para classificação de partículas minerais. A identificação dos subgrupos de hematita foi realizada por meio de classificação supervisionada utilizando parâmetros morfológicos e texturais capazes de discriminar estes subgrupos. Além de discriminar tipos de hematita, partículas de goethita e magnetita também foram identificadas, estas por limiarização. Os autores citados obtiveram com este classificador uma taxa de sucesso global superior a 90%. Trata-se sem dúvidas, de uma grande contribuição para a caracterização microestrutural de minério de ferro. Todavia, é válido ressaltar a complexidade desta rotina, que envolve uma série de medidas de parâmetros complicadas. Em alguns casos estes parâmetros podem não ser capazes de descrever as características das partículas, que são muito semelhantes e comprometer a robustez da metodologia. Portanto, os autores sugeriram que as etapas de classificação desta metodologia poderiam ser refinadas utilizando uma abordagem diferente baseada em *Deep Learning*. Esta nova abordagem possibilitaria o desenvolvimento de sistemas cada vez mais robustos e confiáveis.

Um outro problema clássico na caracterização do minério de ferro é a identificação do quartzo, que apresenta refletância especular muito semelhante à da resina epóxi presente nas amostras. Este fato impossibilita a segmentação por limiarização. E ao longo do tempo tentou-se resolver este problema de algumas formas, como por exemplo combinando imagens de MO e de Microscopia Eletrônica de Varredura (MEV) aproveitando as vantagens de cada técnica. Utilizando o MEV no modo de operação de elétrons retro-espalhados, é possível obter um bom contraste entre estas classes, uma vez que esse tipo de formação de imagem utiliza informações de número atômico e os números atômicos da resina e do quartzo são diferentes, gerando, portanto, diferentes níveis de cinza nas imagens possibilitando a segmentação. A hematita, magnetita e goethita porém, possuem números atômicos próximos, consequentemente apresentam tons de cinza semelhantes nas imagens de MEV, fatores que dificultam a segmentação destas classes nesse tipo de imagem. [30]

Deste modo, essa combinação de técnicas conhecida como Microscopia Colocalizada MO-MEV permite identificar e consequentemente analisar tanto as fases distinguíveis em MEV (resina epóxi e quartzo) quanto as distinguíveis em MO (hematita, magnetita e goethita). Porém, existem algumas dificuldades, os tipos de hematita, por exemplo não são totalmente distinguíveis por limiarização em imagens MO. Além disso, utilizar dois instrumentos diferentes para a aquisição de imagens não é tão simples, além de suscetível a erro humano ao deslocar a amostra, por exemplo. Isso compromete a aquisição de uma mesma região da amostra nas duas técnicas distintas. Embora tenham sido desenvolvidos métodos automáticos para alinhar as imagens obtidas, pode haver também problemas de custo e disponibilidade de equipamento. [31]

Outros métodos foram sugeridos para a resolução deste problema. Delbem et al. [32] desenvolveram uma metodologia automática capaz de fazer a distinção

entre quartzo e resina baseando-se no realce de bordas das partículas de quartzo e se mostrou eficiente para as amostras por eles analisadas. Porém é limitada pelo fato de que na grande maioria das imagens de amostras de minério de ferro, as partículas de quartzo não apresentam bordas bem definidas. Iglesias et al. [31] observaram que, embora a segmentação de quartzo e resina em imagens de MO seja complicada de ser realizada pelos métodos de limiarização conhecidos, o ser humano especializado em reconhecer esse tipo de material consegue, muitas vezes sem tanto esforço, diferenciar as duas classes. Tendo isso em mente, os autores sugeriram uma nova abordagem que parte do princípio de treinamento de um sistema baseado em *Deep Learning* por um operador humano capaz de realizar esta identificação. A possibilidade se mostrou viável neste e em outros estudos que serão abordados com mais detalhes na próxima seção, após a introdução de alguns conceitos importantes.

Na próxima seção serão discutidos conceitos básicos de inteligência artificial para que seja mais compreensível a utilização dessas técnicas no presente trabalho.

### 3.3. Conceitos Básicos de Inteligência Artificial

A IA é uma área de estudo que aplica teorias e técnicas ao desenvolvimento de máquinas capazes de simular a inteligência humana. Se dedica, portanto, a criar sistemas que desempenhem tarefas complexas com nível similar ao que um ser humano seja capaz de realizar. [33] O desenvolvimento desses sistemas é possível por meio de máquinas que aprendem a desempenhar as funções estipuladas e a tomar decisões a partir de dados fornecidos. A Aprendizagem de Máquinas ou *Machine Learning* (ML) se dá exatamente pelo treinamento de sistemas que se tornam capazes de analisar dados e aprender com eles. Dentro da área de ML, uma subárea de IA, está a Aprendizagem Profunda ou *Deep Learning* (DL). [34]

Nos últimos anos, a área de DL ganhou bastante reconhecimento devido a avanços promissores, principalmente no campo de visão computacional. Tecnologias baseadas neste princípio possuem diversas aplicações cotidianas, como aplicativos de celular, redes sociais, reconhecimento facial, mecanismos de busca, fluxo de veículos no trânsito etc. Este campo de pesquisa já existe há bastante tempo, mas devido às limitações computacionais, era uma área que não tinha tanta visibilidade. Só na última década as pesquisas e aplicações cresceram de forma mais rápida, principalmente devido ao desenvolvimento de poderosas unidades de processamento gráfico (GPUs) capazes de processar muitos dados em paralelo. Isso foi fundamental, já que estas redes exigem a utilização de GPUs por utilizarem um grande conjunto de dados. [35]

As redes DL são compostas por algoritmos conectados formando uma Rede Neural Artificial (RNA) complexa semelhante às conexões neurais do cérebro humano. Estas redes são extremamente adaptáveis a várias complexidades de dados e regras. Uma característica muito importante das redes DL é a capacidade de aprender de forma autônoma características necessárias para reconhecer padrões, diferente de ML que depende de um operador para determinar quais características são importantes e desenvolver uma forma de extraí-las e inserir estas informações na rede. É importante ressaltar também que, enquanto no cérebro biológico



qualquer neurônio pode se conectar a qualquer outro neurônio em uma certa distância física, as RNAs possuem diferentes camadas, conexões e direções de propagação de dados. Essa característica é responsável pelo aumento do uso de métodos de redes DL na identificação e classificação de imagens.

Para proporcionar uma melhor compreensão sobre DL, alguns conceitos precisam ser abordados, como o de RNAs, que são algoritmos matemáticos inspirados no funcionamento do cérebro humano. Essas redes aprendem a desempenhar tarefas por meio de treinamentos. Numa RNA, várias unidades de processamento denominadas neurônios artificiais desempenham algumas tarefas de forma aproximadamente análoga aos neurônios biológicos presentes no cérebro humano. [33][8]

A Figura 8 apresenta de forma simplificada as principais partes de um neurônio biológico. São elas: o corpo celular, do qual se ramificam os dendritos que transmitem sinais provenientes das extremidades para o corpo celular; e o axônio que é uma ramificação mais extensa a qual conduz os sinais do corpo celular até as extremidades. Os dendritos são terminais de entrada e o axônio é um terminal de saída de informações. Através das extremidades do axônio os sinais são transmitidos aos outros neurônios (sinapse). [36] [37]

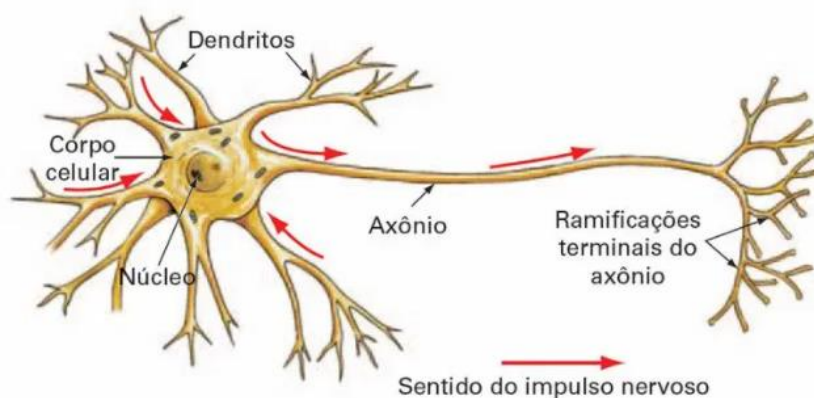


Figura 8 - Demonstração simplificada da estrutura de um neurônio biológico. [37]

O neurônio artificial basicamente recebe um conjunto de entradas, realiza alguns cálculos matemáticos e dá uma saída.[34] Conforme mencionado anteriormente, a estrutura deste neurônio é inspirada no neurônio biológico. A Figura 9 exibe uma representação esquemática do neurônio artificial. Nesta representação as informações que “chegam” de outros neurônios são evidenciadas pelos dados de entrada ( $x$ ). As sinapses são simuladas pelos pesos ( $w$ ), que podem ser excitatórios ou inibitórios sobre os dados de entrada resultando na ativação ou não do neurônio. O corpo celular é representado pelo somador ( $\Sigma$ ) e o axônio pela saída ( $y$ ).

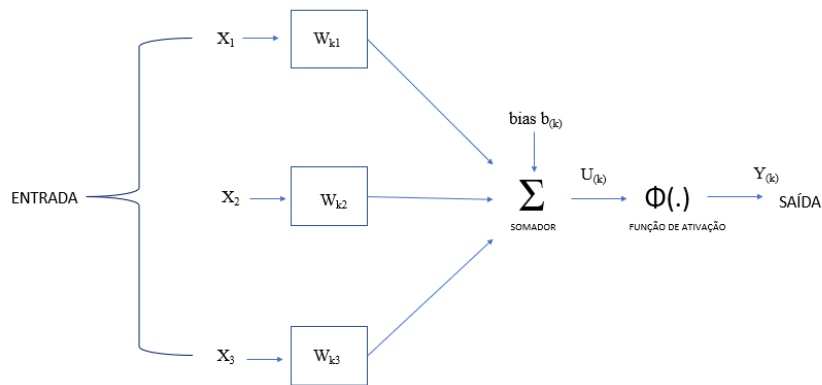


Figura 9 - Demonstração esquemática simplificada de um neurônio artificial. Adaptado de [37]

De forma resumida, a função  $\Sigma$  realiza um somatório dos estímulos ( $x$ ) multiplicado pelos pesos ( $w$ ) e posteriormente uma função de ativação determinará, qual será a saída ( $y$ ). Estes são chamados de parâmetros da rede neural, que são ajustáveis, meio através do qual uma rede neural aprende.[34] A função de ativação determina a saída do neurônio, existem alguns tipos de função de ativação, dentre elas: a função degrau que aproxima de 1 todo valor maior ou igual a 0 (zero), *Rectified Linear Unit* (ReLU), que mantêm os valores de saída, *Softmax* que converte os resultados obtidos fornecendo como saída probabilidades (em %), dentre outras.

Uma RNA é formada pela conexão de vários neurônios artificiais, de forma que a saída de um neurônio é a entrada de um subsequente formando uma rede com camadas para processar informações. A Figura 10 exhibe a estrutura simples de uma RNA com uma camada de entrada, duas camadas ocultas e uma de saída. Nas camadas ocultas são realizadas todas as ações e cálculos com base nos dados de entrada para fornecer dados de saída. As RNAs podem possuir números variados de camadas ocultas.[38]

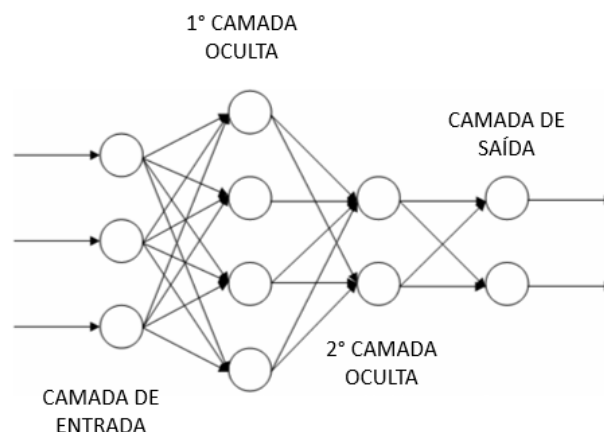


Figura 10 - Representação de uma RNA simples com uma camada de entrada, duas ocultas e uma de saída. Adaptado de [38]

As redes DL têm a capacidade de aprender tarefas complexas a partir de treinamentos utilizando dados de entrada brutos.[2] Supondo que se deseja classificar imagens utilizando DL, a Figura 11 exibe uma demonstração de como funciona esta classificação. A coluna amarela representa alguns pixels da imagem (dados de entrada), o kernel 4x4 representa os pesos, e a coluna magenta os *Bias* que serão somados ao resultado da multiplicação matricial. Neste exemplo, depois de obtidos os resultados, é aplicada a função de ativação *softmax* que transforma os resultados em probabilidade atribuindo a imagem à determinada classe, com base neste valor probabilístico.

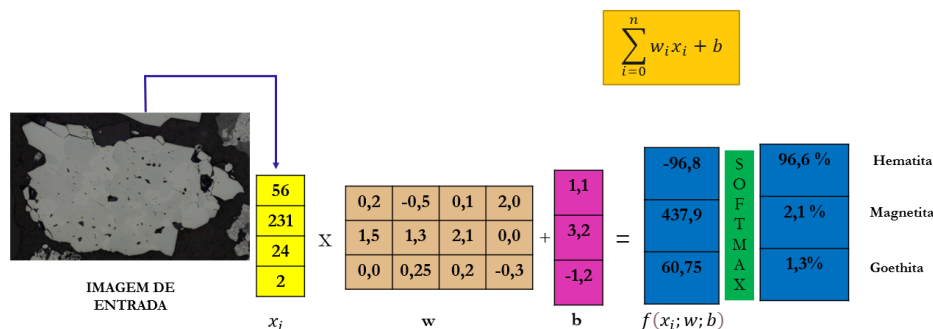


Figura 11 – Representação de Classificação de imagens com RNA

Um exemplo de redes DL muito utilizadas na área de imagens são as Redes Neurais Convolucionais (*Convolutional Neural Networks – CNNs*). As CNNs são um tipo de RNA, que realiza em pelo menos uma de suas camadas, uma operação de convolução. Este é o principal diferencial das CNNs em relação as RNAs que realizam multiplicação matricial comum. As CNNs têm sido aplicadas para desempenhar diversas funções incluindo a classificação de imagens e tem fornecido ótimos resultados na resolução de problemas enfrentados nesta área de estudo. [2][7]

### 3.4. Redes Neurais Convolucionais

A estrutura básica de uma CNN é composta por camadas subsequentes que se repetem, são elas: camada convolucional, camada de *pooling* e camada totalmente conectada.[2] Conforme apresentado na Figura 12.

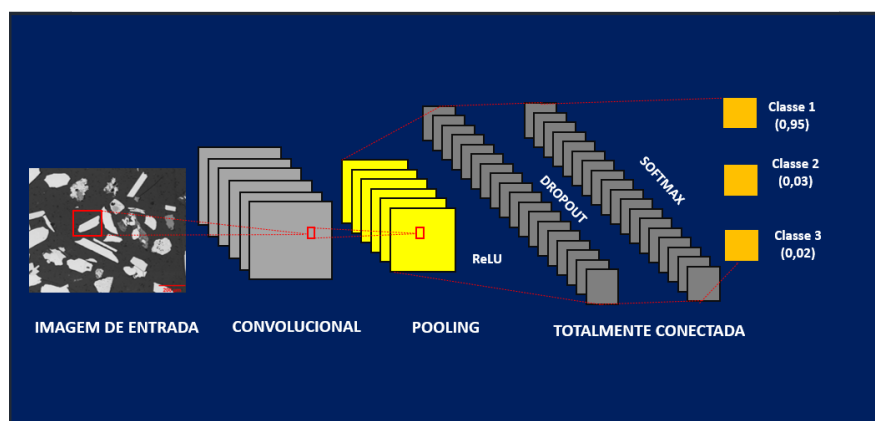


Figura 12 – Representação esquemática das camadas básicas de uma CNN.

Na representação esquemática também é possível notar além das funções de ativação anteriormente mencionadas (ReLU e *Softmax*) a função *Dropout*, geralmente utilizada para resolver problemas de *overfitting* – quando o sistema se adequa excessivamente ao conjunto de dados. Ao se trabalhar com imagens, os dados de entrada fornecidos para esse tipo de rede profunda são as próprias imagens (matrizes de pixels). Na camada convolucional é realizada uma operação matemática de convolução, definida por:

$$(H)_{ij} = (W_k * x)_{ij} + b_k \quad (1)$$

Na equação (1), “x” representa os dados de entrada, “W<sub>k</sub>” são os pesos de filtros lineares e “b<sub>k</sub>” são *bias* para o mapa de características, e a saída é exatamente o resultado da equação. O símbolo “\*” indica que não se trata de uma multiplicação matricial comum, mas de uma convolução. [2] Nesta camada o aprendizado de vários filtros (*kernels*) se dá de maneira que cada um deles é capaz de extrair informações importantes das imagens de entrada e a operação de convolução garante que essas informações serão levadas à próxima camada. Na Figura 13 está exposta de forma gráfica uma operação de convolução entre uma imagem de entrada (I) e um kernel (K), resultando em um mapa de características (I\*K) que carrega informações da imagem de entrada.

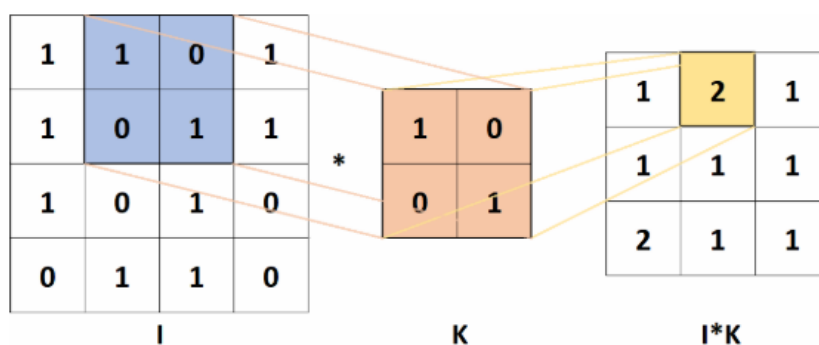


Figura 13 - Exemplo ilustrativo de uma operação de convolução entre uma imagem I e um kernel K gerando um mapa de características I\*K. [38]

A operação de convolução dispõe também de alguns parâmetros que podem ser ajustados a fim de avaliar e/ou melhorar os resultados da rede. São eles: tamanho do kernel e seus coeficientes, os “passos” (*stride*) por meio dos quais o kernel se move sobre a matriz de entrada, e o preenchimento (*padding*) que corresponde a adição de bordas à matriz para aumentar suas dimensões. [38] Conforme observado na Figura 12, após a camada de convolução na CNN existe a camada de *Pooling*. Nesta camada ocorre uma modificação em determinados locais da matriz que sai da camada de convolução por meio de um agrupamento que geralmente diminui suas dimensões. Existem algumas formas de agrupamento, o mais comum é o *max pooling*, onde apenas o valor máximo de uma dada área da matriz é representado na resultante (representação gráfica na Figura 14).

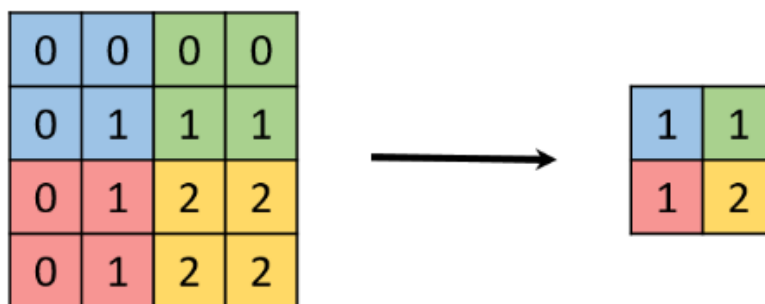


Figura 14 - Representação gráfica da operação de *max pooling* em uma imagem 4x4 utilizando um kernel 2x2.

As saídas das camadas convolucional e *pooling* carregam características importantes das imagens de entrada. E na camada totalmente conectada (Figura 12) estas informações são utilizadas como parâmetros para atribuir essas imagens a classes previamente estipuladas no modelo. Normalmente também se tem uma função de ativação associada à camada totalmente conectada que introduz uma operação não linear ao sistema. A função de ativação mais comum em CNNs é a ReLU, que embora seja não-linear conserva informações lineares úteis.

Para que uma CNN aprenda a desempenhar determinada função é preciso submetê-la ao processo de aprendizado, que no caso deste trabalho será realizado de forma supervisionada. Para o caso específico deste trabalho, essa forma de aprendizado possui duas etapas básicas: treinamento e teste. Ao se trabalhar com imagens, determina-se um conjunto de imagens representativas de todas as classes que se deseja identificar e esse conjunto é fornecido à rede durante o treinamento. Além disso, um conjunto para teste, ambos também contendo imagens representativas das classes, e particularmente, o de teste deve possuir imagens desconhecidas para sistema. Os dois conjuntos possuem a imagem original rotulada com o *ground truth* (valor real conhecido, ou seja, objetos de interesse delineados com suas respectivas classes associadas). O treinamento de uma CNN é iniciado gerando de forma aleatória os valores dos pesos. Posteriormente esses valores são ajustados com o objetivo de melhorar as taxas de acerto da rede (diminuindo o erro). A forma mais comum de realizar esse tipo de ajuste é através do algoritmo *backpropagation*.

O funcionamento do *backpropagation* pode ser descrito nas seguintes etapas:

- (1) Valores de filtros da camada convolucional e pesos da totalmente conectada são gerados aleatoriamente;
- (2) Ao receber uma imagem como entrada, a rede realiza o processo de cálculo através das diferentes camadas, obtendo no final um valor de probabilidade para cada classe. É válido mencionar que as imagens de entrada são objetos rotulados com as respectivas classes;
- (3) Um erro total é calculado na camada de saída, esse erro corresponde à relação entre o valor esperado e o obtido;
- (4) O *backpropagation* calcula os valores de gradiente do erro, e a partir destes, ajusta filtros e pesos na proporção que contribuíram para minimizar o erro total. Esse ajuste é o que permite que o sistema aprenda. O cálculo deste erro pode ser realizado por meio das funções de erro (*Loss Function*). É por meio da minimização do erro que os algoritmos de treinamento são capazes de modificar os pesos da rede

neural, ou seja, de treinar o modelo; (5) As etapas do 2 ao 4 são repetidas para cada imagem do conjunto de treinamento determinado pelo operador que está treinando a rede. [39]

Quando as etapas do 2 ao 4 forem repetidas para todas as imagens do conjunto de treinamento, pode-se dizer que o sistema concluiu um ciclo de treinamento, também conhecido como época. O treinamento da rede é realizado por diversas épocas até que o valor do erro seja menor que um determinado limiar ou até que um número máximo de épocas estipulado seja atingido. E a partir disso, os valores de filtros e pesos estarão ajustados para classificar de forma eficiente novas imagens (como as do conjunto de teste). Existe um algoritmo de otimização denominado Gradiente Descendente, que é responsável por ditar quais variáveis contribuem para a minimização da função de erro, utilizando então o seu gradiente. E esse processo é realizado repetidas vezes até que se atinja o menor valor possível para a função de erro. Desta forma, este algoritmo juntamente com o *backpropagation* é utilizado para compor o método de treinamento de redes neurais com muitas camadas, como é o caso das CNN's.

Quando se trabalha com redes neurais são utilizados hiperparâmetros, que são parâmetros cujos valores são definidos antes do processo de aprendizagem. Estes parâmetros determinam como a rede é treinada e a sua estrutura. Dentre vários hiperparâmetros importantes estão o *learning rate* e *batch size*. O *learning rate* pode ser definido como uma taxa de atualização dos pesos no treinamento. Esta taxa é utilizada para ditar o ritmo no qual o algoritmo regula os pesos da rede em relação a *loss function*. O *batch size* é utilizado pelo fato de as redes DL trabalharem com um extenso conjunto de dados, então estas informações são divididas para simplificação do processo. Este parâmetro, portanto, representa o número de dados utilizados por interação. Quando *batch size* equivale a 1, por exemplo, o número de interações é equivalente ao número de épocas. [40]

É importante ser cauteloso na escolha da quantidade de imagens para a etapa de treinamento, pois pode influenciar no aprendizado da rede. Se o treinamento é realizado com um conjunto pequeno de imagens em muitas épocas, por exemplo, existe a possibilidade de ocorrer *overfitting*. Ou seja, a rede se adapta excessivamente às imagens fornecidas no treinamento e perde um pouco a capacidade de generalização, o que não é desejável visto que se deseja tornar o sistema o mais robusto possível. E se o treinamento for realizado com poucas épocas, o modelo não será complexo o suficiente para representar os dados (*underfitting*).

Na etapa de teste são comparadas as previsões de classes de pixels fornecidas pela rede (valor predito) com imagens desconhecidas do treinamento, mas rotuladas com o *ground truth*. A comparação é avaliada pelo operador por meio de observações qualitativas nas imagens, mas também pode ser realizada de forma quantitativa analisando as métricas obtidas, que serão discutidas no item 3.6 Métricas de Validação.

As CNNs têm apresentado um ótimo desempenho em tarefas de visão computacional, incluindo a classificação de imagens. Conforme mencionado

anteriormente, Iglesias et al.[31] realizaram o treinamento de uma CNN para resolver um problema clássico na caracterização de minério de ferro que é a discriminação de quartzo e resina, essas duas classes possuem tonalidades muito próximas nas imagens, o que é problema para realizar uma limiarização, por exemplo. A dificuldade da limiarização é superada, utilizando uma CNN, uma vez que essa etapa não é necessária, pois a rede desempenha a função de extrair informações importantes das imagens e de classificá-las. O banco de imagens para treinamento e validação da rede do estudo mencionado, foi criado por um operador especializado que realizava recortes em imagens de minério de ferro adquiridas em MO de regiões correspondentes à resina e de partículas inteiras de quartzo. A rede utilizada possuía duas camadas convolucionais: a primeira com 32 kernels 3x3 e a segunda com 64 kernels 3x3, ambas seguidas da função de ativação ReLU. A camada de *pooling* com kernels 2x2 e função de ativação *maxpooling*. Por fim, a camada totalmente conectada possuía dois neurônios de saída representando as duas classes que se desejava discriminar. Após as etapas de treinamento e teste, os autores obtiveram taxas de sucesso global superiores a 95% garantindo que as CNN podem compor um método eficaz para a resolução deste problema.

O estudo citado é sem dúvidas um enorme passo para a resolução deste problema de classificação de imagens minerais utilizando CNNs. No entanto, este estudo realiza apenas uma classificação dos recortes correspondentes às classes. Ou seja, diz a probabilidade de determinado recorte corresponder a classe quartzo ou resina. O sistema não realiza, por exemplo uma detecção dos objetos na imagem inteira ou uma segmentação. Existe uma pesquisa em desenvolvimento pelo Grupo de Análise de Imagens e Microscopia Digital (MicDigi) da PUC-Rio, que busca o aperfeiçoamento deste caso específico (quartzo X resina) utilizando um algoritmo capaz de realizar segmentação de instâncias. Com base nisso e para seguir a presente discussão é preciso definir dois termos importantes: segmentação semântica e segmentação de instâncias.

O termo segmentação semântica é aplicado para algoritmos que são capazes de classificar cada pixel da imagem como pertencente a uma determinada classe, não discriminando objetos, apenas pixels em classes. Já algoritmos que realizam segmentação de instâncias, localizam, classificam e segmentam objetos.[41] A Figura 15 apresenta uma demonstração do resultado esperado para cada tipo de segmentação citado.

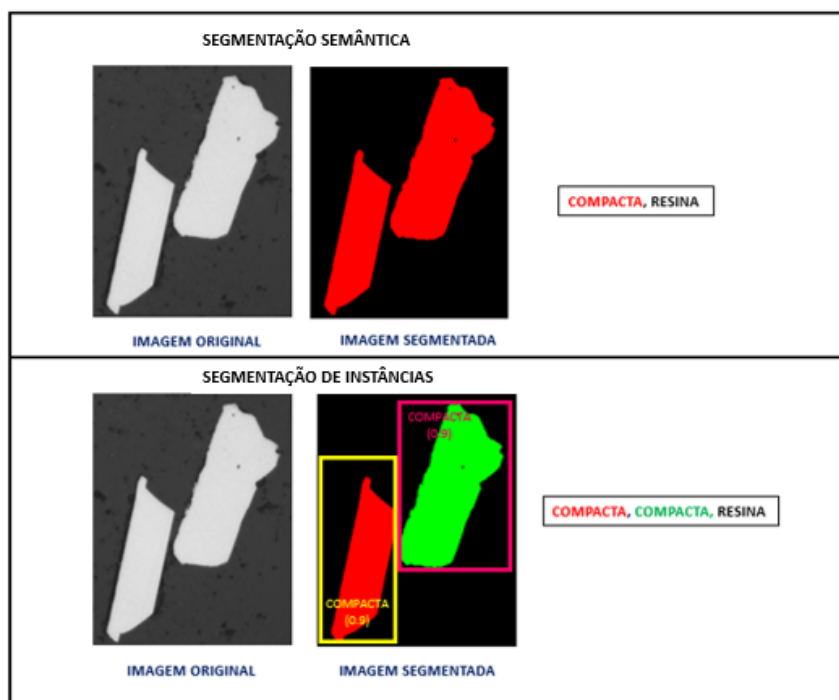


Figura 15 - Exemplificação dos tipos de segmentação semântica e de instâncias.

Santos et al. [42] desenvolveram um método automático baseado em *Deep Learning*, abordando segmentação de instâncias para classificação de diferentes classes de partículas em imagens petrográficas de macerais de carvão. Estes macerais são de difícil segmentação automática por tonalidade pois as principais diferenças entre as classes estão exatamente em informações texturais e morfológicas. Dentre outros, esse foi um dos motivos da escolha da aplicação de segmentação de instâncias para este problema. Os autores construíram um extenso banco de dados de imagens contendo as classes de interesse e treinaram modelos baseados no algoritmo *Mask R-CNN* para identificar estas classes. A partir desta abordagem foi possível obter taxas de acerto superiores a 80% para o reconhecimento das classes dos macerais de carvão. Estas partículas foram identificadas, classificadas e segmentadas nas imagens. O estudo é uma das referências de metodologia para este trabalho.

Com base nas discussões anteriores, o presente trabalho faz uma aplicação de segmentação de instâncias nas imagens de *pellet feed*, ou seja, os minerais de ferro serão identificados, classificados e segmentados nas imagens. Portanto, utilizou-se um algoritmo chamado *Mask R-CNN* que é o estado da arte para segmentação de instâncias [9].

### 3.5. *Mask R-CNN*

O objetivo principal da utilização do algoritmo *Mask R-CNN* neste contexto é tornar o sistema capaz de localizar cada partícula mineral na imagem inteira delimitando suas fronteiras, classificá-los como objetos individuais, e gerar uma máscara de segmentação para cada partícula. Tudo isso se torna possível principalmente a partir da detecção de objetos, que é uma forma de localizar objetos



distintos determinando a menor caixa delimitadora (*bounding box*) que os circunscreve. A partir disso é possível associar cada objeto a uma classe e obter máscaras de segmentação sobre eles.

O algoritmo *Mask R-CNN*, assim como *Faster R-CNN* [43] e *Fast R-CNN* [44] é um aperfeiçoamento da *R-CNN* [45], e apresenta um sistema mais rápido, completo e intuitivo que os anteriores. A *R-CNN* é uma arquitetura de rede focada em detecção de objetos, a qual trabalha de forma que são sugeridas propostas de região ou regiões de interesse (ROIs) nas imagens fornecidas ao treinamento. Estas regiões propostas são regiões relevantes da imagem que podem conter objetos, elas são então avaliadas por CNNs independentes para que sejam discriminadas em classes. Esta arquitetura foi aprimorada para trabalhar com mais rapidez, pois processar cada ROI separadamente, deixava o processo lento.

A *Fast R-CNN*, portanto, processa toda a imagem e as ROIs são obtidas por um mapa de recursos gerado por uma CNN. Estas ROIs são agrupadas, e é obtido um tamanho fixo para cada proposta. Esta mudança aumentou a acurácia da detecção assim como a velocidade de todo o processo. Após este avanço, foi introduzida a *Faster R-CNN* que propõe uma rede de propostas de região (RPN), gerando as propostas diretamente no mapa de recursos da CNN. A RPN compartilha camadas convolucionais com a rede que detecta os objetos, e isso também reduz a sobrecarga computacional. Adicionalmente, a *Mask R-CNN* fornece uma máscara de segmentação sobre as instâncias das imagens. A estrutura do algoritmo em questão, consiste em algumas partes: *Backbone* (este trabalho emprega uma ResNet 101 - 101 camadas convolucionais - em seu *backbone*), *Feature Pyramid Network* (FPN) [46], Classificação de ROI, e Máscara de segmentação. [47]

- *Backbone*: uma CNN padrão que extrai informações importantes das imagens e gera um mapa de características;
- FPN: extrai recursos padrão utilizando as camadas convolucionais do *backbone*, adiciona uma segunda pirâmide que pega os recursos de alto nível da primeira e os passa para as camadas inferiores;
- RPN: uma rede neural encontra rapidamente áreas na imagem nas quais provavelmente existem objetos de interesse;
- Classificação de ROI: a partir das ROIs determinadas na RPN as regiões são categorizadas em classes específicas e a caixa delimitadora é aperfeiçoada de modo a se aproximar do tamanho do objeto. A adaptação da caixa delimitadora ao tamanho do objeto gera saídas de tamanho variável e para não haver problema na etapa seguinte, é realizado um redimensionamento do mapa gerado para que se trabalhe com tamanhos fixos;
- Máscara de segmentação: são geradas máscaras sobre os objetos de interesse classificados.

### 3.6 Métricas de Validação

Para o caso deste trabalho foram utilizadas as métricas: *Precision*, *Recall* e *F1-score*. Para que estas sejam compreensíveis existem alguns conceitos importantes a serem apresentados. As métricas relacionam aquilo que foi rotulado nas imagens (*ground truth*), com aquilo que a rede foi capaz de identificar com o modelo gerado (predição da rede).

Desta forma, se numa comparação o *ground truth* e a predição da rede são verdadeiros, o resultado é Verdadeiro Positivo (VP). Se algo na imagem não foi rotulado e não foi identificado pela rede, ou seja, o *ground truth* e a predição da rede são negativos, o resultado é Verdadeiro Negativo (VN). Se algo foi rotulado, mas não foi identificado pela rede, o resultado é Falso Negativo (FN). E por fim, se algo não foi rotulado, mas identificado pela rede, o resultado é Falso Positivo (FP). A Figura 16 apresenta um exemplo de matriz de confusão relacionando os resultados descritos.

		Valor Real	
		Positivo	Negativo
Valor Predito	Positivo	VP	FP
	Negativo	FN	VN

Figura 16 - Representação esquemática de uma matriz de confusão.

Além da observação da matriz de confusão também é possível avaliar a classificação realizada por meio do cálculo de métricas de validação. Estas métricas podem ser utilizadas para avaliar, desde classificações mais simples, biclasse por exemplo, quanto para mais complexas que é o caso da metodologia em desenvolvimento no presente trabalho. Isto é avaliado em imagens baseando-se na mesma relação que define se um resultado é VP, VN, FP ou FN. Esta relação é fundamentada no princípio da *Intersection over Union* (IoU), que vai relacionar o desempenho da rede a partir da interseção dos objetos segmentados (predição da rede) e *ground truth*, sobre a união dos mesmos fatores, conforme apresentado na equação abaixo:

$$IoU = \frac{\text{objeto segmentado} \cap \text{ground truth}}{\text{objeto segmentado} \cup \text{ground truth}} \quad (2)$$

Desta forma, os acertos e erros da rede são quantificados relacionando os VP, VN, FN, FP com o princípio do IoU. Para este trabalho foi definido um IoU de 0,5. Com base nessas informações, agora é possível definir as métricas utilizadas neste trabalho, são elas:

Tabela 1 - Descrição das métricas usadas para validar os resultados no presente trabalho. Adaptado de [42]

Métrica	Definição	Descrição
<b>Precision</b>	$Precision = \frac{VP}{VP + FP} =$ $= \frac{\text{objetos detectados corretamente}}{\text{todas as detecções feitas}}$	Indica o quanto os objetos detectados pelos modelos pertencem ou se aproximam das classes dadas na rotulagem.
<b>Recall</b>	$Recall = \frac{VP}{VP + FN} =$ $= \frac{\text{objetos detectados corretamente}}{\text{todo o ground truth}}$	A taxa de VP no modelo. Indica quanto do <i>ground truth</i> foi detectada.
<b>F1- score</b>	$F1score = \frac{2 \times (Recall \times Precision)}{Recall + Precision}$	Representa a média harmônica entre <i>Precision</i> e <i>Recall</i> .

Neste trabalho serão consideradas tanto as matrizes de confusão dos treinamentos, quanto as métricas descritas na Tabela 1. As métricas foram importantes para que os resultados obtidos fossem compreendidos e avaliados, desta forma contribuindo para as tomadas de decisão e desenvolvimento da metodologia. É válido ressaltar, também, que as métricas serão apresentadas de forma geral e individual (por classe).

## 4

## Materiais e Métodos

A metodologia deste trabalho foi desenvolvida em algumas etapas que serão descritas nesta seção.

### 4.1. Aquisição de Imagens

As imagens utilizadas neste trabalho foram adquiridas na pesquisa de Iglesias et al., [29]. Segue abaixo uma breve descrição das amostras e do método de aquisição utilizados, com base em seu estudo.

As amostras foram fornecidas pela Vale. São amostras de finos de minério de ferro, todas de *pellet feed*. Estas foram embutidas em resina epóxi gerando amostras com 30 mm de diâmetro. Foram lixadas e polidas de forma que estivessem adequadas para a aquisição de imagens no MO. O instrumento utilizado para a aquisição das imagens foi um MO AxioImager M2m – Zeiss dotado de uma câmera digital Zeiss AxioCam MRc5 (até 2584 x 1936 pixels de resolução) disponível no Laboratório de Microscopia Digital (LMD) do Departamento de Engenharia Química e de Materiais (DEQM) da PUC-Rio. As imagens foram adquiridas utilizando uma lente objetiva de 20X (EC Epiplan 20 X/0,40) em RGB e 24 bits obtendo uma resolução espacial de 0,53  $\mu\text{m}$ /pixel.

A aquisição das imagens foi realizada por meio da formação de mosaicos de forma automática controlada pelo software AxioVision 4.9.1. Neste processo de aquisição, o operador apenas posiciona a amostra no microscópio e informa ao software o tamanho do mosaico. As condições de iluminação são previamente definidas na rotina, podendo ou não ser ajustadas. Os mosaicos BF foram formados por imagens *full frame* 6x3 (1292x968 pixels) e os mosaicos CPOL por *subframes* 49x25 (200x200 pixels).

### 4.2. Descrição das Classes

Neste trabalho foram desenvolvidos dois modelos principais de classificação das partículas minerais. Um que utiliza imagens BF, o Modelo 1, e um que utiliza imagens CPOL, o Modelo 2. As classes a serem estudadas inicialmente no Modelo 1 seriam: Hematita Compacta, Martita, Goethita, Microcristalina e Quartzo. Mas percebeu-se que para caso específico das amostras de *pellet feed* estudadas neste trabalho não estavam presentes partículas de hematita Microcristalina e havia uma quantidade muito pequena de partículas da classe Quartzo. Então, para este trabalho, decidiu-se trabalhar com: Hematita Compacta, Martita e Goethita.

Um fator decisivo para a organização da hierarquia dos modelos, foi a observação das partículas de Martita. Inicialmente, todas as Hematitas que tinham aparência mais porosa nas imagens BF, eram incluídas na classe Martita, mas

observando os pares de imagens percebeu-se que quando observadas em CPOL algumas destas partículas consideradas Martita na BF, eram na verdade Hematitas Policristalinas. Por esse motivo, para o Modelo 1 decidiu-se dividir estas partículas em duas classes: Porosa e Martita. A Figura 17 exibe variados exemplos destas duas classes.

Em estudos iniciais realizados com base na observação destas classes, partículas de Porosas e Martitas pertencem à mesma classe. Entendia-se como partículas de Hematita Porosa as partículas de Hematita que passaram/passam por processo de martitização, no qual a Magnetita inicia uma oxidação e dá origem à Martita. Porém nem todas as partículas Porosas devem ser classificadas como Martita, pois muitas já passaram pelo processo completo de martitização, apresentando na imagem CPOL características de uma Hematita Policristalina. Em outros casos, as partículas Porosas ainda apresentam trama de Martita quando observadas na imagem CPOL. Por isso decidiu-se organizar estas partículas nas seguintes classes mostradas na Figura 17 para o Modelo 1 (imagens BF) e na Figura 18 as classes de interesse do Modelo 2 (imagens CPOL). A automatização da diferenciação destas classes é de extrema importância, uma vez que anteriormente essa análise visual era realizada pela VALE de forma manual a partir da avaliação dos pares BF/CPOL.





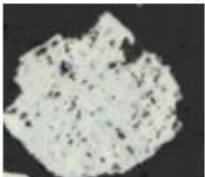
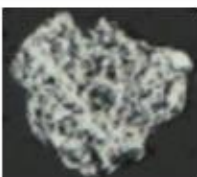

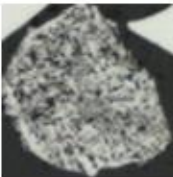




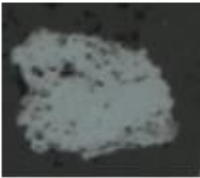



Classe (Modelo 1)	Descrição	Exemplos			
Compacta	Partículas de hematita muito brilhantes e de aspecto homogêneo.				
Porosa	Partículas de hematita com poros e sem presença perceptível de coloração de magnetita.				
Martita	Partículas porosas de hematita com forte presença de magnetita.				
Goethita	Partículas porosas de cor predominantemente acinzentada.				

Figura 17 - Exemplos e descrição das características das classes de partículas minerais de interesse para o Modelo 1 (imagens BF).






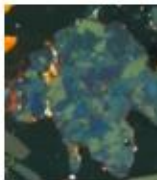
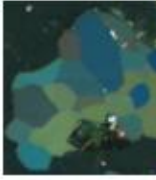

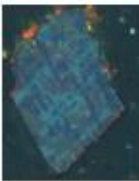
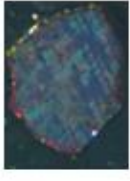
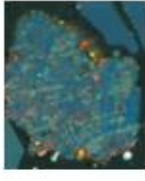
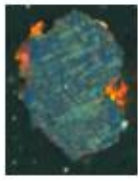
Classe (Modelo 2)	Descrição	Exemplos			
Monocristalina	Partículas de hematita em que um único cristal pode ser observado.				
Policristalina	Partículas de hematita nas quais é possível observar dois ou mais cristais.				
Martita	Partículas que apresentam traços característicos de martização.				

Figura 18 - Exemplos e descrição das características das classes de partículas minerais para o Modelo 2 (imagens CPOL).

É importante ressaltar que para desenvolver esta metodologia as características mineralógicas das partículas foram respeitadas, mas o que se levou em consideração foi principalmente as características de suas imagens: padrões que se repetem em determinadas partículas como textura, cor e formato. Estes padrões foram essenciais para a divisão das classes. Com isso, foi possível fornecer informações suficientemente processáveis pela rede DL utilizada para que o aprendizado fosse eficiente. Conforme mencionado anteriormente, foram utilizados dois tipos de imagem, BF e CPOL, e em cada tipo é possível observar características distintas das partículas. Com base nessas observações foi proposta uma hierarquia de classes que relacionasse os dois modelos, descrita a seguir.

### 4.3. Descrição da Hierarquia dos Modelos

A hierarquia descrita nesta seção foi pensada de acordo com os testes realizados e resultados obtidos. Os detalhes de construção serão descritos com mais detalhes na seção de Resultados e Discussão. Dito isto, nesta seção será descrita a hierarquia que relaciona os dois modelos, exemplificando quais tipos de imagem foram usadas em cada etapa. De acordo com o estudo das imagens e a observação das classes, entendeu-se que o Modelo 1 deveria ser dividido em dois submodelos, 1A e 1B. No Modelo 1A é realizada a identificação, classificação e segmentação (inferência) de partículas de Hematita Compacta em imagens BF utilizando imagens sem qualquer pré-processamento, conforme mostrado na Figura 19A. Uma vez que as partículas de Hematita Compacta tenham sido identificadas pelo Modelo 1A, as áreas ocupadas por elas nas imagens foram colocadas em preto para a próxima inferência no Modelo 1B (Figura 19B). Portanto, o Modelo 1B realiza inferência das partículas Porosas e Martitas em imagens BF, com áreas em preto anteriormente ocupadas pelas Hematitas Compactas. A Goethita também é identificada no Modelo 1, mas não foi desenvolvido um submodelo para inferir esta classe. Foi realizada uma segmentação por limiar para identificá-la em imagens originais sem pré-processamento como a apresentada na Figura 19a.

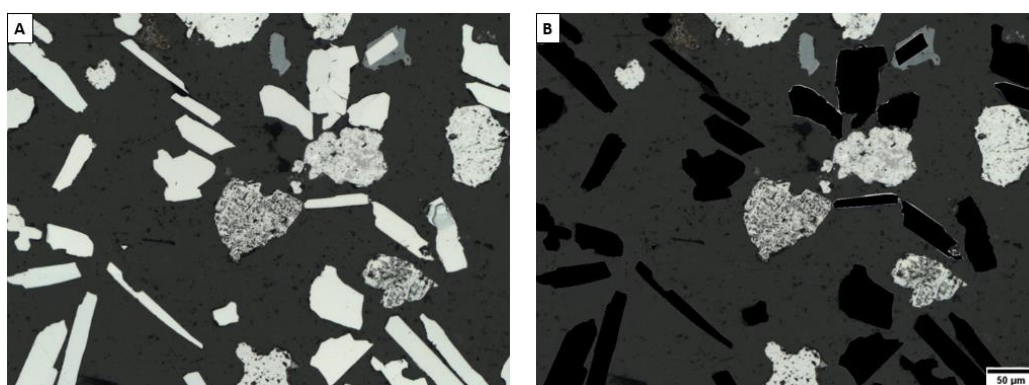


Figura 19 - Exemplo de imagem de entrada utilizada para A) Modelo 1A que tem as partículas de Hematita Compacta como classe de interesse e B) para o Modelo 1B que tem as partículas Porosas e Martitas como classes de interesse.

Diferente do Modelo 1, o Modelo 2 foi construído em uma só etapa de inferência para as classes Monocristalina, Policristalina e Martita. As imagens utilizadas nesta inferência passaram por um pré-processamento para retirar o fundo



muito ruidoso presente nas imagens CPOL e partículas muito pequenas que causavam problemas. A descrição destas etapas será feita mais a frente (5.2. Resultados Modelo 2), mas a imagem resultante utilizada como *input* para Modelo 2 é apresentada na Figura 20.

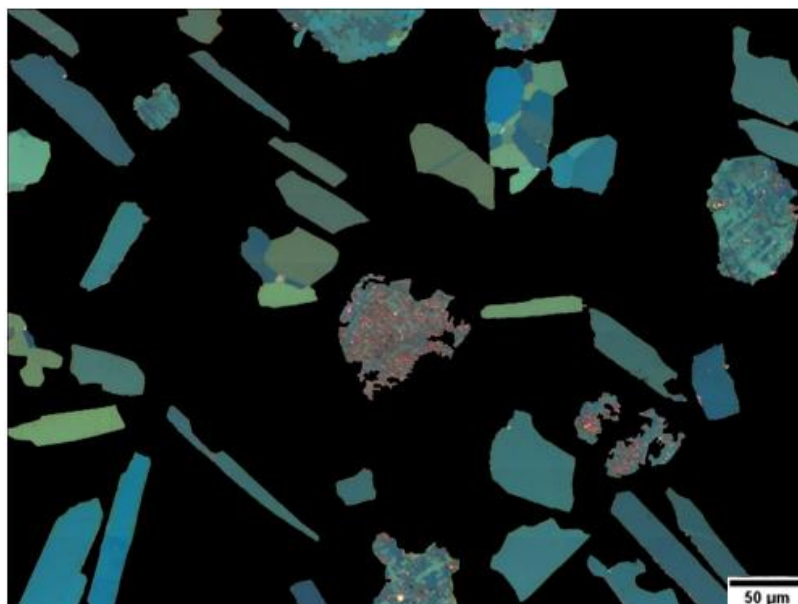


Figura 20 - Exemplo de imagem de entrada utilizada para o Modelo 2 que tem as partículas Monocristalinas, Policristalinas e Martitas como classes de interesse.

Deste modo, o Modelo 1 se baseou nas Imagens BF e identifica as Classes Compacta, Porosa, Martita e Goethita. Já o Modelo 2 utilizou as Imagens CPOL para discriminar Monocristalinas, Policristalinas e Martitas. Depois da identificação, foi realizada uma limiarização nas partículas de Martita (Modelo 1) para quantificar a presença de Magnetita. Esta lógica está representada no fluxograma da Figura 21, que apresenta um esquema detalhando todo o processamento desenvolvido para os dois modelos.

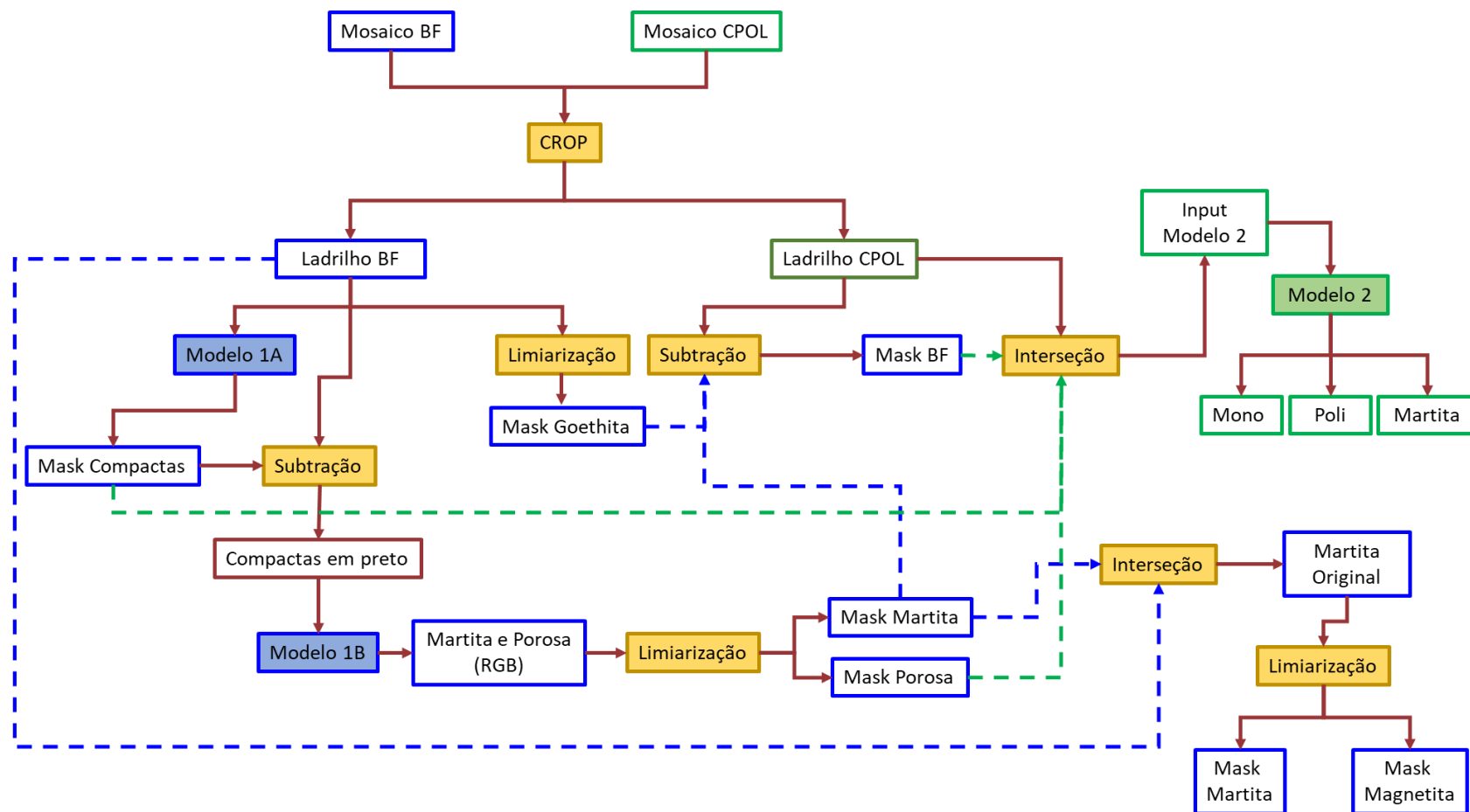


Figura 21 - Fluxograma representativo da hierarquia dos modelos desenvolvidos.

Observando o fluxograma é possível perceber que as etapas dos dois modelos estão interligadas. Conforme apresentado no fluxograma, os mosaicos BF e CPOL são divididos em ladrilhos e cada ladrilho passa por processamentos diferentes. O ladrilho BF original é segmentado pelo Modelo 1A e a partir desta identificação são obtidas máscaras das partículas da classe Compacta (*Mask Compacta*). A *Mask Compacta* é utilizada para subtrair estas partículas identificadas do ladrilho original, e gerar uma imagem com esta classe em preto. A imagem do ladrilho com Compactas em preto é segmentada pelo Modelo 1B e então são identificadas as classes Martita e Porosa, obtendo uma imagem RGB com uma cor específica para cada classe. Realiza-se então uma limiarização e são obtidas as *Masks* para cada classe. Então, é feita uma interseção da *Mask* da classe Martita com o ladrilho BF original, e são obtidas imagens das Martitas originais, nesta é possível realizar uma limiarização e obter *mask* de Martita e Magnetita. Além disso, o ladrilho original passa por uma limiarização e uma *Mask* da classe Goethita é obtida. Realiza-se uma interseção das classes Compacta e Porosa com o ladrilho CPOL e subtrai-se a *Mask* Goethita e Martita do ladrilho CPOL (obtidos na imagem BF). Obtendo então a imagem que é *input* para o Modelo 2. O Modelo 2 infere as classes Monocristalina, Policristalina e Martita.

Para que fosse possível o desenvolvimento destes modelos, foi preciso criar um banco de dados representativo de todas estas classes. Na próxima subseção está descrito como foi realizada esta construção.

#### 4.4. Construção do Banco de Dados

A criação do banco de dados é uma etapa essencial para o treinamento de uma rede DL. Nesta etapa o objetivo foi obter uma boa representatividade de todas as classes que se deseja identificar nas amostras de *pellet feed*. A construção foi feita a partir do delineamento ou marcação do perímetro de cada uma das partículas de interesse presentes nas imagens adquiridas no MO e atribuição de suas respectivas classes. Para se referir a este delineamento e atribuição de classes para as partículas no presente trabalho, será utilizado o termo rotulagem. O objetivo da rotulagem é que o maior número de partículas individuais representando cada classe seja fornecido ao sistema a ser treinado.

Para fins de organização, o banco de imagens foi criado em duas etapas gerando dois subconjuntos, com imagens referentes a:

- (1) Classes minerais em imagens BF: Compacta, Martita, Porosa e Goethita;
- (2) Classes texturais de hematita em imagens CPOL: Monocristalina, Policristalina e Martita.

##### 4.4.1. Rotulagem Imagens BF

Na presente metodologia, o processo de rotulagem foi realizado no site *VGG Image Annotator* (VIA) [48] no qual foi possível rotular as partículas e gerar um arquivo de texto contendo as coordenadas das marcações realizadas e as classes correspondentes a cada uma das partículas. Antes de realizar a rotulagem no VIA os mosaicos adquiridos em MO, que são arquivos grandes, foram divididos em

várias imagens de dimensões menores (1292 x 968). Posteriormente, as imagens resultantes foram preparadas para a rotulagem. A preparação foi realizada de forma automática por uma rotina que:

- (1) Segmenta as imagens originais de maneira a incluir a maioria das partículas presentes;
- (2) Gera imagens binárias;
- (3) Obtém as coordenadas dos objetos das imagens binárias;
- (4) A partir das coordenadas, constrói-se um arquivo de texto no formato *json* contendo informações necessárias para que o arquivo seja legível no VIA.

Ao carregar o arquivo *json* com as informações mencionadas e as imagens correspondentes no VIA, a grande maioria das partículas já aparecem delineadas e incluídas na classe Compacta por *default*, por ser a classe com maior representatividade nas amostras. Depois disso, foi necessário atribuir a classe correta a cada uma das partículas que não fazem parte da classe Compacta. A Figura 22 exibe a interface do VIA com todas as partículas de uma imagem BF rotuladas e a caixa de diálogo que possibilita ao operador realizar a atribuição das classes.

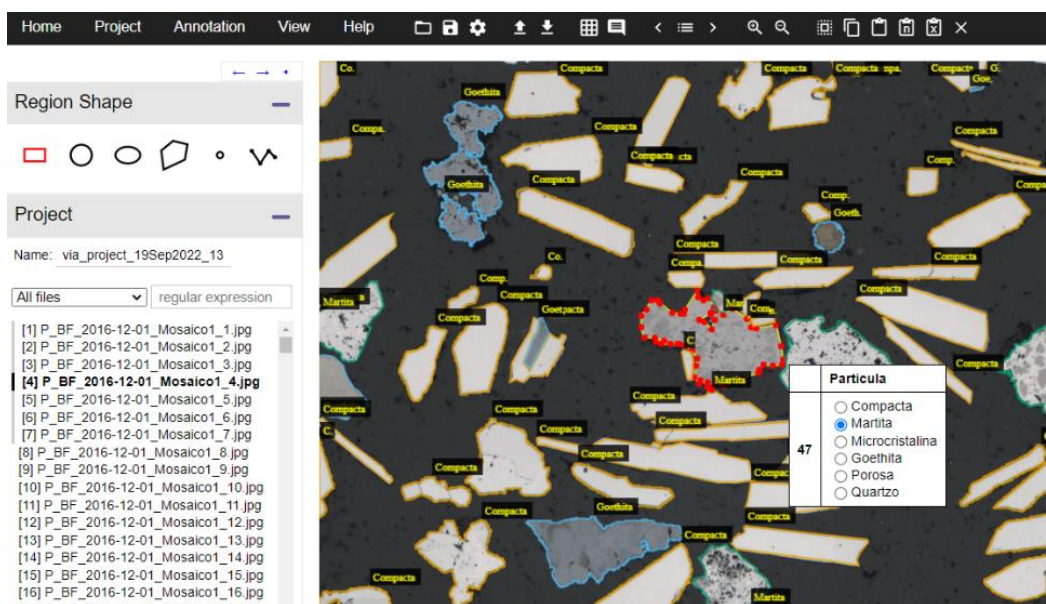


Figura 22 - Interface do VIA com uma imagem rotulada e com as respectivas classes atribuídas pelo operador.

Muitas vezes, o delineamento gerado automaticamente também precisa ser feito manualmente pelo operador com as ferramentas de desenho da interface, que podem ser observadas na parte superior esquerda da imagem acima (*region shape*). A Figura 23 apresenta um exemplo de um dos principais problemas que ocorrem: várias partículas incluídas numa mesma rotulagem. Em casos como este o operador precisa excluir a rotulagem e refazer manualmente para cada partícula. Terminada a rotulagem correta das imagens, exporta-se o arquivo *json* com as informações e

esse arquivo é fornecido juntamente com as imagens correspondentes para iniciar o treinamento da rede.

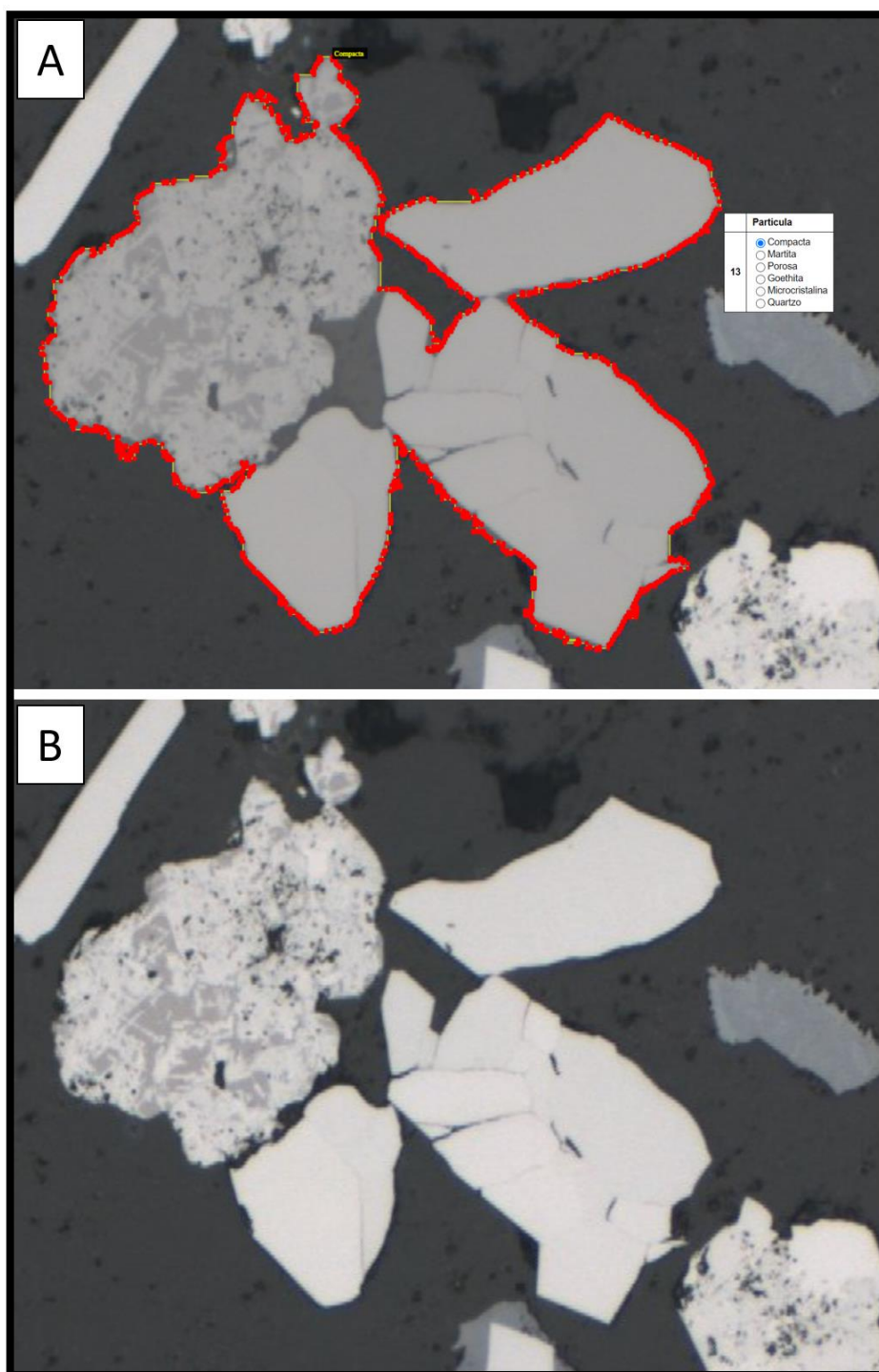


Figura 23 - Representação de um dos possíveis erros da rotulagem automática. Em A) a rotulagem incorreta e em B) visualização das partículas, que são separadas e foram incluídas em um mesmo rótulo.

Quando se iniciou o processo de rotulagem, não se tinha conhecimento de quantas imagens precisariam ser rotuladas para obter bons resultados nos treinamentos. Então, cada grupo de imagem utilizado é proveniente de um mosaico adquirido de determinada amostra. Trabalhou-se com um arquivo *json* para cada

grupo de imagem citado. E à medida que eram feitos testes, iam sendo adicionados mais grupos de imagens seguido do arquivo *json* correspondente ao grupo adicionado. Então para realizar o treinamento foi preciso unir todos esses arquivos *json* em um só e depois exportar. Essa parte do trabalho é delicada e precisa de bastante atenção, pois esses arquivos são grandes e contém informações essenciais para o treinamento. Essa junção dos arquivos de texto foi feita de forma manual, utilizando um editor de texto de código livre. Ainda será preciso, no futuro, automatizar esta etapa. Após a rotulagem das Imagens BF, foi possível trabalhar com a rotulagem das imagens CPOL que será descrita na subseção a seguir.

#### 4.4.2. Rotulagem Imagens CPOL

A partir das imagens CPOL corretamente alinhadas com as imagens BF, foi possível realizar uma transferência da rotulagem já realizada nas imagens BF para as imagens CPOL. Desta forma, não foi necessário repetir todo o trabalho exaustivo de rotulagem das partículas nas imagens que seriam utilizadas no Modelo 2. Após essa transferência foram realizadas alterações no arquivo de texto, já que as classes a serem observadas no Modelo 2 são diferentes das do Modelo 1. A associação de classes dos modelos foi feita a partir da observação de suas distribuições. A classe Compacta no Modelo 1 é a mais presente, enquanto no Modelo 2 a Monocristalina é a mais presente. Então uma associação foi feita entre essas duas classes para que no Modelo 2 fosse preciso fazer menos modificações no arquivo de texto referente a rotulagem, otimizando o tempo. Tendo em vista estas explicações, foram feitas as seguintes modificações no arquivo de texto: as partículas Compactas do Modelo 1 foram modificadas para Monocristalinas e as partículas Porosas para Policristalinas. A Figura 24A apresenta uma imagem com estas associações já realizadas.

Um outro detalhe a ser observado nestas associações é a classe Martita (Modelo1). Para que não houvesse confusão com a Classe Martita (Modelo 2), elas foram excluídas das imagens (colocadas em preto) para a rotulagem do Modelo 2 (Figura 24B - indicada com a seta vermelha). Após a retirada da Martita, as partículas foram analisadas e atribuiu-se cada partícula à classe correta de acordo com a hierarquia proposta.



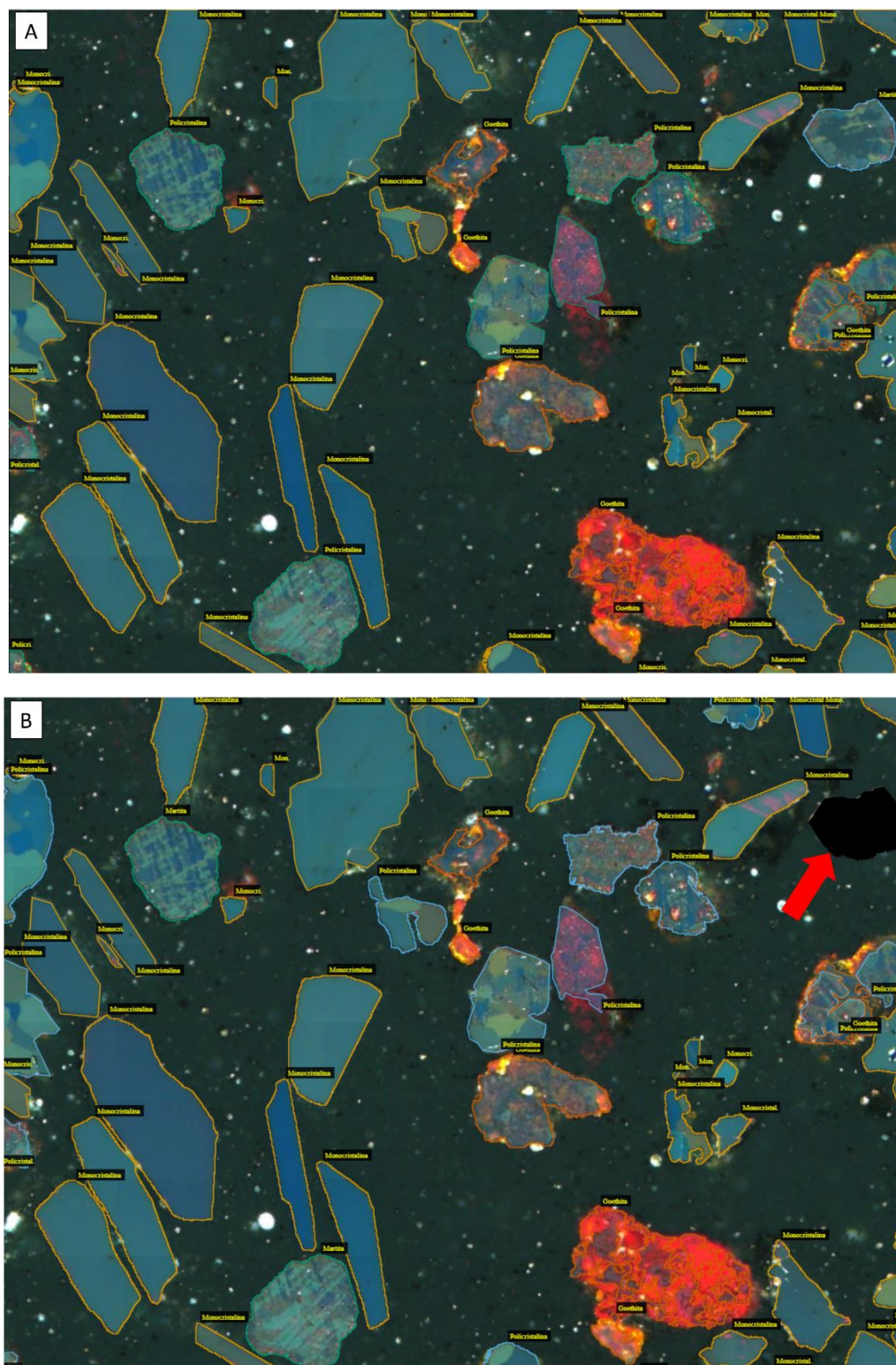


Figura 24 – Processo de rotulagem realizado nas imagens CPOL com base na rotulagem das imagens BF. A) rotulagem transferida da imagem BF para a CPOL com as respectivas associações de classes realizadas, e Martita (Modelo1) em preto – mostrada com a seta vermelha. B) imagem CPOL corretamente rotulada para o Modelo 2.

#### 4.5. Descrição do Banco de Dados

A rotulagem foi realizada em paralelo aos testes. Assim, trabalhou-se com diversos *datasets* que foram sendo incrementados à medida que os testes eram realizados e se percebia a necessidade de aumentar a quantidade de dados de entrada. Por esse motivo, será apresentada uma tabela enumerando estes *datasets* para que a discussão seja mais compreensível e clara. A Tabela 2 apresenta esse detalhamento para o Modelo 1.

Tabela 2 - Quantidades de imagens e partículas dos *datasets* utilizados no Modelo1.

<i>Dataset</i>	Quantidade de Imagens	Partículas Compacta	Partículas Porosa	Partículas Martita	Partículas Goethita
<b>1</b>	430	18326	1377	1069	1778
<b>2</b>	620	30842	1823	2386	3086
<b>3</b>	740	36533	2415	2769	3707
<b>4</b>	900	60375	4959	2492	5999

Os *datasets* utilizados para o Modelo 2 estão apresentados na Tabela 3 abaixo.

Tabela 3 - Quantidades de imagens e partículas de cada classe utilizadas no Modelo 2.

<i>Dataset</i>	Quantidade de Imagens	Partículas Monocristalina	Partículas Policristalina	Partículas Martita
<b>1</b>	300	14015	2651	186
<b>2</b>	900	54944	9147	1210

Observando as quantidades de partículas é possível perceber um desbalanceamento entre classes. Para o Modelo 1, por exemplo, a quantidade de partículas da classe Hematita Compacta é bastante superior ao restante. Desta forma, entendeu-se que o minério hematítico estudado era composto majoritariamente por hematitas compactas. Tendo em vista este desbalanceamento, decidiu-se então gerar imagens sintéticas nas quais mais partículas das classes menos abundantes estivessem presentes. Assim como Ekbatani et.al [49] afirmaram, uma das principais limitações de se trabalhar com algoritmos baseados em *Deep Learning* está em quando se possui poucos dados. Uma solução que se mostrou viável para tal dificuldade foi exatamente a geração de dados sintéticos. Pensando nisso, neste trabalho foi realizada uma geração de imagens sintéticas para aumentar os dados de treinamento. Esta está descrita na próxima seção.



#### 4.5.1. Criação de Imagens Sintéticas

O objetivo da criação das imagens sintéticas foi gerar imagens com as mesmas características de *background* das imagens obtidas no MO, mas que contivessem apenas as classes menos representadas no Modelo 1 (Hematita Porosa e Martita).

A criação das imagens sintéticas se deu em alguns passos:

- Foi extraído o *background* das próprias imagens obtidas no MO, excluindo as partículas das imagens.
- As partículas menos representadas foram extraídas das imagens originais e foi criada uma imagem para cada partícula contendo apenas a própria partícula.
- As imagens individuais das partículas foram incluídas no *background* criado dispostas em proporções definidas por classe. Estas partículas foram rotacionadas e colocadas em posições diferentes das originais.
- A rotulagem de cada partícula já anteriormente delineada foi resgatada para que o conjunto de imagens sintéticas não precisasse ser completamente rotulado novamente. A partir disto, foi gerado um arquivo de texto que continha as coordenadas das anotações das imagens sintéticas já com as classes corretas atribuídas. A Figura 25 apresenta um exemplo de imagem sintética criada contendo majoritariamente partículas de Martita, pois era a classe menos representada no *dataset*.

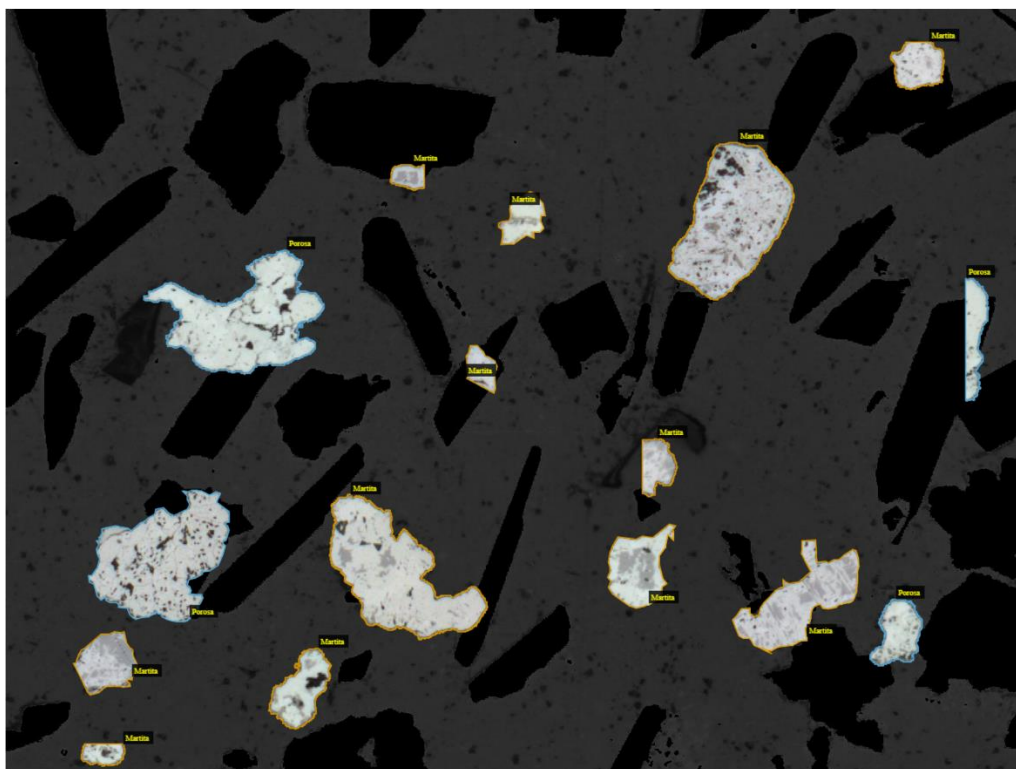


Figura 25 - Exemplo de imagem sintética criada e completamente rotulada.

Foram criadas 200 imagens sintéticas. Nestas imagens foram geradas 564 partículas da classe Porosa e 1891 da classe Martita. Desta forma, foi obtido mais um *dataset* a ser descrito nesta discussão. A inclui as imagens sintéticas.

Tabela 4 apresenta todos os *datasets* utilizados para o Modelo 1 novamente, desta vez incluindo o *dataset* 5, que inclui as imagens sintéticas.

Tabela 4 - Descrição dos *datasets* utilizados para o Modelo1 incluindo as imagens sintéticas geradas.

<b>Dataset</b>	<b>Quantidade de Imagens</b>	<b>Partículas Compacta</b>	<b>Partículas Porosa</b>	<b>Partículas Martita</b>	<b>Partículas Goethita</b>
<b>1</b>	430	18326	1377	1069	1778
<b>2</b>	620	30842	1823	2386	3086
<b>3</b>	740	36533	2415	2769	3707
<b>4</b>	900	60375	4959	2492	5999
<b>5</b>	1100	60375	5523	4383	5999

A partir destes *datasets* foram realizados treinamentos e observados os resultados.

#### 4.6. Descrição das etapas de treinamento

Todos os treinamentos deste trabalho foram realizados em uma máquina com quatro unidades de processamento gráfico (GPUs GTX 1080-TI), 2 CPUs Intel Xeon E5-2630 v4 e 192GB de memória RAM, com uma imagem treinada em cada GPU por vez. As rotinas de treinamento foram programadas em linguagem *Python*. E as etapas de pós-processamento, operação entre imagens e morfológicas foram todas realizadas no software FIJI.[50] Os treinamentos foram realizados em algumas etapas. Os *datasets* foram divididos proporcionalmente: 80% das imagens para treinar (*train*) a rede e 20% para validar os resultados (*val*). Então, com estas informações o treinamento foi realizado. Os seguintes parâmetros foram definidos para o treinamento:

- As imagens de entrada foram redimensionadas para um tamanho padronizado;
- Foram executadas 1500 épocas de treinamento;
- Foi processada 1 imagem por GPU utilizada, e foram utilizadas 4 GPUs;
- O *batch size* foi definido pelo número de imagens processadas por GPU multiplicado pela quantidade de GPUs utilizadas =  $1 \times 4 = 4$ ;
- O número de etapas de validação executadas a cada época de treinamento foi calculado pelo número total de imagens de treinamento dividido pelo *batch size*;
- O *learning rate* foi de 0.001 e foram definidas 640 âncoras para a etapa de RPN;
- Uma ROI é considerada positiva em relação ao *ground truth* quando apresenta  $\text{IoU} \geq 0,5$ ;
- A variação de um treinamento para outro envolve o número de classes consideradas, a quantidade de imagens utilizadas e o pixel médio calculado para o conjunto *train*.

Após a realização do treinamento, o modelo gerado foi aplicado ao conjunto *val* equivalente. Por meio desta etapa foram obtidas as métricas e a matriz de confusão com resultados quantitativos do desempenho do modelo. Foi realizada também uma inferência nas imagens do conjunto *val*, com base no modelo treinado, com parâmetros equivalentes às outras duas etapas. Na etapa de inferência obtém-se uma imagem com as partículas identificadas por meio de caixas envoltórias (*bounding boxes*) delimitando cada partícula, classificadas com um *score* que define a certeza de que cada partícula pertence a cada classe e segmentadas com uma máscara com uma cor específica para cada classe sobre a imagem original. A Figura 26 apresenta uma imagem de inferência do Modelo 1B, que se refere à classificação de Porosas e Martitas em imagens BF.

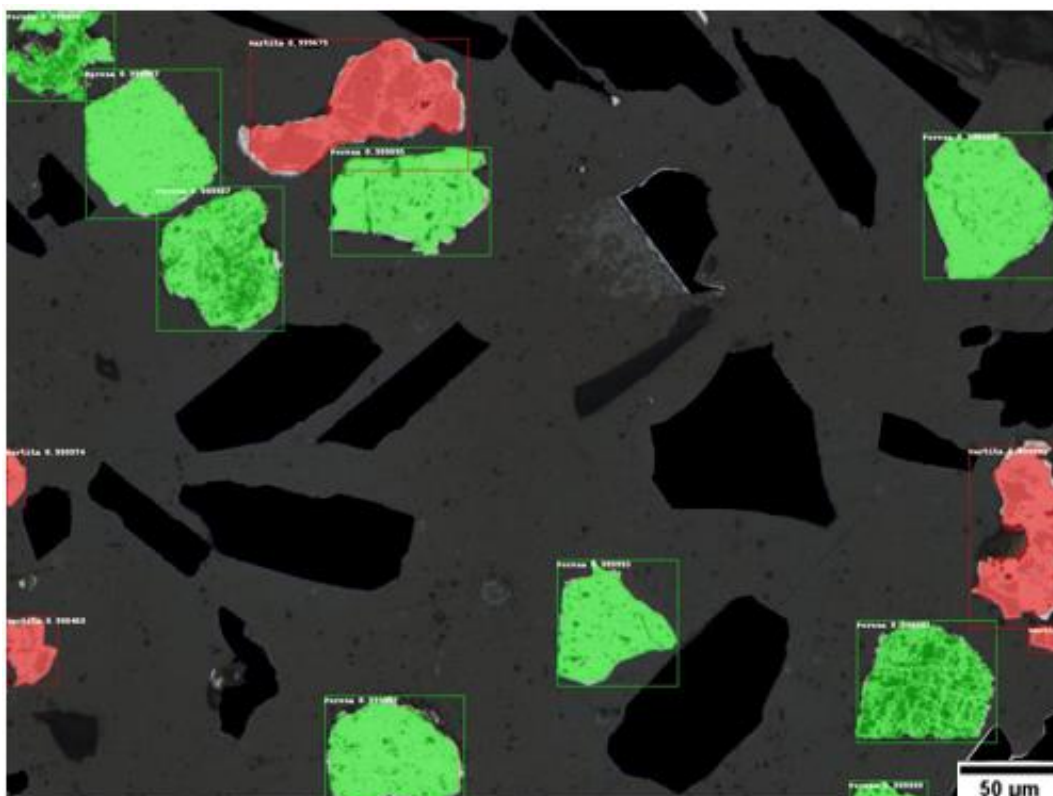


Figura 26 - Exemplo de imagem de inferência obtida a partir do Modelo 1B.

Desta forma cumpre-se o objetivo de identificar, classificar e segmentar as partículas nas imagens. É possível também obter apenas a máscara de segmentação, sem *score* ou *bounding boxes*, a fim de separar estas partículas para realizar medidas posteriormente. Esta etapa será apresentada na seção de resultados.

## 5

### Resultados e Discussão

Para que a discussão dos resultados seja mais organizada e coerente, esta seção foi dividida em resultados por modelo desenvolvido. Portanto, uma subseção para o Modelo 1, que utiliza as imagens BF e uma subseção para o Modelo 2 que utiliza as imagens CPOL.

#### 5.1. Resultados Modelo 1

##### 5.1.1. Desenvolvimento do Modelo

Vários testes foram realizados para o desenvolvimento do Modelo 1. Por isso, a Tabela 5 apresenta um guia discriminando cada teste por ordem de apresentação nesta seção.

Tabela 5 - Descrição das principais características dos treinamentos realizados para o desenvolvimento do Modelo 1.

Treinamento	Classes de Interesse	Característica das Imagens	Número de Imagens
1	Compacta, Martita, Porosa e Goethita	Originais sem pré-processamento	430
2	Compacta, Martita, Porosa	Goethitas em preto por limiarização	430
3	Martita e Porosa	Originais sem pré-processamento	430
4	Martita e Porosa	Compactas em preto por rotulagem e Goethitas em preto por limiarização	430
5	Martita e Porosa	Compactas em preto por rotulagem	430
6	Martita e Porosa	Compactas em preto por rotulagem	620
7	Martita e Porosa	Compactas em preto por rotulagem	740
8	Martita e Porosa	Compactas em preto por rotulagem	900
9	Martita e Porosa	Compactas em preto por rotulagem	1100
10	Martita e Porosa	Compactas em preto por rotulagem	1062
11	Compacta	Originais sem pré-processamento	740
12	Compacta	Originais sem pré-processamento	900
13	Goethita	Originais sem pré-processamento	620

É importante ressaltar que a ordem em que estes treinamentos são mostrados não necessariamente é a ordem na qual eles foram realizados. Foi escolhida a forma mais didática e compreensível para os leitores, que mostrasse a lógica das tomadas de decisão no desenvolvimento do modelo. Além disso, foram realizados muitos outros testes que não se considerou importante apresentar.

Desta forma, inicialmente foi realizado o Treinamento 1 incluindo todas as classes do Modelo 1: Compacta, Martita, Porosa e Goethita. A quantidade de imagens e partículas de cada classe utilizada foi a descrita no *dataset* 1 (430 imagens). Estas quantidades foram então divididas em *train* e *val*, de acordo com as proporções definidas na seção anterior. Este treinamento foi realizado então com 4 classes + 1 *background*, 1500 épocas – tempo aproximado de 5 dias. Todos os treinamentos realizados neste trabalho consideraram 1500 épocas e o tempo de processamento varia um pouco em relação à quantidade de dados por treinamento. As imagens para este treinamento foram utilizadas sem nenhum pré-processamento. Na Tabela 6 estão apresentadas as métricas gerais e individuais calculadas para a validação do modelo obtido a partir do Treinamento 1.

Tabela 6 - Métricas gerais e individuais calculadas com base nos resultados obtidos para o Treinamento 1.

Classe\Métricas	Precision (%)		Recall (%)		F1-score (%)	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Compacta</b>	73,10	89,11	76,27	87,19	74,65	88,14
<b>Martita</b>		17,70		11,20		13,72
<b>Porosa</b>		68,42		15,89		25,79
<b>Goethita</b>		20,26		53,98		29,46

As métricas gerais foram razoavelmente boas para um primeiro teste, mas é válido ressaltar que estes resultados são mascarados pelos excelentes resultados da classe Compacta, já que as métricas das outras três classes não correspondem a uma boa quantidade de acertos. Isso foi associado às características da classe Compacta já mencionadas anteriormente e sua grande representatividade no *dataset*, enquanto as demais classes ainda estavam bem pouco representadas no *dataset* 1 e têm características texturais mais complicadas.

Neste treinamento, foi possível perceber uma excelente identificação da classe Compacta, F1- *score* individual igual a 88,14%. Trata-se de um bom resultado, e esperado, dado que as partículas desta classe são partículas de característica uniforme, tonalidade bem mais brilhante que as demais e uma representatividade muito maior no *dataset* utilizado. Por outro lado, além da boa identificação, foi possível observar que muitas partículas da classe Compacta não foram inferidas pelo modelo, ou seja, foram observados muitos FN de Compacta. Na Figura 27 é possível observar uma associação entre imagens rotuladas e inferidas contendo FN de compacta.

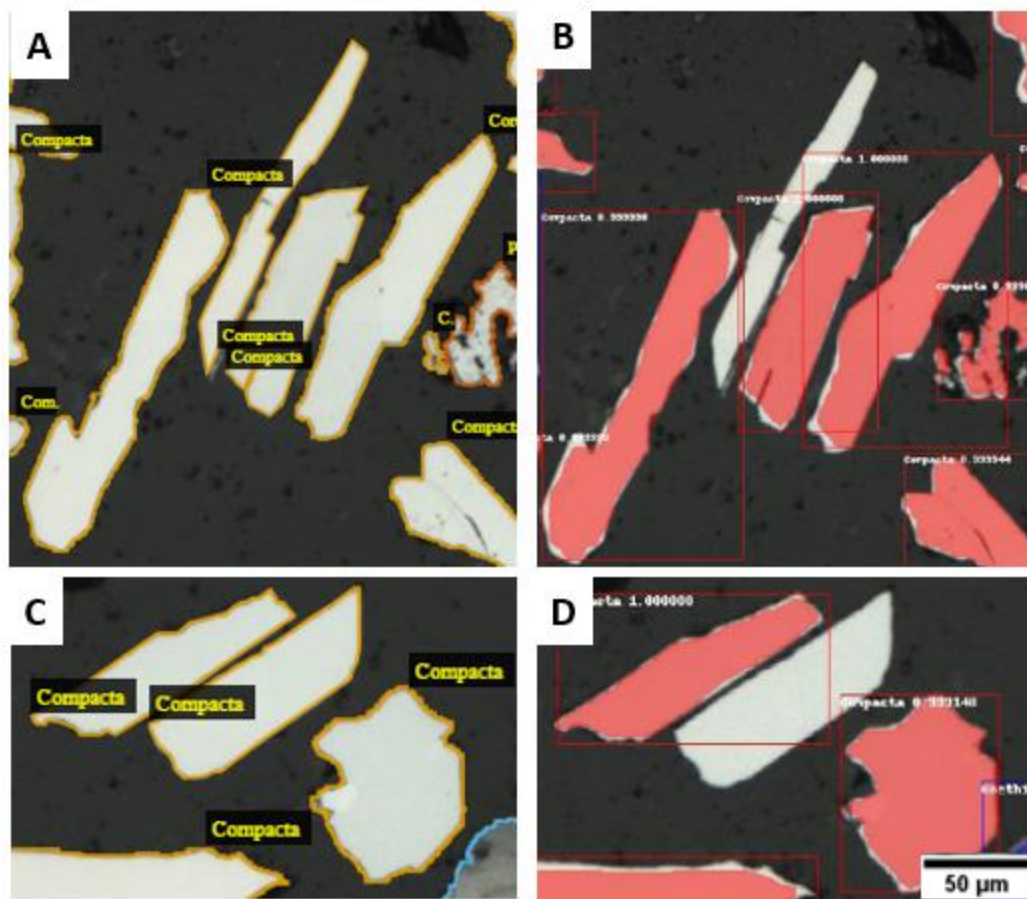


Figura 27 - Exemplos de FN da classe Compacta. Em A está apresentado o *ground truth* referente à inferência em B e em C está apresentado o *ground truth* referente à inferência em D.

Os FN apresentados ocorrem principalmente em imagens que têm muitas partículas próximas umas das outras, e algumas vezes em partículas na borda da imagem. Este fato foi associado à quantidade de propostas de região geradas para cada imagem - essa quantidade pode ser pequena para algumas imagens que têm maior número de partículas. Porém, foi difícil identificar uma quantidade de propostas de região apropriada para o problema. Esse resultado impacta principalmente no *Recall* e consequentemente, no *F1-score*. Não se buscou resolução imediata para este problema dado que existiam outros mais urgentes, referente às outras classes com métricas individuais mais baixas.

Para as classes Martita e Porosa, as métricas foram realmente muito baixas, e isso está associado principalmente à baixa representatividade e dificuldade na diferenciação das duas classes que, conforme mencionado anteriormente, são bem similares. Além disso, têm texturas mais complexas e variadas que o aspecto textural da Compacta, aumentando a dificuldade de aprendizado do sistema. Um outro problema encontrado neste treinamento foram os FP da classe Goethita. Esta identificação foi muito importante e determinante para os passos seguintes de desenvolvimento desta metodologia. Foi possível perceber que todas as classes tinham partículas que foram erroneamente inferidas na classe Goethita. É possível constatar este fato na coluna central da matriz de confusão do Treinamento 1

(Figura 28), que relaciona cada uma das classes com o que foi inferido como Goethita.

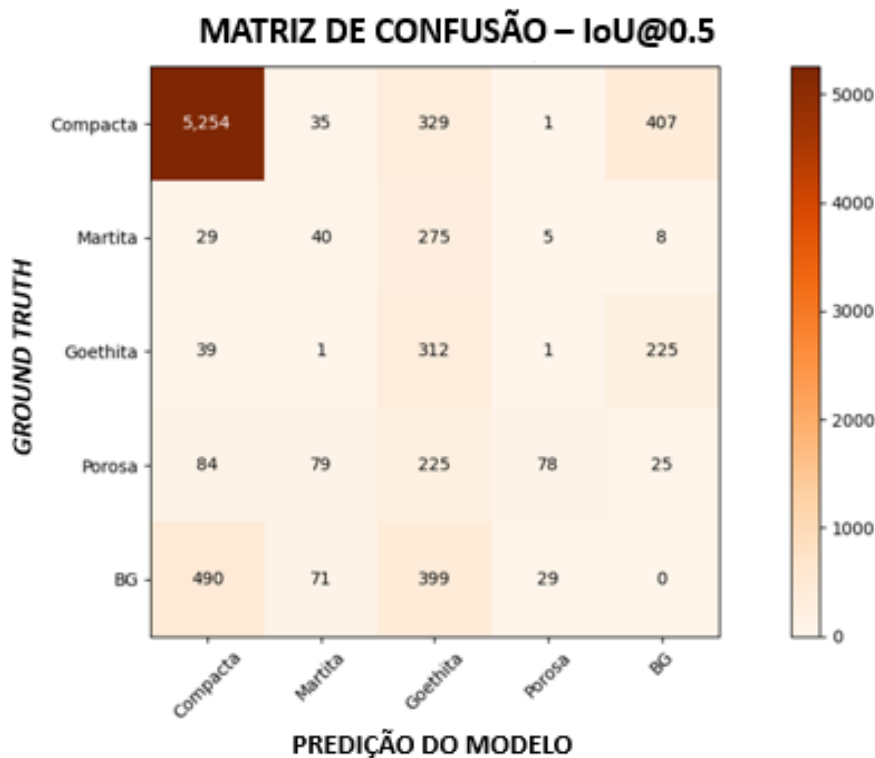


Figura 28 - Matriz de Confusão que relaciona a predição do modelo com o *ground truth* do Treinamento 1.

Desta forma, entendeu-se que era preciso ter bastante atenção com a classe Goethita, pois esta classe aparentemente gerava problemas para a identificação de todas as outras. Esse resultado foi surpreendente, uma vez que a tonalidade da Goethita é bem distinta das demais, tanto que é possível segmentar esta classe por limiarização sem tantos problemas. Ao analisar as imagens inferidas, percebeu-se que partículas de qualquer classe que contivessem “buracos” eram inferidas como Goethita, que é o caso tanto de Porosas e Martitas, quanto de algumas Compactas que provavelmente são policristalinas (vistas na CPOL) e na imagem BF possuem alguns “buracos”. Associou-se isso ao fato da tonalidade da Goethita muitas vezes ser similar à do *background* e, uma vez que os “buracos” tornam visível a resina por baixo da partícula que faz parte do *background*, essa confusão pode ser gerada.

Outra questão sobre esta classe é que suas partículas nem sempre têm a mesma tonalidade, exatamente porque existem algumas subclassificações tais como as goethitas terrosas e vítreas[16], por exemplo. Como não fazia parte do escopo deste trabalho realizar estas subclassificações, trabalhou-se apenas com a Classe Goethita, que engloba todas as suas subclasses. Na Figura 29 é possível observar exemplos de identificação de cada uma das classes, nos quais todas as partículas foram erroneamente identificadas como Goethita. As imagens inferidas estão apresentadas em pares juntamente com as imagens rotuladas para que seja possível compará-las.



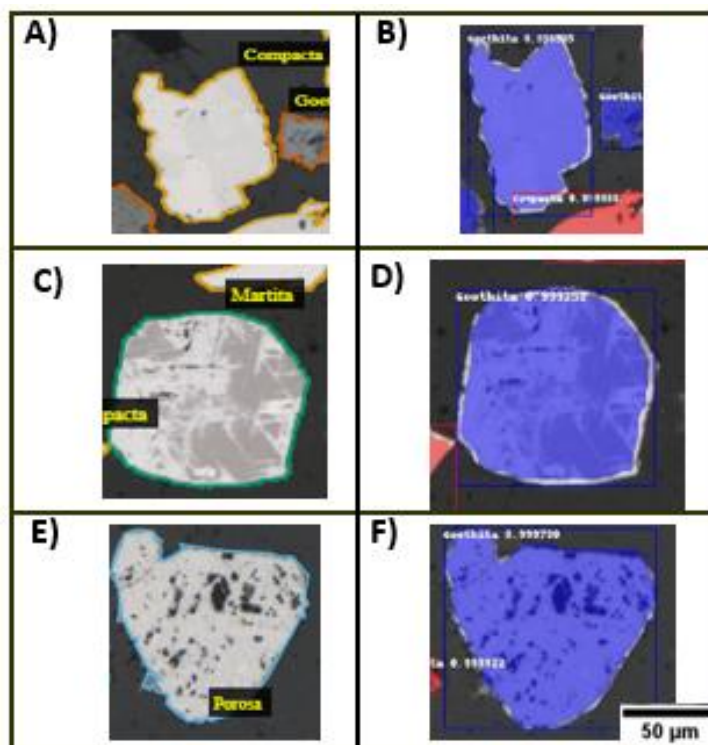


Figura 29 - Exemplos de FP da classe Goethita. A coluna da esquerda apresenta o *ground truth* e a da direita apresenta a inferência.

Por todos os motivos discutidos, decidiu-se realizar um teste excluindo a Goethita das imagens, por limiarização. Portanto as imagens utilizadas para o Treinamento 2 passaram por esta etapa de pré-processamento. As imagens utilizadas tinham todos os pixels na faixa de limiar da Goethita colocados em preto. A exclusão dessas partículas poderia ser realizada também por rotulagem, ou seja, tudo aquilo que foi rotulado como Goethita poderia ser colocado em preto. Porém, no caso específico da Goethita, existem partes entre ou ao redor de outras partículas que possuem coloração de Goethita e, caso a exclusão fosse feita por rotulagem, esses fragmentos não seriam excluídos pois não haviam sido rotulados. Pensando no modelo como um todo, caso a exclusão fosse feita por rotulagem, a extração da Goethita deveria estar garantida em alguma etapa do Modelo. Portanto, decidiu-se excluir a Goethita por limiarização. A Figura 30 apresenta um exemplo do pré-processamento realizado para o Treinamento 2.

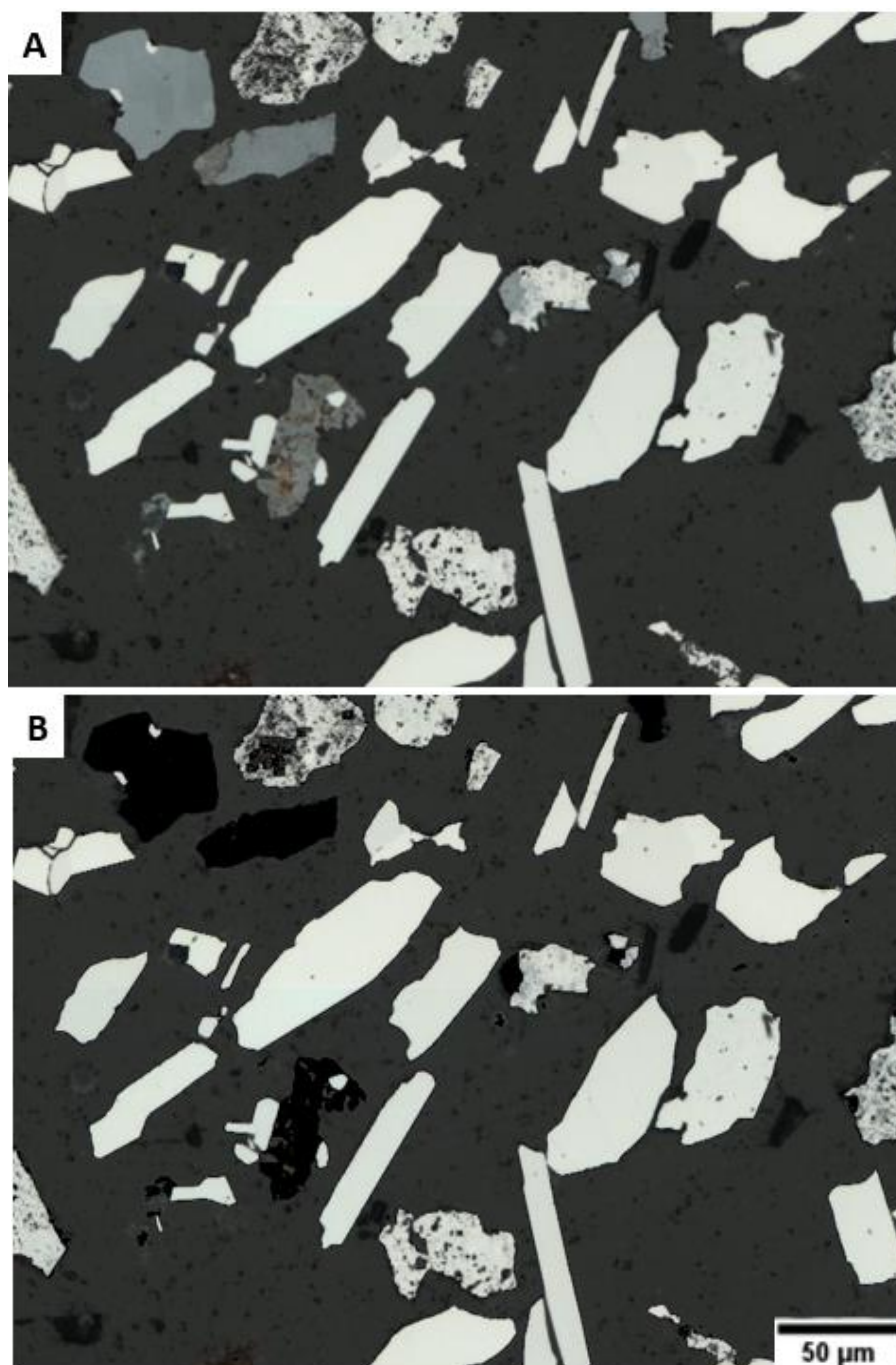


Figura 30 - Exemplo do pré-processamento realizado para a retirada da Goethita. Em A está apresentado um exemplo de imagem original e em B a imagem pré-processada com todos os pixels com tonalidade de Goethita colocados em preto.

Deste modo, o Treinamento 2 foi realizado considerando 3 classes: Compacta, Martita e Porosa + 1 *background*, utilizando o *dataset 1* (430 imagens). Duração de treinamento aproximada de 4 dias. As imagens de entrada foram pré-processadas com o objetivo de colocar as goethitas em preto. Observando os resultados da validação do Treinamento 2 (Figura 31), foi possível observar que o número de acertos para as classes Martita e Porosa aumentou bastante com a retirada da Goethita (comparando com a matriz do treinamento anterior na Figura 28).

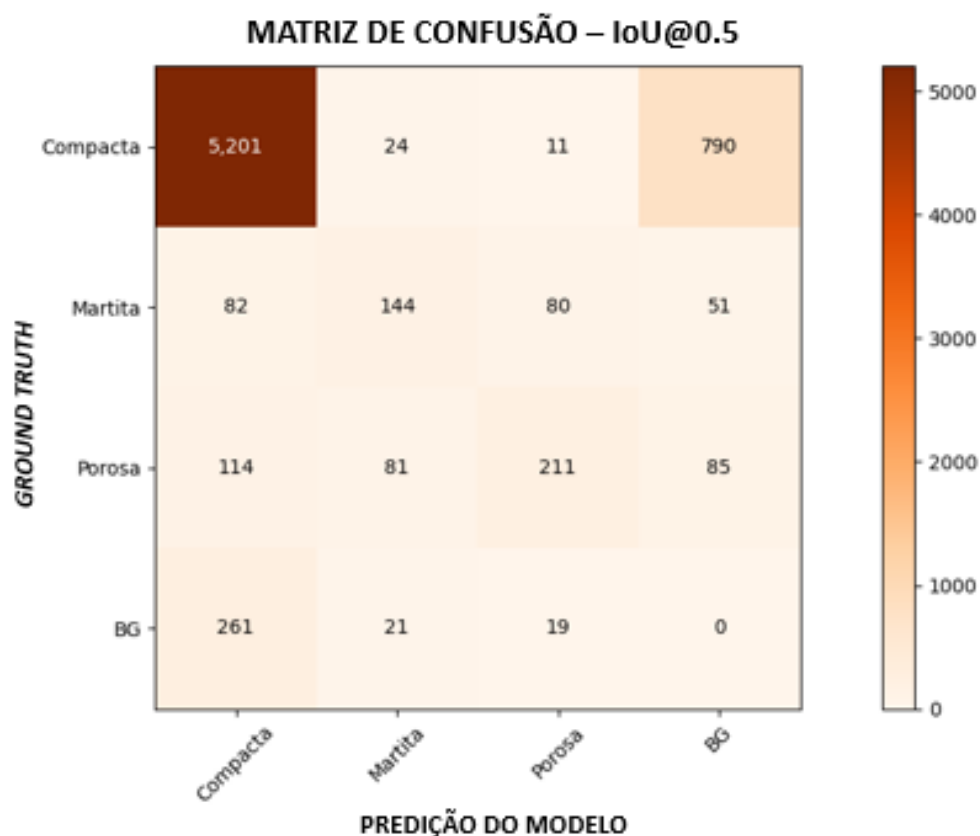


Figura 31 - Matriz de confusão obtida para a validação do Treinamento 2.

Apesar de ainda existirem erros para estas classes, as métricas melhoraram bastante. O *F1-score* no caso da Martita subiu de 13,72% para 45,93%, e o da Porosa de 25,79% para 51,97% (Tabela 7). Um fato relevante é que estas classes ainda são pouco representadas em relação à classe Compacta. Quanto à classe Compacta foi possível perceber uma diminuição no número de acertos e um aumento no número de FN. O aumento dos FN foi mais impactante para as métricas, reduzindo um pouco o valor do *Recall* (comparando as métricas do Treinamento 2 na Tabela 7 com as métricas anteriores na Tabela 6). Mesmo com a diminuição do *Recall* individual da Compacta, as métricas desta classe ainda são muito boas em relação às demais.

Tabela 7 – Relação das métricas gerais e individuais obtidas para o Treinamento 2.

Classe\Métricas	Precision (%)		Recall (%)		F1-score (%)	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Compacta</b>		91,92		86,31		89,03
<b>Martita</b>	88,91	53,33	80,83	40,34	84,68	45,93
<b>Porosa</b>		65,73		42,97		51,97

Com relação às métricas gerais, todas aumentaram e chegaram em números considerados bons. Porém, conforme dito anteriormente, esses números são reflexo principalmente dos resultados da Compacta. Sabendo disso, pensou-se em separar o Modelo1 em alguns submodelos, que posteriormente foram denominados como Modelo 1A e Modelo 1B.

Para que os resultados sejam bem compreendidos, será apresentado primeiro o desenvolvimento do Modelo 1B. O primeiro teste realizado foi um treinamento considerando apenas as classes Martita e Porosa. Esta escolha foi realizada com base na representatividade das classes. O Treinamento 3, então, considerou 2 classes + 1 *background*, utilizando o *dataset* 1. A duração de treinamento foi de aproximadamente 3 dias. As imagens utilizadas neste treinamento não passaram por nenhum pré-processamento. Apesar do conhecimento sobre a problemática com a classe Goethita, decidiu-se não retirar as partículas desta classe das imagens neste caso, para observar o comportamento do modelo com a quantidade de dados equilibrada (Martita e Porosa) e com a presença da Goethita e Compactas no *background* das imagens.

Neste treinamento a identificação das partículas típicas de Martita, ou seja, aquelas que têm bastante Magnetita visível foi realizada corretamente, quando identificadas. E o mesmo ocorreu para as Porosas com muitos poros. A questão é que a grande maioria das partículas não foi identificada pelo modelo, gerando uma enorme quantidade de FN. Avaliando a matriz de confusão, a classe Martita, por exemplo, teve 84 VP e 261 FN, enquanto a classe Porosa teve 69 VP e 411 FN. As métricas obtidas para o Treinamento 3 foram bem baixas, com F1-score abaixo de 50%. Ao analisar as imagens, também percebeu-se que apenas 2 ou 3 partículas foram identificadas por imagem na maioria dos casos, mesmo contendo muitas partículas das classes de interesse. Considerando que isso poderia ter acontecido dada a grande quantidade de dados no *background*, decidiu-se realizar outro teste.

Neste novo teste para o desenvolvimento do Modelo 1B (Treinamento 4), decidiu-se realizar um pré-processamento diferente. Foi retirado das imagens tudo o que estava rotulado como Compacta, ou seja, as partículas da classe Compacta foram colocadas em preto por rotulagem. As Goethitas permaneceram em preto por limiarização. O Treinamento 4 então considerou 2 classes (Martita e Porosa) + 1 *background*. As imagens do *dataset* 1 foram pré-processadas conforme mencionado e o tempo de processamento aproximado de 2 dias. Neste treinamento as métricas de validação de ambas as classes consideradas foram muito boas. A Figura 32 apresenta a matriz de confusão referente a este treinamento.

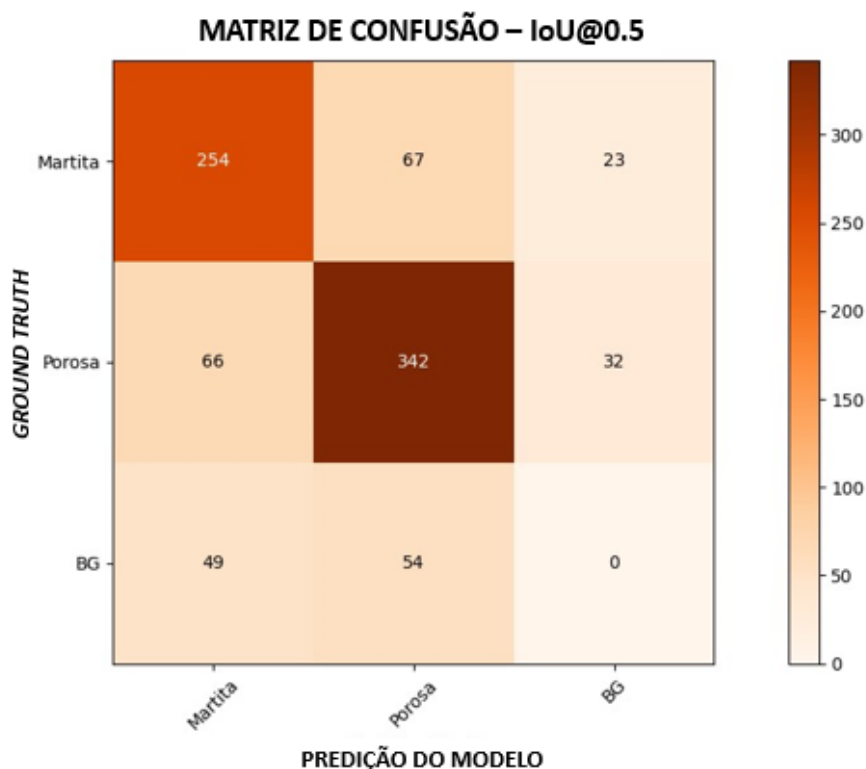


Figura 32 - Matriz de Confusão referente à validação do Treinamento 4.

Ao observar as imagens inferidas pelo modelo obtido a partir do Treinamento 4, notou-se que a grande maioria das partículas das classes de interesse foi corretamente identificada, classificada e inferida, conforme refletido na matriz de confusão e nas métricas. (Tabela 8).

Tabela 8 - Métricas gerais e individuais calculadas a partir da validação do modelo obtido do Treinamento 4.

Classe\Métricas	Precision (%)		Recall (%)		F1-score (%)	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Martita</b>	71,63	68,83	76,02	73,87	73,76	71,25
<b>Porosa</b>		73,87		77,73		75,75

A fim de refinar ainda mais os resultados e melhorar as métricas de validação na tentativa de obter um modelo mais confiável, foi realizada uma análise cuidadosa nas imagens. Percebeu-se que algumas partículas, principalmente de Martita que ficaram “esburacadas” com a retirada da Goethita por limiarização, ou que possuíam resquícios de Goethita ao seu redor, não foram inferidas corretamente, e muitas vezes nem foram identificadas. A Figura 33 apresenta um exemplo de imagem inferida pelo modelo gerado a partir deste treinamento. Nesta imagem é possível observar uma partícula de Martita não identificada, e que ao redor desta partícula existia uma coloração de Goethita que foi retirada por limiarização.

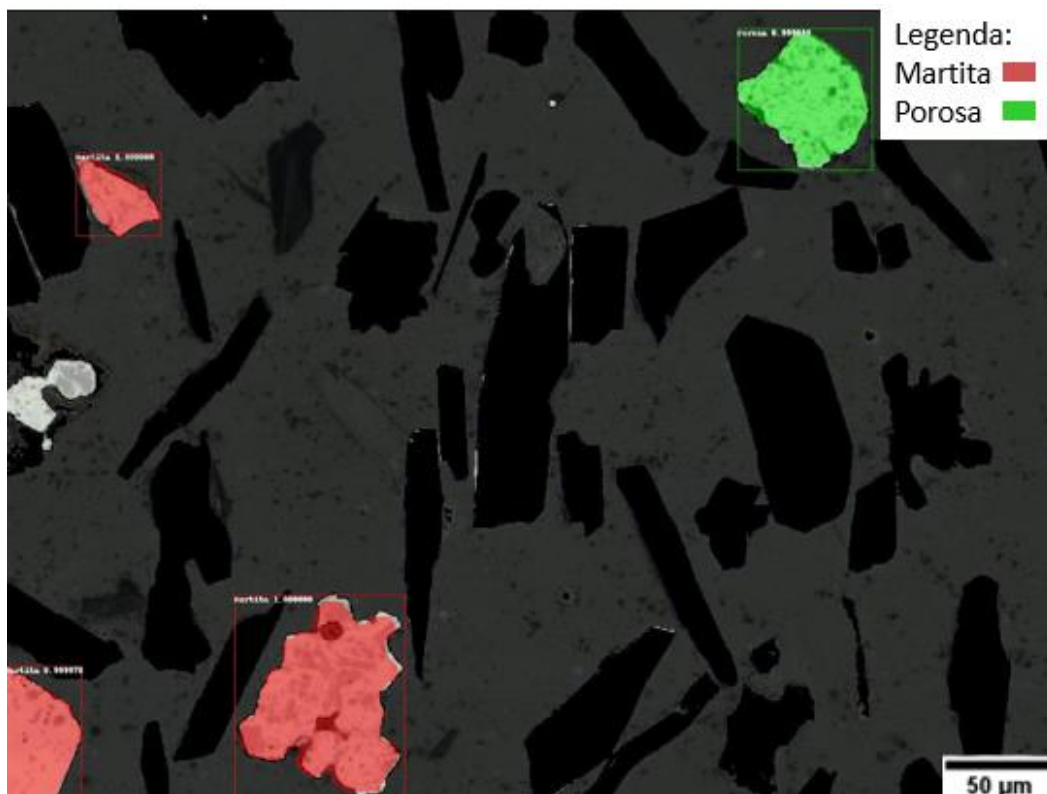


Figura 33 - Exemplo de imagem inferida pelo modelo obtido a partir do Treinamento 4, contendo uma partícula de martita não identificada

A partir da observação de casos como este, decidiu-se realizar um novo treinamento, mantendo as Goethitas nas imagens para observar o impacto. Desta forma, o Treinamento 5 foi realizado considerando as classes Martita e Porosa +1 *background*, com as Compactas em preto por rotulagem e as Goethitas no *background*. Foram utilizadas as imagens do *dataset* 1 (430 imagens) e o treinamento levou cerca de 3 dias para ser concluído. Os resultados observados para o Treinamento 5 foram bem parecidos com os do Treinamento 4, porém um pouco melhores. É possível observar na Figura 34 a mesma imagem apresentada na Figura 33, mas inferida com o modelo obtido pelo Treinamento 5. Neste caso, a partícula de Martita rodeada por Goethita foi inferida corretamente. Além disso, é notável que o sistema tenha entendido que a Goethita não faz parte das classes de interesse.



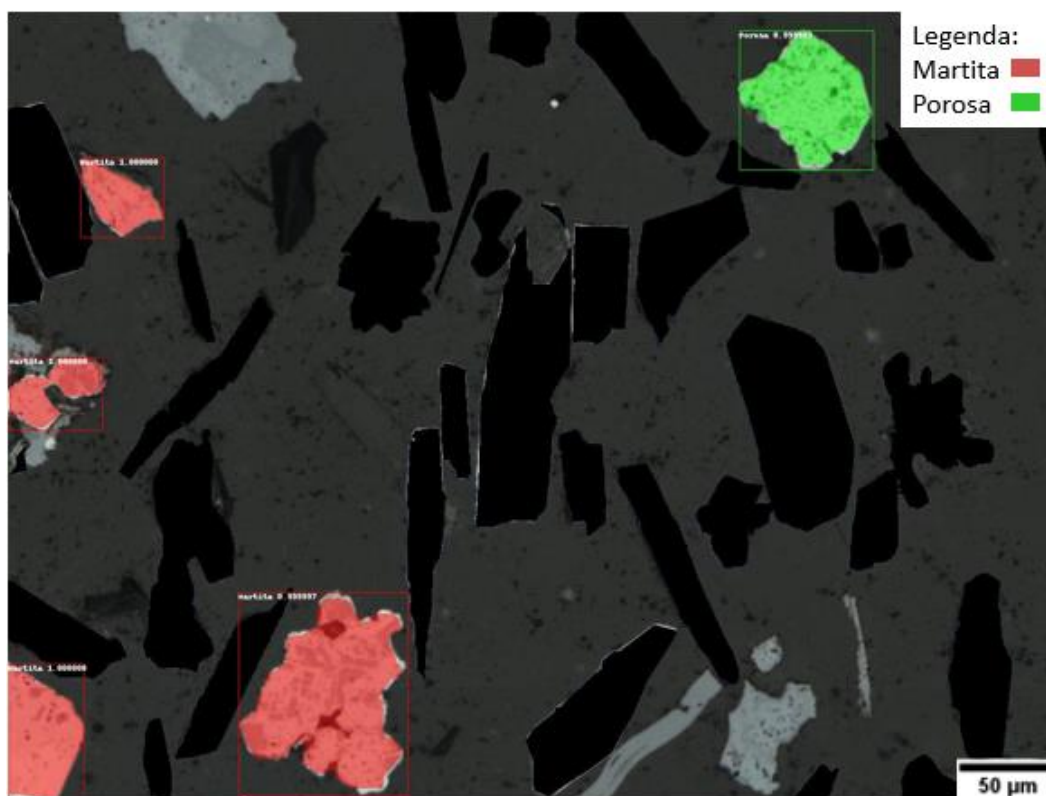


Figura 34 - Exemplo de imagem inferida pelo modelo obtido a partir do Treinamento 5.

O caso exemplificado na Figura 33 e na Figura 34 aconteceu com várias imagens que continham partículas próximas a resquícios de Goethita ou “esburacadas” com sua retirada, confirmando que o sistema se comportou melhor com a estrutura de imagens utilizadas no Treinamento 5. As métricas gerais e individuais estão apresentadas na Tabela 9.

Tabela 9 - Métricas gerais e individuais calculadas a partir da validação do modelo obtido do Treinamento 5.

Classe\Métricas	Precision (%)		Recall (%)		F1-score (%)	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Martita</b>	76,15	72,49	79,09	80,57	77,58	76,32
<b>Porosa</b>		79,36		77,93		78,64

A Figura 35 apresenta um exemplo de uma imagem inferida pelo modelo obtido no Treinamento 5. Este exemplo é bem representativo da maioria dos casos obtidos nestes resultados. A observação deve ser feita da seguinte forma: o contorno das partículas em amarelo e nomes das classes é referente ao *ground truth* fornecido para o treinamento, as máscaras representadas nas cores observadas na legenda são as máscaras de segmentação com respectivos *bouding boxes*, *scores* e partícula segmentada, gerados a partir da inferência de imagens pelo modelo já treinado. Desta forma é possível, por meio da comparação, perceber os acertos do modelo.

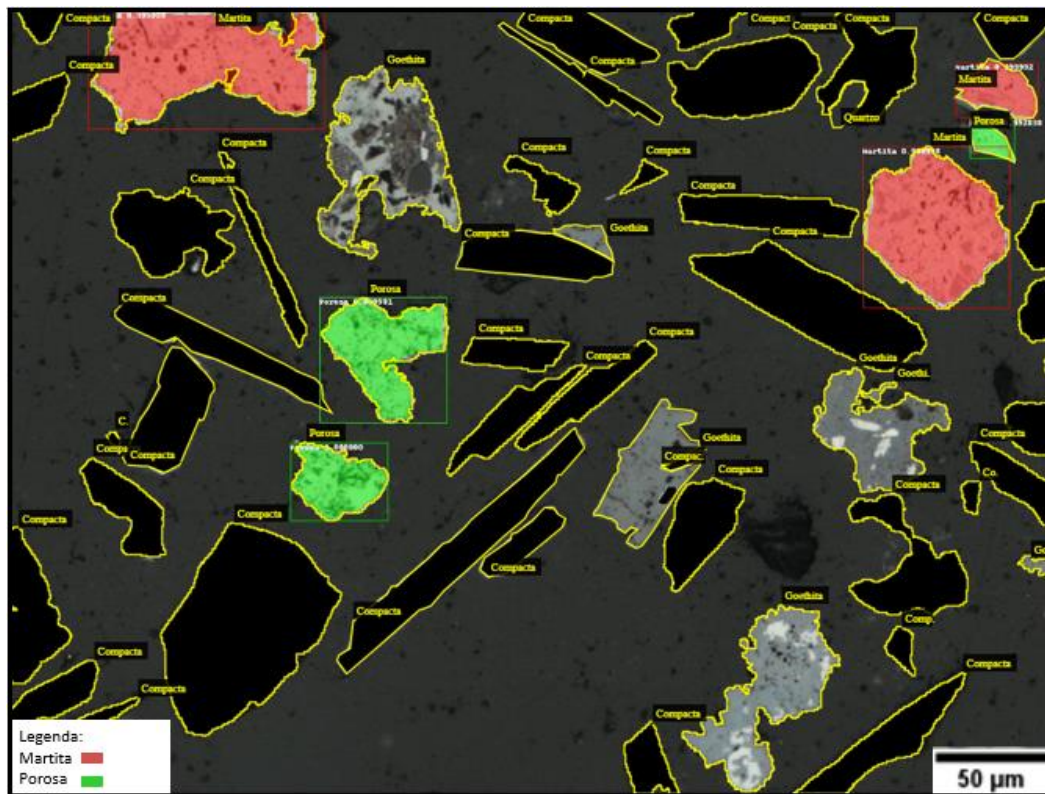


Figura 35 – Imagem de validação inferida pelo modelo obtido a partir do Treinamento 5. Comparação por meio da sobreposição da inferência sobre a rotulagem.

Devido aos bons resultados obtidos no Treinamento 5, decidiu-se que esse submodelo seria mantido para compor o Modelo 1, posteriormente definido como Modelo 1B. Entendeu-se que, para refinar estes resultados, era preciso aumentar o banco de dados. Então, o *dataset* foi incrementado algumas vezes e realizados treinamentos com esta mesma configuração de classes e pré-processamento nas imagens à medida que o *dataset* foi aumentando. A Figura 36 apresenta um comparativo entre as métricas gerais obtidas para o Treinamento 5, realizado com o *dataset* 1 (430 imagens), e os Treinamentos 6 e 7, realizados com as mesmas configurações, mas com os *datasets* 2 (620 imagens) e 3 (740 imagens), respectivamente.

Ao observar as comparações no gráfico da Figura 36 é possível perceber uma diminuição nas métricas à medida que o banco de dados foi sendo incrementado. A única métrica que aumentou um pouco no último incremento foi o *Precision*, ou seja, o modelo estava classificando e consequentemente segmentando bem aquilo que ele conseguia identificar. Porém estava havendo algum problema na identificação, dados os resultados das demais métricas. De acordo com o que foi discutido neste trabalho, as redes *Deep Learning* melhoram o desempenho proporcionalmente ao aumento do conjunto de treinamento [34]. Então existia um problema neste caso. Portanto, investiu-se na investigação da causa deste problema observando as imagens inferidas, para uma análise qualitativa, sempre comparando-as com o *ground truth*.



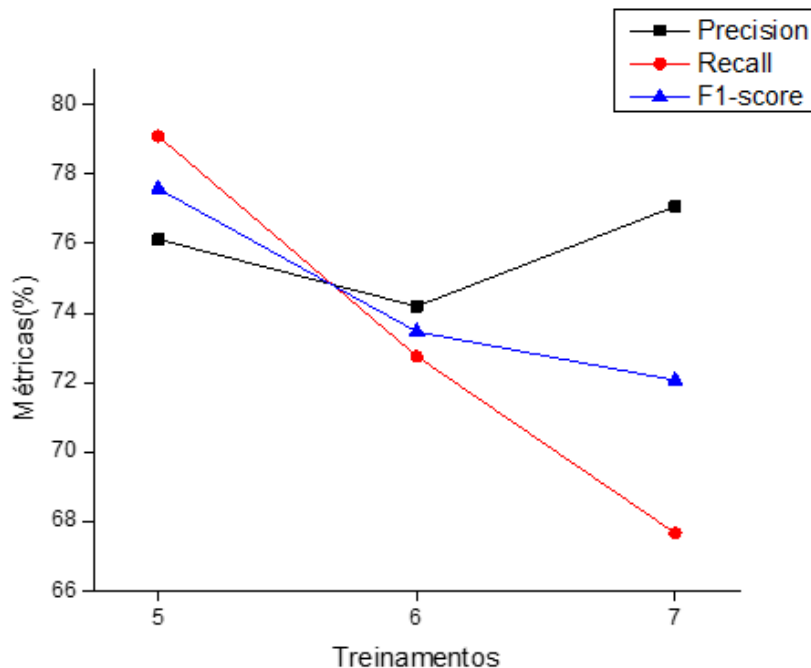


Figura 36 - Gráfico comparativo das métricas gerais obtidas para as validações dos treinamentos 5, 6 e 7.

A partir desta observação, percebeu-se alguns problemas na rotulagem. Um deles foi a dificuldade de definir um limiar de porosidade e presença de Magnetita que fosse capaz de definitivamente separar as classes Martita e Porosa. Pois existiam partículas que causavam dúvidas ao operador no momento da rotulagem, partículas como as exemplificadas na Figura 37. Observando-se, por exemplo o caso mais didático mostrado na Figura 37e, é possível notar muitos poros que a colocaria na classe Porosa mas também percebe-se “manchas” com coloração de Magnetita, que a colocaria na classe Martita, e o mesmo ocorre nas demais partículas apresentadas na Figura 37. A problemática estava em encontrar um equilíbrio de separação destas classes que fosse claro e organizado para o aprendizado do sistema.

A rotulagem precisava ser muito criteriosa para que, o sistema não fosse induzido ao erro por uma rotulagem errada. Depois de uma análise cuidadosa de casos de partículas que poderiam trazer este problema de classificação, todas as dúvidas foram tiradas com um especialista da Vale que pôde orientar e fornecer informações importantes para esta diferenciação. Durante esta revisão foi possível observar que, além das características de poros e quantidade de Magnetita presente, estas classes também se diferenciam pela forma. As partículas de Martita possuem, na maioria das vezes ângulos mais retos e formato mais característico, como é possível observar nos exemplos mostrados na Figura 37. Já as partículas da classe Porosa, possuem ângulos mais indefinidos e formatos diferentes, sem um padrão muito definido.

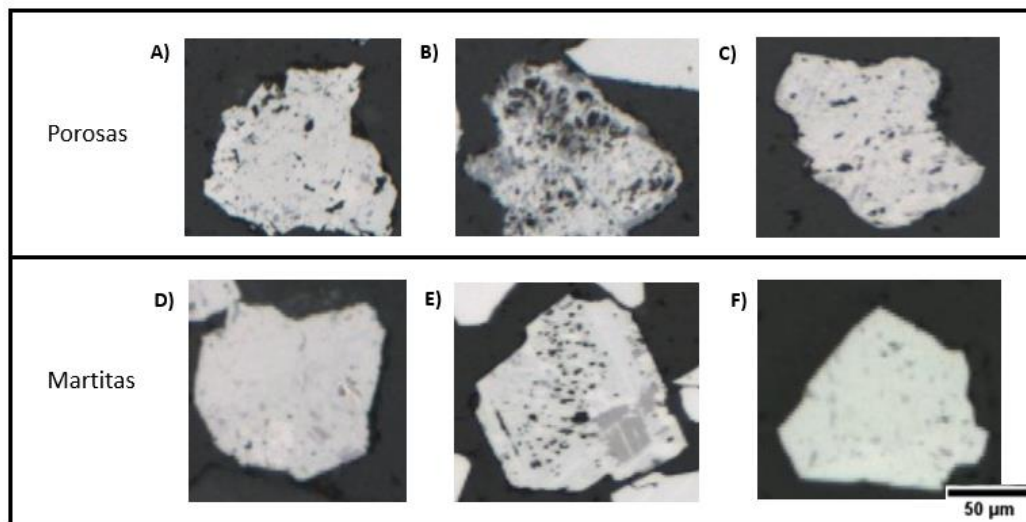


Figura 37 - Exemplos de partículas das classes Porosa e Martita que causaram dúvidas durante a rotulagem das imagens.

Após a revisão e correção de toda a rotulagem, o *dataset* foi incrementado mais uma vez e foi realizado o Treinamento 8. O Treinamento 8 foi realizado com as mesmas configurações descritas nos anteriores; classes de interesse foram Martita e Porosa, as partículas da classe Compacta foram colocadas em preto por rotulagem e a classe Goethita compõe o *background*. Foi utilizado o *dataset* 4 (900 imagens). A Figura 38 apresenta uma imagem inferida pelo modelo gerado a partir do Treinamento 8, a imagem mostra tanto a máscara de segmentação gerada na inferência quanto a rotulagem em amarelo no perímetro de cada partícula (*ground truth*). Nesta comparação *ground truth* x inferência é possível observar que as partículas que foram identificadas foram inferidas corretamente, mas que algumas partículas não foram identificadas, e isso se repete em muitas imagens do banco de dados.



Figura 38 - Imagem de validação inferida pelo modelo obtido a partir do Treinamento 8. Comparação por meio da sobreposição da inferência sobre a rotulagem.

Com relação às métricas, foi possível observar que os VP aumentaram, mas os FP continuaram em uma quantidade muito grande, assim como observado nas imagens, as métricas refletiram que muitas partículas não estavam sendo identificadas, mesmo após o incremento e todas as revisões na rotulagem. Um fato importante a ser comentado é que após a última revisão da rotulagem, as classes Martita e Porosa ficaram com as quantidades de partícula um pouco desbalanceadas, durante a revisão, a maioria das partículas que geraram dúvidas eram da classe Porosa e foram incluídas nesta classe. Desta forma o número de partículas da classe Porosa se tornou consideravelmente maior que o da classe Martita. Pensando nesse desbalanceamento, decidiu-se gerar imagens que contivessem apenas estas duas classes, balanceando de tal forma a aproximar o número de partículas da classe Martita à quantidade de partículas da classe Porosa e ainda, aumentar a quantidade de ambas, de forma proporcional.

Então, após esse treinamento foram criadas as imagens sintéticas, conforme descrito na metodologia deste trabalho e realizou-se o Treinamento 9 utilizando o *dataset 5* (1100 imagens). A configuração do treinamento foi a mesma citada anteriormente, com classes de interesse sendo Martita e Porosa, partículas de Compacta em preto e Goethita no *background*. A Tabela 10 apresenta as métricas obtidas a partir da validação do modelo obtido do Treinamento 9.

Tabela 10 - Métricas gerais e individuais calculadas a partir da validação do modelo obtido do Treinamento 9.

Classe\Métricas	<i>Precision (%)</i>		<i>Recall (%)</i>		<i>F1-score (%)</i>	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Martita</b>	84,29	79,71	71,05	59,93	77,11	68,42
<b>Porosa</b>		84,76		76,75		81,23

As métricas apresentadas são muito boas, principalmente *Precision*, porém ainda existem problemas de identificação que se refletem nas demais métricas. Após a observação destes resultados, decidiu-se inferir todo o *dataset* (*train* e *val*) com esse modelo obtido e observar os principais problemas que ocorriam. A partir destas observações foi feito um refino no banco de dados. Foi realizada então uma revisão no banco de dados a fim de manter apenas as imagens que contivessem imagens bem representativas de cada classe. Após esta revisão, o treinamento foi repetido (Treinamento 10). A Tabela 11 apresenta as métricas obtidas da validação deste modelo.

Tabela 11 - Métricas gerais e individuais calculadas a partir da validação do modelo obtido do Treinamento 10.

Classe\Métricas	<i>Precision (%)</i>		<i>Recall (%)</i>		<i>F1-score (%)</i>	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Martita</b>	84,76	84,76	73,79	74,58	78,89	79,34
<b>Porosa</b>		86,27		73,12		78,51

As métricas gerais melhoraram após o refino do banco de dados, e isso se dá principalmente pelo aumento do *Precision* de ambas as classes. O gráfico comparativo apresentado na Figura 39, mostra cada uma das métricas gerais em cada um dos treinamentos realizados neste formato Martita x Porosa com Compactas em preto e Goethita no *background*. É possível observar que *Precision* aumentou bastante para o Treinamento 10 em relação aos primeiros treinamentos, isso reflete o quanto o sistema aprendeu e passou a acertar cada vez mais as partículas que foram identificadas. Já o *Recall* chegou a ter valores bem baixos no decorrer do processo, principalmente no Treinamento 8, mas também alcançou um valor consideravelmente bom no último treinamento. O F1-score é reflexo das duas anteriores e alcançou bons resultados.

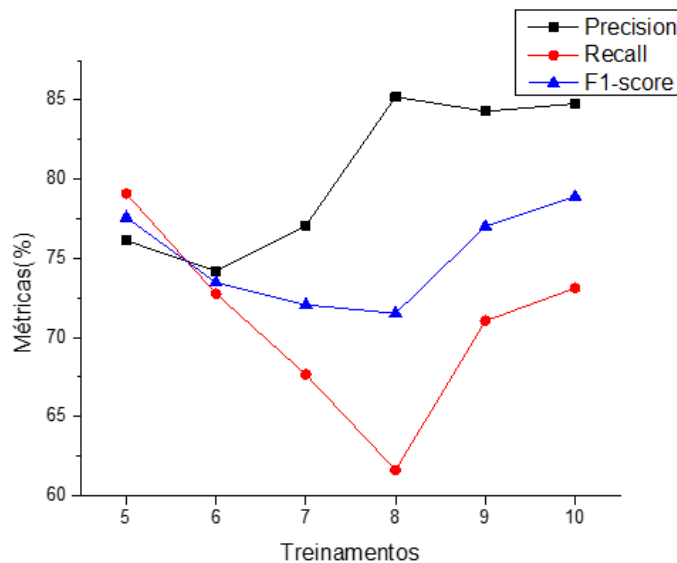


Figura 39 - Gráfico comparativo das métricas gerais obtidas para as validações dos treinamentos 5, 6, 7, 8, 9 e 10.

Deste modo, o Modelo1B foi utilizado para identificar, classificar e segmentar as classes Martita e Porosa. No entanto, ainda era preciso obter submodelos para identificação das classes Compacta e Goethita. Portanto, agora será discutido o desenvolvimento do Modelo 1A. Foi realizado então, o Treinamento 11 que considerou a classe Compacta e o *background* (todo o resto dos objetos contidos nas imagens). As imagens utilizadas não passaram por nenhum pré-processamento, e o *dataset* utilizado para este primeiro teste foi *dataset 3* (740 imagens). O treinamento durou cerca de 8 dias, os treinamentos que envolvem a classe Compacta foram bem mais demorados pois trata-se de uma classe que possui um grande número de partículas.

O que se observou nas imagens inferidas neste treinamento foi que a identificação da grande maioria das partículas da classe Compacta foi muito boa. Ocorreram alguns FN e FP, mas comparados a quantidade de acertos esses erros são quase irrelevantes, porém não se deve ignorá-los. Decidiu-se então observar cuidadosamente todos os resultados da inferência do modelo treinado comparando com o *ground truth* para encontrar possíveis causas para estes problemas. Depois desta revisão, foi realizado o Treinamento 12, com as mesmas configurações do Treinamento 11, mas com *dataset* incrementado - *dataset 4* (900 imagens). Este treinamento durou cerca de 12 dias para ser concluído. A Tabela 12 apresenta a comparação das métricas de validação dos modelos obtidos a partir dos Treinamentos 11 e 12.

Tabela 12 - Comparação entre métricas de validação dos modelos obtidos nos Treinamentos 11 e 12.

Classe	Treinamento\Métricas	<i>Precision</i> (%)	<i>Recall</i> (%)	<i>F1-score</i> (%)
Compacta	Treinamento 11	97,01	79,26	87,36
Compacta	Treinamento 12	97,30	85,66	90,98

Observando a Tabela 12 é possível observar um aumento considerável nas métricas no Treinamento 12 em especial no *Recall*. Foi um excelente resultado, uma vez que esta métrica tem sido a mais baixa devido a problemática dos FN. A Figura 40 apresenta uma imagem inferida pelo modelo obtido deste treinamento sobreposta ao seu *ground truth* para que se possa visualizar melhor os acertos e possíveis erros presentes. Este é um resultado típico do que ocorre na maioria das imagens, muitas inferidas corretamente. É possível observar também olhando para a imagem que o modelo consegue discriminar muito bem o *background* da classe de interesse.

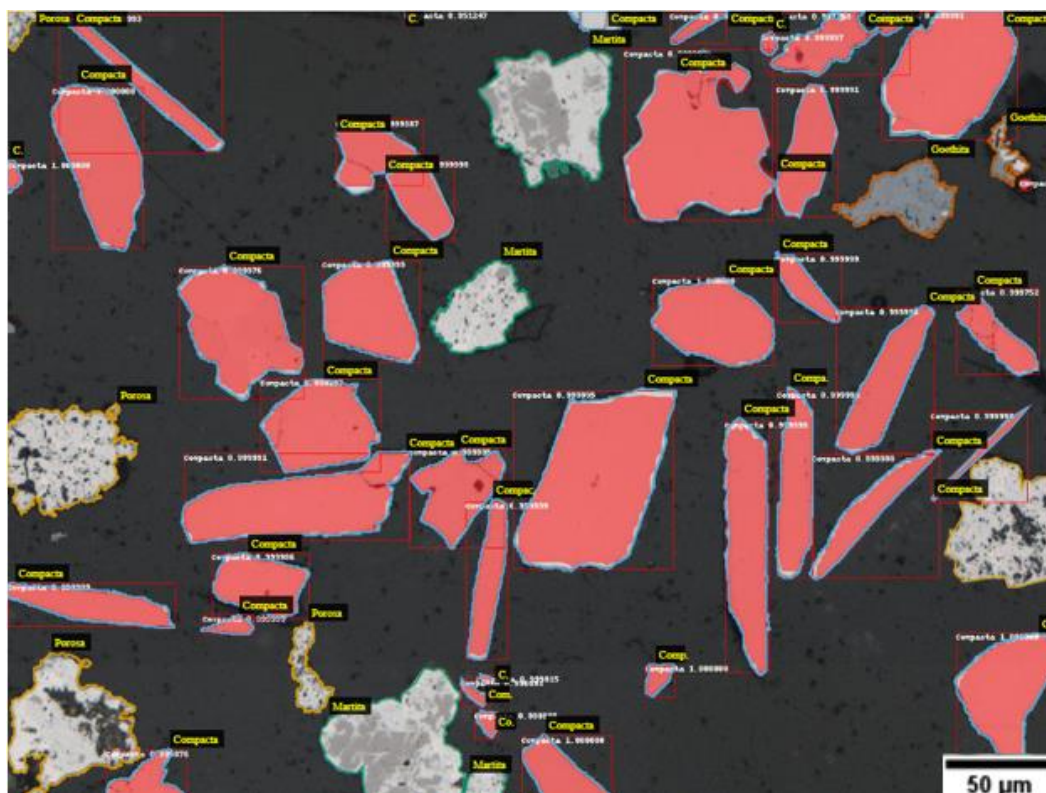


Figura 40 - Imagem inferida pelo modelo obtido a partir do Treinamento 12 para observação da quantidade de acertos referentes a classe de interesse do modelo: Compacta.

Um fato importante a ser comentado é que mesmo em imagens com problemas de foco, como a mostrada na Figura 41a o resultado é muito bom (Figura 41b), a inferência das partículas de interesse é realizada com sucesso. Isso fornece uma informação importante sobre a capacidade de adaptação do sistema.



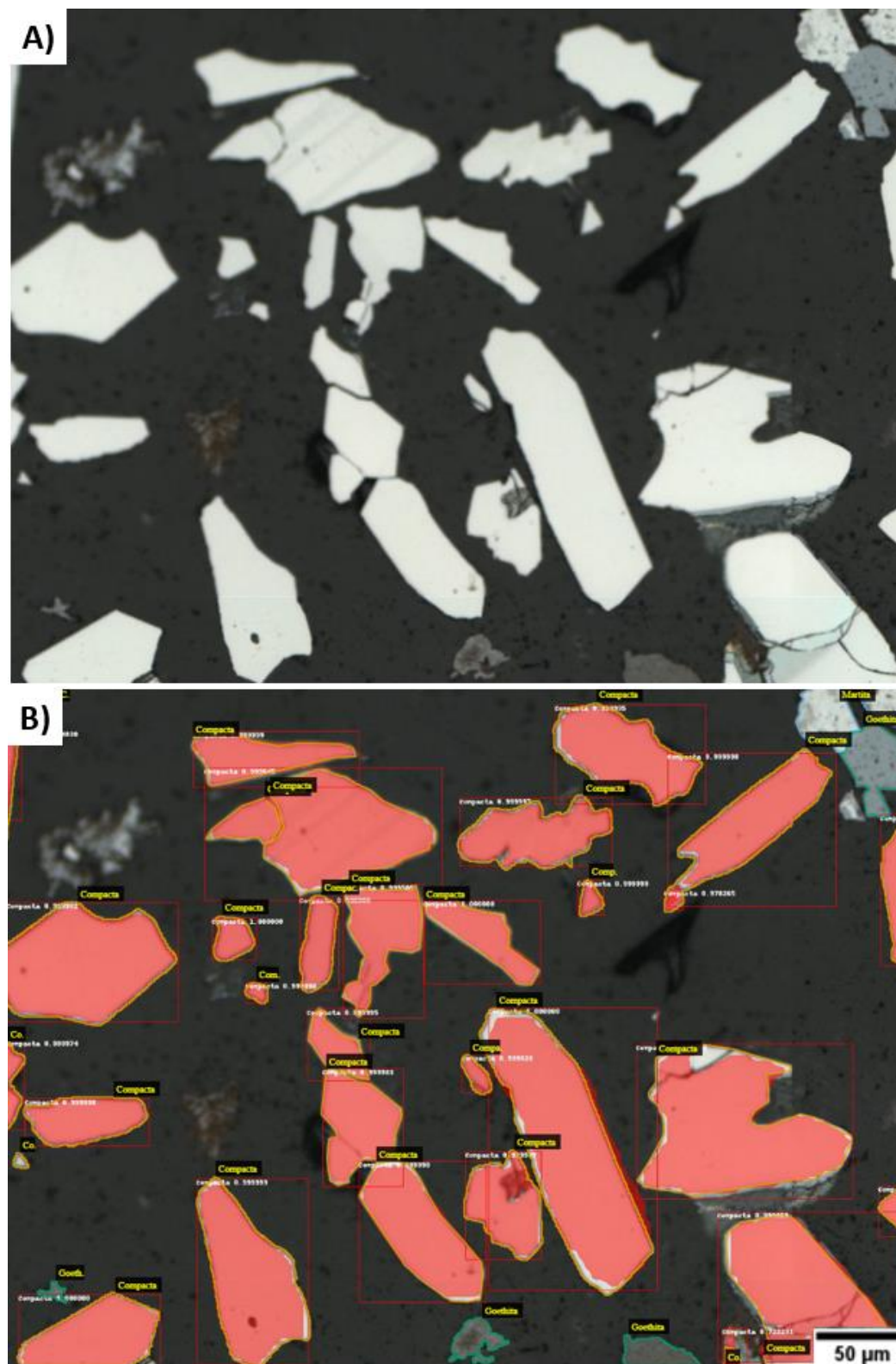


Figura 41 - Exemplo de inferência realizada pelo modelo obtido a partir do Treinamento 12 em uma imagem com problema de foco em (a), e o resultado da imagem inferida sobreposta ao *ground truth* em (b).

Conclui-se a partir das métricas e da observação das imagens que o Modelo1A estaria pronto para inferir as partículas da classe Compacta. Desta forma, as classes Compacta, Martita e Porosa já podiam ser identificadas no Modelo 1, restando apenas a Goethita. Então realizou-se mais testes como o Treinamento 13 que considerou as classes Goethita e *background*. As imagens não passaram por nenhum pré-processamento e foi utilizado o *dataset 2* (620 imagens). Os resultados obtidos não foram satisfatórios, as métricas foram muito baixas e percebeu-se que o modelo treinado não conseguia diferir a classe Goethita das demais classes contidas no *background*. Desta forma, obteve-se uma quantidade muito grande de FP, ou seja, quase todas as partículas contidas nas imagens eram inferidas como Goethita.

A Figura 42 exibe um exemplo bem representativo do que se pôde perceber da inferência utilizando este modelo. Praticamente todas as partículas são identificadas, mas classificadas erroneamente, e algumas partículas da classe de interesse do modelo nem foram identificadas neste caso. Avaliando as imagens e os resultados quantitativos obtidos neste teste e nos demais realizados que não foram expostos aqui, foi possível confirmar que a Goethita é realmente um caso complicado para o Modelo 1. Desta forma, entendeu-se que a melhor saída seria a limiarização desta classe.

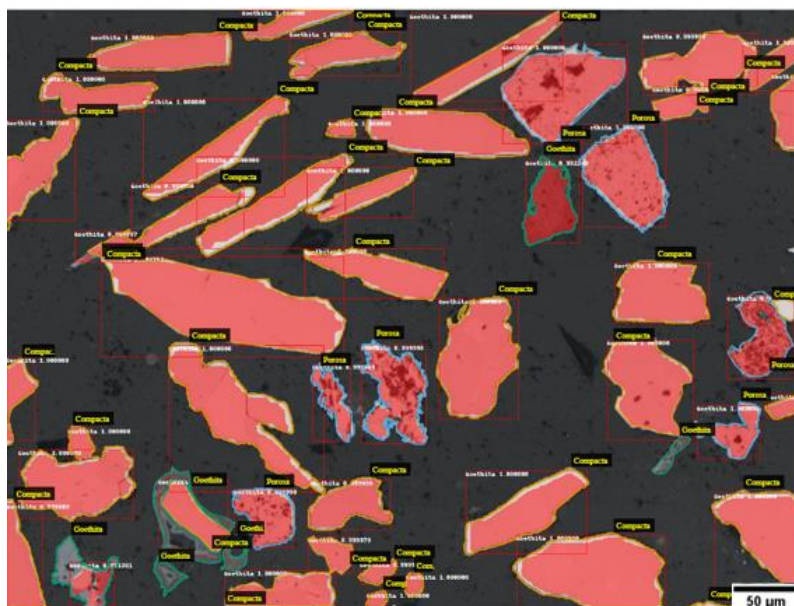


Figura 42 - Imagem inferida pelo modelo obtido a partir do Treinamento 13 que considerou as classes Goethita e *background*.

Fundamentando-se em todas as discussões apresentadas dos resultados obtidos para o Modelo 1, a hierarquia do Modelo ficou definida da seguinte forma: Modelo 1A identifica a classe Compacta nas imagens originais, as partículas identificadas são colocadas em preto e o Modelo 1 B identifica as classes Porosa e Martita nestas imagens, e a limiarização da Goethita é realizada nas imagens originais sem pré-processamento.



Para melhor compreensão da hierarquia deste modelo serão apresentadas imagens em sequência do passo a passo deste Modelo. A Figura 43a apresenta a imagem BF original utilizada como *input* pro Modelo 1A (Compacta x BG) que identifica a classe Compacta (Figura 43b), as partículas da classe Compacta já identificadas são então colocadas em preto (Figura 43c) e esta imagem é utilizada como *input* para o Modelo 1B (Martita x Porosa) e as classes Martita e Porosa são identificadas conforme mostrado na Figura 43d. A Figura 43e apresenta a imagem binária obtida para a classe Goethita a partir da limiarização na imagem original. Deste modo obteve-se todas as classes de partículas identificadas para o Modelo1.

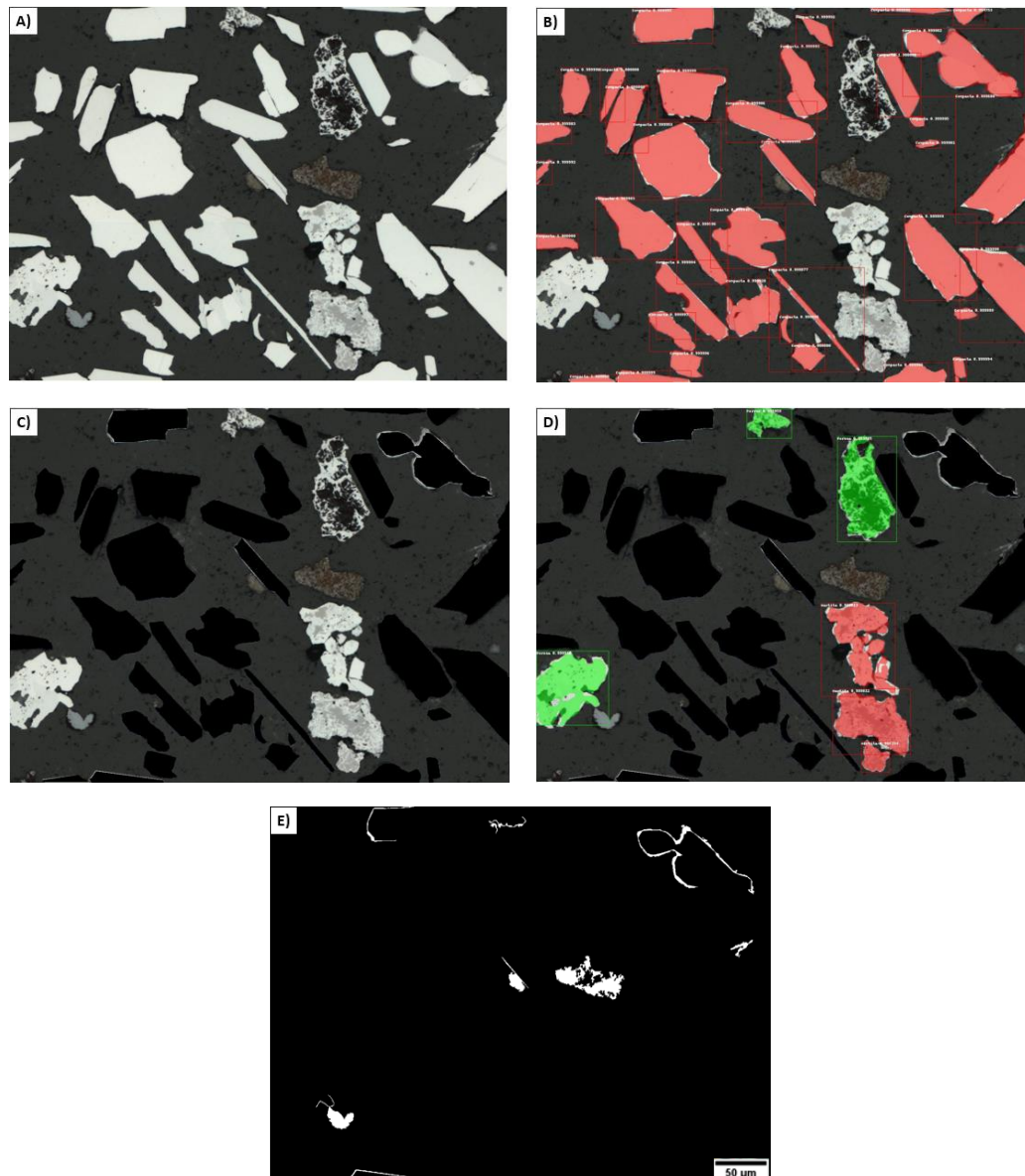


Figura 43 - Exemplificação das etapas do Modelo 1. A) a imagem original BF, B) a imagem inferida pelo modelo 1B identificando a classe Compacta (em vermelho), C) a imagem com as Compactas identificadas em preto que é *input* para Modelo 1B D) imagem inferida pelo Modelo 1B (Porosas em verde e Martitas em vermelho), E) imagem binária da classe Goethita obtida a partir da limiarização na imagem original.

Após o desenvolvimento do modelo que possibilitou a inferência das classes, foi possível obter as máscaras de segmentação para cada uma das classes. A próxima seção apresentará o pós-processamento realizado nestas imagens.

### 5.1.2. Pós-processamento

Após a obtenção dos modelos 1A e 1B e da segmentação da classe Goethita é possível, além de obter uma imagem de inferência, obter máscaras de segmentação de cada classe para que a partir delas possam ser realizadas medidas relevantes para a caracterização do minério. Portanto nesta seção serão apresentados resultados do pós-processamento realizado a partir desta obtenção. Inicialmente para o Modelo 1A foram obtidas imagens como a mostrada na Figura 44B, considerando a imagem original (Figura 44A), nas quais as partículas da classe de interesse (Compacta) estão em vermelho sob um *background*. Neste caso a obtenção da imagem binária é bem simples, uma vez que se tem apenas uma classe na imagem. Portanto esta imagem (RGB) é convertida para 8bits (Figura 44C) e a partir de um *threshold* automático (24-255) é obtida a imagem binária apresentada Figura 44D. Nesta imagem, os objetos (em branco) são as partículas da classe compacta inferidas automaticamente pelo modelo observadas na imagem original.

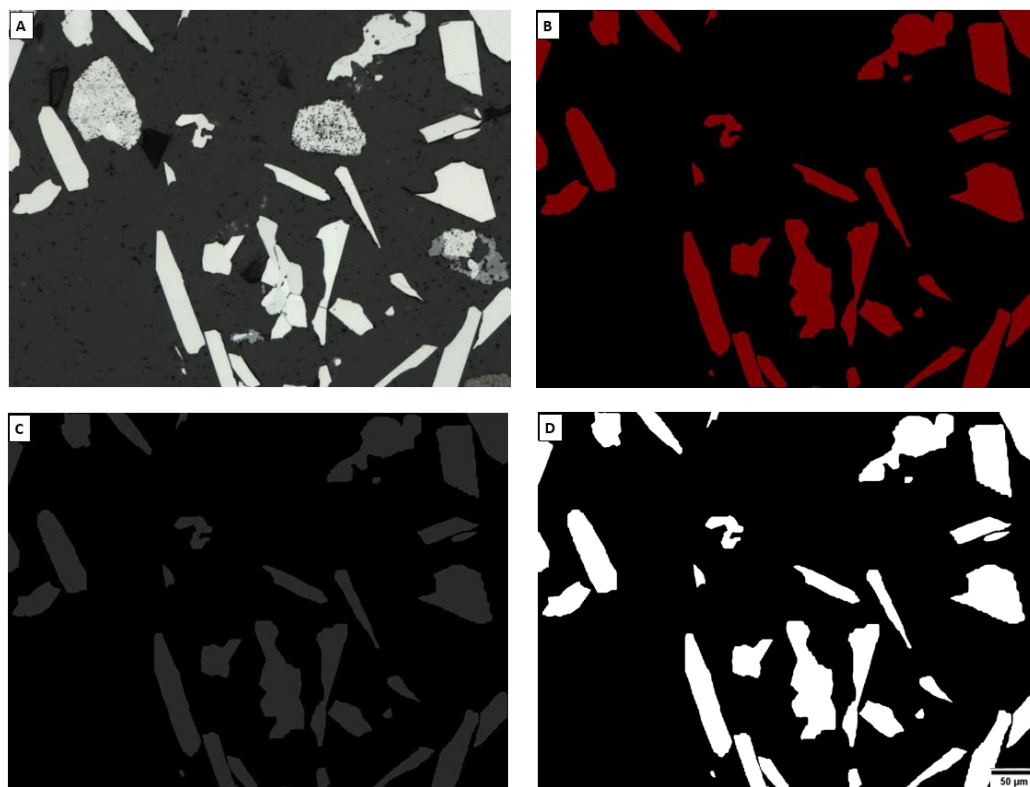


Figura 44 – Processamento realizado após a inferência do Modelo 1A. A) Exemplo de imagem original utilizada para inferência do modelo 1A, B) imagem RGB obtida a partir da inferência do Modelo 1A, em vermelho a classe compacta, C) imagem em 8bits, D) imagem binária com as partículas da classe compacta segmentadas.

Seguindo para o Modelo 1B, e considerando a imagem original que contém as compactas em preto (Figura 45A) é obtida a imagem (RGB) mostrada na Figura 45B, na qual os objetos em vermelho são as partículas da classe Martita e os objetos

em verde são correspondentes às partículas da classe Porosa. Neste caso, uma limiarização comum não funcionaria, pois, ao converter esta imagem para 8 bits todos os objetos possuem uma mesma tonalidade. Então, foi necessário separar em canais R, G e B, excluindo o canal B. É possível observar os canais R e G nas Figura 45C e Figura 45D, respectivamente.

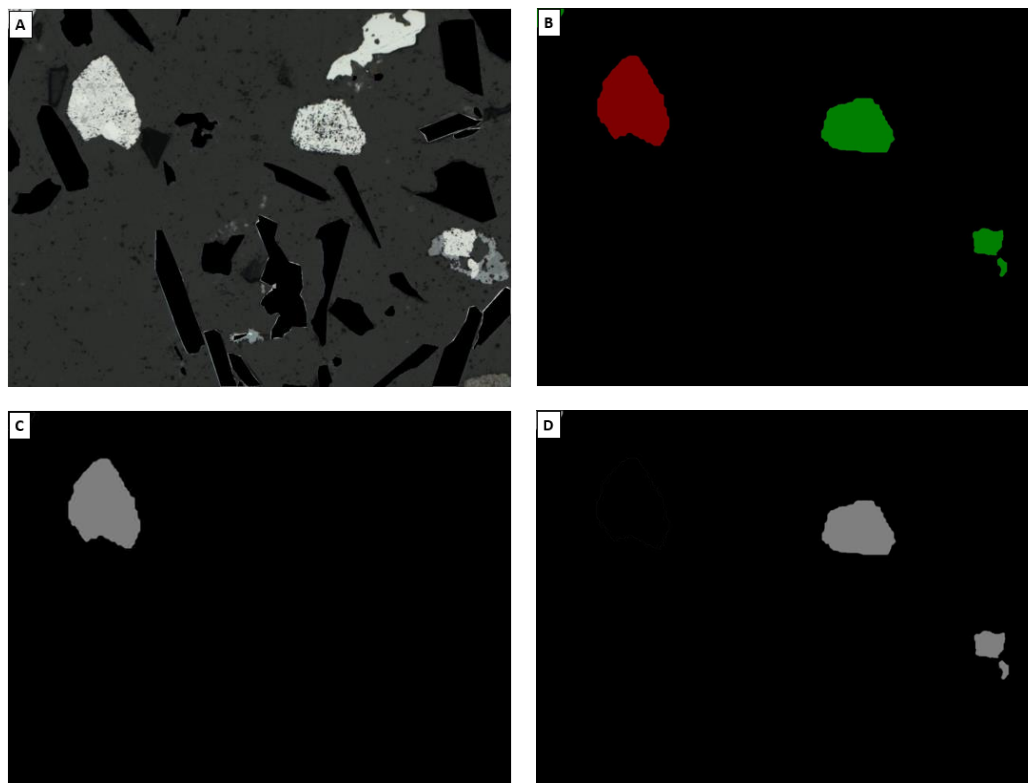


Figura 45 - Processamento realizado após a inferência do Modelo 1B. A) Exemplo de imagem original utilizada para inferência do modelo 1B, B) Imagem RGB obtida a partir da inferência do modelo 1B, C) Imagem em tons de cinza obtida a partir do canal R, D) Imagem em tons de cinza obtida a partir do canal G.

Posteriormente foi possível realizar a segmentação automática da classe Martita no canal R, limiar (68-255) e da Porosa no canal B, limiar (67-255). Depois da obtenção da imagem binária da Martita, foi realizada uma interseção com a imagem original para que fosse possível obter a Martita original e segmentar a magnetita presente nestas partículas. A Figura 46 apresenta esta segmentação focada na partícula de Martita presente na imagem apresentada nesta discussão.

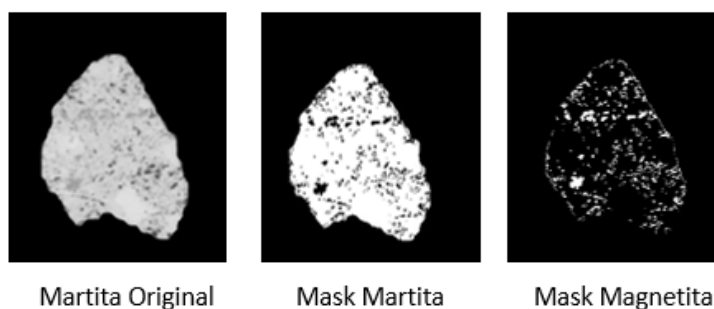


Figura 46 - Demonstração da segmentação da magnetita presente numa partícula de Martita.

Além disso, foi realizada também a limiarização da Goethita na imagem original (Figura 47).

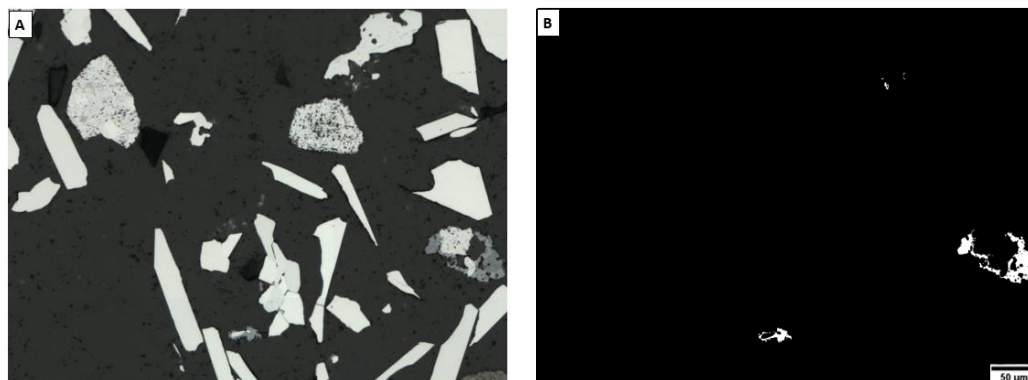


Figura 47 – Comparação entre A) imagem original e B) imagem binária referente a segmentação da Goethita com base em (A).

A partir da obtenção destas imagens binárias foi possível realizar medidas quantitativas de todas as classes de interesse. Por fim, a Figura 48 apresenta uma visualização da segmentação de todas as classes, através do comparativo entre uma imagem rotulada (Figura 48a) e uma figura que une todas as segmentações realizadas pelo Modelo 1. Nesta imagem cada uma das classes de interesse está apresentada com uma *Look Up Table* (LUT) diferente, conforme apresentado na legenda (Figura 48b). A partir desta comparação é possível observar os acertos do modelo para cada uma das classes reconhecidas por ele.

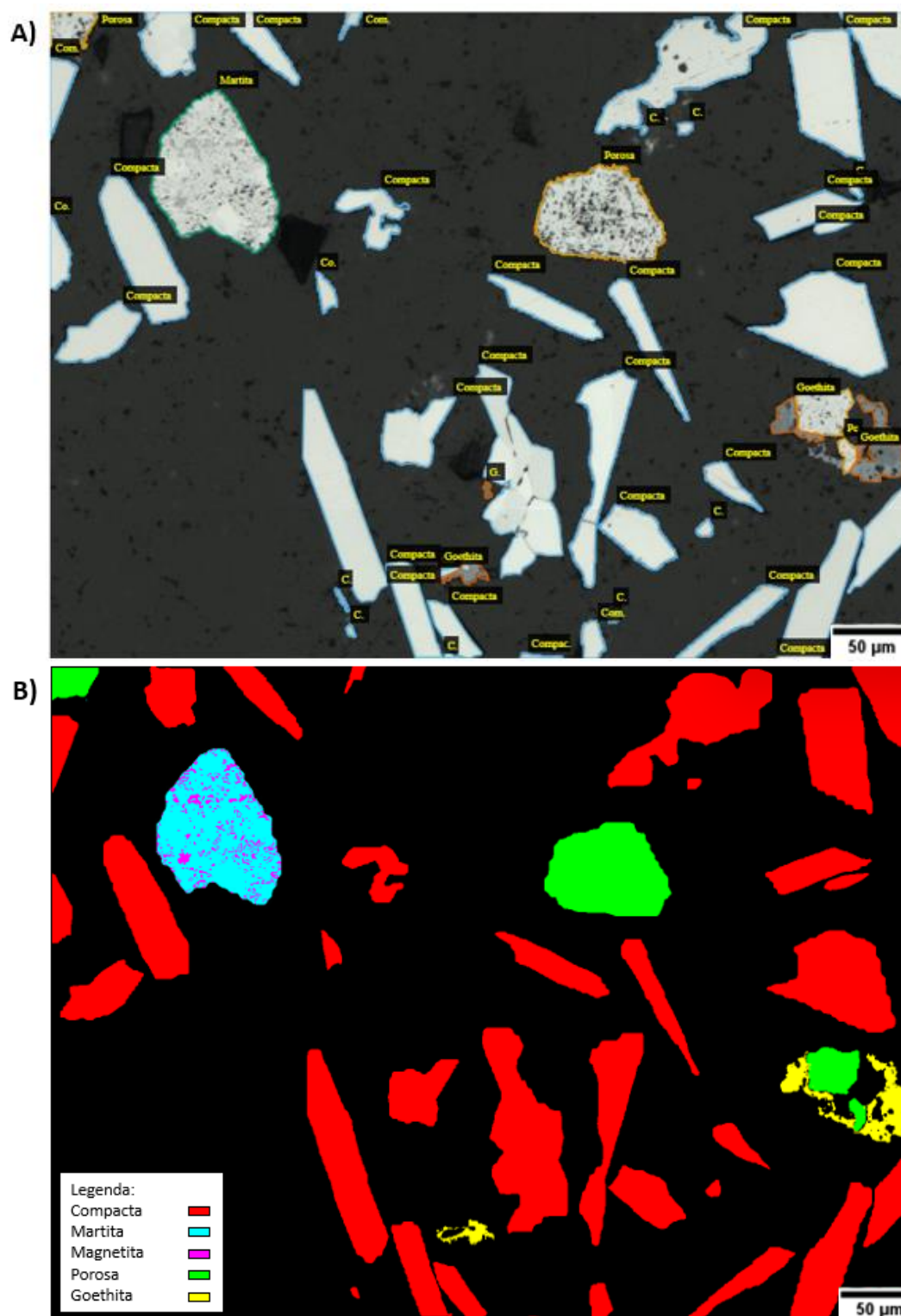


Figura 48 - Comparação entre A) imagem original rotulada e B) imagem contendo todas as classes de interesse do Modelo 1 identificadas por cores diferentes conforme a legenda na imagem.

A partir da obtenção das imagens binárias de cada classe foi realizada a extração de atributos para cada uma das imagens de validação do *dataset*.

### 5.1.3. Extração de Atributos

Nesta seção será apresentado um estudo exploratório sobre a extração de atributos das partículas segmentadas no Modelo 1. Conforme mencionado nesta discussão, a partir da metodologia desenvolvida foi possível segmentar as seguintes classes em imagens BF (Modelo 1):

- Compacta,
- Martita,
- Magnetita,
- Porosa e
- Goethita.

Portanto, o objetivo desta seção é apresentar resultados quantitativos, avaliando possíveis padrões de forma e tamanho que possam existir para estas classes. As imagens foram calibradas com a escala correspondente (tamanho do pixel = 0,53  $\mu\text{m}$ ) e realizou-se as medidas para cada partícula segmentada de cada classe em questão.

Foram escolhidos os seguintes parâmetros:

- Fração de Área (*Area Fraction* - AF),
- Razão de Aspecto (*Aspect Ratio* – AR),
- *Feret* Máximo,
- Circularidade e
- *Roundness*.

A AF representa a razão entre a área de todos os objetos de uma classe e a área total da imagem. O AR representa uma medida de proporcionalidade entre o eixo menor e eixo maior de cada objeto. Neste caso a medida foi feita utilizando como referência as medidas de *Feret* Mínimo e Máximo, respectivamente. O *Feret* Máximo fornece exatamente a medida do maior eixo do objeto.

A circularidade é dada por:  $4\pi \cdot \text{área} / \text{perímetro}^2$ . Objetos mais alongados possuem circularidade mais próxima de zero e um objeto perfeitamente circular, apresenta circularidade igual a 1. Esta medida é muito sensível a irregularidades no contorno dos objetos de forma que mesmo objetos circulares podem apresentar Circularidade  $\ll 1$  se o contorno foi muito irregular.

A medida de *Roundness* é dada por  $4 \cdot \text{area} / (\pi \cdot \text{Feret}^2)$ . Objetos mais alongados possuem *Roundness* mais próxima de zero e um objeto perfeitamente circular, apresenta *Roundness* igual a 1. Esta medida é pouco sensível a irregularidades no contorno dos objetos e mais sensível ao alongamento do objeto.

O primeiro parâmetro explorado foi a AF de cada classe nas diversas imagens. Os gráficos apresentados a seguir seguem o mesmo padrão de cores da imagem segmentada apresentada na Figura 48b da seção anterior.



Observando estes gráficos (Figura 49) é possível perceber que a classe Compacta possui fração de área bem maior que as demais classes. As classes Martita e Magnetita ocupam uma fração abaixo de 7%, a classe Magnetita em especial ocupa uma área muito pequena. Em observações qualitativas das imagens, foi possível notar que a maioria das partículas de Martita encontrava-se no final do processo de martitização, por isso a baixa presença de Magnetita. As partículas da classe Porosa apresentaram frações abaixo de 10%, e as da classe Goethita abaixo de 5%. Todas estas informações são respaldadas inclusive pela representatividade das classes, que correspondem a estas proporções.

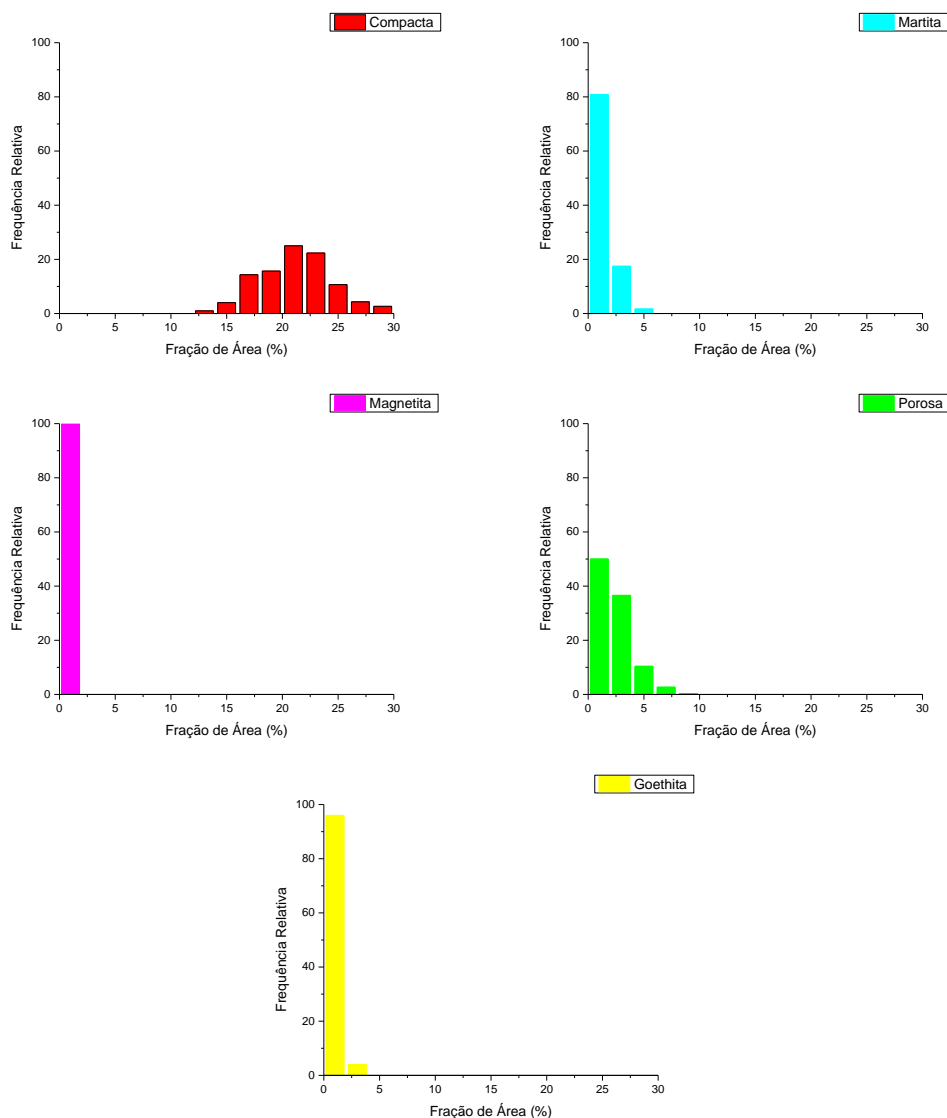


Figura 49 - Distribuição de Fração de Área para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme a legenda dos gráficos.

Para o caso da classe Magnetita, mediu-se apenas a AF. Por não se tratar especificamente de partículas desta classe, não faria sentido realizar medidas de tamanho e forma. Conforme explicado anteriormente, a Magnetita presente nas amostras é na verdade resquício desta classe nas partículas de Martita antes de completar o processo de martitização.

O segundo parâmetro apresentado nesta discussão é o AR (Figura 50). Para as Compactas é possível observar que existe uma variação bem ampla deste parâmetro. Isso indica que a classe compacta possui partículas de formas variadas, o que é perceptível observando as imagens. Já para o caso das Martitas, as maiores frequências estão entre 0,6 e 0,8 de AR, indicando partículas mais próximas de uma forma equiaxial. Para as Porosas a maior frequência está por volta de  $AR = 0,7$ , o que indica semelhança com as partículas de Martita. Para as Goethitas as maiores frequências estão em  $AR = 0,5$ , também apresentando uma distribuição ampla.

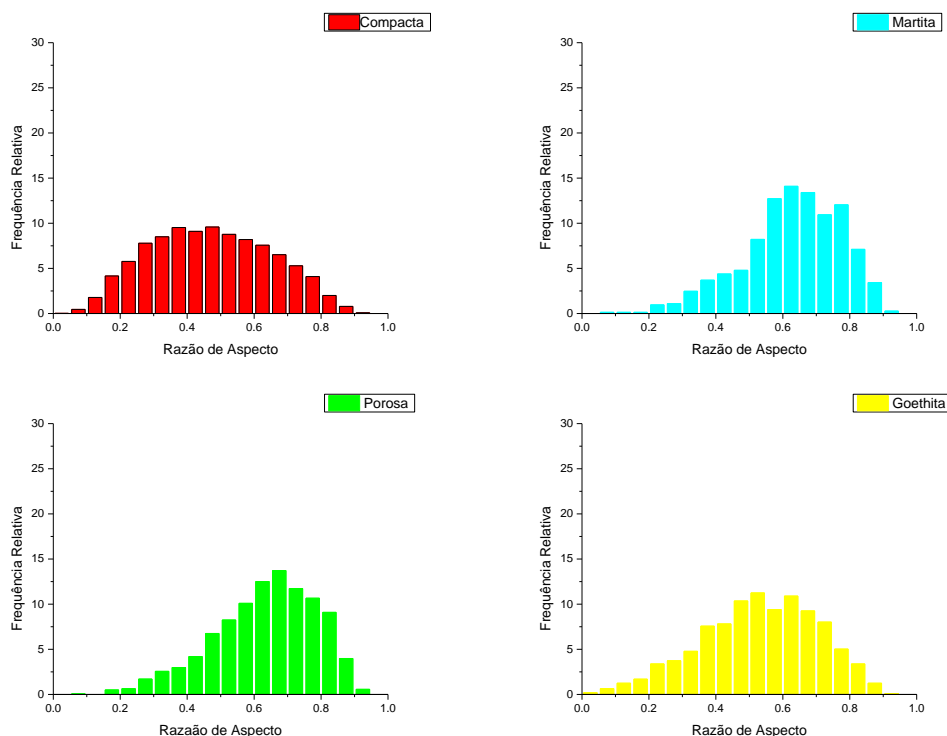


Figura 50 - Distribuição de AR para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme a legenda dos gráficos.

O próximo parâmetro a ser observado é o *Feret* Máximo (Figura 51). Para as partículas da classe Compacta, o eixo máximo variou de valores bem próximos a 0 até 200  $\mu\text{m}$ , com frequências maiores de partículas entre 25  $\mu\text{m}$  e 50  $\mu\text{m}$ . Esta classe realmente possui uma variação no tamanho das partículas, já observada qualitativamente. As maiores frequências para a Martita também estão por volta de 50  $\mu\text{m}$ , enquanto que para a Porosa estão um pouco abaixo de 50  $\mu\text{m}$ . Segundo Liming [16], as hematitas típicas de utilização no *pellet feed* possuem dois principais intervalos de tamanho de 45 a 120  $\mu\text{m}$  e de 45 a 60  $\mu\text{m}$ . Esta informação está de acordo com o obtido nas medidas mostradas de *Feret* Máximo. Para a Goethita os valores estão concentrados abaixo de 50  $\mu\text{m}$ . É válido ressaltar que além das partículas de Goethita existem áreas em volta de partículas de outras classes ou entre partículas que possuem coloração de Goethita. Considerando que estas áreas podem ser muito pequenas, isso pode contribuir bastante para as variações de tamanho.



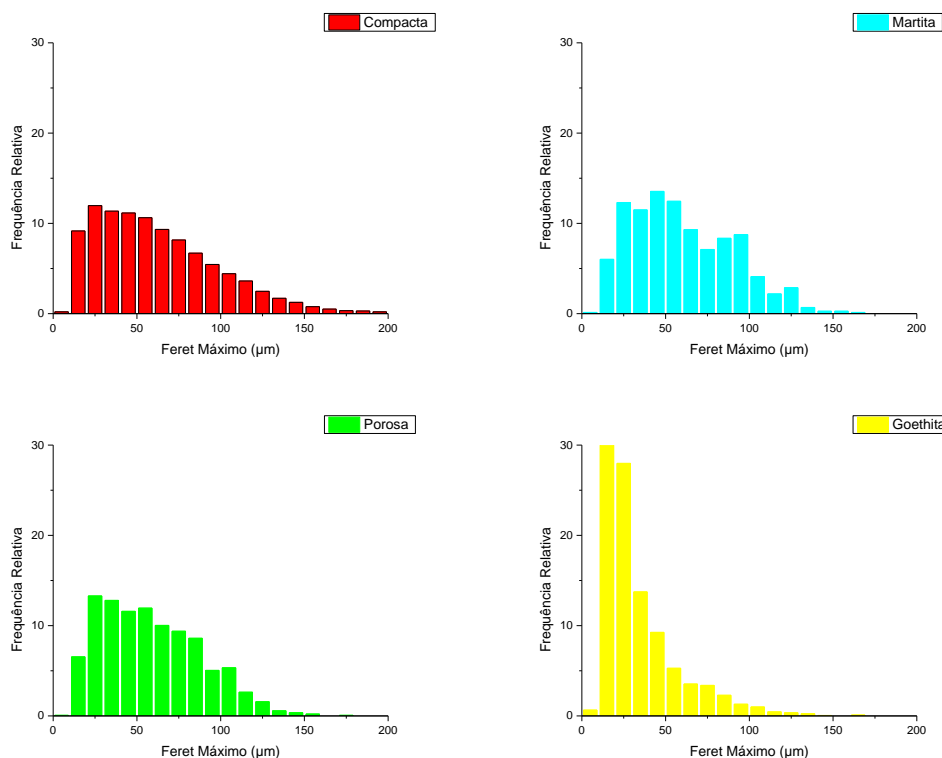


Figura 51 - Distribuição de *Feret* Máximo para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme a legenda dos gráficos.

Quanto aos resultados referentes à circularidade foi possível observar nos gráficos abaixo (Figura 52) as variações para cada classe. Para a classe Compacta, observou-se uma variação desde 0,1 até 0,9, com frequências maiores próximo a 0,7. Estes resultados apontam para partículas mais próximas de um formato circular. Observando qualitativamente estas partículas é possível perceber que a grande maioria das partículas é mais lamelar, apesar de também existirem algumas mais arredondadas, principalmente as menores. Considerando que a circularidade é sensível ao perímetro das partículas, estes resultados também podem estar apontando para partículas com bordas mais regulares. Já para as Martitas, os valores variam um pouco mais para a esquerda no eixo x do gráfico, com frequências maiores próximo a 0,5. Essa classe de partículas já possui bordas mais irregulares, o que pode estar afetando o resultado da circularidade. Para as Porosas, as frequências maiores estão por volta de 0,6 e para a Goethita abaixo de 0,2. Conforme mencionado anteriormente, o tamanho e formato dos objetos segmentados na classe Goethita são bem variados, isso também foi refletido para as medidas de forma.

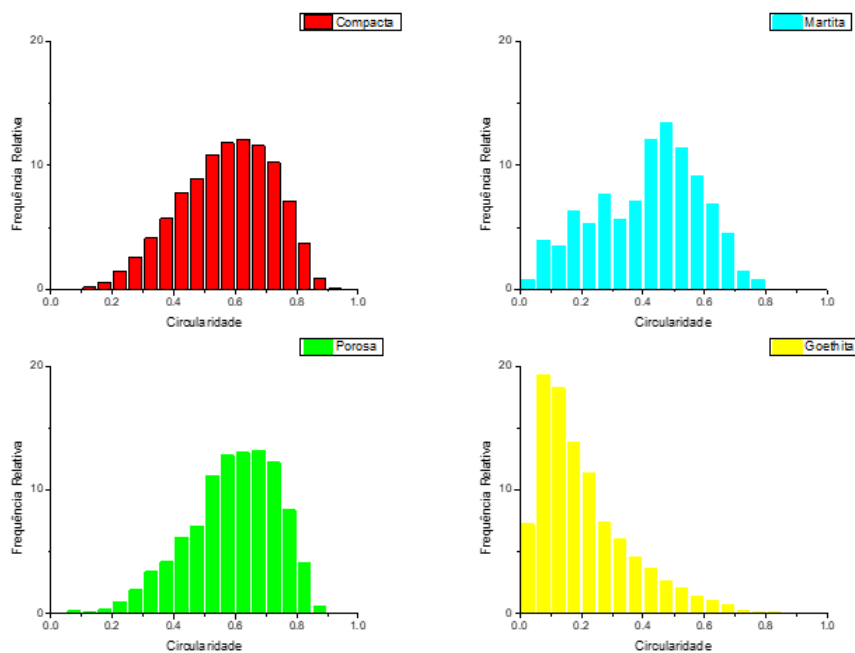


Figura 52 - Distribuição de Circularidade para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme as legendas dos gráficos.

Os resultados para *Roundness* estão apresentados na Figura 53. Para a classe Compacta as frequências maiores ficaram bem mais à esquerda do gráfico (abaixo de 0,5), quando comparados às frequências da circularidade para esta classe, por exemplo. Estes resultados concordam com o que se observa sobre o formato das partículas desta classe analisando as imagens. Para a classe Martita as maiores frequências estão logo acima de 0,6, assim como para as Porosas, indicando mais uma vez semelhança de forma entre estas classes. No caso da Goethita já se percebe uma distribuição bem mais larga, variando de próximo a 0 até 1, diferente do que se observa no gráfico de circularidade.

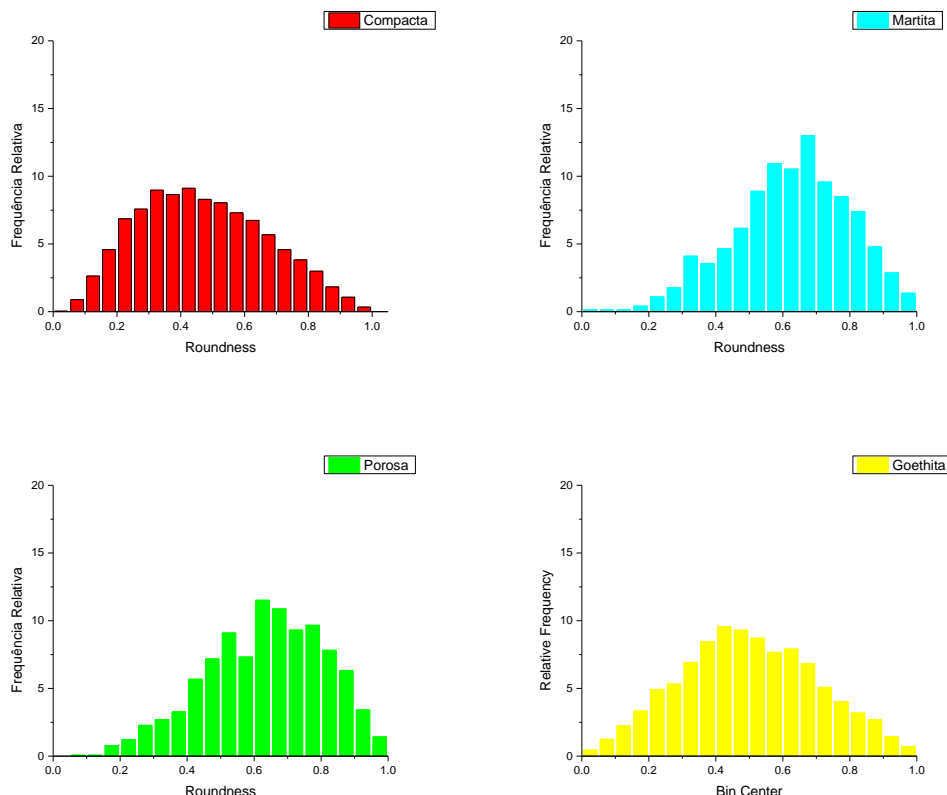


Figura 53 - Distribuição de *Roundness* para as classes Compacta, Martita, Porosa e Goethita, representadas pelas cores conforme as legendas dos gráficos.

A partir deste estudo exploratório, foi possível concluir que a classe Compacta possui partículas com AF maior em relação às demais classes, bordas mais regulares, e com formas variadas, predominantemente mais alongadas. As partículas de Martita ocupam áreas menores nas imagens, possuem contornos irregulares e formato mais equiaxial. A classe Porosa apresenta características bem semelhantes à classe Martita. As Goethitas são partículas, em geral, menores e mais distantes de um formato circular.

## 5.2. Resultados Modelo 2

### 5.2.1. Desenvolvimento do Modelo

Conforme foi descrito anteriormente, os modelos 1 e 2 estão relacionados. E de acordo com a hierarquia apresentada, o Modelo 2 identificará as classes Monocristalina, Policristalina e Martita em imagens CPOL. Neste modelo, as Martitas são relacionadas à classe Porosa do Modelo 1. As Martitas identificadas no Modelo 1, estarão em preto nas imagens utilizadas para o Modelo 2, pois já foram identificadas corretamente no modelo anterior. A Figura 54 apresenta o processamento realizado, no qual as Martitas do Modelo 1 (Figura 54a) são colocadas em preto na imagem CPOL (Figura 54b). Desta forma as imagens originais do Modelo 2 foram utilizadas conforme mostrado na Figura 54b.

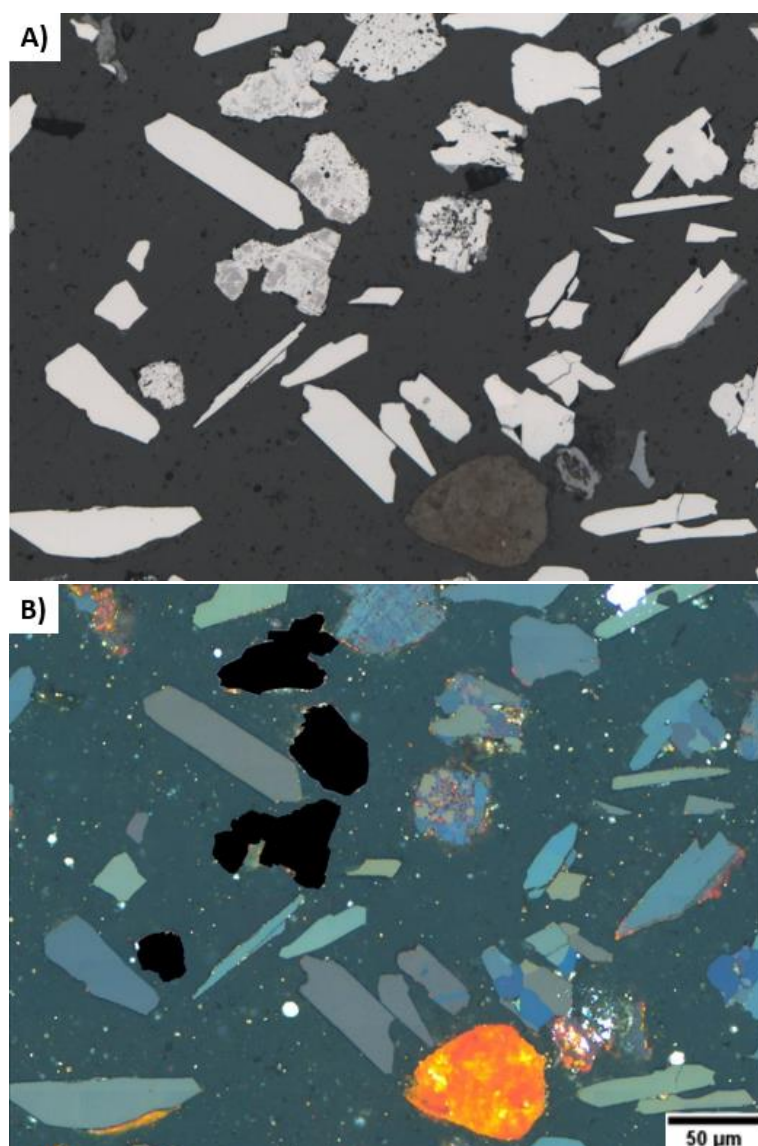


Figura 54 - Representação do processamento realizado no banco de dados para o Modelo 2. Em A) a imagem original BF utilizada no Modelo 1 e em B) a imagem CPOL utilizada para o Modelo 2 com as Martitas identificadas no Modelo 1 em preto.

Todos os treinamentos realizados neste trabalho consideraram 1500 épocas e o tempo de processamento variou um pouco em relação à quantidade de dados por treinamento. Conforme mencionado na discussão do desenvolvimento do Modelo 1, foi escolhida uma ordem mais compreensível para os leitores para apresentar os resultados (eles não necessariamente foram obtidos na ordem apresentada). Os modelos 1 e 2 foram construídos simultaneamente e cada resultado apresentado é importante para justificar as decisões tomadas e explicar a construção dos modelos. Para que a discussão de todos os testes realizados seja mais organizada e didática, a Tabela 13 apresenta uma descrição de cada teste por ordem de apresentação nesta seção.

Tabela 13 - Descrição das principais características dos treinamentos realizados para o desenvolvimento do Modelo 2.

Treinamento	Classes de Interesse	Característica das Imagens	Quantidade de Imagens
1	Monocristalina, Policristalina, Martita, Goethita	Originais sem pré-processamento	300
2	Monocristalina, Policristalina, Martita	Goethitas em preto por limiarização na BF	300
3	Monocristalina, Policristalina, Martita	Goethitas em preto por limiarização na BF e retirada do <i>background</i>	300
4	Monocristalina, Policristalina, Martita	Goethitas em preto por limiarização na BF e retirada do <i>background</i>	900
5	Monocristalina, Policristalina, Martita	Goethitas em preto por limiarização na BF, retirada do <i>background</i> , preenchimento de buracos, exclusão de objetos menores e revisão de contornos da rotulagem	900

Desta forma, o Treinamento 1 considerou as classes Monocristalina, Policristalina, Martita e Goethita + 1 *background* e utilizando as imagens originais do Modelo 2. Este treinamento usou o *dataset* 1 (300 imagens). Neste primeiro treinamento foi considerada a classe Goethita para a qual, que até o momento deste treinamento, ainda não se tinha certeza de que forma identificá-la. Este treinamento durou cerca de 6 dias para ser concluído. As métricas gerais para o Treinamento 1 ficaram em torno de 80%, conforme pode ser observado na Tabela 14. As métricas gerais são muito boas, mas olhando para as métricas individuais foi possível identificar problemas. A classe Monocristalina, por exemplo apresentou métricas altas, e a Policristalina apresentou métricas razoavelmente boas. Mas para as classes Martita e Goethita foram obtidas métricas com valores bem baixos.

Ao observar as métricas e as imagens foi possível perceber que ainda existiam muito erros, principalmente relacionados às classes Martita e Goethita. Um fato relevante é que este primeiro teste foi realizado com poucas imagens, e principalmente a classe Martita ainda era muito pouco representada (Tabela 3). A quantidade de partículas disponíveis para o aprendizado ainda parecia insuficiente, de acordo com o observado. Além disso, existe a questão da semelhança entre partículas da classe Martita e Policristalina, o que torna a identificação mais complicada.

Tabela 14 - Métricas gerais e individuais obtidas a partir da validação do modelo obtido no Treinamento 1.

Classe\Métricas	<i>Precision (%)</i>		<i>Recall (%)</i>		<i>F1-score (%)</i>	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Monocristalina</b>	86,22	91,16	80,48	88,11	83,25	89,61
<b>Policristalina</b>		79,84		65,28		71,83
<b>Martita</b>		46,34		35,85		40,43
<b>Goethita</b>		48,97		41,63		45,01

Com relação à Goethita foram observados erros semelhantes aos observados no Modelo 1. Muitas Goethitas são confundidas com o *background* pela diversidade de características de suas partículas e muitas se confundem com cada uma das outras classes, reduzindo as métricas. Esta foi a última tentativa incluindo a classe Goethita nos modelos. A partir deste ponto decidiu-se realizar sua identificação por limiarização, nas imagens BF, conforme apresentado nos resultados do Modelo 1. Deste modo, para o Treinamento 2 do Modelo 2, foi realizado um pré-processamento nas imagens envolvendo a limiarização da Goethita nas imagens BF e uma interseção com as imagens CPOL para obter as imagens de treinamento para o próximo teste. O Treinamento 2 foi realizado considerando, então, três classes (Monocristalina, Policristalina e Martita + 1 *background*) e as imagens originais foram pré-processadas retirando as Goethitas por limiarização. O *dataset* utilizado foi o *dataset* 1 (300 imagens). A duração deste treinamento foi em torno de 5 dias.

Ao realizar limiarização da Goethita existe um impacto também nas outras partículas que possuem coloração de Goethita em seu interior. Isso ocorre principalmente com as Martitas, que ficam “esburacadas”, conforme exemplificado na Figura 55.

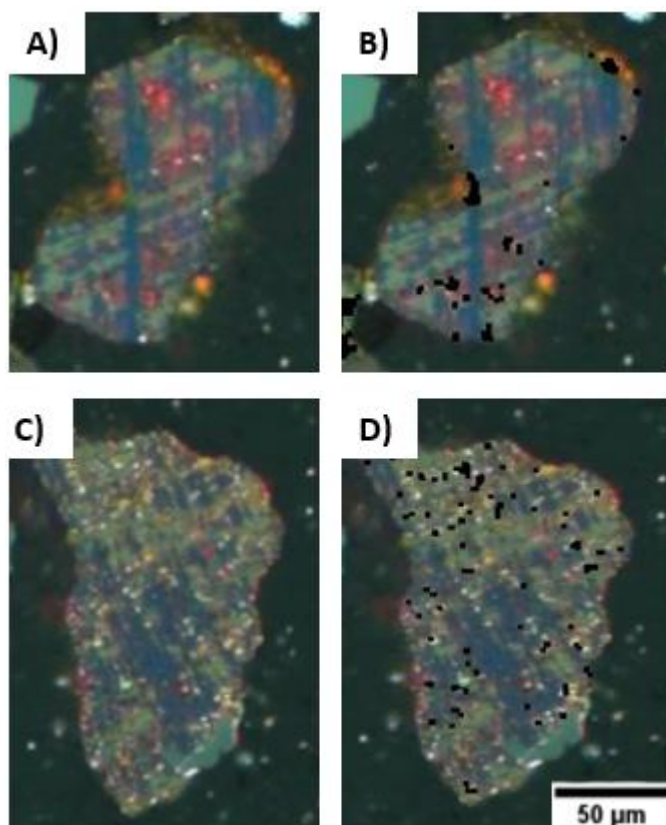


Figura 55 - Exemplos de partículas danificadas após a limiarização da Goethita.  
A) e C) antes e B) e D) depois.

De início não se preocupou com este fato e o Treinamento 2 foi realizado para observar como o sistema se comportaria com esses problemas. De fato, observou-se que com este pré-processamento as métricas melhoraram. É possível observar na Tabela 15, que o F1-score geral obtido desta vez foi de 85,94%. Para as Monocristalinas houve uma pequena diminuição nas métricas em relação ao treinamento anterior, mas não tão impactante, dado que a identificação desta classe é muito boa. Já para as classes Policristalina e Martita houve um pequeno aumento do F1-score.

Tabela 15 - Métricas gerais e individuais obtidas a partir da validação do modelo obtido no Treinamento 2.

Classe\Métricas	Precision (%)		Recall (%)		F1-score (%)	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Monocristalina</b>	89,16	91,32	82,94	86,59	85,94	88,89
<b>Policristalina</b>		79,48		66,30		72,29
<b>Martita</b>		44,23		43,40		43,81

Os resultados já eram muito bons, exceto os da Martita. Então, decidiu-se realizar mais um teste com esse mesmo *dataset* antes de utilizar um *dataset* com mais dados para os próximos testes. Observando as imagens, percebeu-se que alguns objetos espúrios do *background*, inclusive resquícios de limiarização da Goethita, traziam bastante ruído para as imagens e atrapalhavam um pouco as identificações. Desta forma, decidiu-se realizar um processamento que retirasse o fundo das imagens, obtendo apenas as partículas de interesse sobre um fundo preto. A Figura 56 apresenta um exemplo de imagem sem fundo e sem a presença de Goethita.

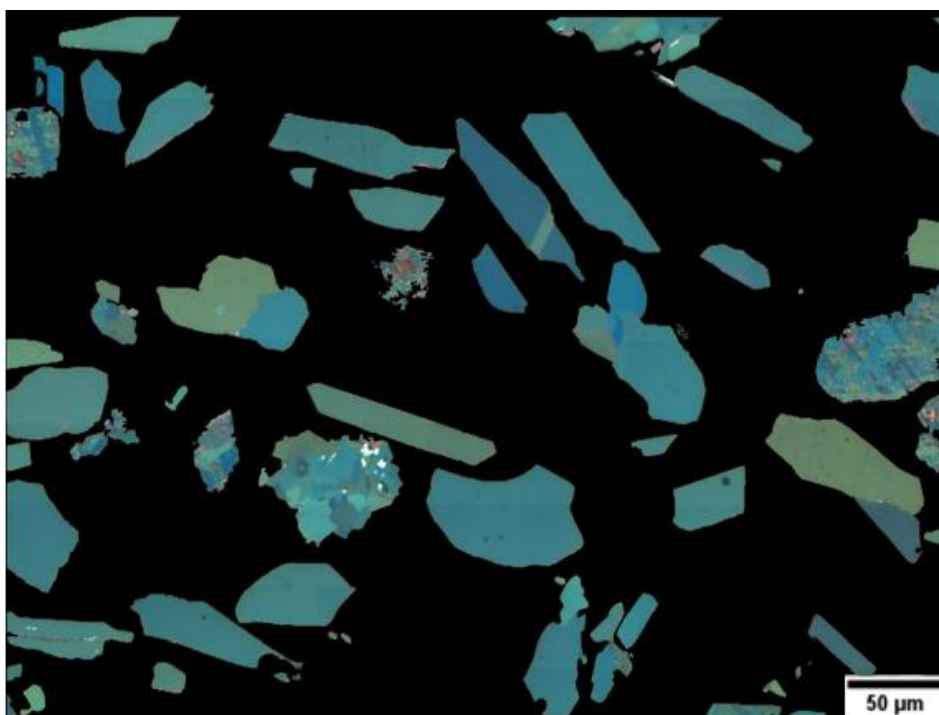


Figura 56 - Exemplo de imagem utilizada no Treinamento 3, após a limiarização da Goethita e a retirada do fundo com objetos espúrios.

Então, o Treinamento 3 foi realizado com estas imagens pré-processadas, utilizando o *dataset* 1 e considerando as três classes de interesse (Monocristalina, Policristalina e Martita). Este treinamento durou cerca de 5 dias para ser concluído. A maioria das métricas aumentou. A Tabela 16 apresenta a comparação de métricas gerais entre os três treinamentos realizados.

Tabela 16 - Comparação entre as métricas gerais obtidas para as validações dos Treinamentos 1,2 e 3 do Modelo 2.

Treinamentos\Métricas	<i>Precision</i> (%)	<i>Recall</i> (%)	<i>F1-score</i> (%)
<b>Treinamento 1</b>	86,22	80,48	83,25
<b>Treinamento 2</b>	89,16	82,94	85,94
<b>Treinamento 3</b>	90,60	87,99	89,28



A partir dos resultados obtidos no Treinamento 3, decidiu-se seguir com esta configuração de treinamento; as imagens com a Goethita limiarizada e o fundo em preto. As métricas da Martita não melhoraram muito como as das demais classes, com estas modificações, mas decidiu-se manter, uma vez que se acreditou que o maior problema com a classe Martita era sua pouca representatividade. Portanto, o treinamento seguinte já utilizou o *dataset* 2 (900 imagens), entendendo que o incremento contribuisse para a melhoria dos resultados. Desta forma, o Treinamento 4 foi realizado exatamente como o Treinamento 3, com a mesma configuração de pré-processamento e considerando as três classes (Monocristalina, Policristalina e Martita + 1 *background*) e desta vez com mais imagens. Este treinamento durou cerca de 12 dias para ser concluído.

Conforme esperado, com o aumento dos dados as métricas da classe Martita aumentaram, F1-score chegou a 73,85 % conforme pode ser constatado na Tabela 17. Isso foi uma excelente confirmação de que a quantidade de dados está diretamente ligada a melhoria do aprendizado do modelo. Porém, neste caso, ao tempo que se inclui mais dados também se inclui variabilidade de informações, partículas diferentes etc. Então houve também diminuição nas métricas das classes Monocristalina e Policristalina. Apesar disso, as métricas de validação continuaram boas, mas foi preciso parar e avaliar com mais cuidado as imagens. Nesta avaliação, foi possível perceber o impacto da limiarização da Goethita, que até aquele momento não parecia prejudicar, mas observou-se que os maiores problemas aconteciam exatamente em partículas “danificadas” pela limiarização.

Conforme dito anteriormente, isto acontecia principalmente nas partículas de Martita, mas não exclusivamente nelas. Então, notou-se que esses “buracos” se tornaram, para o modelo, uma característica da classe. Quando uma partícula de Martita estava esburacada, ele a classificava corretamente, mas quando os buracos apareciam em partículas de outra classe, a partícula era identificada como Martita erroneamente, conforme apresentado nos exemplos da Figura 57. Estes exemplos estão apresentados mostrando a rotulagem (em amarelo) e a inferência do modelo (máscara de segmentação em verde) para cada partícula.

Tabela 17 - Métricas gerais e individuais obtidas a partir da validação do modelo obtido no Treinamento 3.

Classe\Métricas	Precision (%)		Recall (%)		F1-score (%)	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Monocristalina</b>	89,49	92,63	80,92	82,96	84,99	87,53
<b>Policristalina</b>		75,67		72,11		73,85
<b>Martita</b>		63,93		58,50		73,85

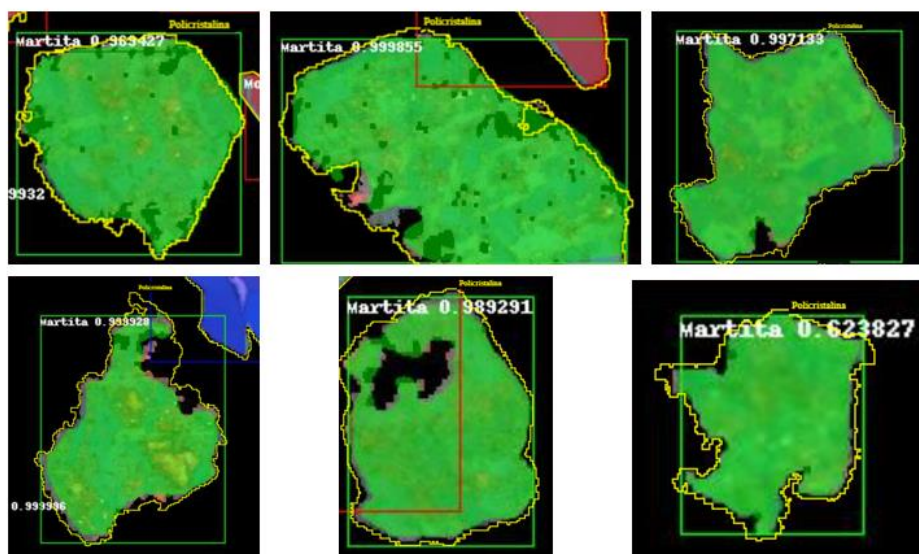


Figura 57 - Exemplos de partículas da classe Policristalina (rotulagem em amarelo) que foram inferidas pelo modelo como Martita (máscara de segmentação em verde).

Tendo isso em vista, o processo de limiarização da Goethita foi revisado a fim de incluir operações que minimizassem os problemas principais que são as partículas esburacadas e os resquícios da limiarização (partículas muito pequenas). Para lidar com as partículas esburacadas foram utilizadas operações como *fill holes*, *close*, *erode* e *dilate*. O resultado deste processamento pode ser observado na Figura 58, que compara o antes (Figura 58a) e depois (Figura 58b) do processamento em uma imagem de exemplo. As partículas indicadas por números e setas estão apresentadas ao lado das figuras com um certo aumento para melhorar a visualização. Quanto a partículas muito pequenas foi realizado um *scrap*, que eliminou partículas abaixo de um certo tamanho.

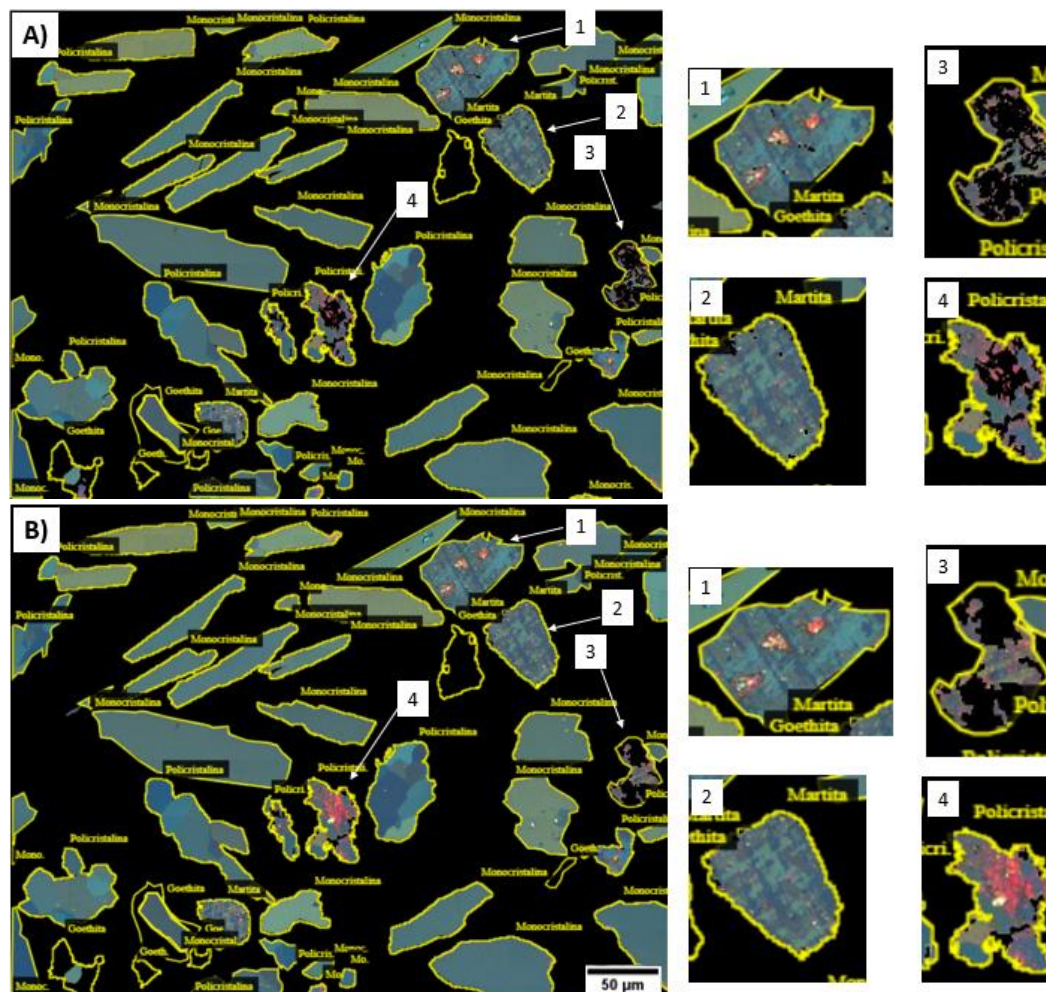


Figura 58 - Exemplo do pós-processamento utilizado para reduzir os ruídos de buracos nas partículas e objetos muito pequenos que são resquícios da limiarização da goethita: A) Antes do pós-processamento e B) imagem pós processada.

Mesmo após o processamento algumas partículas continuaram “deformadas” e isso acontecia mais nas bordas das partículas, o que comprometia a rotulagem. Ou seja, havia espaço em preto circunscrito à rotulagem (observe a Figura 59a e Figura 59c). Pensando nisso, decidiu-se criar uma rotulagem que se adaptasse à forma destas partículas pós-processadas. Mas rotular todas as partículas novamente seria um trabalho demorado e exaustivo. Então, criou-se uma rotina capaz de realizar esse processamento automaticamente, utilizando a rotulagem já existente e as partículas pós-processadas. Deste modo, o primeiro passo foi gerar máscaras das partículas rotuladas em imagens separadas, ou seja, a partir de um conjunto de 900 imagens originais, foram obtidas 900 imagens contendo apenas as máscaras de Monocristalina, 900 imagens contendo apenas as máscaras de Policristalinas e 631 imagens contendo apenas as máscaras de Martita (as Martitas não estão presentes em todas as imagens do *dataset*). Isso foi realizado a partir de uma rotina que cria máscaras de determinada classe a partir da rotulagem. Depois disso, foi realizada uma interseção das imagens pós-processadas contendo as partículas “deformadas” com cada uma destas máscaras, obtendo, em imagens separadas, as partículas

originais de cada classe já pós-processadas. E a partir destas imagens foi gerado automaticamente o arquivo de texto com a rotulagem correta de cada classe.

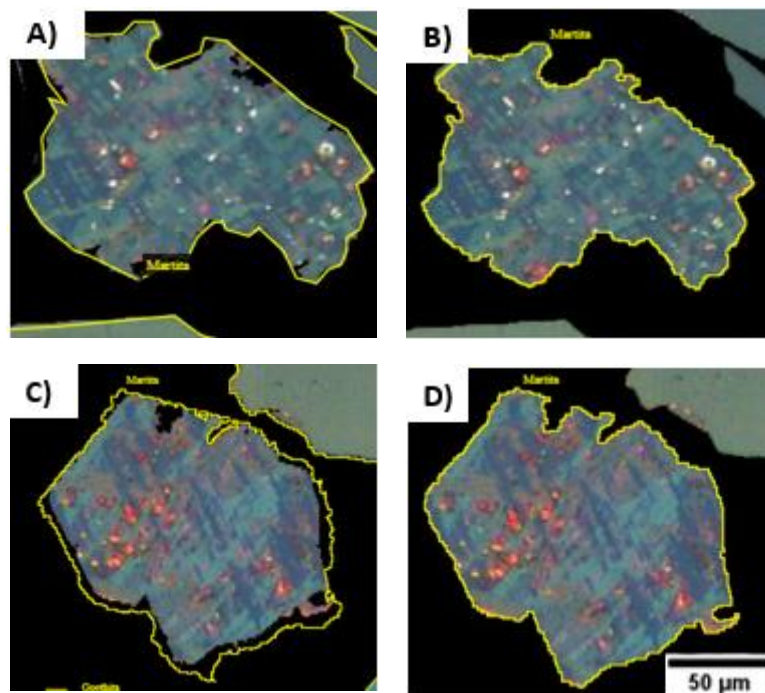


Figura 59 - Comparação de partículas rotuladas. Antes, A) e C), e depois, B) e D), de gerar a nova rotulagem

Em posse das imagens pós-processadas e o novo *json* foi então realizado o Treinamento 5. Utilizando o *dataset 2*, considerando as classes Monocristalina, Policristalina e Martita + 1 *background*. A duração deste treinamento foi de cerca de 11 dias. As métricas gerais e individuais obtidas para a validação deste treinamento estão apresentadas na Tabela 18.

Tabela 18 - Métricas gerais e individuais obtidas a partir da validação do Treinamento 5.

Classe\Métricas	Precision (%)		Recall (%)		F1-score (%)	
	Geral	Individual	Geral	Individual	Geral	Individual
<b>Monocristalina</b>		94,69		92,76		93,71
<b>Policristalina</b>	91,31	80,87	89,61	79,51	90,45	80,18
<b>Martita</b>		73,76		74,38		74,07

As métricas obtidas para o Modelo 2 são realmente impressionantes, considerando a complexidade das classes e a quantidade de variações que elas possuem. Os gráficos apresentados na Figura 60 apresentam a evolução das métricas obtidas para os treinamentos realizados do Modelo 2.

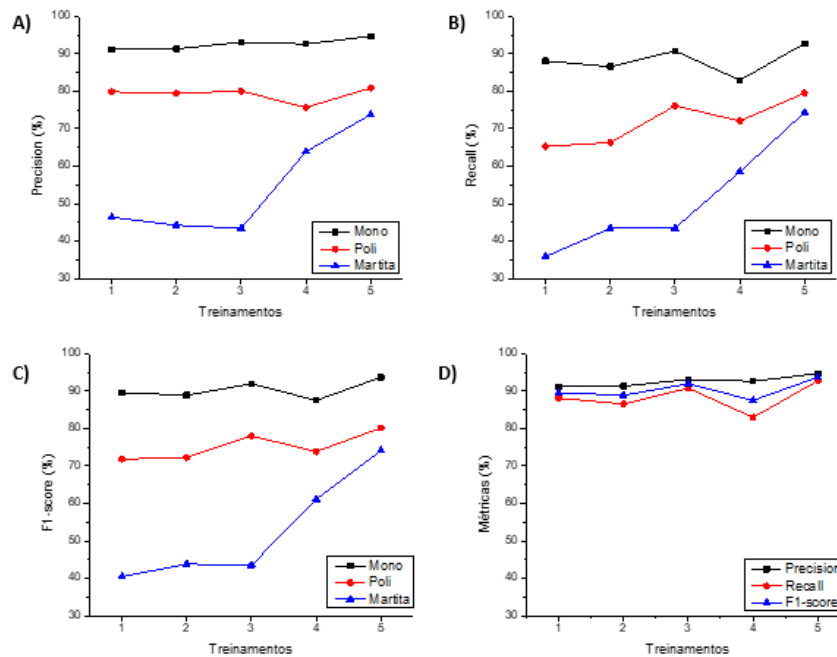


Figura 60 - Gráficos comparativos das métricas obtidas a partir das validações dos Treinamentos 1,2,3,4 e 5. As métricas individuais por classe em A) *Precision* B) *Recall* C) *F1-score* e em D) as métricas gerais.

Observando os gráficos apresentados na Figura 60 é possível observar a evolução das métricas *Precision*, *Recall*, *F1-score* (Figura 60a, Figura 60b e Figura 60c, respectivamente) para cada uma das classes. A medida que os treinamentos foram sendo realizados, é possível notar um aumento considerável das métricas para a Martita quando o fundo das imagens foi retirado e o conjunto foi incrementado (Treinamento 4). Para as demais classes houve uma redução da métrica neste mesmo ponto, mas que foi superada após o pós-processamento focado no preenchimento de buracos e na eliminação de objetos pequenos, além da nova anotação gerada, que se adaptou às partículas deformadas (Treinamento 5).

A Figura 60d, apresenta uma relação das três métricas gerais com os treinamentos realizados. É possível observar que apesar dos problemas iniciais, as métricas gerais sempre se mantiveram acima de 80%. Mas a partir de observações nas imagens e nas métricas individuais é que foi possível perceber que elas precisavam melhorar. Ficou claro que as decisões tomadas foram corretas uma vez que se pode ver nas imagens e nas métricas gerais e individuais os impactos das modificações. Desta forma, considerou-se o modelo obtido a partir do Treinamento 5 adequado para realizar a identificação, classificação e segmentação das três classes de interesse do Modelo2.

Desta forma, o processamento realizado no Modelo 2 inicia-se colocando as Martitas do Modelo 1 em preto (Figura 61a). Em seguida é realizada a limiarização da Goethita na imagem BF e a interseção com a CPOL correspondente, e ainda, o fundo ruidoso é excluído juntamente com partículas espúrias muito pequenas oriundas da limiarização da Goethita, gerando uma imagem como a mostrada na Figura 61b. Na Figura 61c a mesma imagem rotulada exemplificando o *ground*



*truth* e finalmente na Figura 61d está apresentada a imagem de inferência com as classes separadas por cor conforme a legenda na figura.

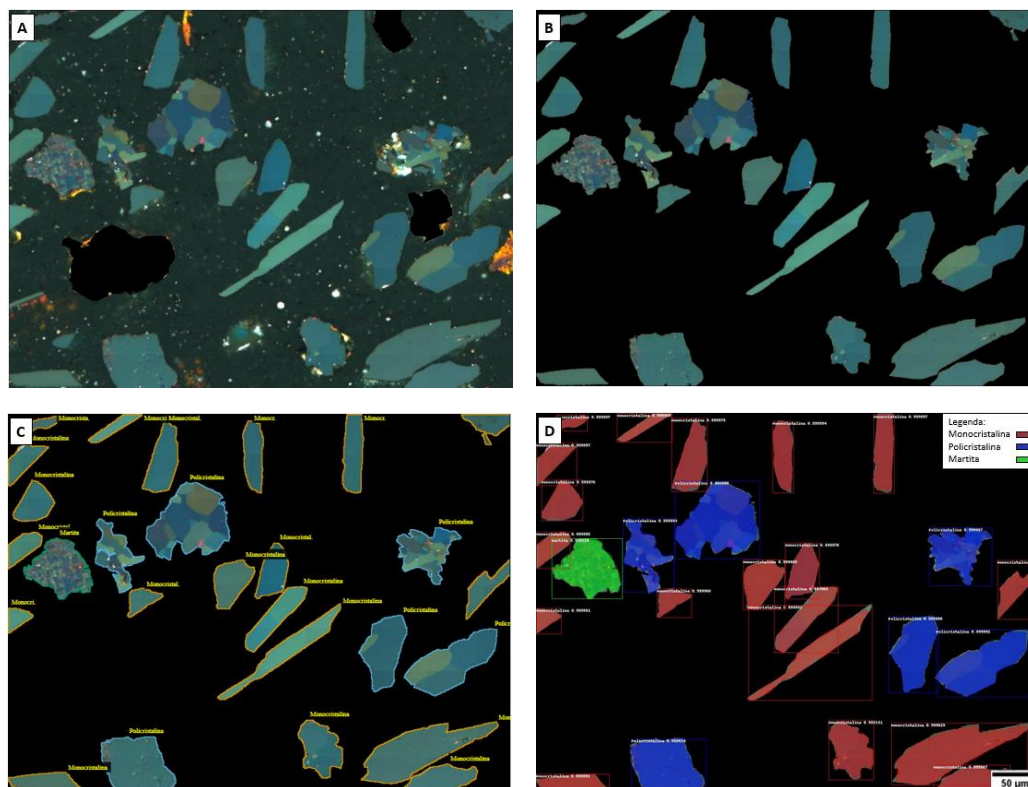


Figura 61 – Exemplificação do processamento realizado no Modelo 2. A) imagem original utilizada como input do modelo (com martitas do Modelo 1 em preto), B) imagem processada com limiarização da Goethita, retirada do fundo para redução de ruído e exclusão de objetos muito pequenos, C) input do modelo contendo as rotulagens correspondentes, D) Imagem de inferência do Modelo 2 com todas as classes de interesse identificadas corretamente conforme mostrado na legenda em cores.

Após o desenvolvimento do modelo que possibilitou a inferência das classes, foi possível obter as máscaras com estas classes segmentadas. A próxima seção apresentará o pós-processamento realizado nestas imagens.

### 5.2.2. Pós-processamento

O Modelo 2 foi construído para ter apenas uma etapa. Então, após realizar a inferência com o modelo desenvolvido numa imagem (Figura 62a), obtém-se uma imagem (RGB) com cada uma das três classes representada por um canal (Figura 62b). Nesta imagem, os objetos em vermelho são as partículas da classe Monocristalina, os objetos em verde são correspondentes às partículas da classe Martita e os objetos em azul são da classe Policristalina. Da mesma forma que ocorre para o Modelo 1, é preciso separar os canais R, G e B para obter as imagens binárias de cada classe. Deste modo, é possível observar os canais R (Figura 62C), G (Figura 62D) e B (Figura 62E).

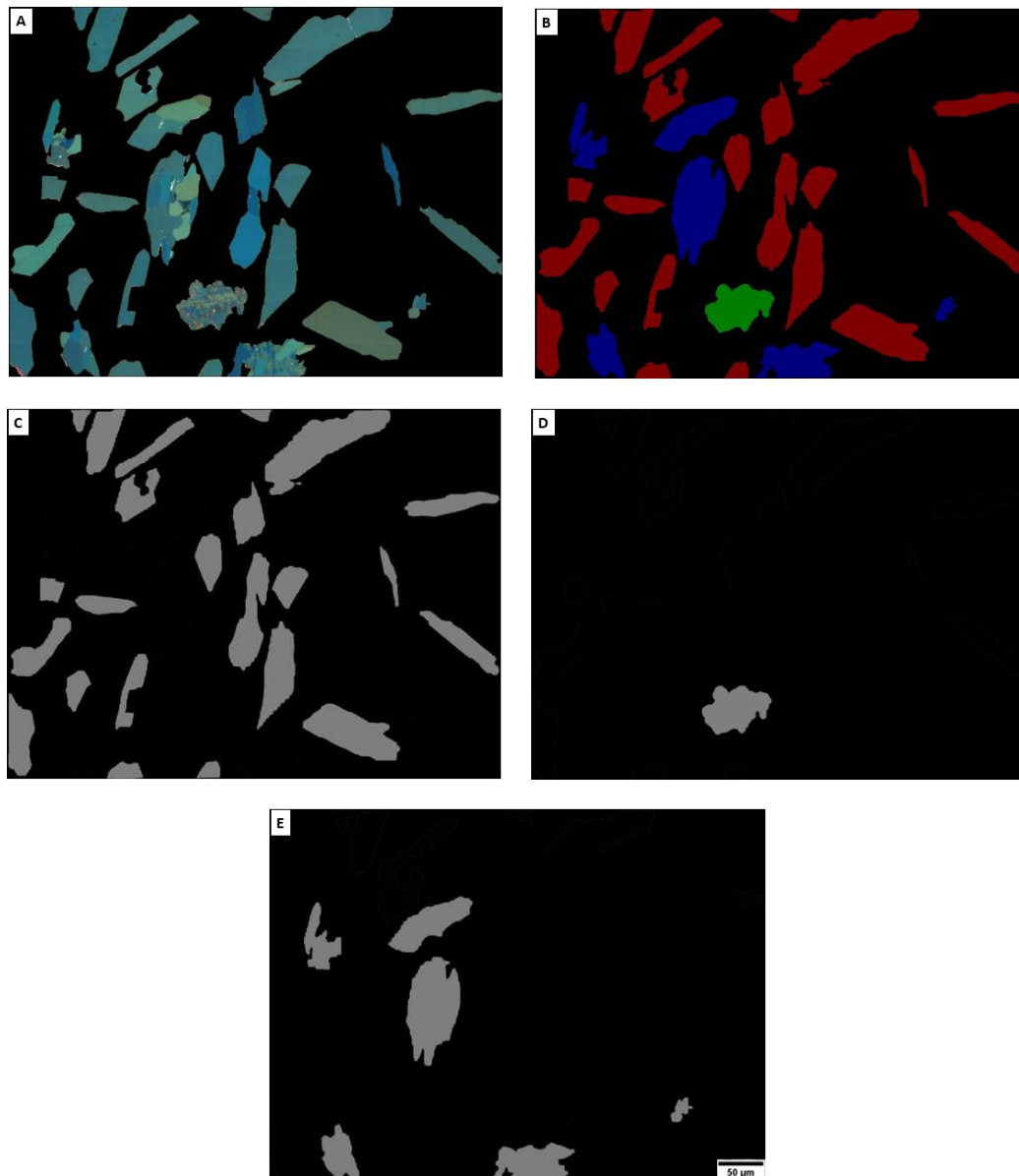


Figura 62 – Exemplo de pós processamento realizado após a inferência do Modelo 2 na imagem A) original, em B) Imagem RGB obtida a partir da inferência do Modelo 2; em vermelho estão as partículas da classe Monocristalina, em verde as da classe Martita e em azul as da classe Policristalina. C) Imagem em tons de cinza obtida a partir do canal R (Monocristalinas), D) Imagem em tons de cinza obtida a partir do canal G (Martitas), E) Imagem em tons de cinza obtida a partir do canal B (Policristalinas).

Posteriormente foi possível realizar a segmentação da classe Monocristalina no canal R, limiar (68-255), da Martita no canal G, limiar (66-255) e da Policristalina no canal B, limiar (66-255). A partir destas imagens binárias foi possível realizar as medidas das classes. Para fins de visualização a Figura 63 apresenta uma comparação entre uma imagem original rotulada (Figura 63A) e uma figura que une todas as segmentações realizadas pelo Modelo 2 (Figura 63B). Nesta imagem cada uma das classes de interesse está apresentada com uma LUT diferente, conforme apresentado na legenda. A partir desta comparação é possível observar

de forma qualitativa os acertos do modelo para cada uma das classes reconhecidas por ele.

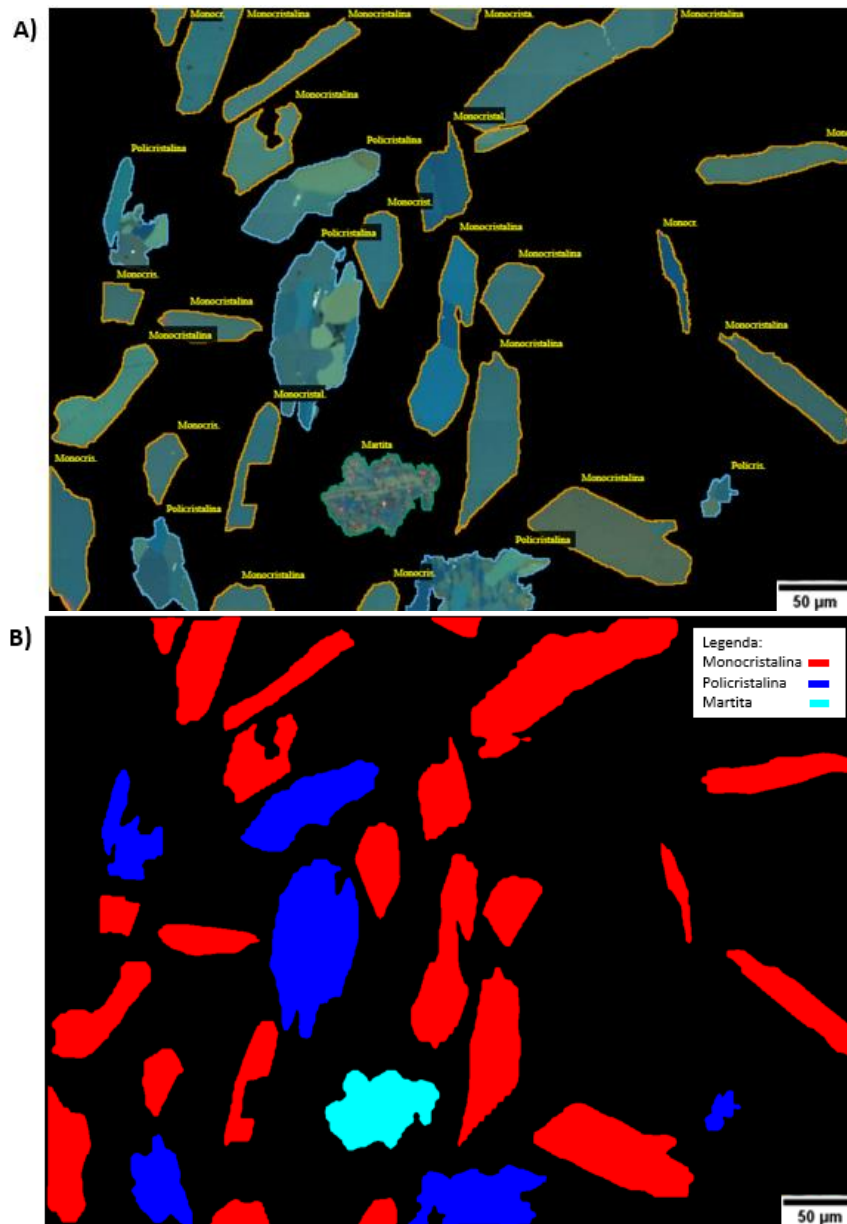


Figura 63 - Comparação entre A) imagem original rotulada e B) imagem contendo todas as classes de interesse do Modelo 2 identificadas por cores diferentes conforme a legenda na imagem.

A partir da obtenção das imagens binárias de cada classe foi realizada a extração de atributos para cada uma das imagens de validação do *dataset*.

### 5.2.3. Extração de Atributos

Nesta seção será apresentado um estudo exploratório sobre a extração de atributos das partículas segmentadas no Modelo 2, da mesma maneira descrita no para o Modelo 1.



Conforme mencionado nesta discussão, a partir da metodologia desenvolvida, foi possível segmentar as seguintes classes em imagens CPOL (Modelo 2):

- Monocristalina,
- Policristalina e
- Martita.

Portanto, o objetivo desta seção é apresentar resultados quantitativos, avaliando possíveis padrões de forma e tamanho que possam existir para estas classes. As imagens foram calibradas com a escala correspondente (tamanho do pixel =  $0,53 \mu\text{m}$ ) e realizou-se as medidas para cada partícula segmentada de cada classe em questão. Foram escolhidos os mesmos parâmetros discutidos anteriormente para o Modelo 1: Fração de Área (*Area Fraction* - AF), Razão de Aspecto (*Aspect Ratio* - AR), *Feret* Máximo, Circularidade e *Roundness*. Os gráficos a seguir apresentam a mesma escala de cores representada na imagem segmentada (Figura 63) da seção anterior.

Observando os gráficos de AF abaixo (Figura 64), é possível notar que as partículas da classe Monocristalina ocupam em geral, uma área maior que as Policristalinas e Martitas. A classe Monocristalina não possui partículas que ocupam frações de área muito abaixo de 5%, enquanto para as Policristalinas as frações estão a um limite próximo a 15%. Para as Martitas as frações são, em geral, ainda menores (abaixo de 8%).

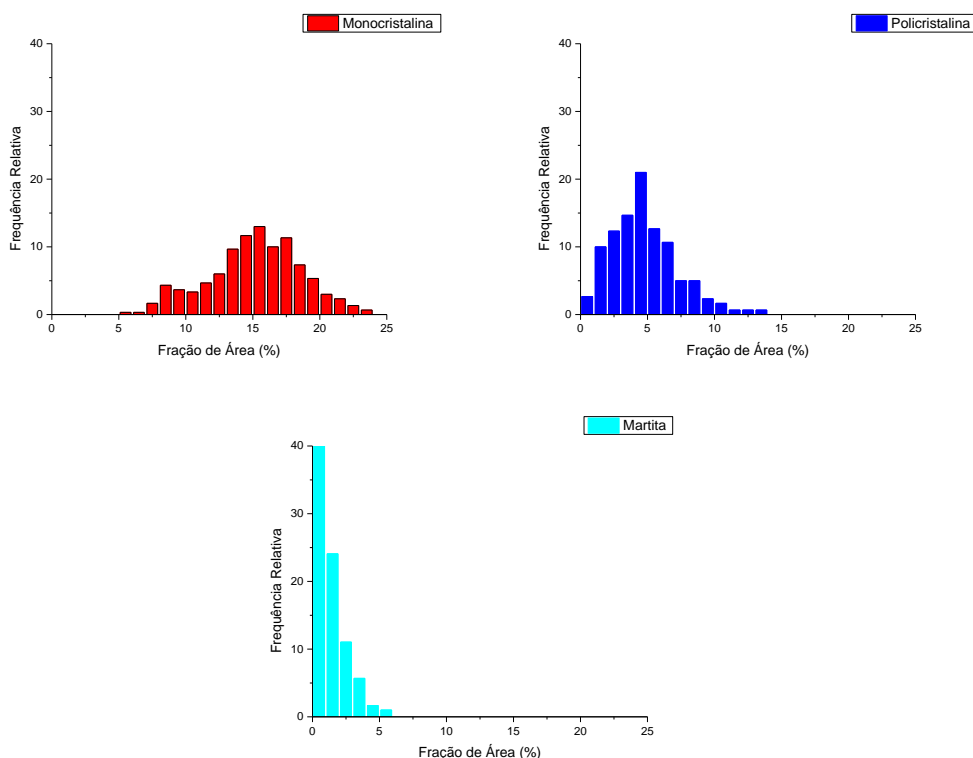


Figura 64 - Distribuição de Fração de Área para as classes Monocristalina, Policristalina e Martita, representadas pelas cores conforme a legenda dos gráficos.

Em relação ao AR (Figura 65), as classes Monocristalina e Policristalina apresentam distribuições que abrangem desde um pouco acima de  $AR=0$  a muito próximo de  $AR=1$ . Sendo a distribuição da classe Monocristalina mais deslocada para a esquerda (eixo x) e a da Policristalina mais para a direita. Isto indica que as partículas monocristalinas tendem a ser mais alongadas do que as policristalinas. A classe Martita também apresentou uma variação grande, mas com frequências bem maiores acima de  $AR=0,6$ , ou seja, estas partículas tendem a ser menos alongadas do que as mono ou policristalinas.

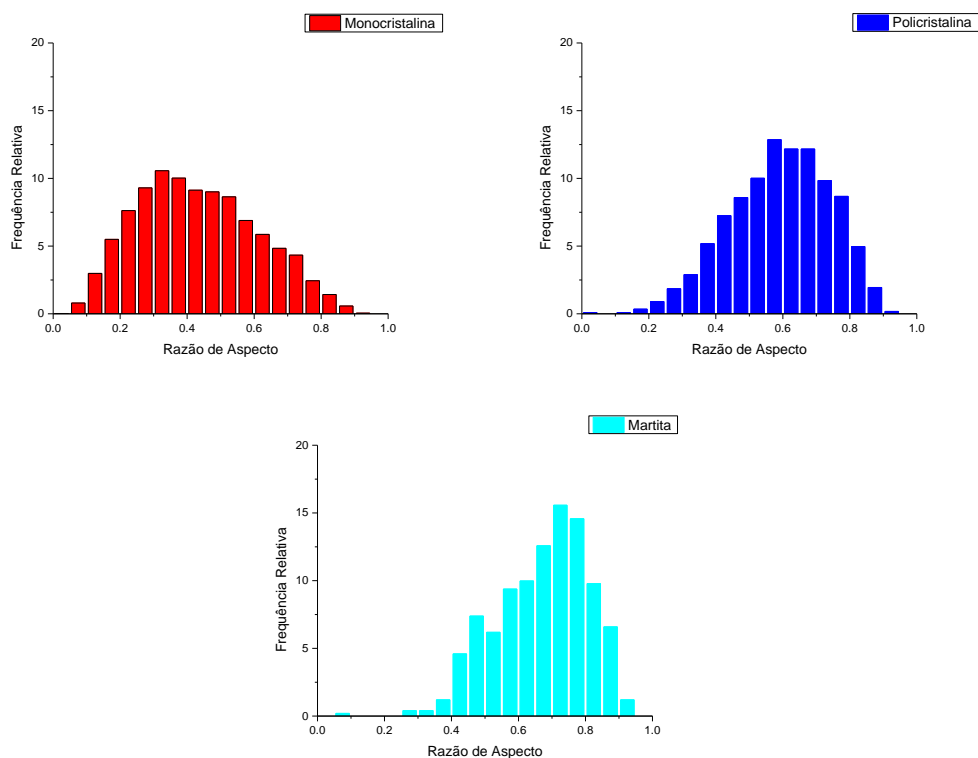


Figura 65 - Distribuição de AR para as classes Monocristalina, Policristalina e Martita, representadas pelas cores conforme a legenda dos gráficos.

Na Figura 66, é possível observar informações sobre a medida de *Feret* Máximo de cada classe de partículas. O comportamento dos gráficos das classes Monocristalina e Policristalina apresentam-se semelhantes, com uma leve tendência para valores maiores de *Feret* Máximo para a classe Monocristalina. Para a classe Martita, o gráfico indica um tamanho entre 50 e 100  $\mu\text{m}$  para maioria das partículas.

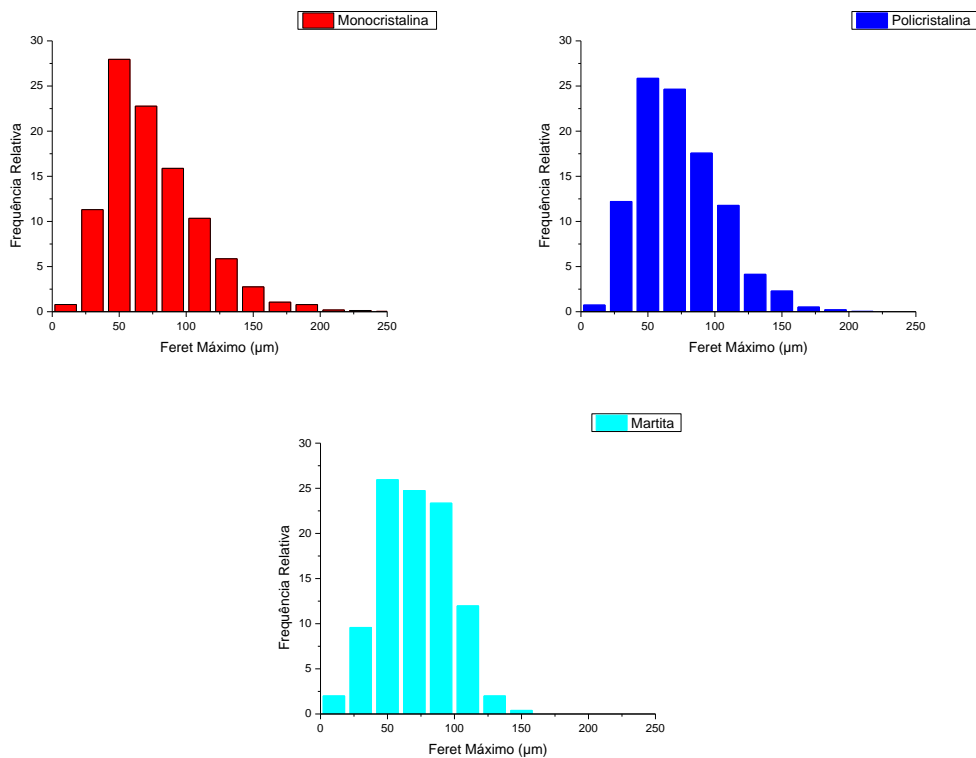


Figura 66 - Distribuição de *Feret* Máximo para as classes Monocrystalina, Polycrystalina e Martita, representadas pelas cores conforme a legenda dos gráficos.

Para a circularidade (Figura 67), as classes Monocrystalina e Polycrystalina também apresentaram comportamentos parecidos, com maiores frequências próximos a 0,6, com uma leve tendência a partículas polycristalinas mais circulares. Para a classe Martita, as maiores frequências estão abaixo de 0,5 indicando partículas mais distantes do formato circular.

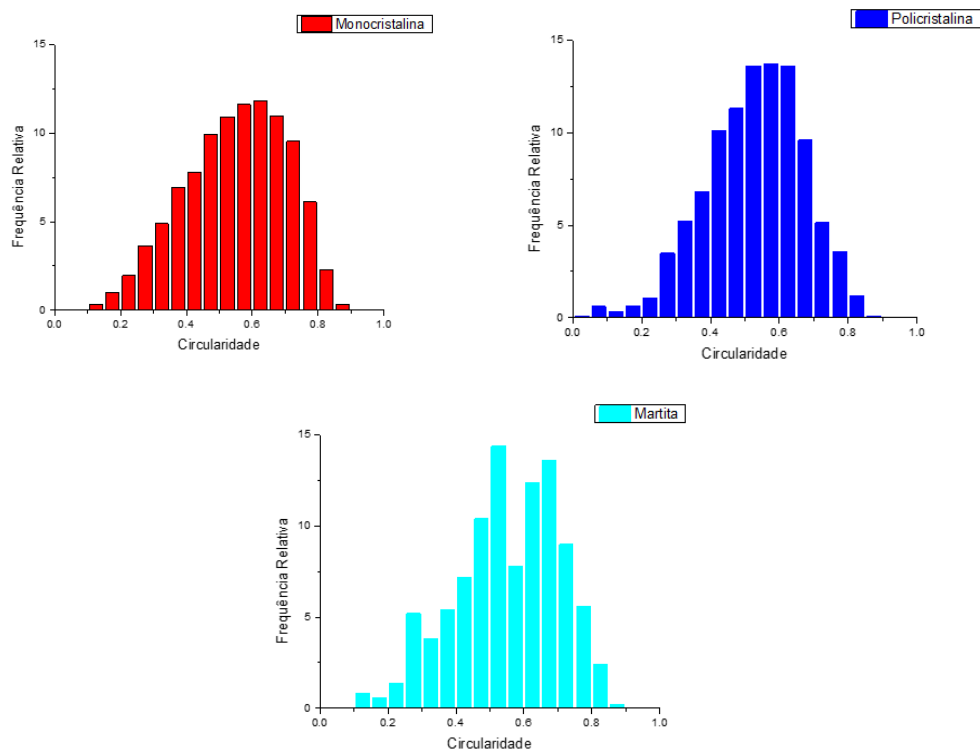


Figura 67 - Distribuição de Circularidade para as classes Monocristalina, Policristalina e Martita, representadas pelas cores conforme a legenda dos gráficos.

As medidas de *Roundness* (Figura 68) indicaram que as partículas monocristalinas tendem a ser mais alongadas do que as policristalinas, que por sua vez são mais alongadas do que as martitas.

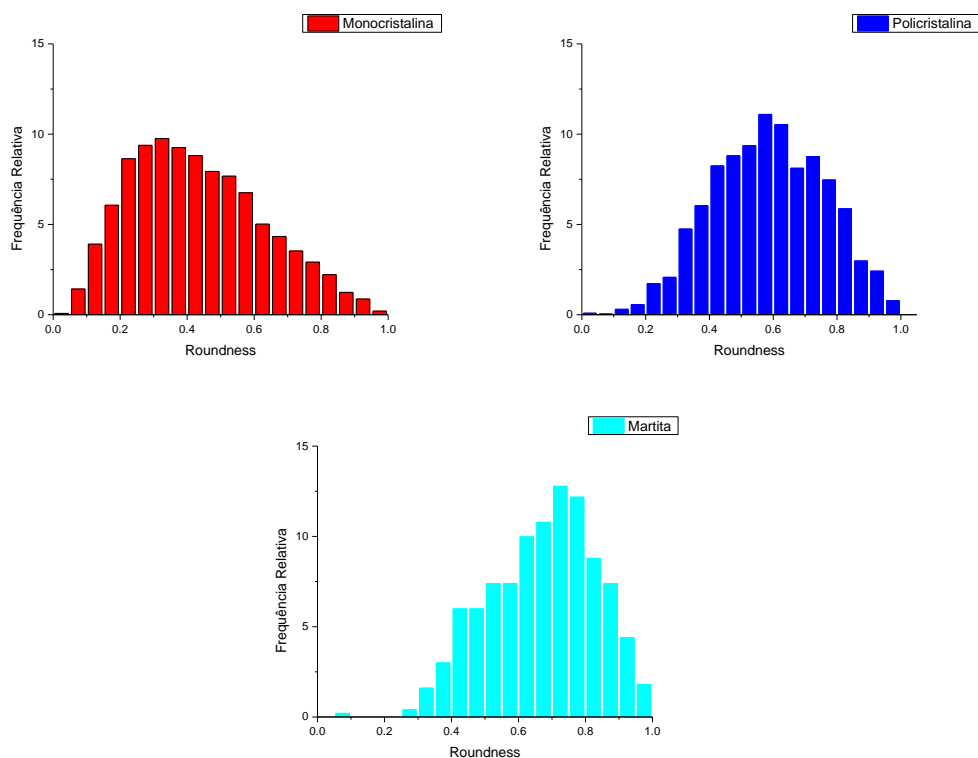


Figura 68 - Distribuição de *Roundness* para as classes Monocristalina, Policristalina e Martita, representadas pelas cores conforme a legenda dos gráficos.

Para as classes do Modelo 2, percebeu-se comportamentos muito semelhantes, o que permite confirmar a dificuldade de definir parâmetros simples que sejam capazes de discriminar estas classes. Na maioria das vezes existirá uma interseção entre características de classes diferentes.

### 5.3. Utilização do Sistema Desenvolvido

Tendo em vista o desenvolvimento destes modelos, um ponto importante a ser esclarecido é a forma como eles serão aplicados. Conforme foi mencionado, os treinamentos dos modelos exigiram uma máquina com alto poder de processamento, mas uma vez que foram obtidos os modelos finais, uma máquina mais simples pode realizar a inferência a partir destes modelos. Em um conjunto de 300 imagens, por exemplo a inferência durou cerca de 20 minutos variando um pouco para menos ou para mais dependendo das classes que se deseja inferir (quantidade de dados). Então, o que se pretende é unir todas as rotinas automáticas e modelos desenvolvidos nesta tese, gerando um sistema automatizado para ser utilizado no CTF da Vale. Este sistema automatizado seria formado por:

- Uma rotina automática que divide os mosaicos BF e CPOL obtidos no MO em ladrilhos de tamanhos específicos,
- Uma rotina que alinha os ladrilhos BF e CPOL,

- Modelo 1 (1A e 1B) e Modelo 2 para inferência das classes respectivas,
- Rotinas de segmentação (Goethita e Martita-Magnetita),
- Rotinas de operações entre imagens, como por exemplo a que elimina as goethitas da CPOL com base na BF,
- Rotinas de pós processamento que geram imagens binárias de cada classe a partir da inferência realizada pelos modelos e
- Rotina de medidas quantitativas.

Todas estas rotinas e modelos serão integrados de forma que o usuário possa aplicar facilmente no cotidiano das análises do CTF. Todas as etapas programadas de forma adaptável, utilizando caixas de diálogo pré-definidas, caso se deseje realizar análises personalizadas priorizando apenas determinadas classes.

## 6 Conclusões

Esta tese teve como objetivo desenvolver um método de identificação, classificação e segmentação de minerais de ferro em amostras de *pellet feed* utilizando ferramentas baseadas em *Deep Learning*. Foram utilizados pares de imagens obtidas em um Microscópio Ótico em modo Campo Claro e em Polarização Circular. Cada tipo de imagem revelava características importantes de cada classe de partícula mineral. Foi construído um banco de dados para cada classe a partir da rotulagem de partículas nas imagens. Foram desenvolvidos modelos utilizando o algoritmo *Mask R-CNN* para identificá-las.

O Modelo 1 identificou as classes Compacta, Porosa, Martita e Goethita em imagens obtidas em Campo Claro. Este modelo foi dividido em Modelos 1A, 1B e Etapa de segmentação. O Modelo 1A identificou as partículas da classe Compacta, apresentando um F1-score de 90,98%, mostrando-se bastante eficaz para a identificação desta classe. O Modelo 1B foi treinado para identificar as classes Porosa e Martita, obtendo F1-score de 78,89%. Um ponto importante no caso do Modelo 1B foi o desenvolvimento de imagens sintéticas para que se obtivesse um aumento no banco de dados das classes Porosa e Martita, que quando comparadas com a classe Compacta, têm uma representatividade muito baixa. Isso ocorreu exatamente por conta do tipo de minério utilizado que era muito rico em hematita compacta. A inserção das imagens sintéticas no conjunto de treinamento ajudou a melhorar o reconhecimento das classes, o que se refletiu em maiores valores para as métricas de validação. No entanto, apesar dos bons resultados, entende-se que este conjunto ainda precisa ser aumentado para que este modelo se torne mais confiável.

A partir da identificação das partículas de Martita pelo Modelo 1B foi possível realizar uma segmentação por tonalidade para separar a Magnetita presente nestas partículas. Ainda na etapa de segmentação foi realizada a limiarização das partículas de Goethita, pois até o momento não foi possível desenvolver um modelo para identificação desta classe. Isto se deve ao fato de que a Goethita possui uma diversidade muito grande de partículas, com diferentes formas, tamanhos, texturas etc. Desta forma, entende-se que seria preciso um conjunto muito maior para que o sistema aprendesse a diferenciar todas estas partículas. Após a obtenção das imagens binárias de todas estas classes foram realizadas medidas quantitativas nas partículas, o que permitiu identificar tendências gerais de tamanho e forma.

O Modelo 2 identificou as classes Monocristalina, Policristalina e Martita. Para este modelo foi obtido um F1-score 90,45%. Assim como no Modelo 1, neste modelo também existe uma diferença na representatividade das classes. Foi possível observar que este minério possuía muito mais partículas monocristalinas.



Consequentemente, as outras classes estão menos representadas, principalmente a Martita. Então, apesar dos resultados serem bem promissores ainda será preciso incrementar este banco de dados para que esta classificação seja mais robusta e confiável. Após a obtenção das imagens binárias destas classes também foi possível realizar medidas quantitativas, o que permitiu identificar tendências gerais de tamanho e forma destas partículas no minério.

Dada a clássica dificuldade de segmentação das partículas minerais, principalmente relacionada aos tipos de hematita, a metodologia desenvolvida mostrou-se eficaz para tratar esta dificuldade. A abordagem utilizando o algoritmo *Mask R-CNN* para identificação, classificação e segmentação das partículas foi de grande contribuição para as boas métricas de validação obtidas para ambos os modelos.

Como trabalhos futuros sugere-se que o banco de dados seja incrementado para que a identificação e as métricas se tornem cada vez melhores. Uma outra etapa futura seria o desenvolvimento de um modelo que fosse capaz de identificar os diferentes tipos de Goethita, pois incluir todos eles em uma só classe não se mostrou eficiente para o momento, razão pela qual foi realizada uma limiarização comum. Uma outra etapa futura importante é aprimorar estes modelos para que também sejam capazes de diferenciar partículas minerais no *sinter feed*. Este fino de minério é amplamente utilizado e de grande interesse para a indústria.

- [1] DONSKOI, E. et al. **Automated optical image analysis of goethitic iron ores**. *Mineral Processing and Extractive Metallurgy: Transactions of the Institute of Mining and Metallurgy*, [s.l.], v. 0, n° 0, p. 1–11, 2020. ISSN: 2572665X, DOI: 10.1080/25726641.2019.1706375..
- [2] AZIMI, S. M. et al. **Advanced steel microstructural classification by deep learning methods**. *Scientific Reports*, [s.l.], v. 8, no 1, p. 1–14, 2018. ISBN: 4159801820037, ISSN: 20452322, DOI: 10.1038/s41598-018-20037-5.
- [3] GOMES, O. D. F. et al. **Classification of hematite types in iron ores through circularly polarized light microscopy and image analysis**. *Minerals Engineering*, [s.l.], v. 52, p. 191–197, 2013. ISSN: 0892-6875, DOI: 10.1016/j.mineng.2013.07.019..
- [4] IGLESIAS, J. C. ; GOMES, O. F. .; PACIORNIK, S. **Automatic recognition of hematite grains under polarized reflected light microscopy through image analysis**. *Minerals Engineering*, [s.l.], v. 24, no 12, p. 1264–1270, 2011. ISSN: 08926875, DOI: 10.1016/j.mineng.2011.04.015
- [5] IGLESIAS, J. C, **Desenvolvimento de um sistema de microscopia digital para classificação automática de tipos de hematita em minério de ferro**. Tese (Doutorado em Engenharia Química e Materiais) 136 p. - Pontifícia Universidade Católica do Rio de Janeiro, 2012.
- [6] GONZALEZ, R. .; WOODS, R. . **Digital Image Processing. Digital Image Processing**. [s.l.]: [s.n.], 2008. v. 3rd Edition, 0–943 p. ISBN: 013168728x9780131687288, ISSN: 1098-6596.
- [7] HU, W. et al. **Deep Convolutional Neural Networks for Hyperspectral Image Classification**. *Journal of Sensors*, [s.l.], v. 2015, 2015.
- [8] YEGNANARAYANA, B. **Artificial Neural Networks**. *Annals of Statistics*. 1 ed. New Delhi: [s.n.], 2005. 227–248 p. ISBN: 9789533071886.
- [9] HE, H.; XU, H.; ZHANG, Y.; GAO, K.; LI, H.; MA, L.; LI, J. **Mask R-CNN based automated identification and extraction of oil well sites**. *International Journal of Applied Earth Observation and Geoinformation* 112. <https://doi.org/10.1016/j.jag.2022.102875>, 2022
- [10] MOURÃO, M. . **Introdução a Siderurgia**. São Paulo: [s.n.], 2007. 428 p
- [11] A. . LUZ, J. . SAMPAIO, and C. A. . FRANÇA, “Tratamento de Minérios,” Rio de Janeiro, 2010.
- [12] CARVALHO, P. S. . et al. **Biblioteca Digital BNDES - Minério de ferro**. In: Bndes Setorial 39. [s.l.]: [s.n.], 2014. Disponível em:

<<http://www.bndes.gov.br/bibliotecadigital>>.

- [13] ASSUNÇÃO, F. C. R.; MEDEIROS, E. F. **Siderurgia no Brasil 2010-2025; subsídios para tomada de decisão**. Brasília: [s.n.], 2010. Disponível em: <<http://www.cgee.org.br/atividades/redirect.php?idProduto=6831>>.
- [14] LUZ, Adão Benvindo; LINS, F. A. F. **Introdução ao Tratamento de Minérios**. Tratamento de Minérios. 6o ed. Rio de Janeiro: [s.n.], 2018. p. 1-23. ISBN: 22790704.
- [15] **CSN Mineração**. [s.d.]. Disponível em: <<https://ri.csnmineracao.com.br/a-companhia/minerio-de-ferro/>>. Acesso em: 14/jun./23..
- [16] LIMING, L. **Iron Ore: Mineralogy, Processing and Environmental Sustainability**. [s.l.]: [s.n.], 2015. v. 3, 49–58 p. ISBN: 9780128202265.
- [17] MORAES, S. L. DE; RIBEIRO, T. R. **Brazilian iron ore and production of pellets**. Mineral Processing and Extractive Metallurgy Review, [s.l.], p. 8, 2018. ISSN: 0882-7508, DOI: 10.1080/08827508.2018.1481056.
- [18] PASCOAL, A. . et al. **Estudo da produção de sinter feed e pellet feed de acordo com a granulometria**. Revista Brasileira de Aplicações de Vácuo, [s.l.], v. 35, no 3, p. 151, 2017. ISSN: 0101-7659, DOI: 10.17563/rbav.v35i3.1038..
- [19] CAXITO, F.; DIAS, T. **Recursos Minerais de Minas Gerais - Ferro**. Minas Gerais:[s.n.],2018.Disponívelem:<<http://recursomineralmg.codemge.com.br/wp-content/uploads/2018/10/Ferro.pdf>>.
- [20] VILELA, R. . et al. **Petrografia do minério hematita compacta da Mina do Tamanduá (Quadrilátero Ferrífero, MG)**. Rem: Revista Escola de Minas, [s.l.], v. 57, no 3, p. 157–164, 2004. DOI: 10.1590/s0370-44672004000300004..
- [21] VARAJÃO, C. A. C. et al. **Microporosity of BIF hosted massive hematite ore, iron quadrangle, Brazil**. Anais da Academia Brasileira de Ciencias, [s.l.], v. 74, no 1, p. 113–126, 2002. ISSN: 00013765, DOI: 10.1590/S0001-37652002000100008.
- [22] DONSKOI, E. et al. **Iron ore textural information is the key for prediction of downstream process performance**. Minerals Engineering, [s.l.], v.86, p.10–23, 2016. ISSN:08926875, DOI: 10.1016/j.mineng.2015.11.009.
- [23] DONSKOI, E. et al. **Automated optical image analysis of goethitic iron ores**. Mineral Processing and Extractive Metallurgy: Transactions of the Institute of Mining and Metallurgy, [s.l.], v. 0, no 0, p. 1–11, 2020. ISSN: 2572665X, DOI: 10.1080/25726641.2019.1706375.
- [24] GIANFRANCESCO, A. DI. **Technologies for chemical analyses, microstructural and inspection investigations**. Materials for Ultra-Supercritical and Advanced Ultra-Supercritical Power Plants. [s.l.]: Elsevier Ltd, 2017. p. 197–245. ISBN: 9780081005583, DOI: 10.1016/B978-0-08-100552-1.00008-7.
- [25] SUONINEN, E. . **Optical microscopy**. Surface Characterization: A User's

- Sourcebook, [s.l.], p. 54–56, 2007. ISBN: 9783527612451, DOI: 10.1002/9783527612451.ch1
- [26] HATIBOGLU, C. .; AKIN, S. **A new computerized moving stage for optical microscopes**. Computers and Geosciences, [s.l.], v. 30, no 5, p. 471–481, 2004. ISSN: 00983004, DOI: 10.1016/j.cageo.2004.03.008.
- [27] IGLESIAS, J. C. **Uma metodologia para caracterização de sínter de minério de ferro: Microscopia Digital e Análise de Imagens**. - Dissertação (Mestrado em Engenharia Química e Materiais) 136 p. - Pontifícia Universidade Católica do Rio de Janeiro, 2008.
- [28] GOMES, O. F. .; LIMA, M. P. .; MOTA, G. L. . **Microscopia Multimodal para a caracterização de minériod e ferro: automatizando registro**. In: 43° Ironmaking and Raw Materials Seminar 14° Brazilian Symposium on Iron Ore 1°Brazilian Symposium on Agglomeration of Iron Ore. Belo Horizonte: [s.n.], 2014.
- [29] IGLESIAS, J. C. Á. et al. **Automatic characterization of iron ore by digital microscopy and image analysis**. Journal of Materials Research and Technology, [s.l.], v. 7, no 3, p. 376–380, 2018. ISSN: 22387854, DOI: 10.1016/j.jmrt.2018.06.014.
- [30] GOMES, O. F. .; PACIORNIK, S. **Co-site microscopy-combining reflected light microscopy and scanning electron microscopy to perform ore mineralogy**. Australasian Institute of Mining and Metallurgy Publication Series, [s.l.], no September, p. 695–698, 2008. ISBN: 9781920806873..
- [31] IGLESIAS, J. C. .; SANTOS, R. B. .; PACIORNIK, S. **Deep learning discrimination of quartz and resin in optical microscopy images of minerals**. Minerals Engineering, [s.l.], v. 138, no April, p. 79–85, 2019. ISSN: 0892-6875, DOI: 10.1016/j.mineng.2019.04.032.
- [32] DELBEM, I. D. et al. **Semi-automated iron ore characterisation based on optical microscope analysis: Quartz/resin classification**. Minerals Engineering, [s.l.], v. 82, p. 2–13, 2015. ISSN: 08926875, DOI: 10.1016/j.mineng.2015.07.021.
- [33] OSORIO, F.; BITTENCOURT, J. **Sistemas Inteligentes baseados em Redes Neurais Artificiais aplicados ao Processamento de Imagens**. In: I Workshop de Int. Artificial – UNISC. [s.l.]: [s.n.], 2014. Disponível em: <[https://www.researchgate.net/profile/Fernando\\_Osorio2/publication/228588719\\_Sistemas\\_Inteligentes\\_baseados\\_em\\_redes\\_neurais\\_artificiais\\_aplicados\\_ao\\_processamento\\_de\\_imagens/links/0912f51001cc71ad2b000000/Sistemas-Inteligentes-baseados-em-redes-neurais-ar](https://www.researchgate.net/profile/Fernando_Osorio2/publication/228588719_Sistemas_Inteligentes_baseados_em_redes_neurais_artificiais_aplicados_ao_processamento_de_imagens/links/0912f51001cc71ad2b000000/Sistemas-Inteligentes-baseados-em-redes-neurais-ar)>.
- [34] AMER, M. **A Visual Introduction to Deep Learning**. [s.l.]: [s.n.], 2016. 1–32 p.
- [35] SHRESTHA, A.; MAHMOOD, A. **Review of deep learning algorithms and architectures**. IEEE Access, [s.l.], v. 7, p. 53040–53065, 2019. ISSN: 21693536, DOI: 10.1109/ACCESS.2019.2912200.
- [36] KRUG, Á. B. et al. **Análise e reconhecimento de padrões usando**

- processamento de imagens e inteligência artificial.** Revista de Iniciação Científica da ULBRA, [s.l.], v. 7, no 7, p. 145–154, 2008. ISSN: 2317-4323..
- [37] **Deep Learning Book.** Data Science Academy. 2021. Disponível em: <<https://www.deeplearningbook.com.br>>. Acesso em: 05/maio/21
- [38] SVENSSON, T. **Semantic segmentation of iron ore pellets with Neural Networks.** v. Independen, 70 p. - Lulea University of Technology, 2019. Disponível em: <<http://ltu.divaportal.org/smash/get/diva2:1322752/FULLTEXT01.pdf%0Ahttp://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-74352>>.
- [39] ARAÚJO, F. H. . et al. **Redes Neurais Convolucionais com Tensorflow: Teoria e Prática.** SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. III Escola Regional de Informática do Piauí. Livro Anais - Artigos e Minicursos. [s.l.]: [s.n.], 2017. v. 1, p. 382–406. ISBN: 9788576693956.
- [40] NOURI, S. **Data Science: Interview.** [s.l.]: [s.n.], [s.d.].
- [41] WENCHAO, S.; BAI, S.; KONG, L. **A review on 2D instance segmentation based on deep neural networks.** Image and Vision Computing, [s.l.], v. 120, p. 104401, 2022. ISSN: 02628856, DOI: 10.1016/j.imavis.2022.104401.
- [42] SANTOS, R. B. . et al. **A deep learning system for collotelinite segmentation and coal reflectance determination.** International Journal of Coal Geology, [s.l.], v. 263, no July, 2022. ISSN: 01665162, DOI: 10.1016/j.coal.2022.104111.
- [43] REN, S. et al. **Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.** IEEE Transactions on Pattern Analysis and Machine Intelligence, [s.l.], v. 39, no 6, p. 1137–1149, 2017. ISSN: 01628828, DOI: 10.1109/TPAMI.2016.2577031.
- [44] GIRSHICK, R.; DONAHUE, J.; DARRELL, T.; MALIK, J. **Region-based Convolutional Networks for Accurate Object Detection and Segmentation,** arXiv:1311.2524v5, 2014.
- [45] BHARATI, P.; PRAMANIK, A. **Deep Learning Techniques—R-CNN to Mask R-CNN: A Survey.** Computational Intelligence in Pattern Recognition Advances in Intelligent Systems and Computing. [s.l.], p. 657–668, 2019..
- [46] LIN, T, Y. et al. **Feature Pyramid Networks for Object Detection.** [s.l.], 2017.
- [47] ABDULLA, W. **Splash of Color : Instance Segmentation with Mask R-CNN and Tensorflow.** Matterport Engineering, [s.l.], p. 1–17, 2020..
- [48] DUTTA, A.; ZISSERMAN, A. **The VIA annotation software for images, audio and video.** arXiv, [s.l.], p. 1–8, 2019. ISSN: 23318422.
- [49] EKBATANI, H. .; PUJOL, O.; SEGUI, S. **Synthetic data generation for deep learning in counting pedestrians.** ICPRAM 2017 - Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods, [s.l.], v. 2017-Janua, no Icpam, p. 318–323, 2017. ISBN: 9789897582226, DOI: 10.5220/0006119203180323.

- [50] RUEDEN, C. T. .; SCHINDELIN, J.; HINER, M. C. . **“ImageJ2: ImageJ for the next generation of scientific image data”**. BMC Bioinformatics, [s.l.], no 18:529, 2017. DOI: 10.1186/s12859-017-1934-z.