



APRENDIZADO PROFUNDO PARA IDENTIFICAÇÃO DE LITOLOGIAS EM ROCHAS CARBONÁTICAS DO PRÉ-SAL

Carlos Eduardo Menezes dos Anjos

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientadores: José Manoel de Seixas
Alexandre Gonçalves Evsukoff

Rio de Janeiro
Março de 2020

APRENDIZADO PROFUNDO PARA IDENTIFICAÇÃO DE LITOLOGIAS EM
ROCHAS CARBONÁTICAS DO PRÉ-SAL

Carlos Eduardo Menezes dos Anjos

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU
DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Orientadores: José Manoel de Seixas
Alexandre Gonçalves Evsukoff

Aprovada por: Prof. José Manoel de Seixas
Prof. Alexandre Gonçalves Evsukoff
Profa. Mariane Rembold Petraglia
Profa. Marley Maria Bernardes Rebuzzi Vellasco
Dr. Rodrigo Surmas

RIO DE JANEIRO, RJ – BRASIL
MARÇO DE 2020

Anjos, Carlos Eduardo Menezes dos

Aprendizado profundo para identificação de litologias em rochas carbonáticas do pré-sal/Carlos Eduardo Menezes dos Anjos. – Rio de Janeiro: UFRJ/COPPE, 2020.

XV, 93 p.: il.; 29, 7cm.

Orientadores: José Manoel de Seixas

Alexandre Gonçalves Evsukoff

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2020.

Referências Bibliográficas: p. 59 – 70.

1. Aprendizado Profundo. 2. Transferência de Aprendizado. 3. Rochas Carbonáticas. 4. Micro-tomografia. I. Seixas, José Manoel de *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

Agradecimentos

Gostaria de agradecer a todos que tornaram possível a produção deste trabalho.

Inicialmente, gostaria de ressaltar que este trabalho somente se tornou viável em razão do suporte da Petrobras, que disponibilizou recursos e financiou o trabalho aqui apresentado. Também gostaria de agradecer a NVIDIA por conceder a placa de vídeo que possibilitou o treinamento dos modelos de forma mais rápida e eficiente. Além dessas instituições, gostaria de agradecer ao Laboratório de Métodos Computacionais (LAMCE/Lab2M) e ao Laboratório de Processamento de Sinais (LPS) pelo auxílio prestado.

Desejo, ainda, agradecer à minha família por toda torcida e suporte nesses anos de Mestrado, à minha namorada e, por fim, aos meus professores da UFRJ e ao colegas de trabalho, sem os quais todo o ensinamento aprendido e demonstrado na presente dissertação não teria sido possível.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

APRENDIZADO PROFUNDO PARA IDENTIFICAÇÃO DE LITOLOGIAS EM ROCHAS CARBONÁTICAS DO PRÉ-SAL

Carlos Eduardo Menezes dos Anjos

Março/2020

Orientadores: José Manoel de Seixas

Alexandre Gonçalves Evsukoff

Programa: Engenharia Elétrica

A caracterização de reservatórios do pré-sal ainda é um desafio na indústria do petróleo devido às peculiaridades geológicas e à heterogeneidade das rochas carbonáticas. Essas dificuldades estimulam o desenvolvimento de novos métodos, como a tomografia computadorizada das rochas para geração de imagens. Assim como na caracterização, na área computacional novos métodos vêm sendo desenvolvidos, sendo um deles o aprendizado profundo, que é considerado estado da arte em diversas tarefas, com destaque para a visão computacional. Com isso, apresenta-se, nesta dissertação, a aplicação do mesmo para identificação de padrões em imagens de microtomografia de amostras cilíndricas de rocha (plugue) para a classificação litológica. Dois modelos de redes neurais convolucionais foram propostos e implementados, um modelo base e esse mesmo modelo modificado adicionando-se uma camada de *spatial pyramid pooling* (estratégia diferenciada de *pooling*), onde ambos foram comparados com o uso da técnica de transferência de aprendizado no modelo Inception v3. Como referência para esses modelos, métodos clássicos para extração de características de imagens foram utilizados como entrada para um perceptron multicamadas. O conjunto de dados de interesse possui amostras de 60 plugues disponibilizados pela Petrobras, os quais estão distribuídos em três classes diferentes, anotadas por especialistas. Cada plugue possui 100 imagens disponíveis para uso, totalizando assim 6.000 imagens. Os resultados obtidos mostraram que o modelo Inception v3 superou, na média, os outros modelos, atingindo mais de 80% de média de acurácia. Com isso, foi desenvolvido um fluxograma de trabalho para automatizar e acelerar o processo de classificação litológica de rochas carbonáticas do pré-sal.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

DEEP LEARNING FOR LITHOLOGY IDENTIFICATION ON PRE-SALT CARBONATE ROCKS

Carlos Eduardo Menezes dos Anjos

March/2020

Advisors: José Manoel de Seixas

Alexandre Gonçalves Evsukoff

Department: Electrical Engineering

The characterization of pre-salt reservoirs is still a challenge in the oil industry due to the geological peculiarities and the heterogeneity of carbonate rocks. These difficulties encourage the development of new methods, such as computed tomography of the rocks for image generation. As with characterization, new methods have been developed in the computational field, one of which is deep learning, which is considered state of the art in several tasks, with emphasis on computer vision. With this, it is presented, in this dissertation, its application for the identification of patterns in microtomography images of cylindric rock samples (plug) for the lithological classification of rock samples. Two convolutional neural network models were proposed and implemented, a base model and this same model modified by adding a layer of spatial pyramid pooling (pooling strategy), where both were compared with the use of the technique of transfer learning with the model Inception v3. As a reference for these models, classic methods for extracting characteristics from images were used as input to a multilayer perceptron. The dataset of interest has samples of 60 plugs made available by Petrobras, which are distributed in three different classes, labeled by specialists. Each plug has 100 images available for use, thus summing up to 6,000 images. The results obtained showed that the Inception v3 model surpassed, on average, the other models, reaching more than 80% of average accuracy. With that, a workflow was developed to automate and accelerate the lithological classification process of pre-salt carbonate rocks.

Sumário

Lista de Figuras	x
Lista de Tabelas	xiii
1 Introdução	1
1.1 Motivação	2
1.2 Objetivo	5
1.3 Organização da Dissertação	5
2 Rocha Digital	6
2.1 Amostragem de Rochas	6
2.2 Tomografia Computadorizada	9
2.3 Fluxograma de Trabalho Atual e Proposto	10
3 Revisão bibliográfica	15
3.1 Métodos Clássicos	15
3.1.1 Matriz de Coocorrência dos Tons de Cinza e Haralick	17
3.1.2 Dimensão Fractal	18
3.1.3 Filtros de Gabor	18
3.1.4 Combinação dos Métodos	18
3.1.5 Identificação de Textura em Imagens de Rochas	19
3.2 Aprendizado Profundo	20
3.2.1 Rede Neural Convolucional	20
3.2.2 Aplicações	26
3.2.3 Aprendizado Profundo para Classificação de Imagens de Rochas	26
4 Materiais e Métodos	28
4.1 Dados	28
4.2 Modelos	32
4.2.1 Perceptron Multicamadas	32
4.2.2 Modelos Convolucionais	33
4.2.3 Modelo Inception V3	35

4.2.4	Comparação dos Modelos	35
4.3	Avaliação	36
4.3.1	Figuras de Mérito	36
4.3.2	Validação Cruzada	38
4.3.3	Abordagens de avaliação	39
4.4	Experimentos	39
4.4.1	Pré-processamentos	39
4.4.2	Divisão dos Dados	41
4.4.3	Treinamento	42
5	Resultados e Discussões	44
5.1	Resultados Validação Cruzada	44
5.1.1	Resultados por Imagem	45
5.1.2	Resultados por Plugue	48
5.2	Comparação dos Modelos	52
5.2.1	Matrizes de Confusão por Imagem	53
5.2.2	Matrizes de Confusão por Plugue	54
5.2.3	Análise de quadrante	55
6	Conclusões	57
6.1	Trabalhos Futuros	58
	Referências Bibliográficas	59
A	Ferramentas	71
A.1	Python	71
A.1.1	Numpy	72
A.1.2	Pandas	72
A.1.3	Opencv	72
A.1.4	Matplotlib	72
A.2	Anaconda	72
A.2.1	Jupyter	73
A.3	Pytorch	73
B	Escolha da Topologia da Rede Neural Artificial	74
C	Plugues Utilizados e Grupos de Validação Cruzada	77
C.1	Fold 1	77
C.2	Fold 2	78
C.3	Fold 3	79
C.4	Fold 4	81

C.5	Fold 5	82
C.6	Fold 6	83
D	Resultados	85
D.1	Treino	85
D.1.1	Avaliação por Imagem	85
D.1.2	Avaliação por Plugue	86
D.2	Validação	87
D.2.1	Avaliação por Imagem	88
D.2.2	Avaliação por Plugue	89
D.3	Teste	90
D.3.1	Avaliação por Imagem	90
D.3.2	Avaliação por Plugue	91
E	Produção Científica	93
E.1	Congressos	93
E.1.1	Artigo - Sociedade Brasileira de Geofísica - 2019	93

Lista de Figuras

1.1	<i>Exemplo de testemunho, plugue e lâmina.</i>	3
1.2	<i>Fluxo de trabalho atual e proposto.</i>	4
2.1	<i>Testemunho extraído de um poço e armazenado em uma caixa, escala na lateral direita representando dez centímetros para cada mudança de cor.</i>	8
2.2	<i>Plugue retirado de um testemunho.</i>	8
2.3	<i>Lâmina vista por um microscópio.</i>	9
2.4	<i>Diagrama do fluxo da microtomografia computadorizada. (A) É uma ilustração do processo de aquisição de imagens do microtomógrafo, apresentando a fonte de raio-x, o detector e a amostra. Imagens são projetadas ao se rotacionar a amostra, enquanto raios-x são emitidos. (B) Após o termino do processo anterior, é gerado um conjunto de imagens com ângulos diferentes. (C) Projeções são reconstruídas digitalmente através de algoritmos como retroprojeção filtrada ou reconstrução iterativa. (D) A amostra pode ser visualizada como um volume 3D ou fatias 2D.</i>	10
2.5	<i>Fluxo de trabalho atual e proposto.</i>	12
2.6	<i>Comparação entre lâminas (coluna da esquerda) e microtomografia (coluna da direita) das litologias esferulítico(A), estromatólito(B) e grainstone(C).</i>	13
2.7	<i>Exemplos de cada uma três litologias em imagens de micro-CT de plugues.</i>	14
3.1	<i>Fluxo para classificação de imagens.</i>	16
3.2	<i>Exemplo da estrutura convolucional.</i>	21
3.3	<i>Exemplo da aplicação de convolução em uma matriz de entrada 4×4, preenchimento por 0 de tamanho 1, janela convolucional 3×3 e o stride 1, camada de ativação ReLU e Max Pooling 2×2 com padding 0 e stride 2.</i>	23
3.4	<i>Camada SPP.</i>	24

4.1	<i>Análise dos plugues anotados por especialistas.</i>	29
4.2	<i>Exemplos de cada uma três classes do conjunto de imagens de micro-CT de plugues.</i>	30
4.3	<i>Exemplo de corte na imagem de micro-CT de um plugue com objetivo de maximizar a região de rocha.</i>	31
4.4	<i>Imagens de micro-CT de plugues visualmente parecidas, porém, provenientes de classes diferentes.</i>	31
4.5	<i>Imagem de micro-CT de plugue com artefato.</i>	31
4.6	<i>Imagens de micro-CT de plugues com diferentes condições luminosas.</i>	31
4.7	<i>Estrutura do perceptron multicamadas proposto.</i>	32
4.8	<i>Estrutura do modelo convolucional proposto.</i>	34
4.9	<i>Estrutura do modelo convolucional proposto sem a camada de SPP.</i>	34
4.10	<i>Estrutura do modelo Inception v3.</i>	35
4.11	<i>Exemplo de k-folds aninhados com 6 folds.</i>	39
4.12	<i>Redimensionamento em uma imagem de 718×604 para uma imagem de 256×256.</i>	41
4.13	<i>Exemplo de uma imagem de cada classe pertencente a cada fold.</i>	41
5.1	<i>Diagrama de caixa das acurácias ao avaliar por imagem individual no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul.</i>	46
5.2	<i>Diagrama de caixa do recall da classe Esferulítico ao avaliar por imagem individual no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul.</i>	47
5.3	<i>Diagrama de caixa do recall da classe Estromatólito ao avaliar por imagem individual no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul.</i>	47
5.4	<i>Diagrama de caixa do recall da classe Grainstone ao avaliar por imagem individual no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul.</i>	48
5.5	<i>Diagrama de caixa das acurácias ao avaliar por plugue no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul.</i>	50
5.6	<i>Diagrama de caixa do recall da classe Esferulítico ao avaliar por plugue no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul.</i>	50
5.7	<i>Diagrama de caixa do recall da classe Estromatólito ao avaliar por plugue no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul.</i>	51

5.8	<i>Diagrama de caixa do recall da classe Grainstone ao avaliar por plugue no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul.</i>	52
B.1	<i>Estrutura da rede neural artificial proposta.</i>	75
B.2	<i>Comparação do índice SP de diferentes topologias no conjunto de treino. 75</i>	
B.3	<i>Comparação do índice SP de diferentes topologias no conjunto de validação.</i>	75
B.4	<i>Comparação do índice SP de diferentes topologias no conjunto de teste. 76</i>	
C.1	<i>Exemplos da classe esferulítico do fold 1.</i>	77
C.2	<i>Exemplos da classe estromatólito do fold 1.</i>	78
C.3	<i>Exemplos da classe grainstone do fold 1.</i>	78
C.4	<i>Exemplos da classe esferulítico do fold 2.</i>	79
C.5	<i>Exemplos da classe estromatólito do fold 2.</i>	79
C.6	<i>Exemplos da classe grainstone do fold 2.</i>	79
C.7	<i>Exemplos da classe esferulítico do fold 3.</i>	80
C.8	<i>Exemplos da classe estromatólito do fold 3.</i>	80
C.9	<i>Exemplos da classe grainstone do fold 3.</i>	80
C.10	<i>Exemplos da classe esferulítico do fold 4.</i>	81
C.11	<i>Exemplos da classe estromatólito do fold 4.</i>	81
C.12	<i>Exemplos da classe grainstone do fold 4.</i>	82
C.13	<i>Exemplos da classe esferulítico do fold 5.</i>	82
C.14	<i>Exemplos da classe estromatólito do fold 5.</i>	83
C.15	<i>Exemplos da classe grainstone do fold 5.</i>	83
C.16	<i>Exemplos da classe esferulítico do fold 6.</i>	84
C.17	<i>Exemplos da classe estromatólito do fold 6.</i>	84
C.18	<i>Exemplos da classe grainstone do fold 6.</i>	84

Lista de Tabelas

4.1	Distribuição das classes anotadas no conjunto de imagens de plugues	30
4.2	Comparação da quantidade de parâmetros treináveis dos modelos. . .	36
4.3	Matriz de Confusão	36
4.4	Tabela com tempos de treinamento para cada modelo utilizado	43
5.1	Tabela com acurácia e índice sp dos modelos ao avaliar por imagem individual no conjunto de teste	46
5.2	Tabela com recall avaliando por imagem da classe Esferulitito no conjunto de teste	46
5.3	Tabela com recall avaliando por imagem da classe Estromatólito no conjunto de teste	47
5.4	Tabela com recall avaliando por imagem da classe Grainstone no conjunto de teste	48
5.5	Tabela com acurácia e índice SP dos modelos ao avaliar por plugue no conjunto de teste	49
5.6	Tabela com recall avaliando por plugue da classe Esferulitito no conjunto de teste	50
5.7	Tabela com recall avaliando por plugue da classe Estromatólito no conjunto de teste	51
5.8	Tabela com recall avaliando por plugue da classe Grainstone no conjunto de teste	51
5.9	Matriz de confusão por imagem do modelo Inception v3 na rodada de número 24 no conjunto de teste	53
5.10	Matriz de confusão por imagem do modelo CNN+SPP na rodada de número 24 no conjunto de teste	54
5.11	Matriz de confusão por imagem do modelo Inception v3 na rodada de número 23 no conjunto de teste	54
5.12	Matriz de confusão por imagem do modelo CNN+SPP na rodada de número 23 no conjunto de teste	54
5.13	Matriz de confusão por plugue do modelo Inception v3 na rodada de número 24 no conjunto de teste	54

5.14	Matriz de confusão por plugue do modelo CNN+SPP na rodada de número 24 no conjunto de teste	55
5.15	Matriz de confusão por plugue do modelo Inception v3 na rodada de número 23 no conjunto de teste	55
5.16	Matriz de confusão por plugue do modelo CNN+SPP na rodada de número 23 no conjunto de teste	55
5.17	Concordância ou discordância dos modelos Inception v3 e CNN+SPP na rodada de número 24 no conjunto de teste	56
5.18	Concordância ou discordância dos modelos Inception v3 e CNN+SPP na rodada de número 23 no conjunto de teste	56
D.1	Tabela com acurácia e índice SP dos modelos ao avaliar por imagem individual no conjunto de treino	85
D.2	Tabela com recall da classe Esferulitito avaliando por imagem individual no conjunto de treino	86
D.3	Tabela com recall da classe Estromatólito avaliando por imagem individual no conjunto de treino	86
D.4	Tabela com recall da classe Grainstone avaliando por imagem individual no conjunto de treino	86
D.5	Tabela com acurácia e índice SP dos modelos ao avaliar por plugue no conjunto de treino	86
D.6	Tabela com recall da classe Esferulitito avaliando por plugue no conjunto de treino	87
D.7	Tabela com recall da classe Estromatólito avaliando por plugue no conjunto de treino	87
D.8	Tabela com recall da classe Grainstone avaliando por plugue no conjunto de treino	87
D.9	Tabela com acurácia e índice SP dos modelos ao avaliar por imagem individual no conjunto de validação	88
D.10	Tabela com recall da classe Esferulitito avaliando por imagem individual no conjunto de validação	88
D.11	Tabela com recall da classe Estromatólito avaliando por imagem individual no conjunto de validação	88
D.12	Tabela com recall da classe Grainstone avaliando por imagem individual no conjunto de validação	88
D.13	Tabela com acurácia e índice SP dos modelos ao avaliar por plugue no conjunto de validação	89
D.14	Tabela com recall da classe Esferulitito avaliando por plugue no conjunto de validação	89

D.15 Tabela com recall da classe Estromatólito avaliando por plugue no conjunto de validação	89
D.16 Tabela com recall da classe Grainstone avaliando por plugue no conjunto de validação	89
D.17 Tabela com acurácia e índice SP dos modelos ao avaliar por imagem individual no conjunto de teste	90
D.18 Tabela com recall avaliando por imagem da classe Esferulítico no conjunto de teste	90
D.19 Tabela com recall avaliando por imagem da classe Estromatólito no conjunto de teste	90
D.20 Tabela com recall avaliando por imagem da classe Grainstone no conjunto de teste	91
D.21 Tabela com acurácia e índice SP dos modelos ao avaliar por plugue no conjunto de teste	91
D.22 Tabela com recall avaliando por plugue da classe Esferulítico no conjunto de teste	91
D.23 Tabela com recall avaliando por plugue da classe Estromatólito no conjunto de teste	91
D.24 Tabela com recall avaliando por plugue da classe Grainstone no conjunto de teste	92

Capítulo 1

Introdução

Os combustíveis fósseis, maior matriz energética do mundo [1, 2], são originados da decomposição de organismos animais e vegetais durante milhares de anos em camadas profundas do solo ou do fundo do mar, sendo seu uso iniciado, principalmente, em meados do século XVIII, com o advento da Revolução Industrial. Atualmente, os combustíveis fósseis geram cerca de 80% de toda energia consumida no mundo, sendo o petróleo, o gás natural e o carvão os mais utilizados [1, 2].

Neste sentido, hoje em dia, o combustível fóssil considerado de maior aplicação comercial é o petróleo. Isso ocorre, principalmente, pelo uso dos seus derivados, como gasolina, o Gás Liquefeito de Petróleo, o querosene, o óleo diesel, além de plásticos e borrachas.

No Brasil, verifica-se a importância do petróleo e seus derivados na matriz energética do país, chegando a quase 37% de toda oferta de energia brasileira, enquanto, no restante do mundo, a oferta chega ao patamar de 32% [1, 2].

Apesar do cenário atual demonstrar um aumento do uso de fontes de energia renováveis, enquanto o percentual de uso dos combustíveis fósseis vem diminuindo, de acordo com o estudo Potencial dos Recursos Energéticos no Horizonte 2050 [3], disponibilizado pela Empresa de Pesquisa Energética. As perspectivas para o ano de 2050 são de um aumento na demanda e produção de petróleo, incentivando-se, assim, o investimento e pesquisa na área. Principalmente pelo fato da maior parte do valor estipulado na produção ser proveniente de poços ainda não descobertos, dando um espaço ainda maior para pesquisa nas etapas de exploração e exploração de poços.

Corroborando com a previsão, ao final de 2007, uma extensa reserva de petróleo e gás natural foi encontrada em uma camada abaixo do sal, no que é conhecido atualmente como pré-sal, consistindo em um reservatório imenso de petróleo e gás natural encontrada sob uma profunda camada de rocha salina no fundo do mar. Existindo, no Brasil, alguns campos de exploração nesse tipo de camada [4]. Nota-se que as reservas de petróleo existentes nesta camada estão em profundidades

abaixo de 5 mil metros, o que aumenta o custo de exploração, demandando altos investimentos.

Desta forma, conclui-se que o cenário futuro corrobora com a grande importância que o petróleo vem ganhando ao longo do tempo, como matriz energética da população e do setor industrial mundial. Assim, sendo possível esperar descobertas de novos reservatórios ainda não explorados, que, apesar de custosos, serão de grande importância no desenvolvimento energético, industrial e tecnológico dos países.

1.1 Motivação

Um desafio importante da indústria do Petróleo é a determinação das propriedades petrofísicas de rochas reservatório. Técnicas modernas, como modelagem numérica, são usadas para simular o comportamento de reservatórios através do cálculo do escoamento em meios porosos; contudo, dependem muito da determinação correta das características das texturas que a compõem [5, 6].

A tomografia computadorizada (do inglês, *Computed Tomography* - CT) e microtomografia computadorizada (do inglês, *Computed Microtomography* - micro-CT) de rochas têm sido utilizadas extensivamente como ferramentas adicionais para auxiliar na determinação das características texturais em amostras de rochas, visando melhorar o desenvolvimento da indústria petrolífera [7–9]. A análise digital de imagens tem sido amplamente aplicada para caracterizar rochas de reservatório heterogêneas, como é o caso das rochas carbonáticas, fornecendo informações detalhadas sobre a textura e a composição das rochas em várias escalas, quantificando a forma, o arredondamento e o tamanho de grãos e poros existentes [6].

No desenvolvimento da exploração petrolífera, o espaço poroso e a mineralogia da rocha podem representar a heterogeneidade do corpo rochoso, sendo, a princípio, calibradores confiáveis para selecionar amostras representativas [6]. Nesse sentido, apesar das amostras com mais alta resolução de imagens enfatizarem as diferentes geometrias de poros e melhorarem a caracterização de tipos de rochas [10], a Física de Rocha Digital (do inglês, *Digital Rock Physics* - DRP) [5] necessita, ainda, da incorporação de vários parâmetros geológicos aos modelos digitais. Contudo, a natureza complexa das rochas e a dificuldade de se definir texturas ainda permanecem como desafio [11].

Uma das dificuldades de análises DRP diz respeito as diferentes escalas de estudo, que variam desde o testemunho (amostras de rocha com noventa centímetros de comprimento) até mini-plugues (amostras cilíndricas de rochas com alguns milímetros), exemplos na Figura 1.1. Isso ocorre pelo fato de uma mesma rocha poder apresentar texturas de várias escalas misturadas espacialmente, que podem ser identificadas como grupos de padrões dentro de uma imagem digital, quando as mesmas são me-

nores que a resolução da imagem em estudo [12].

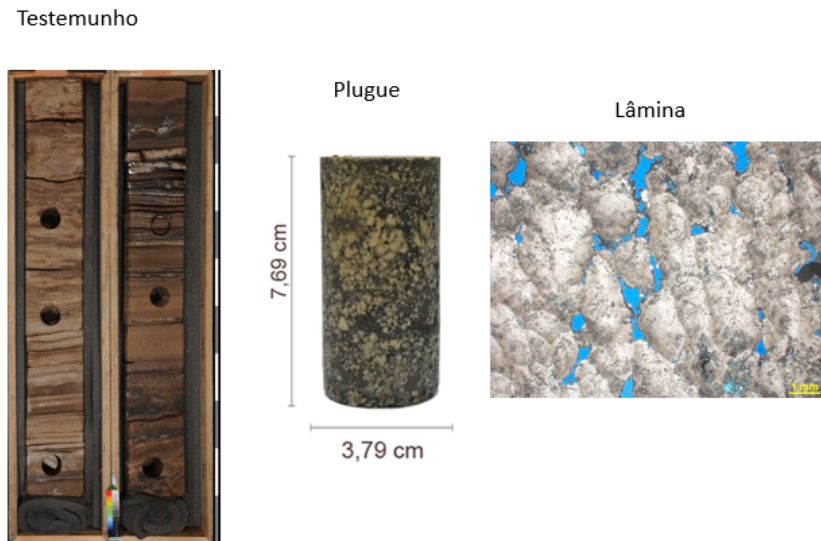


Figura 1.1: Exemplo de testemunho, plugue e lâmina

Uma etapa importante no processo de identificação de tipos de rocha é a verificação de texturas nas diferentes escalas, Figura 1.2, sejam elas lâmina (amostra fina de rocha retirada do plugue - Figura 1.1), plugue ou testemunho. Este processo, atualmente, é feito por especialistas, baseando-se em características visuais como a rugosidade, a granularidade, a distribuição dos tamanhos dos poros, dentre outras [13]. Entretanto, devido ao volume de imagens que devem ser analisadas e da definição ambígua das características visuais, este processo consome tempo, é cansativo e enviesado pelo especialista, de forma que as técnicas de automatização que empregam algoritmos de inteligência computacional surgem como uma alternativa para eliminar esta dependência [14], reduzindo o tempo do processo, custos e deixando os especialistas focarem em tarefas mais importantes. Este trabalho se situa dentro do fluxograma da caracterização reservatório, como visto pela região marcada em vermelho na Figura 1.2, e utiliza imagens de microtomografia para identificação de texturas na escala dos plugues.

Fluxograma de trabalho para caracterização

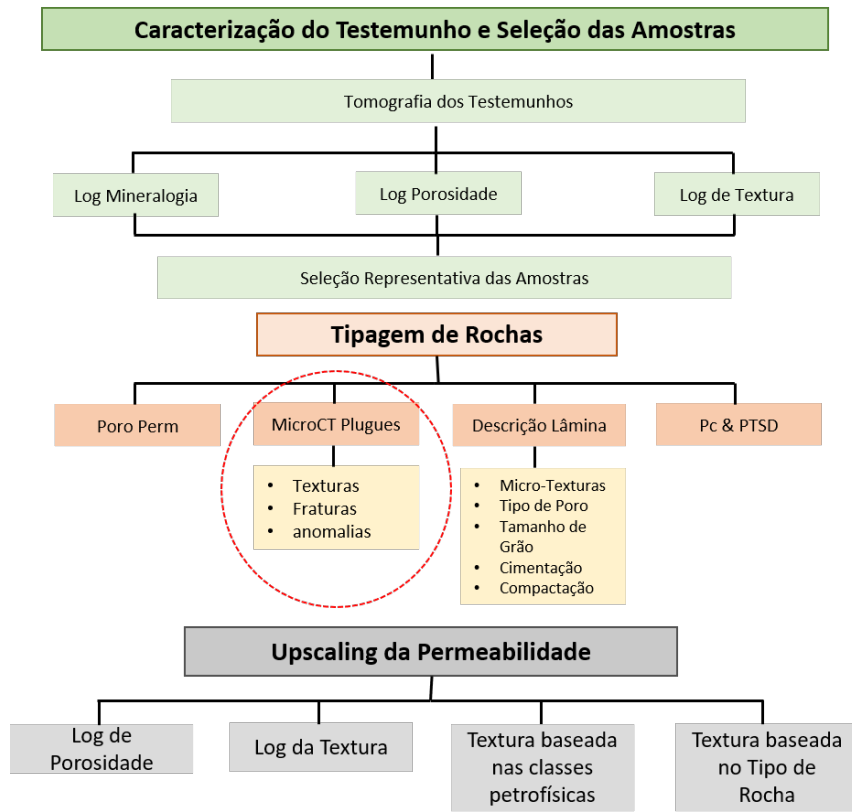


Figura 1.2: Fluxo de trabalho atual e proposto (Adaptado de [15])

Recentemente, o aprendizado profundo tem se destacado na área de inteligência computacional, consistindo no treinamento de modelos de redes neurais artificiais de múltiplas camadas de processamento, os quais conseguem extrair informações a partir de dados de entrada sem pré-processamentos complexos, reduzindo a necessidade de um especialista para analisar os dados e extrair o conjunto de características [14]. Estes modelos têm apresentado excelente desempenho em tarefas de reconhecimento de padrões, classificação e segmentação de imagens [16–18]. Motivando-se assim a aplicação da técnica de aprendizado profundo para identificação de padrões em imagens de rochas carbonáticas do pré-sal para automatizar o processo de identificação de litologias em amostras de rocha.

O problema de identificação de texturas em imagens é muito abrangente, visto que não existe uma definição estritamente matemática para textura, desta forma, para mitigar os problemas de interpretação que podem ser gerados por esse fato, no presente trabalho, foi feito o uso da litologia da amostra como etiqueta para classificação. As litologias das rochas foram escolhidas como linha de base, visto que elas possuem padrões visuais que podem ser vistos nas imagens e definições na literatura [13]. Contudo, mesmo com definições alguns especialistas ainda divergem nas classi-

ficações das amostras, para contornar esse problema, apenas amostras concordadas por dois ou mais especialistas foram utilizadas. O uso dessas etiquetas foi motivado também pela possibilidade de extração de estatísticas de desempenho dos modelos.

1.2 Objetivo

O objetivo deste trabalho é criar modelos de aprendizado profundo para apoiar a caracterização de reservatório, através da classificação litológica de imagens de microtomografia de plugues de rochas carbonáticas do pré-sal. Para atingi-lo, duas propostas foram feitas: a primeira baseia-se na concepção e implementação de modelos convolucionais, treinando-os somente com as imagens de microtomografia, enquanto a segunda utiliza-se da técnica de transferência de aprendizado. O modelo Inception v3 [19] pré-treinado no conjunto de dados ImageNet [20] foi o modelo escolhido para o uso dessa técnica devido a acurácia de 77,5% no ImageNet [20] e seu sucesso em aplicações de transferência de aprendizado [21–24]. Como referência para os modelos de aprendizado profundo, perceptrons multicamadas [25] são treinados utilizando métodos clássicos de processamento de imagens para extração de atributos.

Assim, este trabalho analisa e compara os resultados dos modelos de aprendizado profundo e métodos clássicos para identificação de padrões em imagens de rochas. Considerando-se a revisão de literatura realizada, constatou-se que, este é o primeiro trabalho no qual são aplicados modelos de aprendizado profundo para classificação litológica de imagens de microtomografia de rochas carbonáticas do pré-sal.

1.3 Organização da Dissertação

Esta dissertação está estruturada da seguinte forma: no segundo capítulo, é feita uma introdução à caracterização de reservatórios. No terceiro capítulo encontra-se a revisão bibliográfica de métodos computacionais para identificação de texturas em imagens. No quarto capítulo apresentam-se os dados, os modelos propostos e os métodos usados neste trabalho. No quinto capítulo realiza-se a apresentação e discussão dos resultados obtidos e, por fim, no capítulo 6, são trazidas as conclusões e trabalhos futuros, seguidos da bibliografia e dos apêndices.

Capítulo 2

Rocha Digital

Uma das tarefas mais importantes antes do início de qualquer processo de produção consiste na caracterização do campo [26]. A obtenção de medidas de propriedades petrofísicas do reservatório é crucial para modelagem geológica da acumulação de hidrocarbonetos e também para apoiar cenários de fluxo de fluido em uma rotina de produção ocasional [27]. A caracterização dos reservatórios carbonáticos envolve um conjunto de métodos que ajudam a entender os processos físicos e geológicos que compõem o sistema, incluindo técnicas de modelagem geológica, física, química e matemática para descrever a composição e o arranjo das camadas rochosas, bem como suas propriedades físicas [28].

Três propriedades principais estão ligadas entre si e relacionadas à qualidade do reservatório e à viabilidade de seu desenvolvimento: porosidade, permeabilidade e litologia [29]. Uma rocha, para ser denominada reservatório, deve ter um espaço vazio para armazenar volumes comerciais de hidrocarbonetos, devendo esse sistema poroso estar interconectado para ser capaz de fornecer os hidrocarbonetos aos poços de extração, definindo sua característica de alta permeabilidade. Por fim, a litologia, desempenha um papel essencial nas propriedades permo-porosas e mecânicas, que, por sua vez, são duas das principais propriedades ligadas à viabilidade do reservatório de exploração [29]. A discriminação precisa da litologia ajuda a reduzir o erro na previsão da permeabilidade e do volume de hidrocarbonetos, bem como a descrição dos processos deposicionais e diagenéticos, que também estão relacionados às implicações nas propriedades do fluxo de fluido [30].

2.1 Amostragem de Rochas

A determinação dos tipos de rochas de reservatório é um dos principais passos no processo de modelagem e simulação de reservatório. Em reservatórios carbonáticos, a identificação de tipos de rocha é desafiadora, em razão da heterogeneidade das estruturas ali contidas.

Para a determinação dos tipos de rochas pertencentes ao reservatório é necessário fazer um processo conhecido como testemunhagem, o qual consiste na extração, durante a perfuração do poço, de testemunhos de rocha. Esses testemunhos são amostras de subsuperfície da região de interesse, que são, então, cerradas a cada noventa centímetros de comprimento para facilitar sua locomoção. Após o processo de corte, elas são movidas das plataformas para os centros de pesquisa, onde são processadas. A Figura 2.1 apresenta a imagem de um testemunho dentro de uma caixa de armazenamento. De cada testemunho são extraídas amostras a cada 30 cm, conhecidas como plugues, conforme exemplo na Figura 2.2. Esses plugues são, então, preparados para uso em diversos processos, sendo a microtomografia computadorizada um dos processos não destrutivos, o qual gera, a partir de imagens de raio-x da amostra, um modelo 3d do seu conteúdo.

Outro processo aplicado aos plugues é a denominada extração de lâminas, as quais são seções finas de rocha usualmente visualizadas em microscópios (Figura 2.3) tratando-se de um processo destrutivo. A petrografia bidimensional dessas amostras forma a base da definição litológica [31], sendo esse procedimento subjetivo, pois depende dos atributos escolhidos para a realização da classificação das amostras [29]. Pode-se, por exemplo, focar em cenários biológicos ou enfatizar características petrológicas, contabilizando a granulometria e a mineralogia, na medida em que a quantidade de informação a ser levada em consideração em cada caso, ao analisar um reservatório específico, depende significativamente de uma intuição subjetiva e não de medidas quantificáveis [32]. Essas amostras, por serem seções finas de rocha, podem não ser representativas de todo conteúdo do plugue, visto que é apenas uma subamostragem do topo do plugue, sendo a microtomografia mais representativa do volume.



Figura 2.1: Testemunho extraído de um poço e armazenado em uma caixa, escala na lateral direita representando dez centímetros para cada mudança de cor.



Figura 2.2: Plugue retirado de um testemunho

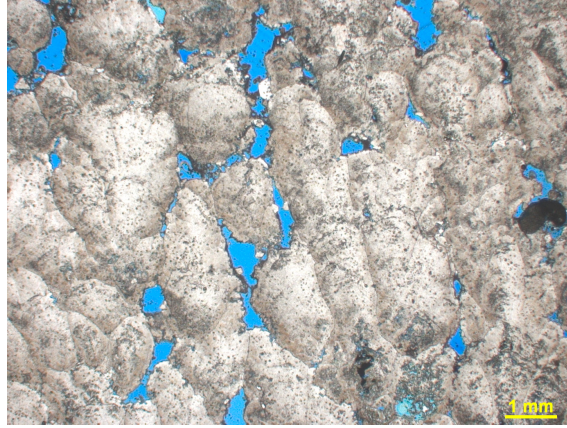


Figura 2.3: Lâmina vista por um microscópio

2.2 Tomografia Computadorizada

A tomografia computadorizada por raios-X é uma ferramenta bem conhecida para gerar modelos tridimensionais sem a destruição da amostra e foi, inicialmente, usada para fins medicinais na década de 1970. Nos anos seguintes, foi aplicada para investigar materiais de terra, como aqueles relacionados à geologia do petróleo [33–35]. Atualmente, é considerada o método mais prático para obter a estrutura interna tridimensional de meios porosos [36]. A aplicação dessa técnica para geração de volumes 3D consiste na criação de imagens usando tomografia de raios-X de uma amostra, quando geralmente coletam-se entre 500 e 3000 projeções 2D em vários ângulos diferentes, e, após essa aquisição, o volume 3D é reconstruído computacionalmente [37].

Neste sentido, a principal limitação da técnica de imagem por tomografia computadorizada estava relacionada à resolução alcançada [38, 39], uma vez que não era possível observar estruturas internas com mais detalhes nas rochas carbonáticas, visto que as tomografias possuíam baixa resolução, ou seja, granularidade mais grossa. No entanto, os avanços tecnológicos abriram possibilidades de estudos em microescala e avaliações mais detalhadas, como a distribuição de tamanho de grãos e poros, surgindo assim a técnica conhecida como microtomografia computadorizada (micro-CT), a qual permite tomografias de alta resolução (granularidade mais fina) em amostras pequenas (Figura 2.4). A disseminação da física das rochas computacionais para obter propriedades elásticas e de transporte das rochas se deveu principalmente a sistemas tridimensionais de imagem de alta resolução e melhoria de hardware e software de computador [40]. Como consequência, microtomografia computadorizada tornou-se parte comum dos fluxos de trabalho de caracterização de reservatórios [7–9], uma vez que permite uma descrição representativa da microestrutura e auxilia no entendimento dos fenômenos físicos do fluxo de fluidos [41, 42]

e propriedades mecânicas [5, 43].

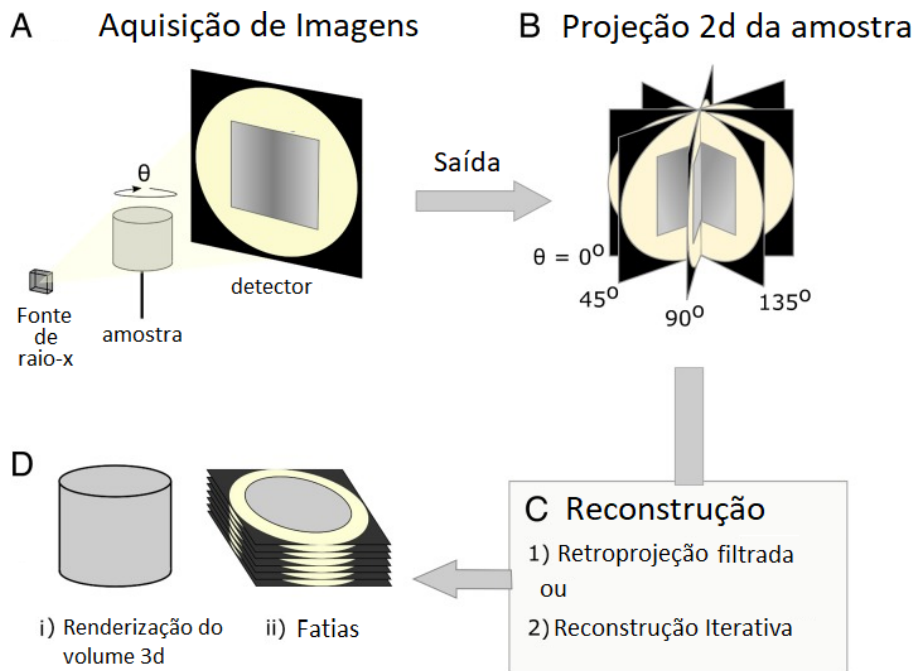


Figura 2.4: Diagrama do fluxo da microtomografia computadorizada. (A) É uma ilustração do processo de aquisição de imagens do microtomógrafo, apresentando a fonte de raio-x, o detector e a amostra. Imagens são projetadas ao se rotacionar a amostra, enquanto raios-x são emitidos. (B) Após o termino do processo anterior, é gerado um conjunto de imagens com ângulos diferentes. (C) Projeções são reconstruídas digitalmente através de algoritmos como retroprojeção filtrada ou reconstrução iterativa. (D) A amostra pode ser visualizada como um volume 3D ou fatias 2D. (Adaptado de [37])

2.3 Fluxograma de Trabalho Atual e Proposto

O método padrão para classificação litológica é baseado na análise de lâminas individuais, levando a interpretações subjetivas de rochas da mesma natureza de acordo com o especialista, mesmo havendo as definições na literatura [13]. Além disso, apenas uma pequena parte da amostra inteira pode ser estudada, posto que, enquanto a lâmina é só uma fatia, as amostras cilíndricas têm cerca de 6 cm de comprimento e 3,8 cm de diâmetro. A imagem de tomografia computadorizada é uma técnica relativamente rápida e barata que pode ser incorporada no fluxo de trabalho de caracterização de reservatório antes de qualquer teste destrutivo, produzindo imagens que correspondem à diversas lâminas em sequência, contudo, com menor resolução. Esses dados podem ser úteis para tarefas de visualização e caracterização quantitativa e qualitativa dos recursos internos, delimitados por variações na densidade e na

composição atômica. Isso motiva as experiências que propomos neste trabalho, que buscam utilizar imagens de microtomografias para apoiar a classificação litológica em vez de lâminas. Ressalta-se que, de acordo com Ketcham e Carlson [44], os dados extraídos do modelo de rocha tridimensional a partir de microtomografia estão intimamente relacionados aos dados de lâminas. Os procedimentos convencionais e os sugeridos são ilustrados na Figura 2.5.

A Figura 2.5.A mostra as etapas usuais para a definição litológica, iniciando pela extração da amostra, seguida pela extração da lâmina e, por fim, a análise petrográfica da mesma. As fotos de lâmina possibilitam uma avaliação mineralógica, uma vez que a resolução alcançada é maior do que a das imagens de microtomografia com o mesmo tamanho de amostra, por exemplo. No entanto, a petrografia fornece apenas uma análise bidimensional e é dependente da orientação de extração da lâmina. Além disso, certos tipos de rochas não são viáveis para a extração de seções finas devido a propriedades mecânicas, levando à necessidade de, para se obter dados representativos, se extrair um conjunto de seções finas paralelas e equidistantes, tornando a amostra inutilizável.

O fluxo de trabalho sugerido para classificar as litologias de amostras de plugue com base em imagens de microtomografias encontra-se na Figura 2.5.B. A proposta utiliza imagens de microtomografia computadorizada de plugues ao invés de lâminas, sendo a vantagem do procedimento recomendado a capacidade de ser utilizada uma porção mais representativa das amostras de rocha. Além de não destruir as amostras após a aquisição das imagens, permite que o mesmo conjunto de dados seja usado para outras análises. Em contrapartida, este procedimento possui uma menor resolução e dificulta a identificação de minerais devido à natureza policromática da energia dos raios-X.

Apesar das imagens de microtomografia apresentarem resoluções menores quando comparadas com as lâminas, ou seja, possuírem menos detalhes ou uma granularidade mais grossa, ao examinar as fotos de lâminas e imagens de microtomografia na Figura 2.6, verifica-se uma semelhança entre elas nas imagens de cada litologia. Ressalta-se que as imagens de lâminas apresentadas são apenas pedaços da imagem de microtomografia com resoluções maiores, podendo ser verificado pela barra de escala, onde na lâmina são 2 mm na microtomografia são 5 mm, ou seja, a imagem da lâmina possui mais detalhes. Ressalta-se que, apesar dessa diferença de resolução, ainda é possível realizar a discriminação entre litologias a partir das imagens de microtomografia computadorizada de plugues.

Terra et al. [13] define 18 macro litologias do pré-sal brasileiro, nomeadamente, mudstone, wackstone, packstone, grainstone, floatstone, rudstone, bioacumulado, brecha, boundstone, estromatólito, trombolito, dendrolito, leolito, esferulito, travertino, laminito, calcário cristalino e dolomito. Como este é um trabalho inicial na

área, as escolhidas para este trabalho foram esferulítico, estromatólito e grainstone, onde exemplos de cada litologia na escala da microtomografia podem ser vistos na Figura 2.7.

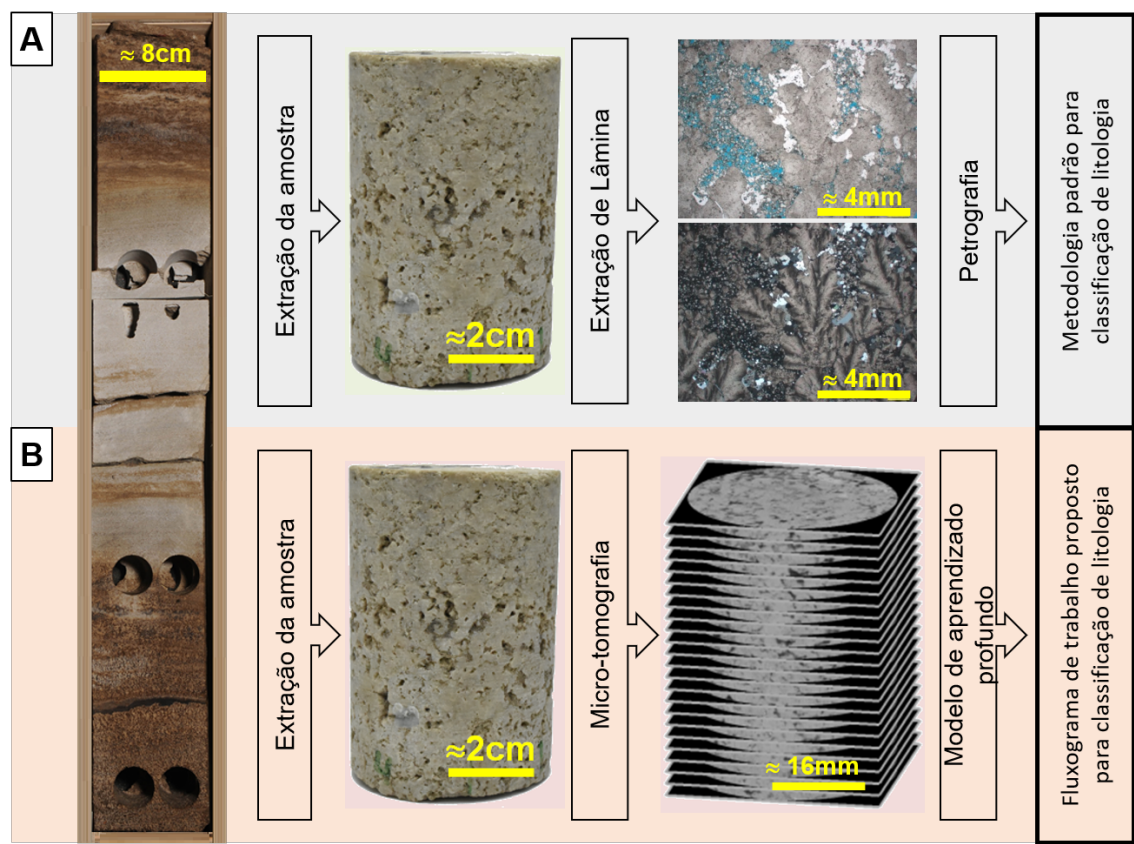


Figura 2.5: Fluxo de trabalho atual e proposto

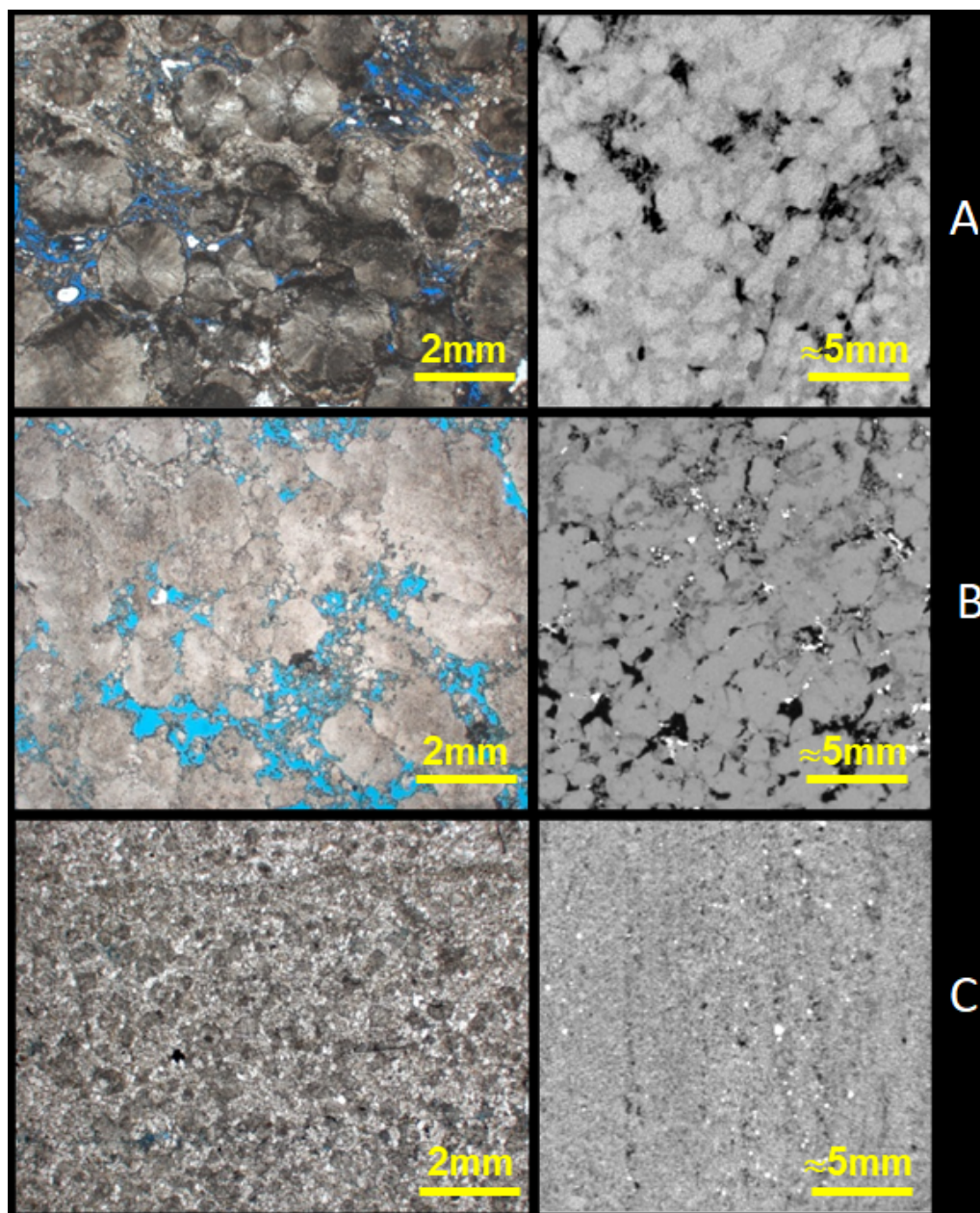


Figura 2.6: Comparação entre lâminas (coluna da esquerda) e microtomografia (coluna da direita) das litologias esferulítico(A), estromatólito(B) e grainstone(C).

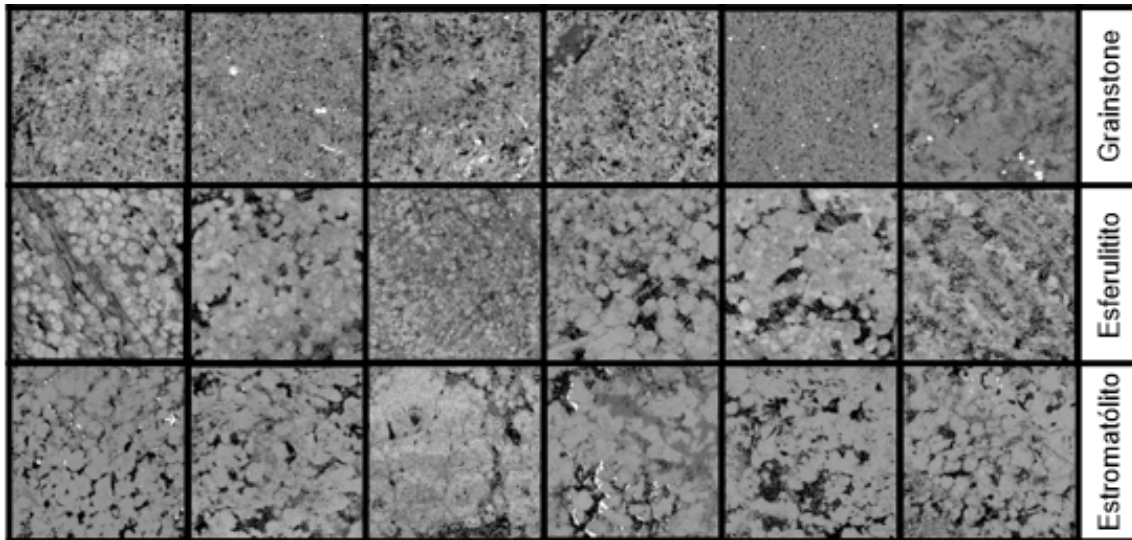


Figura 2.7: Exemplos de cada uma três litologias em imagens de micro-CT de plugues.

Capítulo 3

Revisão bibliográfica

Neste capítulo, será feita a revisão bibliográfica, apresentando-se os métodos clássicos de análise de textura em imagens e aprendizado profundo aplicado para identificação de padrões em imagens.

3.1 Métodos Clássicos

Um grande desafio na análise de textura em imagens é o fato de que textura não possui uma definição matemática, sendo atribuída a padrões visuais complexos, na medida em que são compostas por entidades espaciais organizadas que possuem luminosidade, cor, formato e tamanho [45].

Segundo Erondy, existem três diferentes tipos de texturas [46]: determinísticas, estocásticas e quase-determinísticas. O primeiro tipo refere-se às texturas compostas por um único elemento, sendo muito difícil disto acontecer em imagens reais. O segundo tipo considera as texturas compostas por padrões não regulares. Por fim, o terceiro tipo refere-se às texturas compostas por diversos padrões, que podem ser similares entre si, mas não idênticos. As imagens analisadas neste trabalho podem se encaixar na segunda ou terceira definição, devido à heterogeneidade das amostras de rochas, algumas com padrões não regulares e outras com padrões similares, mas não iguais.

Na literatura [14, 47–49], os problemas de aprendizado de máquina para classificação de imagens são abordados seguindo o fluxo apresentado na Figura 3.1. Esse fluxo é composto de duas etapas: a primeira refere-se à extração de características da imagem, sendo um dos passos mais importantes na maioria dos sistemas de reconhecimento de padrões. Já a segunda um algoritmo que tem a função de identificar cada uma das classes com base nessas características extraídas. Ressalta-se que o problema de identificação de texturas em imagens segue esse mesmo fluxo.

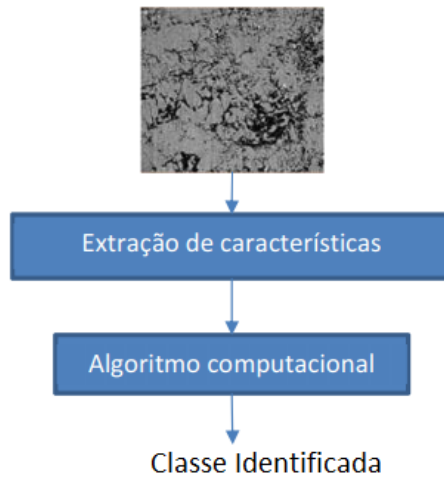


Figura 3.1: Fluxo para classificação de imagens.

O problema de caracterização de texturas, assim como o de classificação de imagens, pode ser abordado de algumas formas, as duas principais são supervisionada ou não supervisionada [14]. A supervisionada refere-se ao processo de ajuste de um modelo de aprendizado de máquina, empregando-se uma etiqueta ou alvo. Já a segunda é usada encontrar padrões nos dados sem o uso de etiquetas. Entre os modelos mais populares que utilizam a abordagem supervisionada destacam-se as redes neurais perceptron de múltiplas camadas (do inglês, *Multilayer Perceptron* - MLP) [25] e as máquinas de vetor suporte (do inglês, *Support Vector Machine* - SVM) [50]. Por outro lado, em relação aos modelos da abordagem não supervisionada, destacam-se as redes neurais do tipo Kohonen [51], métodos de agrupamento como o k-means [52], modelos de misturas de gaussianas [53], entre outros. Neste trabalho foi usado o método supervisionado para o ajuste de parâmetro dos modelos.

A ideia central da extração de características de imagens, neste caso, é descrever as texturas a partir de valores numéricos que possam ser interpretados por um algoritmo para possibilitar sua identificação [54].

Richards e Polit [55] identificaram quatro tipos de abordagens para a extração, nomeadamente, estatístico, geométrico, baseado em modelo e baseado em frequência. Os métodos estatísticos são os mais antigos de extração de características, funcionando com base na premissa de que uma textura pode ser caracterizada pela distribuição de tons de cinza na imagem, entre os quais um dos mais utilizados é o método de extração de características de Haralick, a partir da matriz de coocorrência de tons de cinza (do inglês, *Gray Level Cooccurrence Matrix* - GLCM) [56]. Os métodos do tipo geométrico, por sua vez, tentam caracterizar uma dada textura por um padrão que se repete ao longo da imagem, sendo extremamente úteis em identificação de texturas determinísticas. Entretanto, em texturas reais, poucas são as ocasiões nas quais é possível encontrar um único padrão que caracterize uma textura.

Já os métodos baseados em modelos constroem um modelo artificial da textura em questão, uma vez que, se a textura artificial for a mais próxima possível da textura desejada, os parâmetros do modelo podem ser utilizados como características para uma posterior classificação ou segmentação das texturas, entre os quais destaca-se o modelo fractal [57]. Finalmente, os métodos baseados na frequência consideram que uma textura pode ser caracterizada pelas análises da repetição de um ou mais padrões, de acordo com as diferentes frequências espaciais, sendo que o filtro de Gabor [58] é um dos mais populares.

A escolha de atributos a serem utilizados como entrada para um modelo é um passo importante na etapa de extração de características. Com isso, as características de Haralick, dimensão fractal e filtros de Gabor foram escolhidas como atributos para uso neste trabalho.

3.1.1 Matriz de Coocorrência dos Tons de Cinza e Haralick

A matriz de coocorrência dos tons de cinza é um método estatístico utilizado para examinar texturas, considerando a relação espacial dos píxeis. Trata-se de um histograma de segunda ordem que mede a proporção de ocorrência de pares de píxeis para uma dada distância e orientação entre eles, tendo sido inicialmente proposta em [56] com a intenção de descrever atributos de textura na aplicação de classificação de imagens. A partir da matriz de coocorrência dos tons de cinza, são extraídas as quatorze características definidas por Haralick [56], nomeadamente, segundo momento angular, contraste, correlação, soma dos quadrados, inversa da diferença do momento, soma das médias, soma das variâncias, soma da entropia, entropia, diferença da variância, diferença da entropia, duas medidas de informação de correlação e coeficiente máximo de correlação.

O conjunto de características de Haralick foi usado para segmentar tecidos moles em imagens de tomografias do cérebro [49]. Os autores utilizaram o algoritmo conhecido como BAM, que é um modelo baseado em redes neurais recorrentes, chegando a atingir 93% de acurácia. Já em [59], o problema abordado também consiste na análise de texturas em imagens de tomografia, onde feita a comparação das características extraídas dos conjuntos de imagens com e sem doença, através do teste ANOVA [60]. Verificando, assim, que as características homogeneidade, entropia e energia poderiam ser usadas para discriminação das imagens com e sem doença. Verificou-se também que a entropia seria uma boa característica para diferenciar edema pulmonar e câncer.

3.1.2 Dimensão Fractal

Fractais foram introduzidos por Mandelbrot em 1983 [57] para caracterizar irregularidades e descrever o comportamento natural ou artificial de objetos complexos, sendo o cálculo de sua dimensão importante, na medida que possibilita a computação da similaridade entre diferentes imagens a partir dos seus atributos fractais.

Atualmente, existem diferentes métodos de estimação da dimensão fractal propostos para análise de texturas [61, 62]. Entretanto, apesar de sua proposta, a dimensão fractal sozinha é insuficiente para discriminar texturas, visto que duas imagens distintas podem possuir a mesma dimensão fractal [63, 64].

Mahal Lehamel [65] apresenta o uso da dimensão fractal e outras características para classificação de texturas, realizando testes com diferentes estimadores e avaliando que a dimensão fractal sozinha tem um desempenho inferior no conjunto de dados escolhido. Em [66], os autores utilizam a técnica de *local binary patterns* em conjunto com a dimensão fractal para classificação de texturas de imagens do conjunto de dados Brodatz [67], alcançando acurácia de 97%.

3.1.3 Filtros de Gabor

Os filtros de Gabor foram inspirados em sistemas biológicos e propostos com o intuito de incorporar propriedades dos campos receptivos de células simples no córtex visual [68]. Um conjunto de filtros formam um banco, o qual possui capacidade de descrição espacial e frequencial de atributos de textura em múltiplas resoluções. Esses filtros permitem variar um alto conjunto de parâmetros, o que dificulta sua utilização.

Como uma das principais dificuldades de se utilizar o método da filtragem de Gabor consiste na necessidade de escolher os parâmetros para compor o conjunto de filtros, Sandler et al. [69] apresenta uma metodologia para estimar estes parâmetros e avalia essa escolha em tarefas de segmentação e classificação de texturas.

Al-Kadi [48], por sua vez, apresenta o uso da filtragem de Gabor em um problema de classificação de imagens de tecidos com o propósito de identificar tecidos cancerígenos. O autor explora o uso de características complementares à energia resultante da filtragem de Gabor, que são extraídas de modelos como dimensão fractal, *Gaussian Markov Random Fields*, entre outros. Concluindo que a melhor de combinação de modelos para este problema específico é o filtro Gabor e a dimensão fractal, atingindo acurácia de 95%.

3.1.4 Combinação dos Métodos

A combinação de métodos de extração de características pode se mostrar superior ao uso de um único método [47, 48, 65, 66]. Em [47], os autores estudam dife-

rentes combinações de características texturais em um problema de classificação de texturas. As características texturais são extraídas empregando-se métodos como autocorrelação, estatísticas (média, desvio padrão e mediana), filtro de Gabor e Haralick. Uma vez extraídas, as características são concatenadas para compor um único vetor, o qual é fornecido como entrada para um algoritmo de classificação (SVM), que faz a identificação da textura. Os testes foram realizados nas bases de dados Brodatz e Outex [70], as quais são compostas por imagens de texturas naturais, sendo que os resultados obtidos apontaram que a combinação de características leva a uma melhora significativa de desempenho, superando 95% de acurácia, enquanto as abordagens que empregam apenas um conjunto de características não passaram de 90%.

Sendo assim, a combinação de métodos se mostra importante na escolha de atributos a serem utilizados como entrada para um modelo de aprendizado de máquina. Dessa forma, a combinação das características de Haralick, filtros de Gabor e dimensão fractal foram escolhidas como atributos para obtenção de uma referência para classificação litológica de imagens de microtomografia computadorizada de rochas.

3.1.5 Identificação de Textura em Imagens de Rochas

Nesta subseção serão listados alguns trabalhos que aplicam métodos clássicos para identificação de padrão em rochas. Em [71], dois conjuntos de dados são utilizados, sendo o primeiro o KTH-TIPS [72], o qual consiste em imagens de texturas e o segundo um conjunto de imagens de rochas da categoria *bera sandstone* disponibilizadas pela Petrobras. Ressalta-se que esse tipo de rochas são arenitos, os quais possuem um comportamento mais homogêneo do que as rochas carbonáticas, facilitando, assim, a identificação de padrões nessas imagens. Os autores seguem a estratégia da combinação de atributos, de forma que os métodos escolhidos para a extração de características foram: Haralick, *Maximum Probability*, *Cluster Shade* e *Cluster Prominence* [73]. Após o processo de extração de características, um classificador bayesiano [74] é treinado para realizar a classificação das imagens. Os resultados apresentados conseguiram atingir 90% de acurácia no conjunto de teste.

Os autores em [75] utilizam fotos de rochas, as quais estão divididas em três categorias: ígneas, metamórficas e sedimentares [75]. Ressaltando-se que as imagens utilizadas são fotos de rochas e cada rocha possui formas e tamanhos diferentes, o que facilita a discriminação das classes. O modelo escolhido para classificação foi o baseado nos vizinhos mais próximos (do inglês, *K-nearest neighbours* - KNN) [76] e os atributos usados como entrada foram extraídos a partir de alguns métodos como Haralick, *edgefrequency*, autocorrelação, entre outros, obtendo-se resultados

que chegaram a alcançar 94% de acurácia.

Shang e Barnes [77] comparam árvores de decisão [78], máquinas de vetor suporte e o algoritmo de vizinhos mais próximos na tarefa de classificação de imagens de texturas de rochas. As imagens utilizadas são fotos coloridas de texturas de rochas e os atributos usados como entrada para os modelos foram estatísticas de cor das imagens, estatísticas extraídas a partir da matriz de coocorrência dos tons de cinza e do histograma das imagens. O modelo que obteve melhor desempenho no conjunto de dados utilizado foi a máquina de vetor suporte que atingiu aproximadamente 94% de acurácia nas classes propostas pelos autores.

Perez et al. [79] usam fotos de minerais obtidos na mina de Raglan [80] para classificação litológica. Os autores comparam dois métodos de extração de atributos em imagens, filtros de Gabor e transformada Wavelet. O modelo escolhido para classificação foi a máquina de vetor suporte e o melhor cenário apresentado chegou a obter aproximadamente 85% de acurácia no conjunto de dados utilizado.

Nenhum trabalho foi encontrado que utilizasse um conjunto de imagens de microtomografia de rochas carbonáticas do pré-sal. Ressalta-se que a maioria dos trabalhos apresentados nesta subseção utilizaram-se de fotos de rochas, que são imagens muito distintas de imagens de microtomografias de rochas.

3.2 Aprendizado Profundo

Todos os trabalhos anteriormente mencionados baseiam-se no fluxo descrito na Figura 3.1, no qual o processo de identificação de texturas é feito em duas etapas: a extração de características e a análise automática destas para classificação. Contudo, existe uma limitação nesta abordagem, posto que é necessário escolher corretamente os métodos de extração de características e eles, geralmente, possuem um conjunto de parâmetros que devem ser ajustados, fazendo-se necessário o uso de um especialista em análise de imagens.

O surgimento do aprendizado profundo possibilitou a extração de características sem a necessidade do especialista para fazer a escolha dos métodos de extração de característica e o ajuste de seus parâmetros. Isso só é possível porque os parâmetros existentes nas redes neurais profundas são ajustados de forma automática pelo algoritmo de retropropagação do erro e métodos de otimização. O aprendizado profundo tem demonstrado excelentes resultados em diversas áreas [81–84].

3.2.1 Rede Neural Convolucional

Uma das arquiteturas de rede neural mais utilizadas para análises de imagens é a rede neural convolucional (do inglês, *Convolutional Neural Network* - CNN) [85], que

compartilha várias ideias com métodos clássicos de extração de características em imagens, incluindo bancos de filtros, análises multiescala e classificação no espaço das características, diferenciando-se ao empregar um método de otimização para ajustar os parâmetros dos bancos de filtros. Essa rede se tornou popular quando usada no concurso do ImageNet [81], em razão de, além de possuir um desempenho superior comparada com os outros métodos, sua implementação ser eficiente em unidades de processamento gráfico (do inglês, *Graphics Processing Unit* - GPU) , o que possibilitou treinamentos mais rápidos dos modelos.

Existem várias ferramentas disponíveis online para implementação de redes neurais artificiais, a escolhida para esse trabalho foi o Pytorch [86]. Nesta ferramenta as redes são implementadas adicionando-se camada a camada. No contexto das redes convolucionais, existe uma estrutura comumente utilizada, a qual, em sua forma simples, é a sequência de três camadas consecutivas, denominadas convolucional, ativação e *pooling*, de forma que é possível a repetição das camadas e o empilhamento dessa sequência em quantas vezes desejar. Essa parte do modelo é geralmente conhecida como extração de características e espera-se que modelos com mais camadas possam extrair características mais complexas, visto que o campo receptivo da rede aumenta à medida que o dado de entrada passa por cada camada, permitindo a interpretação do espaço de estruturas em diferentes escalas [14]. Após esse processo, os atributos extraídos são fornecidos como entrada a uma rede totalmente conectada que faz o mapeamento entre os atributos e a saída desejada. A estrutura convolucional pode ser vista na Figura 3.2 e os detalhes dos elementos são fornecidos abaixo.

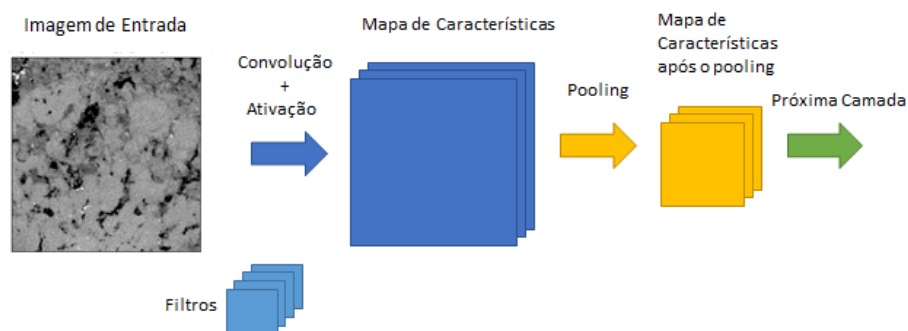


Figura 3.2: Exemplo da estrutura convolucional.

Entradas

Geralmente as entradas das redes convolucionais são imagens, as quais podem ser representadas por matrizes. Uma imagem 2D colorida equivale a uma matriz 3D, onde uma dimensão é da largura, outra do comprimento e a última dos canais de cor RGB. No caso apresentado na Figura 3.2, a entrada é uma imagem 2D em

tons de cinza, o que significa que possui apenas um canal de cor. Essa imagem pode ser representada como uma matriz 2D, onde cada posição da matriz possui a intensidade do pixel. Contudo, as entradas da camada convolucional podem ser vetores 1D, imagens 2D com mais canais e imagens 3D com um ou mais canais, sendo possível que tais entradas também sejam um mapa de características gerados por uma camada convolucional anterior ou uma camada de *pooling* [14].

Camada Convolucional

A camada convolucional consiste em um conjunto de filtros treináveis que possuem um campo receptivo, o qual depende do tamanho do filtro. Cada filtro é convolvido pelas dimensões da entrada, e, então, computa-se o produto escalar entre os parâmetros do filtro e a seção de entrada correspondente, obtendo-se, como resultado desta operação, um mapa de características. É importante observar que os parâmetros do filtro são compartilhados em todas as imagens de entrada, reduzindo-se a quantidade total de parâmetros, quando comparados a uma camada totalmente conectada. Esta operação pode ser representada como:

$$\mathbf{o}_l = f((\mathbf{o}_{l-1} * \mathbf{W}_l) + \mathbf{b}_l), \quad (3.1)$$

onde \mathbf{o}_l é o tensor de saída da camada l ; \mathbf{o}_{l-1} é o tensor de saída da camada anterior ou, quando $l = 1$, a entrada; $*$ é o operador de convolução; \mathbf{W}_l é o tensor de pesos dos filtros da camada l ; \mathbf{b}_l é o vetor de bias dessa mesma camada; e f é a função de ativação. Essa operação permite preservar a informação espacial e a análise estrutural multidimensional, o que é geralmente desejável em problemas de visão computacional para detectar texturas, formas, estruturas complexas, entre outras [14].

Esta operação pode ser interpretada de forma que dado um filtro $n \times n$, que é uma matriz, percorre a entrada, que também é uma matriz, e faz o produto elemento a elemento, seguido de uma soma. Ao analisarmos essa operação por esse ponto de vista, facilita o entendimento do *framework* Pytorch, visto que alguns dos parâmetros disponíveis são o tamanho do filtro, o deslocamento e o preenchimento da entrada. Esse preenchimento é feito para modificar o tamanho da saída dessa operação, usualmente é feito o preenchimento com zeros. Na Figura 3.3, vemos um exemplo de como um filtro convolucional funciona.

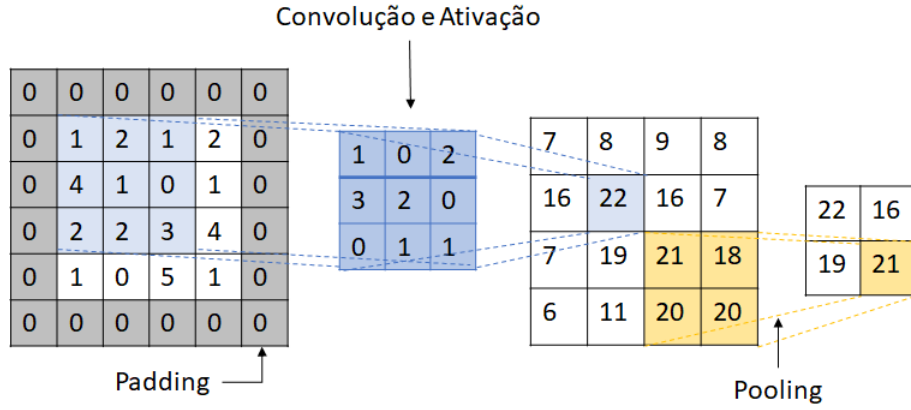


Figura 3.3: Exemplo da aplicação de convolução em uma matriz de entrada 4×4 , preenchimento por 0 de tamanho 1, janela convolucional 3×3 e o stride 1, camada de ativação ReLU e Max Pooling 2×2 com padding 0 e stride 2.

Camada de Ativação

Como apresentado na Equação 3.1, a função de ativação está ligada ao filtro convolucional, sendo aplicada após o operador de convolução e inserindo-se um comportamento linear ou não linear, dependendo da necessidade. Atualmente, a função ReLU (Unidade Linear Retificada - $R(x) = \text{Max}(0, x)$) é a mais utilizada, porém outras, como tangente hiperbólica ou função sigmoide, também podem ser usadas [14].

Camada de Pooling

A camada de *pooling* é usada para diminuir a dimensão das entradas com o objetivo de reduzir o custo computacional do modelo, sendo isso realizado através da subamostragem do mapa de características, adicionando robustez ao processo de extração de atributos, na medida em que o torna invariante a pequenas translações [14]. Esse operador é análogo ao de convolução, contudo, este faz uma seleção de elementos de uma janela e não possui parâmetros treináveis, sendo as funções mais comuns para esse operador a média e o máximo.

Camada Totalmente Conectada

Em um problema de classificação de imagens, após a estrutura apresentada na Figura 3.2, geralmente é adicionada uma camada totalmente conectada na rede neural para permitir a predição das classes [81]. Essa camada é adicionada para transformar os mapas de características em um vetor de características ou para normalizá-los em uma distribuição de probabilidades, possibilitando a classificação. Sua formulação matemática pode ser dada por:

$$\mathbf{h}_l = f(\mathbf{h}_{l-1} \mathbf{W}_l + \mathbf{b}_l), \quad (3.2)$$

onde \mathbf{h}_l é o vetor de saída, \mathbf{h}_{l-1} é a versão vetorizada do tensor \mathbf{o}_{l-1} , \mathbf{W}_l é o tensor de pesos da camada l , \mathbf{b}_l o vetor de bias dessa camada e f a função de ativação.

Ressalta-se que para adicionar a camada totalmente conectada é necessário vetorizar o mapa de características, o que faz com que o modelo da rede neural se torne estático quando implementado no *framework* Pytorch, fixando-se assim o tamanho da entrada. Isso ocorre em razão da camada totalmente conectada esperar uma entrada de tamanho fixo, e, para que seja permitido o treinamento, é necessário que o mapa de características gerado pelas camadas convolucionais tenha sempre o mesmo tamanho.

Spatial Pyramid Pooling

O uso da camada totalmente conectada após a estrutura convolucional tem alguns problemas, sendo um deles a fixação do tamanho da entrada. Um problema, ainda mais alarmante, é a quantidade de parâmetros treináveis que essa camada adiciona. Os autores em [87] propuseram o uso, em redes neurais convolucionais, do *Spatial Pyramid Pooling* (SPP) [88], reduzindo, dessa forma, o número de parâmetros treináveis e permitindo-se o uso de entradas com tamanhos diferentes. Essa técnica consiste na aplicação de diferentes *poolings* de acordo com a entrada, cujo exemplo visual pode ser verificado na Figura 3.4.

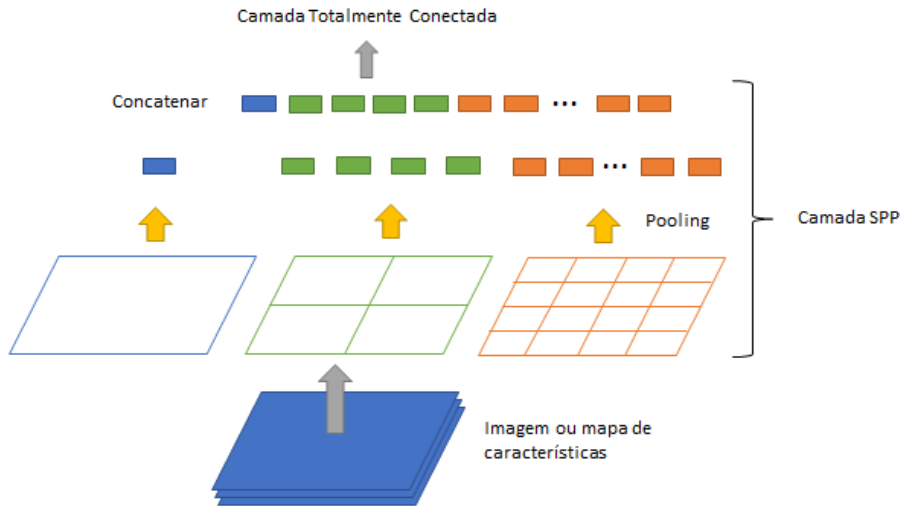


Figura 3.4: Camada SPP

A quantidade de níveis no SPP é arbitrária, posto que quanto mais níveis, mais divisões da entrada serão feitas. No caso apresentado, são usados três níveis de *pooling*, o que faz com que a imagem seja dividida de três formas: o primeiro nível

faz *pooling* na imagem inteira, o segundo divide a imagem em 4 partes, aplicando *pooling* em cada parte, e, finalmente, o terceiro nível divide a imagem em 16 partes, aplicando *pooling* em cada uma das partes. Com isso, é gerado um vetor com $21(1+4+16)$ pontos para cada imagem ou um mapa de características de entrada, possibilitando a camada SPP o uso de imagens de diferentes tamanhos, visto que gera um vetor de mesmo tamanho para cada uma das entradas.

Transferência de Aprendizado

Com a evolução da aplicação dos modelos de aprendizado profundo em reconhecimento de padrões em imagens, tornou-se interessante aplicar a técnica conhecida como transferência de aprendizado nesses tipos de modelos. Essa técnica consiste no treinamento prévio de um modelo em um conjunto de dados diferente do de interesse e do seu retreinamento, para ajuste fino dos parâmetros, no conjunto de dados de interesse [89]. Ela já era aplicada em outros tipos de modelos [89], e, ao ser aplicada em modelos de aprendizado profundo, conseguiu desempenhos satisfatórios, superando os modelos treinados apenas com os dados de interesse [89–91] .

A aplicação dessa técnica em modelos de aprendizado profundo pode ser feita usando-se um modelo pré-treinado em um conjunto de dados específico, como, por exemplo, o ImageNet. Então, substitui-se a última camada do modelo para permitir a predição das classes de interesse, sendo realizado, em seguida, um ajuste fino dos parâmetros no conjunto de dados de interesse. Este ajuste pode ser feito somente nos novos pesos adicionados, ou em todos os pesos da rede [89].

Em [90], os autores desenvolveram um novo tipo de arquitetura de rede neural convolucional ao combinar as transformadas wavelets com os filtros das camadas convolucionais, e, utilizando também modelos pré-treinados para classificar texturas nos conjuntos de dados DTD [92] e KTH-TIPS2 [93]. Os resultados obtidos mostram que a arquitetura proposta é competitiva com modelos de transferência de aprendizado.

Os autores de [91] comparam o modelo AlexNet com o modelo proposto por eles, o T-CNN, caracterizado como um modelo para identificação de texturas. Além de comparar também os resultados usando a técnica de transferência de aprendizado e sem utilizá-la. Sete conjuntos de dados de texturas foram utilizados e os resultados obtidos apontaram que sem o uso da transferência de aprendizado, o modelo T-CNN supera o AlexNet na maioria dos conjuntos de dados. Contudo, ao usar a transferência de aprendizado, o modelo AlexNet obteve desempenho superior na maioria dos casos.

3.2.2 Aplicações

Um dos primeiros modelos convolucionais que ganhou destaque foi o AlexNet [81], no qual o modelo é composto de algumas camadas convolucionais juntamente com camadas de *pooling*, seguidas por camadas totalmente conectadas para a classificação de imagens do conjunto do ImageNet [20]. Com o passar dos anos, mais modelos foram propostos, conseguindo ser capazes de diminuir o erro obtido, sendo um deles o Inception [94], o qual propôs uma maneira diferente do uso das camadas convolucionais e obteve resultados ainda melhores na mesma tarefa de classificação. As CNN apresentam excelentes resultados em problemas de classificação e segmentação de imagens [16–18].

Na área médica o uso de aprendizado profundo vem crescendo nos últimos anos, principalmente com imagens tomográficas [95]. Em [96], a técnica de transferência de aprendizado em modelos convolucionais é feita para classificação de doenças de retina em imagens tomográficas. O conjunto de imagens utilizadas foram 162,721 tomografias de 5,588 pacientes. Os modelos chegaram a atingir 98% de acurácia no conjunto de dados testados. Li et al. [97] fazem a classificação de nódulos pulmonares no conjunto de dados LIDC-IDRI [98]. Os modelos utilizados também fizeram uso da técnica de transferência de aprendizado e chegaram a obter 95% de acurácia no conjunto testado.

As análises de texturas continuam sendo um problema desafiador em razão das texturas frequentemente não possuírem uma forma definida como objetos. Em [99], os autores aplicam modelos convolucionais para a classificação de texturas em três conjuntos de dados diferentes, KTH-TIPS2, FMD, e DTD, mostrando melhores resultados que outros modelos como SVMs e MLPs. Liu et al. [100] fazem uma revisão de artigos aplicados no problema de classificação de texturas e a evolução das técnicas com o tempo até chegar em modelos de aprendizado profundo, mostrando como os modelos mais modernos superaram os métodos clássicos.

3.2.3 Aprendizado Profundo para Classificação de Imagens de Rochas

Apesar do aprendizado profundo ter mostrado bons resultados na identificação de padrões em imagens, poucos trabalhos foram publicados aplicando-se esses modelos em conjuntos de imagens de rochas. Em [101], os autores fazem classificação de amostras de lâminas de arenitos do campo de Ordos [101], usando modelos de redes neurais convolucionais, de forma que três propostas de topologia são testadas, com três, quatro e cinco camadas convolucionais, sendo diferentes taxas de aprendizado também testadas e atingindo-se 98.5% de acurácia no conjunto de teste. Ressalta-se que as imagens utilizadas eram muito distintas entre as classes, facilitando a

discriminação e possibilitando um desempenho de quase 100%.

Os autores em [102] fazem a classificação de fotos de rocha usando dois modelos convolucionais, sendo importante ressaltar que as fotos de rochas são muito diferentes de imagens de microtomografias computadorizadas, posto que, além da resolução ser diferente, as rochas nessas imagens têm formas, tamanhos e cores únicos. O conjunto de dados consiste de 2290 imagens divididas em seis tipos de rochas, das quais 458 foram usadas para teste, obtendo o melhor modelo 97,67% de acurácia.

Em [103], os autores comparam modelos de aprendizado de máquina (SVM, KNN e MLP), modelos convolucionais e transferência de aprendizado para classificação de fotografias de rochas. O conjunto de dados utilizado é composto por 2206 imagens de 12 tipos de estruturas geológicas. Estatísticas das imagens foram usadas como entrada para os modelos de aprendizado de máquina e as imagens foram usadas para os modelos convolucionais e de transferência de aprendizado. Os resultados obtidos apontam o modelo de transferência de aprendizado como o de melhor desempenho no conjunto de dados testado, atingindo acurácia de 91% no conjunto de teste.

Capítulo 4

Materiais e Métodos

Neste capítulo, serão apresentados os dados utilizados, os modelos propostos e os métodos usados para treinamento e avaliação desses modelos.

4.1 Dados

O conjunto de dados deste trabalho é um acervo de imagens de microtomografias de rochas carbonáticas do pré-sal disponibilizadas pela Petrobras para o projeto de pesquisa "Desenvolvimento de técnicas de machine learning para identificação de facies, a partir de tomografia de raios-X de rochas carbonáticas, em multiescala, para análise de escoamento de fluidos".

Os dados disponibilizados foram imagens de 292 plugues e suas respectivas anotações. Essas anotações foram feitas por geólogos especialistas em caracterização de rochas carbonáticas para identificação de litologias nas lâminas seguindo as definições da literatura [13]. Analisando-se a lista com anotações, verificou-se um problema, o qual limitou o uso dos plugues, que foi a discordância entre os especialistas, como pode ser visto na Figura 4.1. Dos 292 plugues anotados, apenas 129 foram anotados por mais de um especialista, onde em 69 deles os especialistas discordaram, ou seja, em 53% dos casos anotados por mais de um especialista ocorreu discordância. Devido a este alto grau de discordância, foi feita a escolha de utilizar apenas os plugues anotados por mais de um especialista e quando estes concordaram, ignorando, assim, os plugues anotados por apenas um especialista. Com isso, apenas 60 plugues puderam ser utilizados. Essa discordância entre os especialistas mostra que existe alguma subjetividade na interpretação das rochas, o que faz com que a automatização desse processo venha a unificar e facilitar as interpretações.

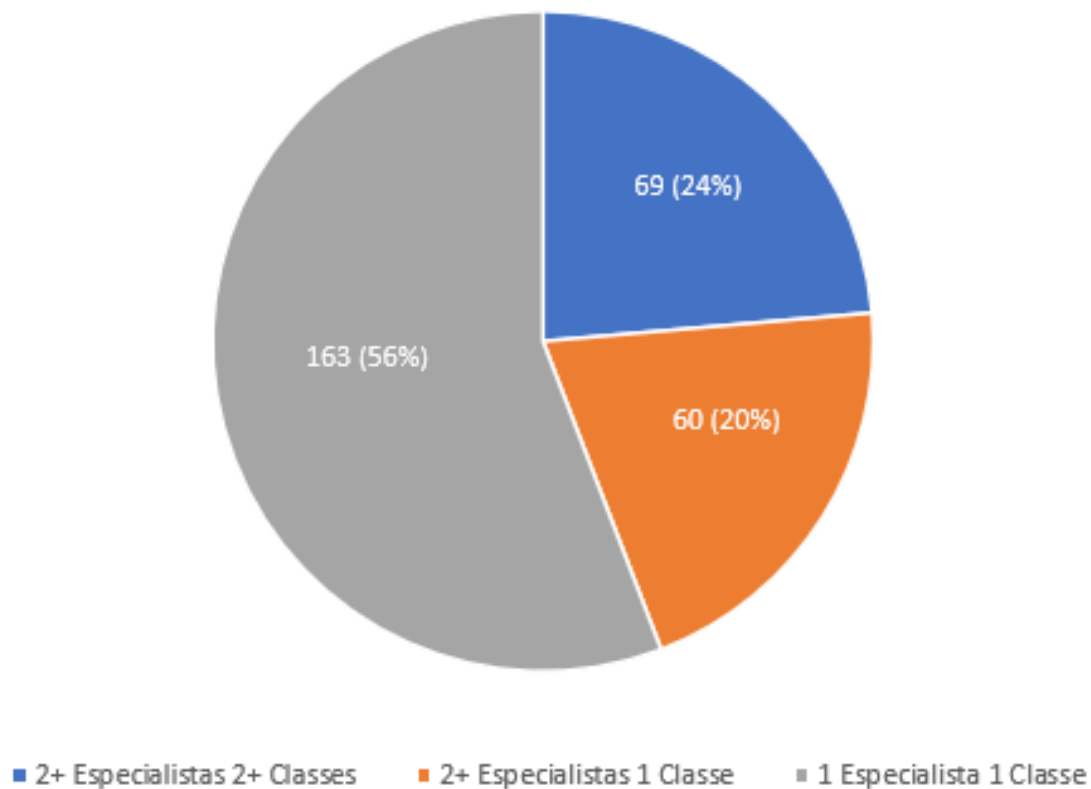


Figura 4.1: Análise dos plugues anotados por especialistas.

Cada um dos 60 plugues escolhidos para uso possuem 100 imagens, totalizando 6.000 imagens com apenas um canal de cor. Ressalta-se que todas as imagens de microtomografia utilizadas possuem resoluções entre $26.8 \mu m$ e $42 \mu m$. Assim como no processo de caracterização das rochas é assumido que as lâminas representam a litologia do plugue, foi assumido que cada um dos plugues é homogêneo, ou seja, a mesma anotação dada para cada lâmina foi usada para todas as imagens da cada plugue. Dessa forma, o conjunto de dados ficou dividido em três litologias: estromatólito, esferulítico e grainstone (exemplos mostrados na Figura 4.2). A distribuição das classes anotadas dos plugues pode ser vista na Tabela 4.1, na qual verifica-se que este conjunto de dados não é balanceado, visto que a classe estromatólito é mais frequente que as demais.

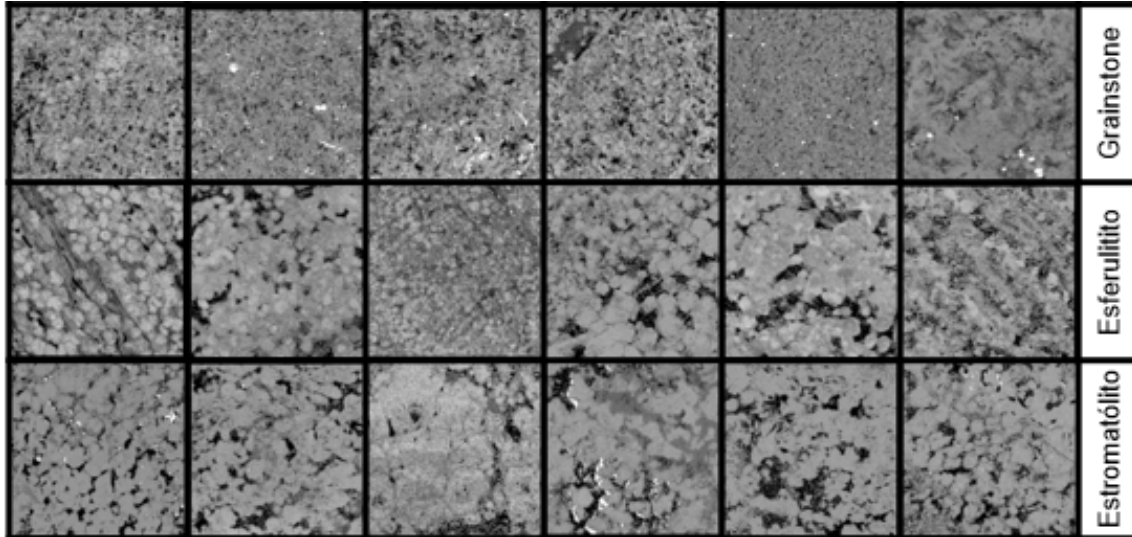


Figura 4.2: Exemplos de cada uma três classes do conjunto de imagens de micro-CT de plugues.

Tabela 4.1: Distribuição das classes anotadas no conjunto de imagens de plugues

	Esferulítico	Estromatólito	Grainstone
plugues	18	24	18
imagens	1800	2400	1800

As imagens de micro-CT de plugues disponíveis possuem diferentes tamanhos, variando de 400×400 a 999×999 . Isso é justificado pelo fato de que, além das resoluções das microtomografias dos plugues não serem as mesmas, as imagens geradas foram cortadas para maximizar a região de rocha, conforme exemplo na Figura 4.3. Contudo, este não é o único desafio deste conjunto, visto que existem alguns casos de imagens que foram classificadas como litologias diferentes, mas são visualmente parecidas na escala do plugue, conforme exemplo na Figura 4.4. Além disso, outras barreiras encontradas foram a existência de artefatos nas imagens, conforme apresentado na Figura 4.5, e as diferentes condições de luminosidade que elas possuem, como pode ser visto na Figura 4.6.

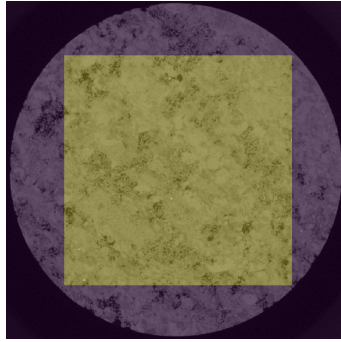


Figura 4.3: Exemplo de corte na imagem de micro-CT de um plugue com objetivo de maximizar a região de rocha.

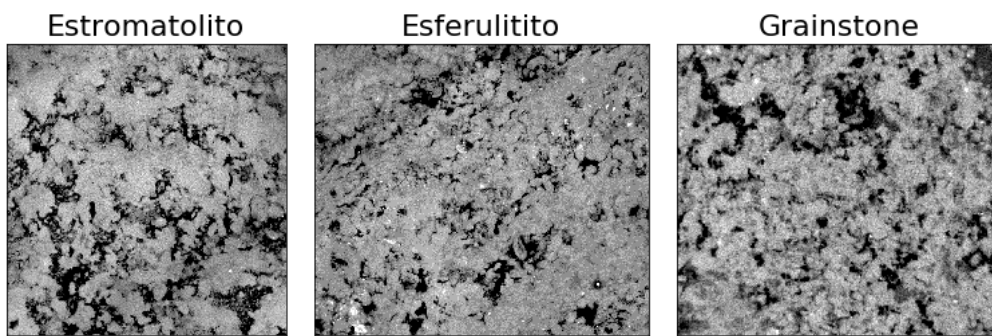


Figura 4.4: Imagens de micro-CT de plugues visualmente parecidas, porém, provenientes de classes diferentes.

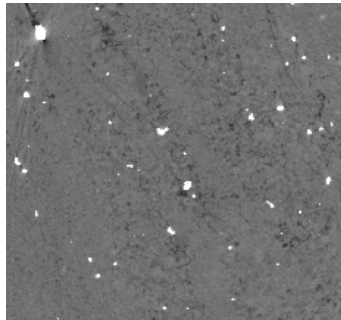


Figura 4.5: Imagem de micro-CT de plugue com artefato.

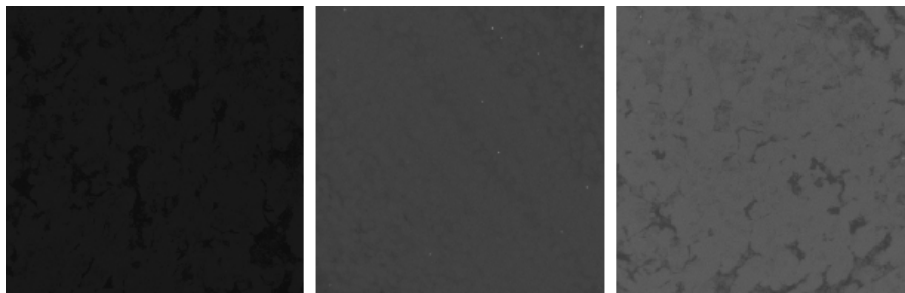


Figura 4.6: Imagens de micro-CT de plugues com diferentes condições luminosas.

4.2 Modelos

Nesta seção, serão descritos os modelos utilizados no trabalho, começando pelos modelos propostos e, em seguida, o modelo escolhido para o uso da técnica de transferência de aprendizado. As ferramentas utilizadas para implementação dos modelos encontram-se no Apêndice A.

4.2.1 Perceptron Multicamadas

O primeiro modelo criado foi um perceptron multicamadas (MLP). Alguns testes para escolha da topologia foram realizados e se encontram no Apêndice B. O modelo implementado possui a estrutura apresentada na Figura 4.7 e o escolhido nos testes realizados foi o que possui 10 neurônios na camada escondida e a taxa de dropout de 0,4.

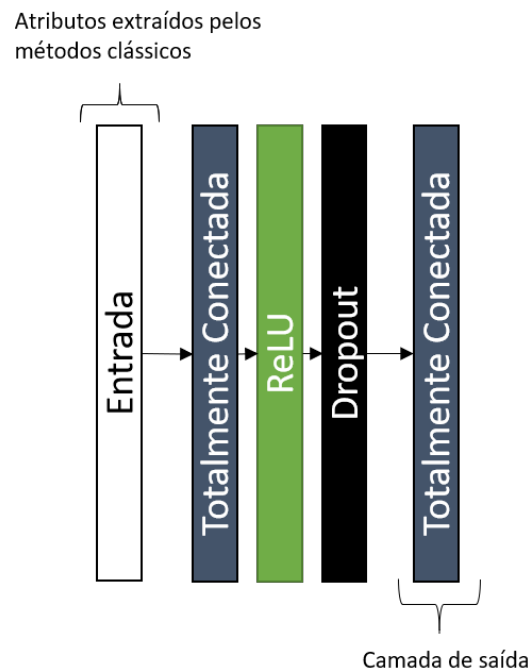


Figura 4.7: Estrutura do perceptron multicamadas proposto

As entradas usadas nesse modelo foram os atributos gerados a partir dos métodos clássicos de extração de características de imagens: características de Haralick, dimensão fractal e filtros de Gabor, os quais geraram um vetor de 47 atributos para ser usado como entrada. Essas técnicas são utilizadas para identificação de texturas e possuem implementações consolidadas disponíveis gratuitamente online.

4.2.2 Modelos Convolucionais

Foram propostos dois modelos convolucionais com estruturas similares, de forma que sua principal diferença é a presença da camada SPP, mostrados nas Figuras 4.8 e 4.9.

Essas duas topologias foram implementadas com o objetivo de comparar os resultados de um modelo (com camada SPP) que pode receber imagens de qualquer tamanho e um modelo (sem camada SPP) que precisa de uma entrada com o tamanho fixo, além de comparar também o uso de uma estrutura que reduz o número de parâmetros, mas ao mesmo tempo extrai informações em multi-escala (modelo com camada SPP). Apesar de ser uma solução simples, a implementação do modelo para tratar imagens de diferentes tamanhos é um desafio, visto que para fazer o cálculo do gradiente, as bibliotecas disponíveis transformam as redes neurais em modelos estáticos para facilitar a otimização dos mesmos, não permitindo assim entradas com tamanhos variáveis. Para solucionar esse problema, foi necessária a criação de uma estrutura de dados e um iterador que possibilitasse treinar os modelos em batches, viabilizando, assim, a paralelização do treinamento, não existindo nenhuma biblioteca no momento que disponibilize essa funcionalidade. Essa estrutura de dados gera um viés no treinamento, visto que agrupa imagens de mesmo tamanho. No caso deste trabalho, as imagens de cada um dos plugues têm tamanhos diferentes, o que fez com que o *batches* de treinamento tivessem apenas imagens de um mesmo plugue, enviesando assim o gradiente durante o treinamento do modelo.

Como visto nas Figuras 4.8 e 4.9, a estrutura convolucional do modelo possui 3 camadas convolucionais com função de ativação ReLU, 2 camadas de *max pooling* e uma camada SPP ou mais uma camada de *max pooling*. Essa estrutura é seguida de uma camada totalmente conectada, uma camada de dropout [104], e, por fim, a camada de saída da rede. Ambos modelos possuem 64 filtros na primeira camada convolucional, 48 na segunda e 32 na terceira, e, já na camada totalmente conectada, há 200 neurônios e a taxa de dropout é de 0.4. Todas camadas convolucionais propostas possuem janela de 3×3 , preenchimento por zeros de tamanho um e deslocamento de um, fazendo com que a saída e entrada de cada camada convolucional possuam o mesmo tamanho.

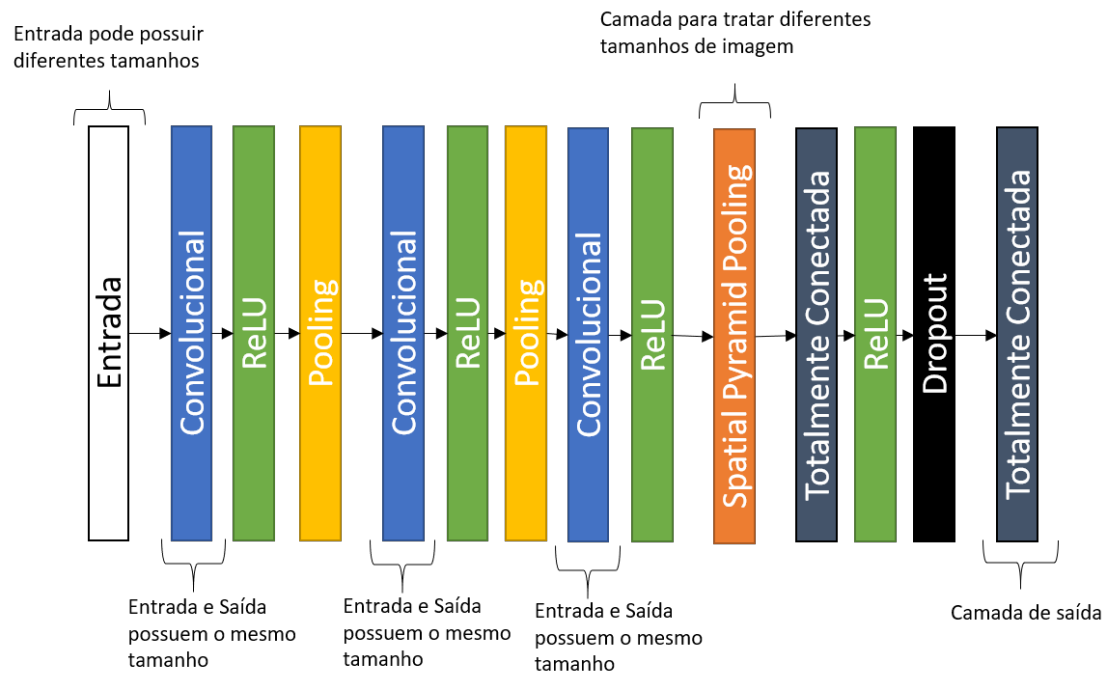


Figura 4.8: Estrutura do modelo convolucional proposto

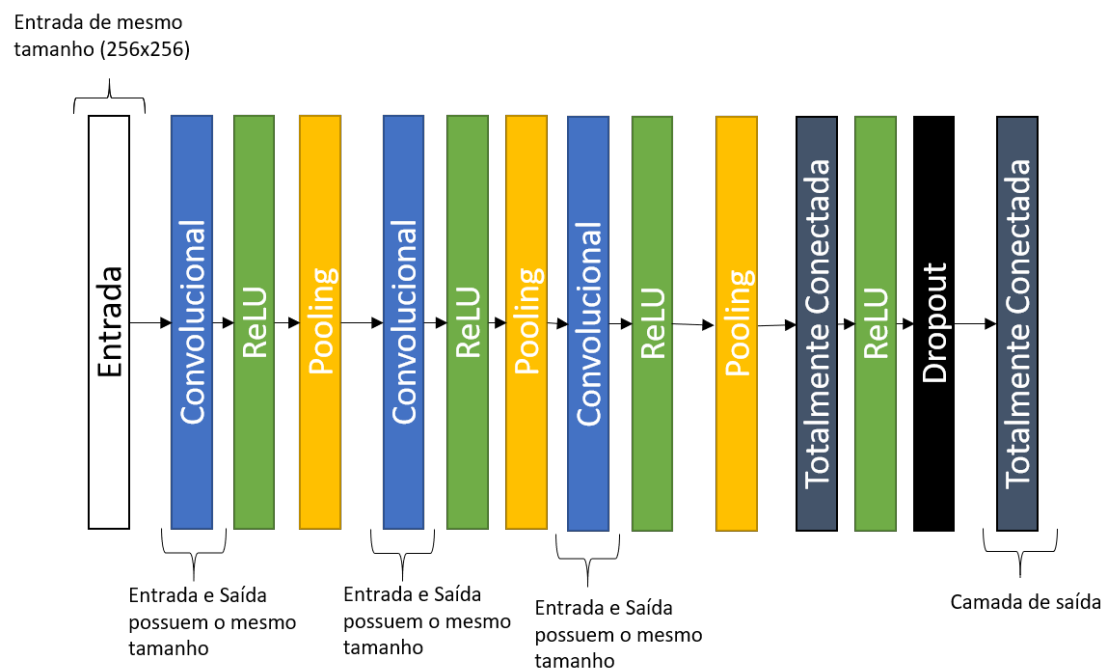


Figura 4.9: Estrutura do modelo convolucional proposto sem a camada de SPP

4.2.3 Modelo Inception V3

O modelo Inception V3 [19] é um modelo de rede neural convolucional utilizado para classificação de imagens. Esse modelo é a terceira versão do modelo Inception proposto por Szegedy [94]. Ele foi escolhido por seu desempenho no conjunto ImageNet [20], com erros de 22.55% no top-1 e 6.44% no top5. Outro motivo para sua utilização consistiu na facilidade de uso e modificação do modelo no framework pytorch, visto que o modelo pré-treinado está disponível online e para utilizá-lo nos dados deste trabalho foi necessário apenas pequenas alterações no código. Na Figura 4.10 é possível ver o diagrama de alto nível do modelo Inception v3 e maiores detalhes sobre sua arquitetura podem ser vistos em [19, 94, 105]. Ressalta-se que neste trabalho todos os parâmetros treináveis da rede foram retreinados e não somente os parâmetros da última camada adicionada. Essa escolha de retreinamento foi feita, devido aos testes preliminares que se mostraram mais promissores que retreinando apenas os pesos da última camada.

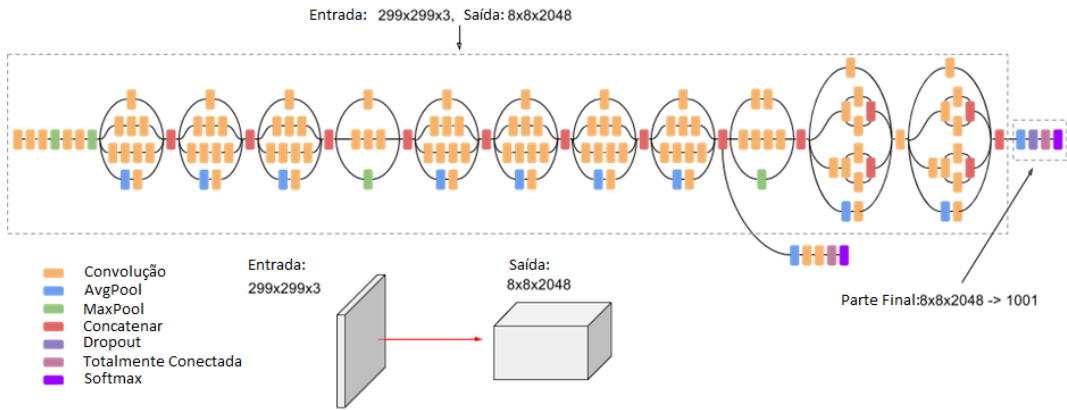


Figura 4.10: Estrutura do modelo Inception v3 (Adaptado de [105])

4.2.4 Comparação dos Modelos

Ao compararmos a quantidade de parâmetros treináveis de cada um dos modelos utilizados neste trabalho, podemos ter uma ideia melhor da diferença entre eles (veja a Tabela 4.2). É notória a diferença da quantidade de parâmetros treináveis, na medida em que é notável que o modelo convolucional sem SPP possui 3,6 vezes menos parâmetros que o modelo Inception v3. Já o modelo com a camada de SPP possui 134 vezes menos que o Inception v3 e 37 vezes menos do que o modelo convolucional sem SPP. O modelo perceptron multicamadas é o que possui menos parâmetros, contudo, se a imagem redimensionada fosse usada como entrada para esse modelo, ele teria aproximadamente 1390 vezes mais parâmetros do que usando os atributos

extraídos pelos métodos escolhidos, o que resultaria em ter mais parâmetros que o modelo convolucional com SPP.

Tabela 4.2: Comparação da quantidade de parâmetros treináveis dos modelos.

Modelo	Parâmetros Treináveis
MLP	513
CNN+SPP	177.395
CNN	6.596.595
Inception v3	23.851.784

4.3 Avaliação

Para realizar a análise dos modelos, é necessário definir figuras de mérito e métodos de extração das estatísticas de desempenho dos mesmos. Nesta seção, serão apresentadas as figuras de mérito escolhidas para este projeto e o método de separação de dados para extração de estatísticas, que foram adotados para análise e treinamento dos modelos.

4.3.1 Figuras de Mérito

Existem diversas maneiras de se avaliar um algoritmo de aprendizado de máquina [60, 106, 107]. As funções de avaliação são conhecidas como figuras de mérito, sendo escolhidas para este trabalho a acurácia e o índice SP [108], as quais são baseadas na matriz de confusão.

A matriz de confusão é a matriz que apresenta visualmente o desempenho de um algoritmo podendo ser construída para todas as classes, ou seja, o número de linhas e colunas é o número de classes, ou pode ser feita uma matriz para cada classe do problema, onde é classificado como a classe ou não classe, gerando, assim, uma matriz de duas linhas e colunas.

Para um problema de três classes, assim como o problema deste trabalho, a maneira escolhida para análise foi o primeiro caso citado, de forma que pode-se usar apenas uma matriz de confusão para visualizar e calcular as funções de avaliação escolhidas, conforme exemplo visto na Tabela 4.3.

Tabela 4.3: Matriz de Confusão			
	\hat{C}_1 (Predita)	\hat{C}_2 (Predita)	\hat{C}_3 (Predita)
C_1 (Real)	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$
C_2 (Real)	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$
C_3 (Real)	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$

Onde $C_{i,j}$ é o número de amostras ou a porcentagem das amostras na posição i,j da matriz na matriz de confusão, no qual i representa a classe real e j representa a classe predita do modelo. A soma de cada linha i da matriz representam todas amostras da classe C_i e a soma de cada coluna j são todas as amostras preditas como da classe C_j . Quando $i = j$ temos o número de amostras classificadas corretamente da classe C_i . Com a definição da matriz de confusão, é possível calcular as funções de avaliação escolhidas.

1. Acurácia

A acurácia consiste em quantas amostras o modelo conseguiu prever corretamente em relação a todas as amostras apresentadas [60, 106, 107]. Sua formulação é dada por:

$$Acurácia = \frac{\sum_{a=1}^N C_{a,a}}{\sum_{a=1}^N \sum_{b=1}^N C_{a,b}} \quad (4.1)$$

onde N é o número total de classes. O numerador é o somatório dos elementos da diagonal principal, enquanto que o denominador é a soma de todos elementos da matriz de confusão.

2. Recall

O recall é a quantidade de amostras de uma classe previstas corretamente para todas as amostras daquela classe [60, 106, 107]. A formulação do recall é dada por:

$$Recall(C_i) = \frac{C_{i,i}}{\sum_{b=1}^N C_{i,b}} \quad (4.2)$$

onde i representa o índice da classe e N é o número total de classes. O numerador é a quantidade de elementos preditos da classe C_i e o denominador corresponde a todos elementos da classe C_i e que pode ser calculado pela soma dos elementos da linha i .

3. Índice Soma-Produto

O índice Soma-Produto foi criado de maneira a equilibrar as eficiências das classes [108]. Ao avaliarmos a Equação 4.3, vemos que, quando a eficiência de uma classe é baixa, o índice tende a ser um valor pequeno, e, quando as eficiências são altas, seu valor se aproxima de 1. O uso desse índice facilita a comparação de modelos, visto que independente do número de classes, teremos apenas um único índice de desempenho.

$$SP = \sqrt{\left(\frac{\sum_{i=1}^N Recall_i}{N}\right) \cdot \left(\sqrt[N]{\prod_{i=1}^N Recall_i}\right)} \quad (4.3)$$

onde N é o número total de classes e o $Recall_i$ representa o recall de cada classe, ou seja, a eficiência de cada classe.

4.3.2 Validação Cruzada

O método de validação cruzada por k-folds é um dos métodos mais populares para estimação do erro em problemas de aprendizado de máquina [60, 106], através da estimação da variância dos resultados de diferentes amostras de teste. Esse método consiste na divisão do conjunto de dados em k subconjuntos de tamanhos iguais ou quase iguais, onde cada subconjunto é conhecido como um fold e são estratificados, ou seja, cada fold possui a mesma distribuição das classes ou quase a mesma distribuição. O modelo é então treinado com $k - 1$ folds e testado no fold restante, sendo este processo repetido k vezes, até que todos conjuntos tenham sido usados para teste uma vez. Dessa forma, o experimento retorna k estimações do erro de classificação do modelo.

O método de k-folds aninhados utilizado neste trabalho consiste na aplicação do método de k-fold adicionando o conjunto de validação, ou seja, primeiro divide-se o conjunto em k folds, onde k-1 folds são usados para treinamento e validação, e o fold restante para teste. Dentro do conjunto de treino e validação, separa-se um fold para validação e o resto para treino, treinando-se o modelo e usando-se o conjunto de validação como critério de parada para o treinamento, de forma que, após o treinamento estar completo, o fold de validação é alterado e um novo modelo é treinado seguindo a mesma lógica. Esse processo é repetido até que todos os folds tenham sido utilizados uma vez para validação. Dessa forma, treinamos k-1 modelos para um conjunto de teste, então altera-se o conjunto de teste, repetindo-se o processo, e, conseqüentemente, treinando no total $k \times (k - 1)$ modelos. A representação visual desse processo pode ser visto na Figura 4.11, onde apresenta-se um caso dessa técnica com $k = 6$, esse exemplo foi escolhido, visto que neste trabalho foram utilizados 6 folds na validação cruzada aninhada.



Figura 4.11: Exemplo de k-folds aninhados com 6 folds

4.3.3 Abordagens de avaliação

Existem duas abordagens de avaliação que podem ser utilizadas neste trabalho. A primeira é analisar os erros e acertos de cada imagem individualmente. Já a segunda é avaliar a classificação de cada plugue, sendo necessário obter uma classificação única por plugue. Para atingir esse objetivo a classe mais frequente entre as classes preditas pelos modelos para todas as imagens de um plugue foi utilizada como classificação para aquele plugue. A vantagem da primeira abordagem é utilizar uma estatística com mais dados, visto que existem 6.000 imagens disponíveis e apenas 60 plugues. No caso da segunda abordagem a vantagem é se aproximar mais de um cenário do mundo real, onde o especialista precisa de uma classificação para o plugue inteiro.

4.4 Experimentos

Nesta seção, serão descritos os pré-processamentos realizados nas imagens, como foi feita a divisão dos dados, como os folds da validação cruzada aninhada foram utilizados e, por fim, como os modelos foram treinados.

4.4.1 Pré-processamentos

Devido à heterogeneidade dos dados e dos modelos propostos, alguns pré-processamentos foram testados. A seguir, os pré-processamentos utilizados nas imagens usadas nos modelos de redes neurais profundas serão descritos: a primeira

proposta foi feita para tentar contornar o problema de diferentes condições luminosas das imagens e o segundo pré-processamento é necessário para que seja possível utilizar o modelo convolucional mais simples e o modelo Inception v3 pré-treinado.

1. Normalização pela média e desvio padrão

A normalização de cada imagem pela sua média e desvio de padrão foi a proposta de pré-processamento para contornar as diferentes condições luminosas das imagens. Como as imagens utilizadas possuem apenas um canal, elas podem ser representadas por uma matriz $\mathbf{M}^{l \times c}$, onde l é o número de linhas da matriz que equivale à largura da imagem em número de pixels e c é o número de colunas que equivale ao comprimento da imagem em número de pixels. A formulação utilizada para essa normalização para cada pixel da imagem pode ser vista na Equação 4.4.

$$M'_{i,j} = \frac{M_{i,j} - \mu}{\sigma}, \quad (4.4)$$

onde M' é a matriz que representa a imagem normalizada, M é a matriz da imagem original, μ é média da intensidade dos pixels da matriz da imagem original e σ o desvio padrão dela. $M'_{i,j}$ é o pixel na posição i,j da matriz da imagem normalizada, onde i é a posição da linha da matriz e j da coluna.

Esse processamento se mostra necessário dada a grande variação de luminosidade das imagens. Se o escalonamento fosse feito pelo mínimo e máximo global, as redes neurais profundas poderiam ficar enviesadas pela luminosidade, visto que existem imagens muito escuras e outras muito claras.

2. Redimensionamento

Em razão de alguns dos modelos utilizados precisarem de entradas de tamanho fixo, fez-se necessário o pré-processamento para redimensionar todas as imagens. A proposta para o redimensionamento foi o uso da função da biblioteca OpenCV [109], sendo esta função uma interpolação linear dos pixels para reduzir a imagem para um tamanho especificado. A técnica de redimensionamento sendo aplicada em uma imagem pode ser vista na Figura 4.12, onde é aplicada em uma imagem de 718×604 transformando-a em uma imagem de 256×256 .

As imagens utilizadas neste trabalho foram redimensionadas para 256×256 para serem usadas como entradas nos modelos convolucionais. Essa escolha se deu pelo fato de que a maioria dos modelos consolidados no conjunto de dados do ImageNet utiliza esse padrão. O único modelo que difere desse valor

para redimensionamento é o modelo Inception v3, o qual foi utilizado neste trabalho, para o qual as imagens foram redimensionadas para 299×299 .

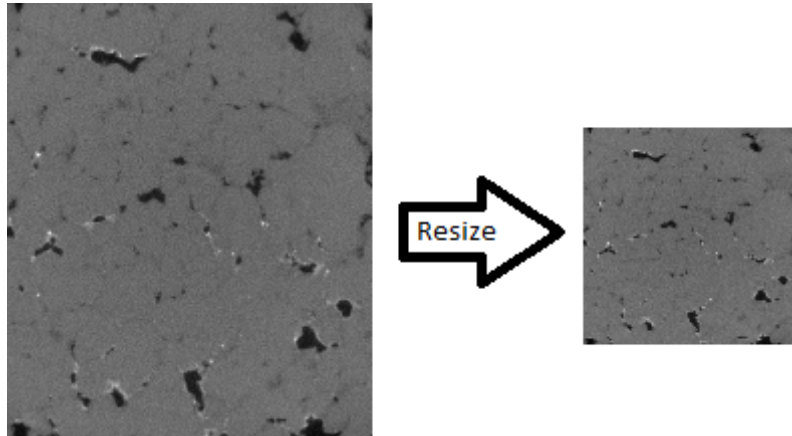


Figura 4.12: Redimensionamento em uma imagem de 718×604 para uma imagem de 256×256

4.4.2 Divisão dos Dados

No conjunto de dados disponível foi aplicado o k-folds aninhado e estratificado, sendo utilizados 6 folds com 10 plugues em cada, onde 3 são da classe esferulítico, 4 da classe estromatólito e 3 da classe grainstone.

A divisão dos folds foi feita no nível dos plugues, visto que as imagens de um mesmo plugue tendem a ser homogêneas, motivo pelo qual a mesma classe é utilizada para todas as imagens de um plugue. Os plugues foram distribuídos aleatoriamente para cada fold, mantendo a distribuição das classes, sendo esses mesmos grupos utilizados para cada um dos modelos. A Figura 4.13 apresenta exemplos de cada classe em cada fold, e, no Apêndice C, é possível verificar um exemplo de cada plugue utilizado em cada um dos folds.

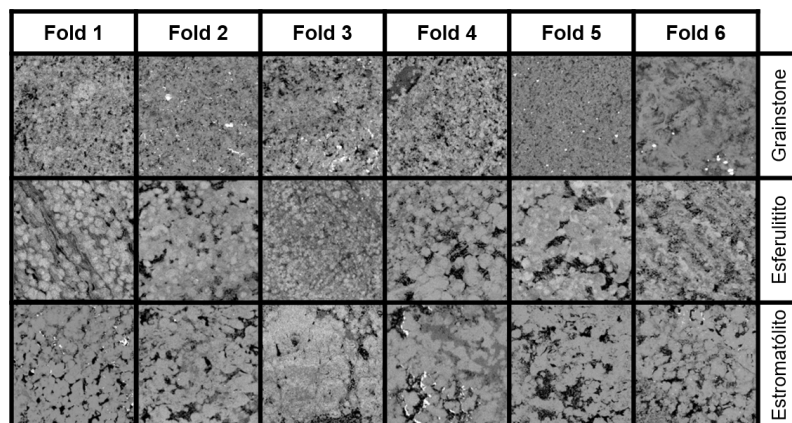


Figura 4.13: Exemplo de uma imagem de cada classe pertencente a cada fold

4.4.3 Treinamento

Todos os modelos usados neste trabalho possuem pesos e bias treináveis, que podem ser representados como um conjunto de matrizes como mostrado a seguir:

$$\theta = \{\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_l, \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_l\}, \quad (4.5)$$

onde l é o número de camadas, \mathbf{W}_i ($i = 1, 2, \dots, l$) denota os pesos de cada camada e \mathbf{b}_i ($i = 1, 2, \dots, l$) seus respectivos bias.

Com o objetivo de atingir a melhor solução para o problema, é necessário atualizar esses pesos, sendo isso realizado através da minimização do erro da função custo entre os valores preditos e reais. A entropia cruzada foi selecionada como função custo dos modelos e sua formulação pode ser representada por:

$$L(y, \hat{y}; \theta) = -\frac{1}{N} \sum_c^C \sum_n^N y_{c,n} \log(\hat{y}_{c,n}(\theta)), \quad (4.6)$$

onde N é o total de amostras, C é o número de classes, $y_{c,n}$ é o valor real para classe c da n -ésima amostra, e $\hat{y}_{c,n}$ é a probabilidade predita para a mesma classe dessa mesma amostra. Os parâmetros da rede são atualizados usando-se o algoritmo de retro-propagação proposto por Lecun [110] e a função custo é otimizada usando-se o algoritmo Adam [111] com taxa de aprendizado de 0,001, $\beta_0 = 0,9$ e $\beta_1 = 0,999$, os quais são indicados como padrão pelo criador do algoritmo. Ao otimizar uma rede neural, é necessário usar um critério de parada do treinamento, esse critério pode ser o número de épocas ou uma parada prematura pelo conjunto de validação. No caso deste trabalho, a acurácia no conjunto de validação foi utilizada para a parada prematura do treinamento.

Ademais, vale ressaltar que, devido ao desbalanceamento das classes no conjunto de dados, utilizou-se a ponderação do gradiente pelo inverso da ocorrência das classes [14], no qual as classes com menor frequência foram multiplicadas por uma constante maior que um, sendo no caso presente utilizado o valor 1,33, de forma a aumentar o valor do gradiente, enquanto a classe de maior frequência teve seu gradiente multiplicado por um. Dessa forma, os gradientes aplicados nos pesos das redes neurais é balanceado, posto que, ao não utilizar uma técnica de balanceamento do gradiente, os modelos treinados ficariam enviesados pela classe com maior ocorrência no conjunto de dados.

A infraestrutura disponível para este trabalho foi uma NVIDIA TITAN V concedida pela NVIDIA. Os tempos de treinamento de cada modelo podem ser vistos na Tabela 4.4. Verifica-se que os modelos treinados com imagens sem redimensionamento levaram 90 segundos por época, isso ocorreu devido a implementação diferenciada para leitura dos dados no processo de treinamento. Já aqueles que usam

imagens redimensionadas levaram 30 segundos por época. Por fim, os perceptrons multicamadas com atributos clássicos levaram 2 segundos por época. Ressalta-se que o modelo Inception v3 levou, na média, menos tempo de treinamento total por fold, isso significa que precisou de menos épocas para treinar.

Tabela 4.4: Tabela com tempos de treinamento para cada modelo utilizado

Tipo de Entrada	Modelo	Tempo por época (s)	Tempo Médio por Fold
Métodos Clássicos	MLP	2s	150
Original	CNN+SPP	90	8178
Redimensionada	CNN+SPP	30	2580
Redimensionada	CNN	30	2275
Redimensionada	Inception	30	1966

Capítulo 5

Resultados e Discussões

Este capítulo tem o propósito de apresentar e discutir os resultados obtidos no conjunto de dados de imagens de micro-tomografia de rochas carbonáticas do pré-sal.

5.1 Resultados Validação Cruzada

Neste trabalho, cinco modelos diferentes foram testados:

1. Perceptron Multicamadas (MLP) o qual utiliza como entrada atributos extraídos das imagens pelos métodos clássicos
2. Modelo convolucional com uma camada SPP (CNN+SPP) o qual utiliza imagens sem redimensionamento como entrada
3. Modelo convolucional com uma camada SPP (CNN+SPP) o qual utiliza imagens com redimensionamento como entrada
4. Modelo convolucional mais simples (CNN) o qual utiliza imagens com redimensionamento como entrada
5. Modelo Inception v3 o qual utiliza imagens com redimensionamento como entrada

Existem duas formas de avaliar os modelos treinados, sendo a primeira a análise do erro de classificação de cada imagem individualmente e a segunda a análise do erro de classificação do plugue como um todo. Esse segundo método designa a classe mais frequente dentre as preditas das imagens de um plugue como a classificação deste. Esse cenário se aproxima mais do caso real, no qual o especialista espera uma informação por plugue e não para cada imagem individual. As tabelas a seguir apresentarão os resultados na forma da média e desvio padrão dos folds da validação

cruzada aninhada, onde podemos considerar o desvio padrão como uma barra de erro do modelo. Nesta seção serão apresentados apenas os resultados obtidos no conjunto de teste, os demais resultados encontram-se no Apêndice D.

Ao final de cada forma de avaliação dois modelos serão escolhidos e comparados utilizando o teste de Wilcoxon. Esse teste foi desenvolvido em 1945 pelo químico e estatístico Frank Wilcoxon e é uma alternativa ao teste-t quando a distribuição da diferença de duas amostras não pode ser assumida como normal. É um teste estatístico de hipótese não-paramétrico usado para verificar se duas amostras provêm da mesma distribuição, sendo válido para dados com quaisquer distribuição e menos sensível a *outliers* que o teste-t [112]. A significância do teste é determinada computando-se o valor p, usualmente quando esse valor é menor que 0,05 considera-se a hipótese nula como rejeitada em favor da hipótese alternativa com confiança de 95%.

5.1.1 Resultados por Imagem

A Tabela 5.1 apresenta a acurácia e o índice SP de cada modelo, avaliando cada imagem individualmente. A Figura 5.1 apresenta o diagrama de caixa das acurácias, onde o modelo marcado com asterisco foi treinado com imagens originais. A diferença de desempenho entre os modelos CNN+SPP utilizando imagens redimensionadas e originais pode ter ocorrido pelo fato de que para viabilizar o treinamento do modelo com imagens originais foi necessário fazer batches com imagens de mesmo tamanho. Como a maioria dos plugues não possuía imagens de tamanhos iguais, quase todos os batches possuíam imagens de um mesmo plugue, podendo ter influenciado de forma negativa no treinamento do modelo.

Na média, os dois melhores modelos foram o Inception v3 e o CNN+SPP com imagens redimensionadas, os quais superaram uma média de 75% nos dois indicadores apresentados, enquanto os outros não conseguiram superar 64%.

O pior modelo foi o que utilizou os atributos clássicos como entrada, o qual ficou muito abaixo dos outros, podendo isso ser justificado pela escolha dos atributos. Talvez, com outros atributos, fosse possível obter um desempenho melhor desse modelo. Contudo, a técnica de aprendizado profundo ganhou notoriedade, visto que seus modelos são capazes de extrair as características necessárias para discriminação das classes sozinhos, apenas utilizando a imagem como entrada.

Nas Tabelas 5.2, 5.3 e 5.4 encontram-se os recalls de cada classe. As Figuras 5.2, 5.3 e 5.4 apresentam os diagramas de caixa dos recalls de cada classe. Verifica-se que os dois modelos, que obtiveram maiores médias na acurácia e no índice SP, obtiveram uma média maior que 70% nos resultados de todas as classes e, na classe Grainstone, chegaram a obter uma média maior que 80%. Ainda analisando a média desses dois modelos, vemos que o modelo Inception v3 supera o CNN+SPP com

imagens redimensionadas em todas as classes, contudo o desvio padrão do Inception v3 é maior na maioria dos casos.

Tabela 5.1: Tabela com acurácia e índice sp dos modelos ao avaliar por imagem individual no conjunto de teste

Tipo de Entrada	Modelo	Acurácia (%)	Índice SP (%)
Métodos Clássicos	MLP	$54,4 \pm 10,1$	$50,0 \pm 11,8$
Imagem Original	CNN+SPP	$64,0 \pm 10,8$	$62,5 \pm 11,7$
Imagem Redimensionada	CNN+SPP	$75,5 \pm 08,7$	$75,1 \pm 08,7$
Imagem Redimensionada	CNN	$61,7 \pm 07,3$	$59,2 \pm 07,5$
Imagem Redimensionada	Inception	$81,0 \pm 11,9$	$79,9 \pm 11,9$

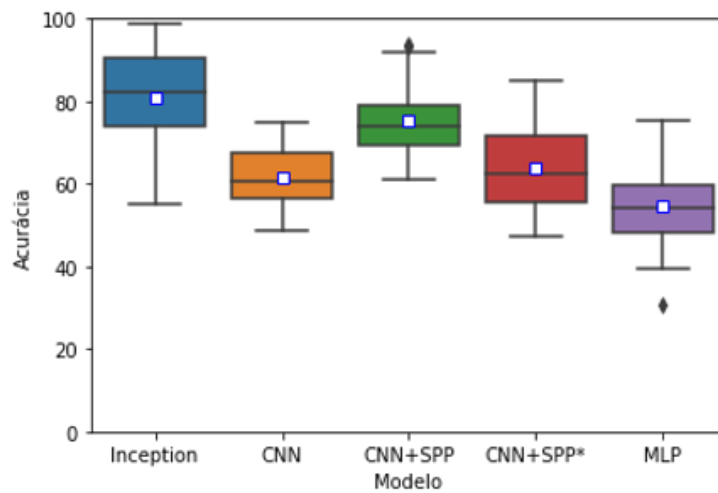


Figura 5.1: Diagrama de caixa das acurácias ao avaliar por imagem individual no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul

Tabela 5.2: Tabela com recall avaliando por imagem da classe Esferulitito no conjunto de teste

Tipo de Entrada	Modelo	Recall - Esferulitito (%)
Métodos Clássicos	MLP	65 ± 26
Original	CNN+SPP	59 ± 21
Redimensionada	CNN+SPP	72 ± 14
Redimensionada	CNN	42 ± 12
Redimensionada	Inception	73 ± 21

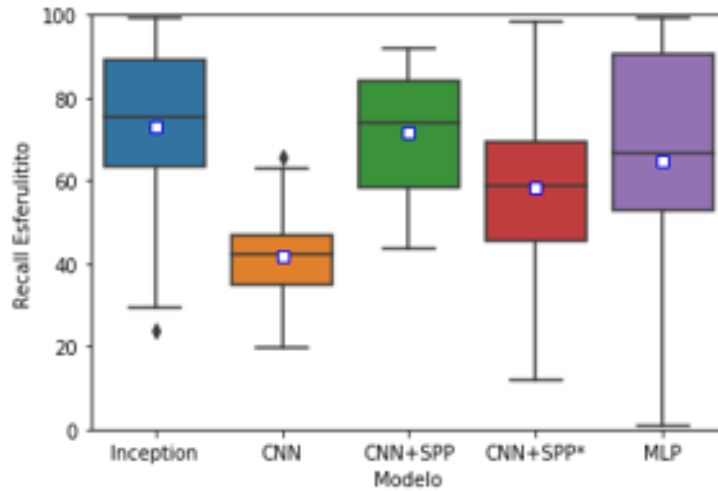


Figura 5.2: Diagrama de caixa do recall da classe Esferulítico ao avaliar por imagem individual no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul

Tabela 5.3: Tabela com recall avaliando por imagem da classe Estromatólito no conjunto de teste

Tipo de Entrada	Modelo	Recall - Estromatólito (%)
Métodos Clássicos	MLP	57 ± 19
Original	CNN+SPP	68 ± 14
Redimensionada	CNN+SPP	73 ± 16
Redimensionada	CNN	71 ± 16
Redimensionada	Inception	82 ± 20

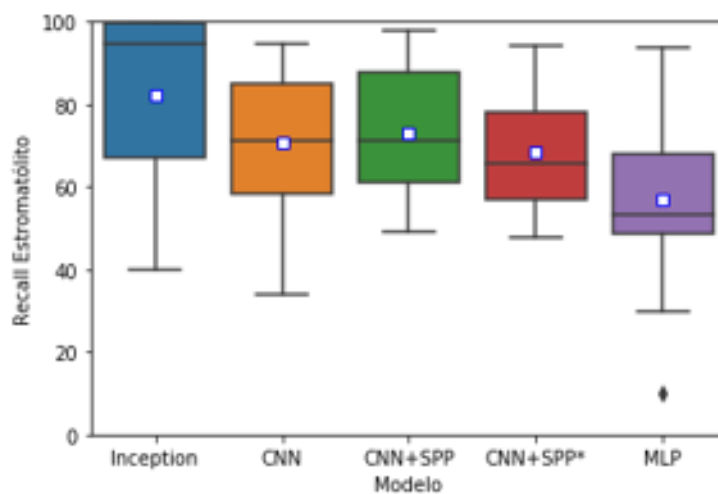


Figura 5.3: Diagrama de caixa do recall da classe Estromatólito ao avaliar por imagem individual no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul

Tabela 5.4: Tabela com recall avaliando por imagem da classe Grainstone no conjunto de teste

Tipo de Entrada	Modelo	Recall - Grainstone (%)
Métodos Clássicos	MLP	41 ± 27
Original	CNN+SPP	64 ± 17
Redimensionada	CNN+SPP	83 ± 18
Redimensionada	CNN	70 ± 14
Redimensionada	Inception	87 ± 13

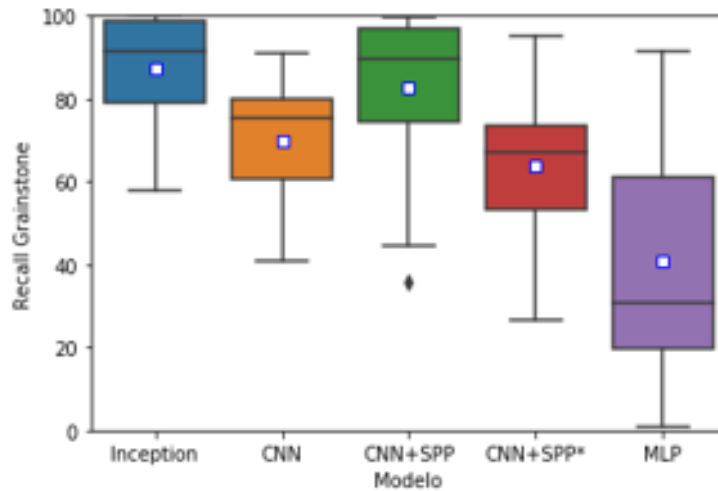


Figura 5.4: Diagrama de caixa do recall da classe Grainstone ao avaliar por imagem individual no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul

Teste Wilcoxon

Comparando os resultados pareados dos folds do modelo CNN+SPP com o Inception V3, obtemos o resultado de $p = 0,024$ ao avaliarmos os resultados do índice SP e $p = 0,009$ nos resultados de acurácia, dando confiança suficiente para negar-se a hipótese de que as amostras provêm da mesma distribuição em ambos os casos. Com isso, o modelo Inception v3 foi escolhido como o principal modelo na avaliação por imagens individuais, pois ele possui uma maior média nos resultados obtidos, tanto na acurácia quanto no índice SP.

5.1.2 Resultados por Plugue

A Tabela 5.5 apresenta a acurácia e o índice SP de cada modelo avaliando a classe mais frequente de cada plugue. A Figura 5.5 apresenta o diagrama de caixa das acurácias, onde o modelo marcado com asterisco foi treinado com imagens originais.

Nota-se que, na média, os dois melhores modelos foram, novamente, o Inception v3 e o CNN+SPP com imagens redimensionadas, ambos modelos superando o valor de 80% nos dois indicadores usados, ressaltando-se que os resultados obtidos por esses modelos está muito mais próximo nesse tipo de avaliação. Verifica-se também que a média das funções de avaliação aumentaram em quase todos os casos, significando que os modelos acertaram a maioria das imagens de cada plugue. Contudo, o desvio padrão dos resultados também aumentou, o que pode ser justificado pelo fato de que cada plugue representa 10% das amostras no conjunto de teste, de forma que o erro de um plugue afeta o resultado final drasticamente.

Ademais, isso é ainda mais acentuado quando se analisam as Tabelas 5.6, 5.7 e 5.8, visto que no caso das classes Esferulítico e Grainstone cada plugue representa 33% das amostras e na classe Estromatólito 25%, fazendo com que um plugue classificado errado afete de forma significativa o desempenho, aumentando o desvio padrão dos resultados. Vale ressaltar que a média dos recalls aumentou na maioria dos casos, como indicado pelo aumento da acurácia e índice SP. As Figuras 5.6, 5.7 e 5.8 apresentam os diagramas de caixa dos recalls de cada classe.

Tabela 5.5: Tabela com acurácia e índice SP dos modelos ao avaliar por plugue no conjunto de teste

Tipo de Entrada	Modelo	Acurácia (%)	Índice SP (%)
Métodos Clássicos	MLP	57 ± 13	45 ± 27
Imagem Original	CNN+SPP	67 ± 14	62 ± 21
Imagem Redimensionada	CNN+SPP	81 ± 11	81 ± 12
Imagem Redimensionada	CNN	70 ± 15	57 ± 31
Imagem Redimensionada	Inception	84 ± 13	83 ± 14

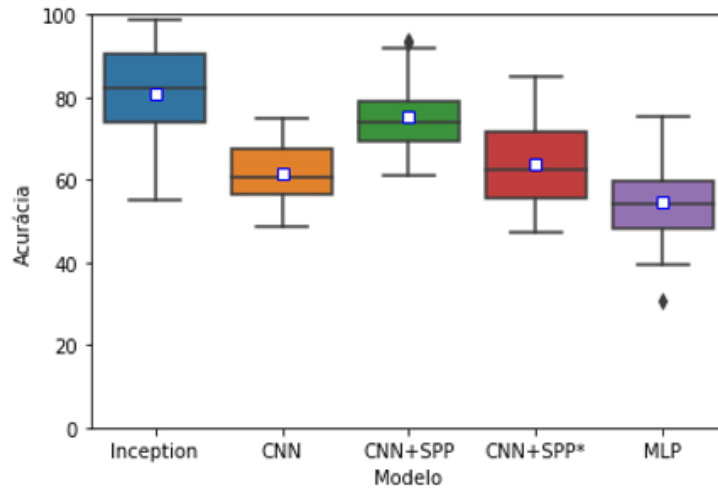


Figura 5.5: Diagrama de caixa das acurácias ao avaliar por plugue no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul

Tabela 5.6: Tabela com recall avaliando por plugue da classe Esferulitito no conjunto de teste

Tipo de Entrada	Modelo	Recall - Esferulitito (%)
Métodos Clássicos	MLP	68 ± 27
Original	CNN+SPP	63 ± 32
Redimensionada	CNN+SPP	81 ± 21
Redimensionada	CNN	38 ± 26
Redimensionada	Inception	77 ± 25

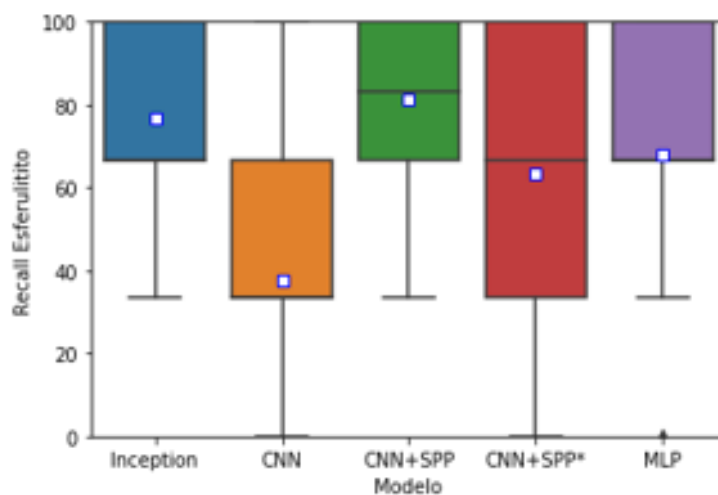


Figura 5.6: Diagrama de caixa do recall da classe Esferulitito ao avaliar por plugue no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul

Tabela 5.7: Tabela com recall avaliando por plugue da classe Estromatólito no conjunto de teste

Tipo de Entrada	Modelo	Recall - Estromatólito (%)
Métodos Clássicos	MLP	59 ± 24
Original	CNN+SPP	69 ± 19
Redimensionada	CNN+SPP	77 ± 20
Redimensionada	CNN	85 ± 20
Redimensionada	Inception	84 ± 21

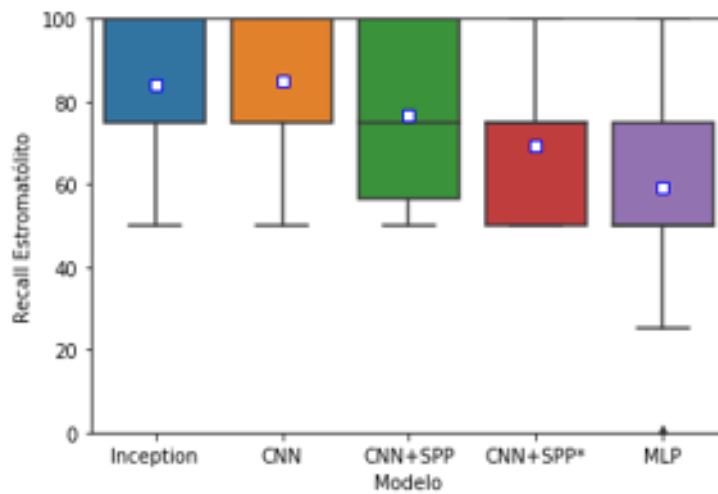


Figura 5.7: Diagrama de caixa do recall da classe Estromatólito ao avaliar por plugue no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul

Tabela 5.8: Tabela com recall avaliando por plugue da classe Grainstone no conjunto de teste

Tipo de Entrada	Modelo	Recall - Grainstone (%)
Métodos Clássicos	MLP	44 ± 32
Original	CNN+SPP	69 ± 26
Redimensionada	CNN+SPP	88 ± 22
Redimensionada	CNN	83 ± 19
Redimensionada	Inception	92 ± 14

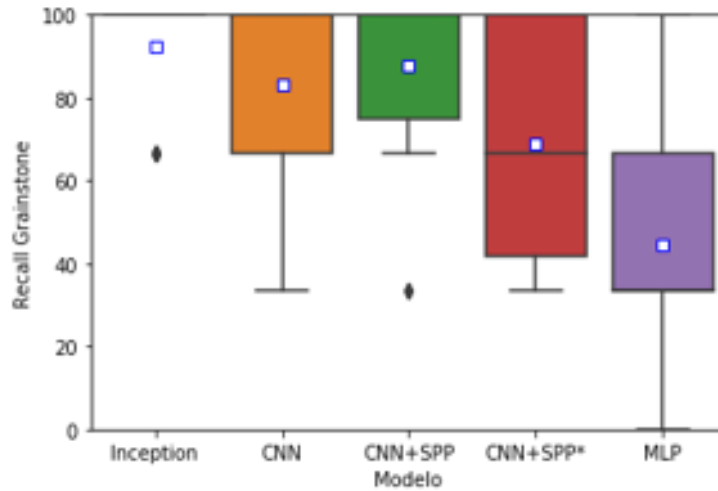


Figura 5.8: Diagrama de caixa do recall da classe Grainstone ao avaliar por plugue no conjunto de teste, onde a média está marcada como um quadrado branco com borda azul

Teste Wilcoxon

No caso da avaliação por plugues, nos resultados pareados dos modelos Inception v3 e CNN+SPP com imagens redimensionados, o teste de Wilcoxon obteve $p = 0,512$ no índice SP e $p = 0,277$ na acurácia, que são valores muito superiores aos obtidos na avaliação por imagem. No caso da acurácia, temos uma confiança de quase 73% para negar a hipótese, contudo, no caso do índice SP, não podemos negar a hipótese de que essas amostras vêm da mesma distribuição, o que dificulta a escolha de um único modelo para esse tipo de avaliação.

5.2 Comparação dos Modelos

Conforme citado na seção anterior, os modelos escolhidos em ambas formas de avaliação foram o modelo Inception v3 e o CNN+SPP com imagens redimensionadas. Para facilitar a comparação entre eles, foi escolhida a rodada da validação cruzada aninhada de melhor desempenho no índice SP entre as trinta rodadas treinadas. Essa escolha foi feita através da comparação dos resultados no índice SP ao aplicarmos o modelo já treinado de cada rodada em todo conjunto de dados e calculando um único índice SP para todo o conjunto de dados.

As matrizes de confusão no conjunto de teste das rodadas escolhidas de cada modelo estão apresentadas a seguir, de forma que os resultados estão divididos em duas seções, uma para cada forma de avaliação. No caso do modelo Inception v3, a melhor rodada foi a de número 23 (fold 5 como fold de validação e fold 4 como fold de teste) e a melhor rodada do modelo CNN+SPP, foi a de número 24 (fold 6 como fold

de validação e fold 4 como fold de teste). Com o objetivo de deixar a comparação mais justa dos modelos, os resultados das duas rodadas serão apresentados para os dois modelos.

5.2.1 Matrizes de Confusão por Imagem

Os resultados apresentados nessa subseção são as matrizes de confusão do conjunto de teste ao avaliarmos as imagens individualmente. A primeira rodada a ser analisada será a de número 24, a qual foi a melhor rodada do modelo CNN+SPP com imagens redimensionadas. As matrizes de confusão do modelo Inception v3 e CNN+SPP podem ser vistas nas Tabelas 5.9 e 5.10. Verifica-se, pelas matrizes de confusão desta rodada, que o modelo Inception v3 atingiu 97,5% na acurácia e 97,3% no índice SP, já o modelo CNN+SPP atingiu aproximadamente 92,0% e 91,6%. Ambos modelos obtiveram bons resultados nessa rodada, superando 90% nas duas medidas. Nota-se também que o modelo Inception v3 superou o modelo CNN+SPP nesta rodada.

As matrizes de confusão da rodada de número 23 (melhores resultados do Inception v3) se encontram nas Tabelas 5.11 e 5.12. Os resultados das funções de avaliação do modelo Inception v3 nesse caso foram 96,1% na acurácia e 95,6% no índice SP, já o modelo CNN+SPP obteve 93,8% e 93,2%. Como no caso de número 24, ambos modelos superaram 90% nas duas medidas, mas o modelo CNN+SPP foi superado pelo modelo Inception v3 nesta rodada.

Verifica-se que a classe mais errada pelo modelo Inception v3 nessas duas rodadas foi a esferulítico e o modelo CNN+SPP também na rodada 23. Contudo, na rodada 24, a classe grainstone foi a mais errada. A classe estromatólito foi a mais acertada em quase todos os casos, menos na rodada 23 pelo modelo CNN+SPP. Ressalta-se que a classe esferulítico foi confundida com todas as outras e, na maioria das vezes, foi confundida com grainstone. Na rodada 23, a classe estromatólito é somente confundida como esferulítico, destacando-se também que os erros das classes não superaram 6%.

Tabela 5.9: Matriz de confusão por imagem do modelo Inception v3 na rodada de número 24 no conjunto de teste

	Esferulítico (Predito)	Estromatólito (Predito)	Grainstone (Predita)
Esferulítico (Real)	95%	2%	3%
Estromatólito (Real)	1%	99%	0%
Grainstone (Real)	0%	2%	98%

Tabela 5.10: Matriz de confusão por imagem do modelo CNN+SPP na rodada de número 24 no conjunto de teste

	Esferulitito (Predito)	Estromatólito (Predito)	Grainstone (Predita)
Esferulitito (Real)	92%	1%	7%
Estromatólito (Real)	5%	95%	0%
Grainstone (Real)	1%	11%	88%

Tabela 5.11: Matriz de confusão por imagem do modelo Inception v3 na rodada de número 23 no conjunto de teste

	Esferulitito (Predito)	Estromatólito (Predito)	Grainstone (Predita)
Esferulitito (Real)	91%	5%	4%
Estromatólito (Real)	0%	100%	0%
Grainstone (Real)	4%	0%	96%

Tabela 5.12: Matriz de confusão por imagem do modelo CNN+SPP na rodada de número 23 no conjunto de teste

	Esferulitito (Predito)	Estromatólito (Predito)	Grainstone (Predita)
Esferulitito (Real)	85%	3%	12%
Estromatólito (Real)	2%	98%	0%
Grainstone (Real)	0%	3%	97%

5.2.2 Matrizes de Confusão por Plugue

Essa subseção apresenta os resultados das matrizes de confusão do conjunto de teste ao avaliarmos a classificação de cada plugue. As Tabelas 5.13 e 5.14 apresentam os resultados da rodada de número 24 dos modelos Inception v3 e CNN+SPP, respectivamente, verificando-se que ambos modelos obtiveram 100% de acurácia nessa abordagem de avaliação para essa rodada. Na rodada de número 23, ambos modelos também obtiveram 100% de acurácia nessa forma de avaliação.

Tabela 5.13: Matriz de confusão por plugue do modelo Inception v3 na rodada de número 24 no conjunto de teste

	Esferulitito (Predito)	Estromatólito (Predito)	Grainstone (Predita)
Esferulitito (Real)	100%	0%	0%
Estromatólito (Real)	0%	100%	0%
Grainstone (Real)	0%	0%	100%

Tabela 5.14: Matriz de confusão por plugue do modelo CNN+SPP na rodada de número 24 no conjunto de teste

	Esferulitito (Predito)	Estromatólito (Predito)	Grainstone (Predita)
Esferulitito (Real)	100%	0%	0%
Estromatólito (Real)	0%	100%	0%
Grainstone (Real)	0%	0%	100%

Tabela 5.15: Matriz de confusão por plugue do modelo Inception v3 na rodada de número 23 no conjunto de teste

	Esferulitito (Predito)	Estromatólito (Predito)	Grainstone (Predita)
Esferulitito (Real)	100%	0%	0%
Estromatólito (Real)	0%	100%	0%
Grainstone (Real)	0%	0%	100%

Tabela 5.16: Matriz de confusão por plugue do modelo CNN+SPP na rodada de número 23 no conjunto de teste

	Esferulitito (Predito)	Estromatólito (Predito)	Grainstone (Predita)
Esferulitito (Real)	100%	0%	0%
Estromatólito (Real)	0%	100%	0%
Grainstone (Real)	0%	0%	100%

5.2.3 Análise de quadrante

Nesta subseção, será feita a análise de concordância e de discordância dos modelos. No caso de concordância dos modelos, existem duas possibilidades, ou ambos os modelos acertam ou erram. Já nos casos de discordância entre os modelos, existem três possíveis casos, o modelo Inception v3 acerta e o CNN+SPP erra, CNN+SPP acerta e o Inception v3 erra ou ambos erram predizendo classes diferentes. Sendo assim, as tabelas a seguir apresentarão a quantidade de amostras de cada um desses cenários para cada rodada escolhida (23 e 24). Devido à taxa de acerto 100% na avaliação por plugue no conjunto de teste, como mostrado na seção anterior, apenas os resultados da avaliação por imagens dos modelos serão comparados.

Começando pela rodada de número 24, a Tabela 5.17 apresenta a quantidade de amostras onde os modelos concordaram e discordaram. Verifica-se que os modelos acertaram juntos 90,2% das 1.000 imagens de teste, o que representa 99% dos casos que os modelos concordaram. Ambos modelos erraram juntos apenas 0,5% das 1.000 imagens de teste. Nas amostras que os modelos discordaram, constata-se que o modelo Inception v3 acertou mais que o modelo CNN+SPP, mas tal fato já era esperado dadas as estatísticas apresentadas anteriormente. Nota-se que não ocorreram casos onde os modelos erraram a classe verdadeira e discordaram.

A Tabela 5.18 apresenta a comparação das predições na rodada de número 23. Nota-se que os modelos acertaram quase a mesma quantidade de amostras que na rodada de número 24. Os casos em que eles erraram juntos dobrou, mas continua sendo um valor pequeno que representa apenas 1% do conjunto de teste. Já quando os modelos discordam entre si, nota-se que o modelo Inception menos amostras que na rodada 24, sendo importante ressaltar que novamente não ocorreram casos onde os modelos erraram a classe verdadeira e discordaram.

Tabela 5.17: Concordância ou discordância dos modelos Inception v3 e CNN+SPP na rodada de número 24 no conjunto de teste

concordaram	acertaram	902
	erraram	5
discordaram	Inception acertou	74
	CNN+SPP acertou	19
	ambos erraram	0

Tabela 5.18: Concordância ou discordância dos modelos Inception v3 e CNN+SPP na rodada de número 23 no conjunto de teste

concordaram	acertaram	904
	erraram	10
discordaram	Inception acertou	55
	CNN+SPP acertou	31
	ambos erraram	0

Capítulo 6

Conclusões

Neste trabalho, foi feita a aplicação do aprendizado profundo, transferência de aprendizado, e de métodos clássicos de extração de características em imagens para identificação de padrões em imagens de microtomografia de rochas carbonáticas do pré-sal visando a classificação litológica. Ao conhecimento do autor, este foi o primeiro trabalho que propôs o uso dessas técnicas para classificação litológica neste tipo de rochas e imagens. A produção científica proveniente da presente dissertação encontra-se no Apêndice E.

Os resultados apresentados compararam o desempenho de cinco modelos de redes neurais de duas formas diferentes: a primeira forma de comparação foi a análise dos resultados obtidos avaliando-se cada imagem individualmente; neste caso, o modelo Inception v3 superou, na média, todos os outros, onde atingiu acurácias com médias próximas a 80%. A segunda forma de avaliação foi a classificação do plugue como um todo, nesse caso, o desempenho médio dos modelos aumentou, o que significa que os modelos foram capazes de acertar a maior parte das imagens na maioria dos plugues. Contudo, o desvio padrão desses resultados também aumentou devido a redução na quantidade de amostras, de modo que uma amostra representa 10% de acurácia no conjunto de teste. Frisa-se que os modelos Inception v3 e CNN+SPP em imagens redimensionadas, na média, superaram todos os outros nessa forma de avaliação, atingindo desempenhos que superaram médias de 80% e nesse tipo de avaliação seus resultados ficaram mais próximos.

Infere-se que o uso de imagens sem redimensionamento não gera uma melhoria de desempenho no conjunto de dados utilizado, visto que além de aumentar em três vezes o tempo de treinamento, o modelo treinado com essas imagens, na média, não superou os resultados do mesmo modelo com imagens redimensionadas. Constatou-se também que a técnica de transferência de aprendizado demonstrou resultados promissores no conjunto de dados disponível, sendo, nesse caso, a escolha do autor para estudos mais aprofundados com uma maior quantidade de dados.

Assim, conclui-se que este trabalho foi capaz de atingir o objetivo proposto, visto

que modelos de aprendizado profundo foram treinados e avaliados no conjunto de imagens de plugues disponibilizado. Este trabalho serve como primeiro passo para a aplicação de aprendizado profundo em imagens de microtomografia de plugues de rochas carbonáticas do pré-sal e a automatização do processo de caracterização dessas imagens.

6.1 Trabalhos Futuros

Existem diversas maneiras de evoluir o trabalho corrente. A principal sugestão seria aumentar o conjunto de dados utilizado, não só no número de amostras, mas na diversidade das classes disponíveis. É possível também usar a técnica de *data augmentation* [113] para aumentar o conjunto de dados, usando tanto modelos generativos como GAN [114], quanto através de técnicas mais simples como cortes, rotações e acréscimo de ruídos na imagens.

Ademais, uma segunda sugestão para evolução do trabalho seria modificar o número de camadas das redes neurais propostas, ou, ainda, utilizar diferentes estruturas de redes convolucionais. Potenciais melhorias de desempenho podem vir ao testarmos outros modelos usando a técnica de transferência de aprendizado. Alguns modelos apresentam desempenho superior ao Inception v3 no conjunto de imagens do ImageNet e estão sendo disponibilizados na internet de forma gratuita.

Por fim, saindo da perspectiva de classificação de imagens, uma outra possível evolução seria a segmentação das texturas dentro de cada uma das imagens; entretanto, trata-se de um problema mais complexo e que necessitaria de dados com anotações pixel a pixel, o que pode ser um empecilho, visto que a tarefa de anotar pixel a pixel é muito trabalhosa sem o uso de ferramentas adequadas.

Referências Bibliográficas

- [1] “EPE - Matriz Energética e Elétrica”. <http://www.epe.gov.br/pt/abcdenergia/matriz-energetica-e-eletrica>, . Acessado em 16-02-2020.
- [2] “British Petroleum - Statistical review of world energy”. <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html>, . Acessado em 16-02-2020.
- [3] “Potencial dos Recursos Energéticos no Horizonte 2050”. http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-227/topico-416/NT04%20PR_RecursosEnergeticos%202050.pdf, . Acessado em 16-02-2020.
- [4] “Petrobrás - Pré-sal”. <http://www.petrobras.com.br/pt/nossas-atividades/areas-de-atuacao/exploracao-e-producao-de-petroleo-e-gas/pre-sal/>, . Acessado em 16-02-2020.
- [5] ARNS, C. H., KNACKSTEDT, M. A., PINCZEWSKI, W. V., et al. “Computation of linear elastic properties from microtomographic images: Methodology and agreement between theory and experiment”, *Geophysics*, v. 67, n. 5, pp. 1396–1405, 2002.
- [6] FLÜGEL, E. *Microfacies of carbonate rocks: analysis, interpretation and application*. Springer Science & Business Media, 2013.
- [7] ARNS, C. H., BAUGET, F., LIMAYE, A., et al. “Pore scale characterization of carbonates using X-ray microtomography”, *Spe Journal*, v. 10, n. 04, pp. 475–484, 2005.
- [8] CLAES, S., SOETE, J., CNUDDÉ, V., et al. “A three-dimensional classification for mathematical pore shape description in complex carbonate reservoir rocks”, *Mathematical Geosciences*, v. 48, n. 6, pp. 619–639, 2016.

- [9] KNACKSTEDT, M. A., ARNS, C. H., LIMAYE, A., et al. “Digital Core Laboratory: Properties of reservoir core derived from 3D images”. In: *SPE Asia Pacific Conference on Integrated Modelling for Asset Management*, p. 14. Society of Petroleum Engineers, 2004.
- [10] DERNAIKA, M., UDDIN, Y. N., KORONFOL, S., et al. “Multi-scale rock analysis for improved characterization of complex carbonates”. In: *SPE Reservoir Characterisation and Simulation Conference and Exhibition*. Society of Petroleum Engineers, 2015.
- [11] HUNTER, S., HOFMANN, R., ESPEJO, I. “Integrating grain-scale geology in digital rock physics”, *The Leading Edge*, v. 37, n. 6, pp. 428–434, 2018.
- [12] ATHANASIOU, L. S., FOTIADIS, D. I., MICHALIS, L. K. *Atherosclerotic Plaque Characterization Methods Based on Coronary Imaging*. Academic Press, 2017.
- [13] TERRA, G., RICARDO SPADINI, A., BARROS FRANÇA, A., et al. “Classificação de rochas carbonáticas aplicável às bacias sedimentares brasileiras”, *Boletim de Geociências - Petrobras*, v. 18, pp. 9–29, 01 2010.
- [14] GOODFELLOW, I., BENGIO, Y., COURVILLE, A. *Deep Learning*. MIT Press, 2016.
- [15] DERNAIKA, M., MANSOUR, B., GONZALEZ, D., et al. “Upscaled Permeability and Rock Types in a Heterogeneous Carbonate Core from the Middle East”. 05 2017. doi: 10.2118/185991-MS.
- [16] GUO, Y., LIU, Y., OERLEMANS, A., et al. “Deep learning for visual understanding: A review”, *Neurocomputing*, v. 187, pp. 27–48, 2016.
- [17] RAWAT, W., WANG, Z. “Deep convolutional neural networks for image classification: A comprehensive review”, *Neural computation*, v. 29, n. 9, pp. 2352–2449, 2017.
- [18] ZHOU, T., RUAN, S., CANU, S. “A review: Deep learning for medical image segmentation using multi-modality fusion”, *Array*, p. 100004, 2019.
- [19] SZEGEDY, C., VANHOUCKE, V., IOFFE, S., et al. “Rethinking the Inception Architecture for Computer Vision”. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pp. 2818–2826, 2016. doi: 10.1109/CVPR.2016.308. Disponível em: <<https://doi.org/10.1109/CVPR.2016.308>>. Acessado em 16-02-2020.

- [20] DENG, J., DONG, W., SOCHER, R., et al. “ImageNet: A Large-Scale Hierarchical Image Database”. In: *CVPR09*, 2009. Acessado em 16-02-2020.
- [21] NGUYEN, L. D., LIN, D., LIN, Z., et al. “Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation”. In: *2018 IEEE International Symposium on Circuits and Systems (IS-CAS)*, pp. 1–5. IEEE, 2018.
- [22] CUI, Y., SONG, Y., SUN, C., et al. “Large scale fine-grained categorization and domain-specific transfer learning”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4109–4118, 2018.
- [23] XIA, X., XU, C., NAN, B. “Inception-v3 for flower classification”. In: *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, pp. 783–787. IEEE, 2017.
- [24] RISHITA, M. V. S., HARRIS, T. A. “Dog Breed Classifier using Convolutional Neural Networks”. In: *2018 International Conference on Networking, Embedded and Wireless Systems (ICNEWS)*, pp. 1–7. IEEE, 2018.
- [25] HAYKIN, S. *Neural Networks and Learning Machines, 3/E*. Pearson Education India, 2010.
- [26] VALENTÍN, M. B., BOM, C. R., COMPAN, A. L. M., et al. “Estimation of permeability and effective porosity logs using deep autoencoders in borehole image logs from the brazilian pre-salt carbonate”, *Journal of Petroleum Science and Engineering*, v. 170, pp. 315–330, 2018.
- [27] CRAFTON, J. W., OTHERS. “Oil and Gas Well Evaluation Using the Reciprocal Productivity Index Method”. In: *SPE Production Operations Symposium*. Society of Petroleum Engineers, 1997.
- [28] BURYAKOVSKY, L., CHILINGAR, G. V., SHIN, S., et al. *Fundamentals of the petrophysics of oil and gas reservoirs*. John Wiley & Sons, 2012.
- [29] VALENTÍN, M. B., BOM, C. R., COELHO, J. M., et al. “A deep residual convolutional neural network for automatic lithological facies identification in Brazilian pre-salt oilfield wellbore image logs”, *Journal of Petroleum Science and Engineering*, v. 179, pp. 474–503, 2019.
- [30] AHR, W. M. *Geology of carbonate reservoirs: the identification, description, and characterization of hydrocarbon reservoirs in carbonate Rocks*. John Wiley & Sons, 2008.

- [31] REMEYSEN, K., SWENNEN, R. “Application of microfocus computed tomography in carbonate reservoir characterization: possibilities and limitations”, *Marine and Petroleum Geology*, v. 25, n. 6, pp. 486–499, 2008.
- [32] TSCHANNEN, V., DELESCLUSE, M., RODRIGUEZ, M., et al. “Facies classification from well logs using an inception convolutional network”, *arXiv preprint arXiv:1706.00613*, 2017.
- [33] VINEGAR, H. J. “X-ray CT and NMR imaging of rocks”, *Journal of Petroleum Technology*, v. 38, pp. 257–259, 1986.
- [34] HONARPOUR, M. M., CROMWELL, V., HATTON, D., et al. “Reservoir Rock Descriptions Using Computed Tomography (CT)”, *60th Annual Technical Conference and Exhibition*, 1985.
- [35] WILDENSCHILD, D., SHEPPARD, A. P. “X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems”, *Advances in Water Resources*, v. 51, pp. 217–246, 2013.
- [36] WANG, Y., ARNS, C. H., RAHMAN, S. S., et al. “Porous structure reconstruction using convolutional neural networks”, *Mathematical Geosciences*, v. 50, n. 7, pp. 781–799, 2018.
- [37] O’SULLIVAN, J. D., BEHNSSEN, J., STARBORG, T., et al. “X-ray micro-computed tomography (μ CT): an emerging opportunity in parasite imaging”, *Parasitology*, v. 145, n. 7, pp. 848–854, 2018.
- [38] DULIU, O. G. “Computer axial tomography in geosciences: an overview”, *Earth-science reviews*, v. 48, n. 4, pp. 265–281, 1999.
- [39] CNUDDÉ, V., MASSCHAELE, B., DIERICK, M., et al. “Recent progress in X-ray CT as a geosciences tool”, *Applied Geochemistry*, v. 25, n. 5, pp. 826–832, 2006.
- [40] DVORKIN, J., DERZHI, N., DIAZ, E., et al. “Relevance of computational rock physics”, *Geophysics*, v. 76, n. 5, pp. E141–E153, 2011.
- [41] DVORKIN, J., DERZHI, N., FANG, Q., et al. “From micro to reservoir scale: Permeability from digital experiments”, *The Leading Edge*, v. 28, n. 12, pp. 1446–1452, 2009.
- [42] MULJADI, B. P., BLUNT, M. J., RAEINI, A. Q., et al. “The impact of porous media heterogeneity on non-Darcy flow behaviour from pore-scale simulation”, *Advances in Water Resources*, v. 95, pp. 329–340, 2016.

- [43] KNACKSTEDT, M. A., LATHAM, S., MADADI, M., et al. “Digital rock physics: 3D imaging of core material and correlations to acoustic and flow properties”, *The Leading Edge*, v. 28, n. 1, pp. 28–33, 2009.
- [44] KETCHAM, R. A., CARLSON, W. D. “Acquisition, optimization and interpretation of X-ray computed tomographic imagery: applications to the geosciences”, *Computers & Geosciences*, v. 27, n. 4, pp. 381–400, 2001.
- [45] BENOIT-CATTIN, H. *Texture analysis for magnetic resonance imaging*. Texture Analysis Magn Resona, 2006.
- [46] ERONDU, O. F. *Medical Imaging*. BoD–Books on Demand, 2011.
- [47] YANG, T., SHI, G., YANG, N., et al. “Combining multiple feature for texture image classification”. In: *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pp. 1–5, Oct 2017. doi: 10.1109/CISP-BMEI.2017.8301967.
- [48] AL-KADI, O. S. “A Gabor Filter Texture Analysis Approach for Histopathological Brain Tumor Subtype Discrimination”, *CoRR*, v. abs/1704.05122, 2017. Disponível em: <<http://arxiv.org/abs/1704.05122>>. Acessado em 16-02-2020.
- [49] E R. SUKANESH, A. P. “Texture Feature Based Analysis of Segmenting Soft Tissues from Brain CT Images using BAM type Artificial Neural Network”, *Journal of Information Engineering and Applications*, v. 1, pp. 34–43, 05 2011.
- [50] VLADIMIR, V. N., VAPNIK, V. *The nature of statistical learning theory*. Springer Heidelberg, 1995.
- [51] KOHONEN, T. *Self-organization and associative memory*, v. 8. Springer Science & Business Media, 2012.
- [52] JAIN, A. K. “Data clustering: 50 years beyond K-means”, *Pattern recognition letters*, v. 31, n. 8, pp. 651–666, 2010.
- [53] REYNOLDS, D. “Gaussian mixture models”, *Encyclopedia of biometrics*, pp. 827–832, 2015.
- [54] JULIEN, O., PAULHAC, L. “3D Medical Image Segmentation: Interactive Texture-Based Approaches”, 2011.
- [55] RICHARDS, W., POLIT, A. “Texture matching”, *Kybernetik*, v. 16, n. 3, pp. 155–162, 1974.

- [56] HARALICK, R. M., SHANMUGAM, K., OTHERS. “Textural features for image classification”, *IEEE Transactions on systems, man, and cybernetics*, , n. 6, pp. 610–621, 1973.
- [57] B. MANDELBROT, B. “The Fractal Geometry of Nature”. v. 51, p. 468 p., 03 1983. ISBN: 0716711869. doi: 10.1119/1.13295.
- [58] TURNER, M. R. “Texture discrimination by Gabor functions”, *Biological Cybernetics*, v. 55, n. 2, pp. 71–82, Nov 1986. ISSN: 1432-0770. doi: 10.1007/BF00341922. Disponível em: <<https://doi.org/10.1007/BF00341922>>. Acessado em 16-02-2020.
- [59] ZAYED, N., ELNEMR, H. “Statistical Analysis of Haralick Texture Features to Discriminate Lung Abnormalities”, *International Journal of Biomedical Imaging*, v. 2015, pp. 1–7, 11 2015. doi: 10.1155/2015/267807.
- [60] FRIEDMAN, J., HASTIE, T., TIBSHIRANI, R. *The elements of statistical learning*, v. 1. Springer series in statistics New York, 2001.
- [61] LEHAMEL, M., HAMMOUCHE, K. “Comparative Evaluation of Various Fractal Dimension Estimation Methods”. In: *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems*, pp. 259–266, Nov 2012. doi: 10.1109/SITIS.2012.46.
- [62] SUN, W., XU, G., GONG, P., et al. “Fractal analysis of remotely sensed images: A review of methods and applications”, *International Journal of Remote Sensing*, v. 27, n. 22, pp. 4963–4990, 2006. doi: 10.1080/01431160600676695. Disponível em: <<https://doi.org/10.1080/01431160600676695>>. Acessado em 16-02-2020.
- [63] KILIC, K. I., ABIYEV, R. H. “Exploiting the synergy between fractal dimension and lacunarity for improved texture recognition”, *Signal Processing*, v. 91, n. 10, pp. 2332 – 2344, 2011. ISSN: 0165-1684. doi: <https://doi.org/10.1016/j.sigpro.2011.04.018>. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0165168411001289>>. Acessado em 16-02-2020.
- [64] ALLAIN, C., CLOITRE, M. “Characterizing the Lacunarity of Random and Deterministic Fractal Sets”, *Physical review. A*, v. 44, pp. 3552–3558, 10 1991. doi: 10.1103/PhysRevA.44.3552.
- [65] LEHAMEL, M., HAMMOUCHE, K. “Texture Classification Using Fractal Dimension and Lacunarity”. 2014.

- [66] BACKES, A. R., DE MESQUITA SÁ, J. J. “Texture Classification Using Fractal Dimension Improved by Local Binary Patterns”. In: *2018 26th European Signal Processing Conference (EUSIPCO)*, pp. 1312–1316, Sep. 2018. doi: 10.23919/EUSIPCO.2018.8553005.
- [67] BRODATZ, P. *Textures: a photographic album for artists and designers*. Dover Pubns, 1966.
- [68] DAUGMAN, J. G. “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters”, *J. Opt. Soc. Am. A*, v. 2, n. 7, pp. 1160–1169, Jul 1985. doi: 10.1364/JOSAA.2.001160. Disponível em: <<http://josaa.osa.org/abstract.cfm?URI=josaa-2-7-1160>>. Acessado em 16-02-2020.
- [69] SANDLER, R., LINDENBAUM, M. “Gabor Filter Analysis for Texture Segmentation”. In: *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW’06)*, pp. 178–178, June 2006. doi: 10.1109/CVPRW.2006.86.
- [70] OJALA, T., MAENPAA, T., PIETIKAINEN, M., et al. “Outex-new framework for empirical evaluation of texture analysis algorithms”. In: *Object recognition supported by user interaction for service robots*, v. 1, pp. 701–706. IEEE, 2002.
- [71] VALENTIN, M. B., BOM, C. R. D., DE ALBUQUERQUE, M. P., et al. “Rock Texture Classification Using Spectral Analysis And Genetically Optimized Texture Features”, *CoRR*, v. abs/1607.01679, 2016. Disponível em: <<http://arxiv.org/abs/1607.01679>>. Acessado em 16-02-2020.
- [72] FRITZ, M., HAYMAN, E., CAPUTO, B., et al. “The kth-tips database”. 2004.
- [73] KOPAL, M., JUNGSMANN, M., BERLAGE, T., et al. “Rock classification based on resistivity patterns in electrical borehole wall images”, *J. Geophys. Eng.*, v. 4, pp. 171–183, 06 2007. doi: 10.1088/1742-2132/4/2/006.
- [74] ZHANG, H. “The optimality of naive Bayes”, *AA*, v. 1, n. 2, pp. 3, 2004.
- [75] SINGH, M., JAVADI, A., SINGH, S. “A comparison of texture features for the classification of rock images”. In: *International Conference on Intelligent Data Engineering and Automated Learning*, pp. 179–184. Springer, 2004.
- [76] ALTMAN, N. S. “An introduction to kernel and nearest-neighbor nonparametric regression”, *The American Statistician*, v. 46, n. 3, pp. 175–185, 1992.

- [77] SHANG, C., BARNES, D. “Support vector machine-based classification of rock texture images aided by efficient feature selection”. In: *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. IEEE, 2012.
- [78] BREIMAN, L. *Classification and regression trees*. Routledge, 2017.
- [79] PEREZ, C. A., SARAIVA, J. A., NAVARRO, C. F., et al. “Rock lithological classification using multi-scale Gabor features from sub-images, and voting with rock contour information”, *International Journal of Mineral Processing*, v. 144, pp. 56 – 64, 2015. ISSN: 0301-7516. doi: <https://doi.org/10.1016/j.minpro.2015.09.015>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0301751615300326>>. Acessado em 16-02-2020.
- [80] TESSIER, J., DUCHESNE, C., BARTOLACCI, G. “A machine vision approach to on-line estimation of run-of-mine ore composition on conveyor belts”, *Minerals Engineering*, v. 20, n. 12, pp. 1129 – 1144, 2007. ISSN: 0892-6875. doi: <https://doi.org/10.1016/j.mineng.2007.04.009>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0892687507001161>>. Acessado em 16-02-2020.
- [81] KRIZHEVSKY, A., SUTSKEVER, I., HINTON, G. E. “ImageNet Classification with Deep Convolutional Neural Networks”. In: Pereira, F., Burges, C. J. C., Bottou, L., et al. (Eds.), *Advances in Neural Information Processing Systems 25*, Curran Associates, Inc., pp. 1097–1105, 2012. Disponível em: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>>. Acessado em 16-02-2020.
- [82] GRAVES, A., MOHAMED, A., HINTON, G. E. “Speech Recognition with Deep Recurrent Neural Networks”, *CoRR*, v. abs/1303.5778, 2013. Disponível em: <http://arxiv.org/abs/1303.5778>>. Acessado em 16-02-2020.
- [83] HERMANS, M., SCHRAUWEN, B. “Training and Analysing Deep Recurrent Neural Networks”. In: Burges, C. J. C., Bottou, L., Welling, M., et al. (Eds.), *Advances in Neural Information Processing Systems 26*, Curran Associates, Inc., pp. 190–198, 2013. Disponível em: <http://papers.nips.cc/paper/5166-training-and-analysing-deep-recurrent-neural-networks.pdf>>. Acessado em 16-02-2020.

- [84] RONNEBERGER, O., FISCHER, P., BROX, T. “U-Net: Convolutional Networks for Biomedical Image Segmentation”, *CoRR*, v. abs/1505.04597, 2015. Disponível em: <<http://arxiv.org/abs/1505.04597>>. Acessado em 16-02-2020.
- [85] LECUN, Y. “Generalization and network design strategies”. In: Pfeifer, R., Schreter, Z., Fogelman, F., et al. (Eds.), *Connectionism in perspective*, Elsevier, 1989.
- [86] “Pytorch: Open Source Deep Learning Library”. <https://pytorch.org/>, . Acessado em 16-02-2020.
- [87] HE, K., ZHANG, X., REN, S., et al. “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition”, *CoRR*, v. abs/1406.4729, 2014. Disponível em: <<http://arxiv.org/abs/1406.4729>>. Acessado em 16-02-2020.
- [88] LAZEBNIK, S., SCHMID, C., PONCE, J. “Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories”. In: *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*, v. 2, pp. 2169–2178, June 2006. doi: 10.1109/CVPR.2006.68.
- [89] ZHUANG, F., QI, Z., DUAN, K., et al. “A Comprehensive Survey on Transfer Learning”, *arXiv preprint arXiv:1911.02685*, 2019.
- [90] FUJIEDA, S., TAKAYAMA, K., HACHISUKA, T. “Wavelet Convolutional Neural Networks for Texture Classification”, *CoRR*, v. abs/1707.07394, 2017. Disponível em: <<http://arxiv.org/abs/1707.07394>>. Acessado em 16-02-2020.
- [91] ANDREARCZYK, V., WHELAN, P. F. “Using filter banks in convolutional neural networks for texture classification”, *Pattern Recognition Letters*, v. 84, pp. 63–69, 2016.
- [92] “Describable Texture Dataset website”. <https://www.robots.ox.ac.uk/~vgg/data/dtd/>, . Acessado em 16-02-2020.
- [93] MALLIKARJUNA, P., TARGHI, A. T., FRITZ, M., et al. “The kth-tips2 database”, *KTH Royal Institute of Technology*, 2006.
- [94] SZEGEDY, C., LIU, W., JIA, Y., et al. “Going deeper with convolutions”, *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,

Jun 2015. doi: 10.1109/cvpr.2015.7298594. Disponível em: <<http://dx.doi.org/10.1109/CVPR.2015.7298594>>. Acessado em 16-02-2020.

- [95] MIOTTO, R., WANG, F., WANG, S., et al. “Deep learning for healthcare: review, opportunities and challenges”, *Briefings in Bioinformatics*, v. 19, n. 6, pp. 1236–1246, 05 2017. ISSN: 1477-4054. doi: 10.1093/bib/bbx044. Disponível em: <<https://doi.org/10.1093/bib/bbx044>>. Acessado em 16-02-2020.
- [96] BHATIA, K. K., GRAHAM, M. S., TERRY, L., et al. “Disease classification of macular Optical Coherence Tomography scans using deep learning software: validation on independent, multi-centre data”, *arXiv preprint arXiv:1907.05164*, 2019.
- [97] LI, W., CAO, P., ZHAO, D., et al. “Pulmonary nodule classification with deep convolutional neural networks on computed tomography images”, *Computational and mathematical methods in medicine*, v. 2016, 2016.
- [98] ARMATO III, S. G., MCLENNAN, G., BIDAUT, L., et al. “The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans”, *Medical physics*, v. 38, n. 2, pp. 915–931, 2011.
- [99] SONG, Y., LI, Q., FENG, D., et al. “Texture image classification with discriminative neural networks”, *Computational Visual Media*, v. 2, n. 4, pp. 367–377, Dec 2016. ISSN: 2096-0662. doi: 10.1007/s41095-016-0060-6. Disponível em: <<https://doi.org/10.1007/s41095-016-0060-6>>. Acessado em 16-02-2020.
- [100] LIU, L., CHEN, J., FIEGUTH, P., et al. “From BoW to CNN: Two decades of texture representation for texture classification”, *International Journal of Computer Vision*, v. 127, n. 1, pp. 74–109, 2019.
- [101] CHENG, G., GUO, W. “Rock images classification by using deep convolution neural network”, *Journal of Physics: Conference Series*, v. 887, pp. 012089, aug 2017. doi: 10.1088/1742-6596/887/1/012089. Disponível em: <<https://doi.org/10.1088%2F1742-6596%2F887%2F1%2F012089>>. Acessado em 16-02-2020.
- [102] RAN, X., XUE, L., ZHANG, Y., et al. “Rock Classification from Field Image Patches Analyzed Using a Deep Convolutional Neural Network”, *Mathematics*, v. 7, n. 8, pp. 755, 2019.

- [103] ZHANG, Y., WANG, G., LI, M., et al. “Automated classification analysis of geological structures based on images data and deep learning model”, *Applied Sciences*, v. 8, n. 12, pp. 2493, 2018.
- [104] SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., et al. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”, *Journal of Machine Learning Research*, v. 15, pp. 1929–1958, 2014. Disponível em: <<http://jmlr.org/papers/v15/srivastava14a.html>>. Acessado em 16-02-2020.
- [105] “Google Cloud Advanced Guide to Inception V3”. <https://cloud.google.com/tpu/docs/inception-v3-advanced>, . Acessado em 16-02-2020.
- [106] JAPKOWICZ, N., SHAH, M. *Evaluating Learning Algorithms: A Classification Perspective*. New York, NY, USA, Cambridge University Press, 2014. ISBN: 1107653118, 9781107653115.
- [107] EVSUKOFF, A. “Ensinando Máquinas”. Material disponibilizado na aula de Inteligência Computacional(COC361) do Centro de Tecnologia da UFRJ.
- [108] DOS ANJOS, A., TORRES, R., SEIXAS, J., et al. “Neural triggering system operating on high resolution calorimetry information”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 559, n. 1, pp. 134 – 138, 2006. ISSN: 0168-9002. doi: <https://doi.org/10.1016/j.nima.2005.11.132>. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168900205022552>>. Acessado em 16-02-2020.
- [109] “OpenCV: Open Source Computer Vision”. <https://docs.opencv.org>. Acessado em 16-02-2020.
- [110] LECUN, Y., BOTTOU, L., BENGIO, Y., et al. “Gradient-based learning applied to document recognition”, *Proceedings of the IEEE*, v. 86, n. 11, pp. 2278–2324, 1998.
- [111] KINGMA, D. P., BA, J. “Adam: A method for stochastic optimization”, *arXiv preprint arXiv:1412.6980*, 2014.
- [112] WILD, C. J. “Chance encounters: A first course in data analysis and inference”, 2000.
- [113] SHORTEN, C., KHOSHGOFTAAR, T. M. “A survey on image data augmentation for deep learning”, *Journal of Big Data*, v. 6, n. 1, pp. 60, 2019.

- [114] GUI, J., SUN, Z., WEN, Y., et al. “A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications”, *arXiv preprint arXiv:2001.06937*, 2020.
- [115] “Python”. <https://www.python.org/>, . Acessado em 16-02-2020.
- [116] “StackOverflow: Learn, Share, Build”. <https://stackoverflow.com/>, . Acessado em 16-02-2020.
- [117] “StackOverflow Developer Survey Results 2018”. <https://insights.stackoverflow.com/survey/2018/#most-loved-dreaded-and-wanted>, . Acessado em 16-02-2020.
- [118] “NumPy: a numerical extension for the computer language Python”. <https://www.numpy.org/>. Acessado em 16-02-2020.
- [119] “Pandas: Python Data Analysis Library”. <https://pandas.pydata.org/>. Acessado em 16-02-2020.
- [120] “Matplotlib: Python Plotting Library”. <https://matplotlib.org/>. Acessado em 16-02-2020.
- [121] “Anaconda: Manage your packages, dependencies and environments”. <https://anaconda.org/>. Acessado em 16-02-2020.
- [122] “Jupyter Notebook”. <http://jupyter.org/>. Acessado em 16-02-2020.

Apêndice A

Ferramentas

Neste apêndice, apresentam-se as tecnologias utilizadas no desenvolvimento desse trabalho.

O projeto consistiu na implementação de redes neurais artificiais para classificação de texturas de imagens. Neste sentido, importante destacar que o processamento de otimização em redes neurais pode ser computacionalmente custoso dependendo do tamanho da rede, sendo a escolha certa de linguagem e ferramentas de desenvolvimento a serem utilizadas algo importante.

A seção será iniciada com a apresentação da linguagem na qual o código foi escrito, junto com algumas das principais bibliotecas utilizadas para processamento, manipulação e visualização dos dados. Em seguida, serão apresentados o gerenciador de dependências e a interface na qual o código foi escrito.

Concluindo a seção, será apresentada a principal ferramenta para o desenvolvimento do projeto, a qual foi usada para implementação das redes neurais artificiais utilizadas.

A.1 Python

Python [115] é uma linguagem de programação de alto nível, interpretada, de script e orientada a objetos. É uma linguagem muito conhecida e difundida, tendo ganhado, inclusive, o terceiro lugar de linguagem mais amada no stackoverflow [116] em 2018 [117].

A linguagem escolhida foi criada priorizando-se a legibilidade do código sobre a velocidade de processamento. Devido a essa escolha, usar a linguagem pura, ou seja, sem nenhum framework, não se mostra eficiente para processamentos mais complexos, como, por exemplo, a multiplicação de matrizes, surgindo, por isso, alguns frameworks para contornar o problema do processamento mais lento.

A.1.1 Numpy

Numpy [118] é um framework de computação numérica muito utilizado como base para outras bibliotecas, algumas inclusive utilizadas neste projeto. Devido à sua implementação diferenciada em C/C++ e FORTRAN, o numpy se mostra muito melhor em processamento de matrizes em comparação ao Python puro, permitindo-se o trabalho com arranjos, vetores e matrizes de N dimensões, que são de grande importância para os ajustes de parâmetros em redes neurais, mantendo-se a facilidade de uma escrita mais simples e limpa, fornecida pelo Python.

A.1.2 Pandas

Pandas [119] é uma biblioteca open source, de fácil implementação e uso para análise e processamento de dados. É de grande importância nesse projeto devido à facilidade que provê para controle e manipulação dos dados em formato tabular, maneira pela qual foi feito o controle dos locais dos arquivos usados e suas anotações.

A.1.3 Opencv

OpenCV [109] foi lançado nos anos 2000 e hoje em dia é uma biblioteca que suporta diferentes linguagens de programação como C++, Python e Java. O OpenCV-Python é a API em python da biblioteca, a qual possui diversos algoritmos relacionados à visão computacional e aprendizado de máquina, sendo usado, neste projeto, principalmente para manipulação e pré-processamento das imagens.

A.1.4 Matplotlib

Em razão do projeto trabalhar com imagens, uma biblioteca para visualização das mesmas se torna necessária, sendo escolhida para o projeto o Matplotlib [120], feita pra criação de gráficos e imagens em uma API de fácil utilização.

A.2 Anaconda

Um grande problema na área de desenvolvimento é o gerenciamento do ambiente de desenvolvimento. Para mitigar os problemas gerados por um ambiente mal gerenciado, no qual as dependências não estão corretamente atualizadas e salvas, foi usado o Anaconda [121]. O Anaconda é um gerenciador de pacotes, ambientes e projetos que facilita o desenvolvimento do trabalho.

A.2.1 Jupyter

O Jupyter [122] é uma aplicação baseada na interface web, que permite o desenvolvimento de software e funcionando como uma IDE. Além das vantagens trazidas por funcionar como uma IDE, facilita a geração de uma documentação limpa e visualmente agradável.

A.3 Pytorch

Existem diversas bibliotecas para desenvolvimento de aplicações de aprendizado de máquina e aprendizado profundo. Uma das principais e com grande destaque por todo o mundo é o Pytorch [86], framework usado amplamente, sendo exemplos de locais que a utilizam a Udacity, Stanford, entre outras.

Isto porque, o Pytorch fornece API de alto nível que facilita a criação e treinamento de modelos, além de permitir o controle de baixo nível para maximizar a flexibilidade e desempenho do sistema, de forma que a escolha por essa biblioteca se deu em razão da possibilidade do uso de grafos dinâmicos no treinamento da rede neural, tornando-se possível a utilização de imagens com diferentes tamanhos como entrada para os modelos.

Em sistemas de grande porte, o desempenho é algo crucial, sendo vital na área de aprendizado de máquina, principalmente em aprendizado profundo, onde existem modelos com milhões de parâmetros, que podem demorar dias para rodar em clusters de computadores. Por isso, o Pytorch foi criado de modo que seja possível a paralelização dos processamentos em CPU e GPU.

Além disso, importante mencionar que o Pytorch é uma biblioteca de software de código aberto para computação numérica de alto desempenho, funcionando em diversos sistemas operacionais, como Linux, Windows e MacOS, havendo inclusive pacotes para aplicações mobile e raspberry.

Apêndice B

Escolha da Topologia da Rede Neural Artificial

Como apresentado anteriormente, a arquitetura do modelo de rede neural artificial proposto pode ser visto na Figura B.1. Testes variando a quantidade de neurônios na camada escondida foram realizados, e, a partir da validação cruzada aninhada, o modelo que obteve melhores resultados no índice SP do conjunto de validação foi escolhido. A quantidade de neurônios na camada escondida variou de 5 a 100, pulando de 5 em 5, totalizando assim 20 modelos. Os resultados obtidos dos modelos com o índice SP no conjunto de treino, validação e teste são apresentados nas Figuras B.2, B.3 e B.4. Como o intuito deste teste foi a escolha da topologia a ser comparada com os modelos de aprendizado profundo, apenas o índice SP foi avaliado. Apesar dos resultados estarem muito próximos ao avaliarmos os resultados do conjunto de validação, verifica-se que o modelo que obteve a maior média foi o de 10 neurônios na camada escondida, sendo este escolhido entre os modelos treinados. Nota-se também que os modelos com mais neurônios obtiveram melhores performances no conjunto de treino, contudo, isso não se refletiu nos conjuntos de validação e teste, o que indica sobre ajuste desses modelos. É importante realçar que o modelo escolhido também apresentou a maior média no conjunto de teste, como visto na Figura B.4.

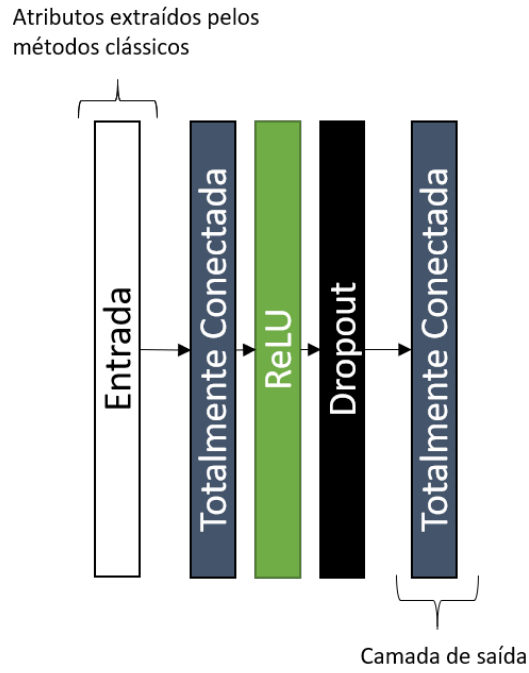


Figura B.1: Estrutura da rede neural artificial proposta

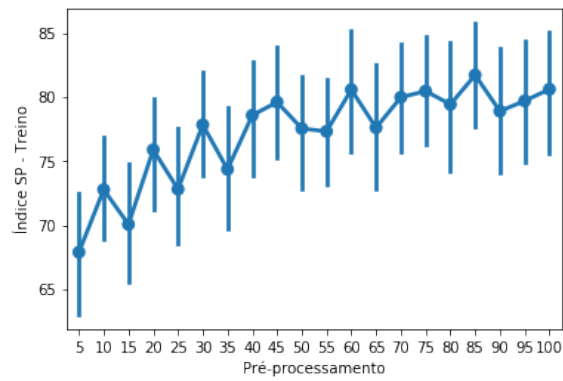


Figura B.2: Comparação do índice SP de diferentes topologias no conjunto de treino

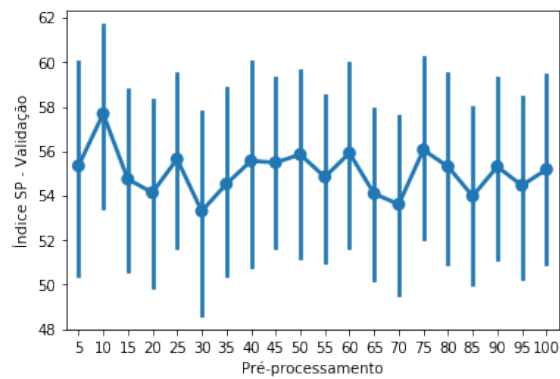


Figura B.3: Comparação do índice SP de diferentes topologias no conjunto de validação

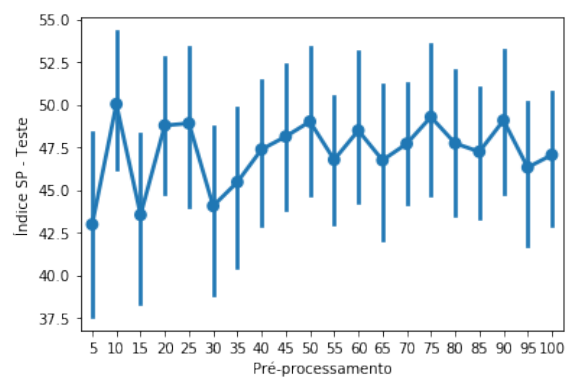


Figura B.4: Comparação do índice SP de diferentes topologias no conjunto de teste

Apêndice C

Plugues Utilizados e Grupos de Validação Cruzada

Neste apêndice, apresenta-se um exemplo de cada plugue separando-os pelos folds da validação cruzada, ressaltando a heterogeneidade das amostras.

C.1 Fold 1

A Figura C.1 apresenta um exemplo de cada plugue da classe esferulítico, a Figura C.2 apresenta um exemplo de cada plugue da classe estromatólito e a Figura C.3 apresenta um exemplo de cada plugue da classe grainstone.

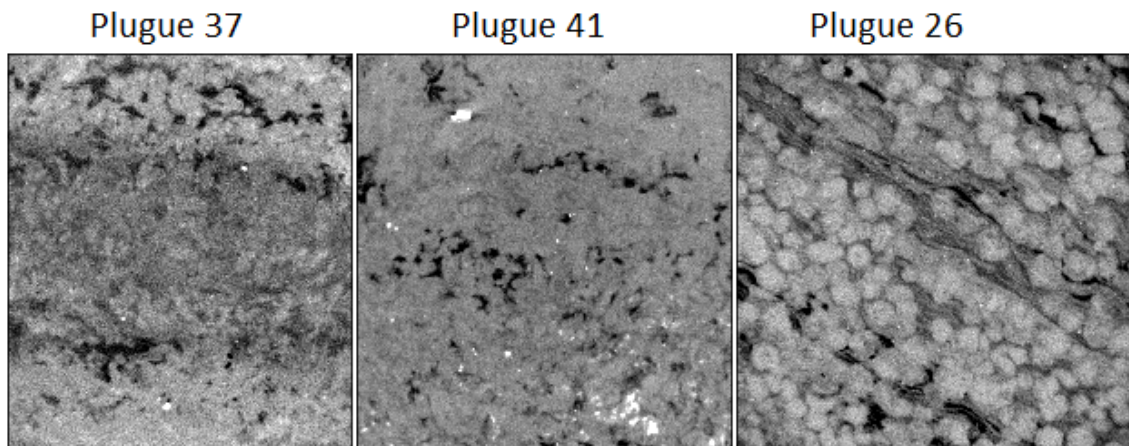


Figura C.1: Exemplos da classe esferulítico do fold 1

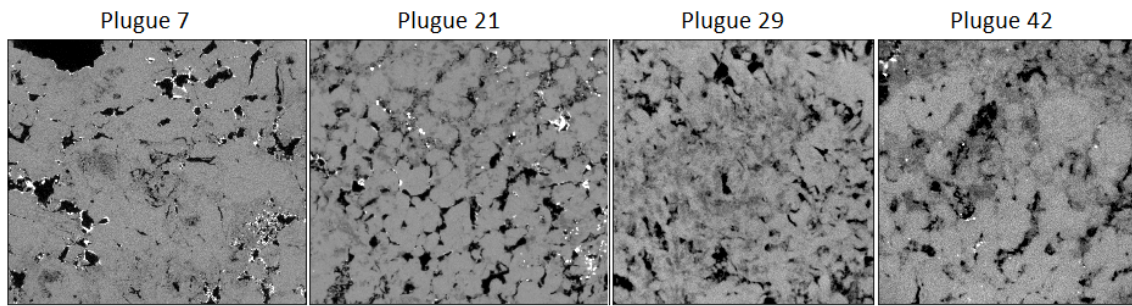


Figura C.2: Exemplos da classe estromatólito do fold 1

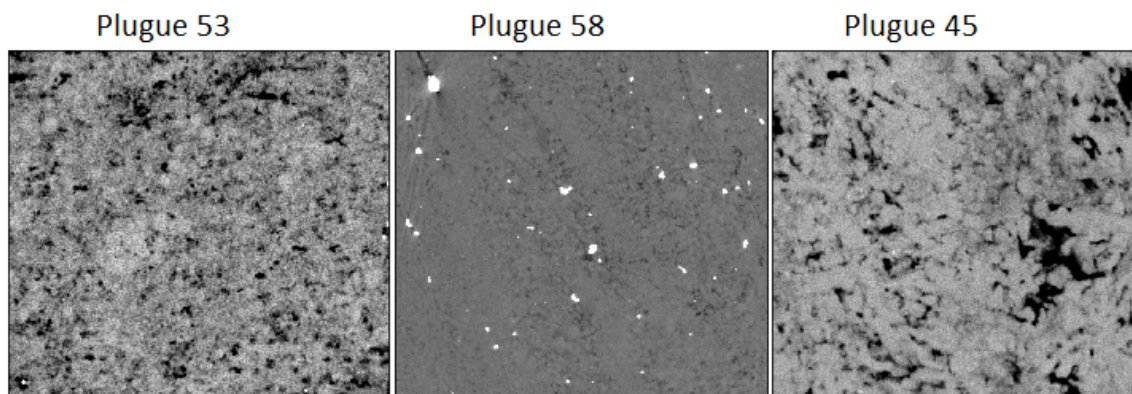


Figura C.3: Exemplos da classe grainstone do fold 1

C.2 Fold 2

A Figura C.4 apresenta um exemplo de cada plugue da classe esferulítico, a Figura C.5 apresenta um exemplo de cada plugue da classe estromatólito e a Figura C.6 apresenta um exemplo de cada plugue da classe grainstone.

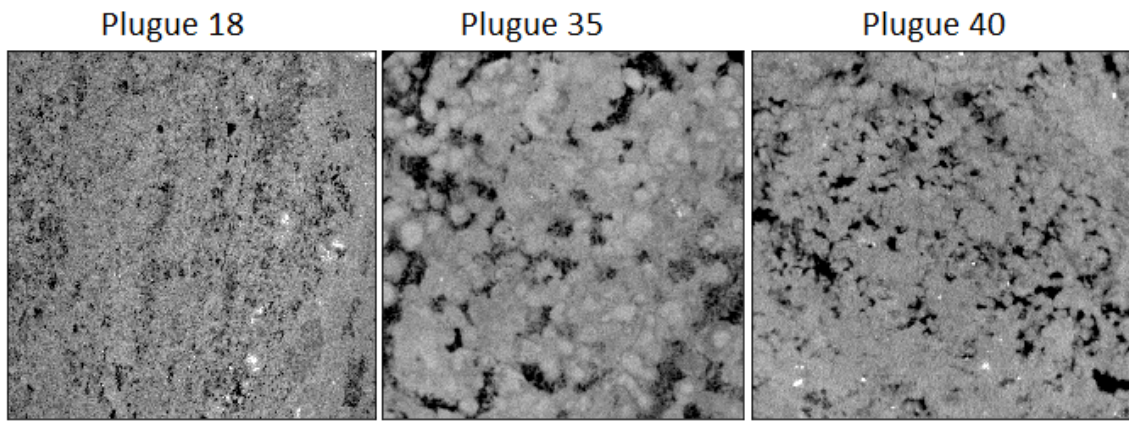


Figura C.4: Exemplos da classe esferulitito do fold 2

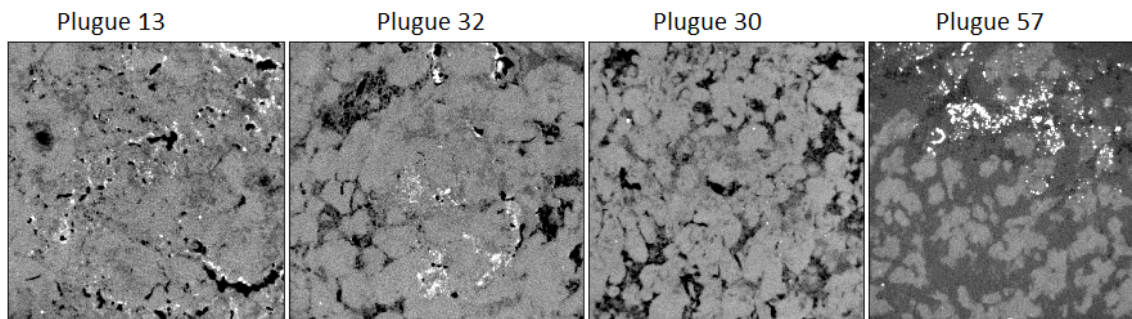


Figura C.5: Exemplos da classe estromatólito do fold 2

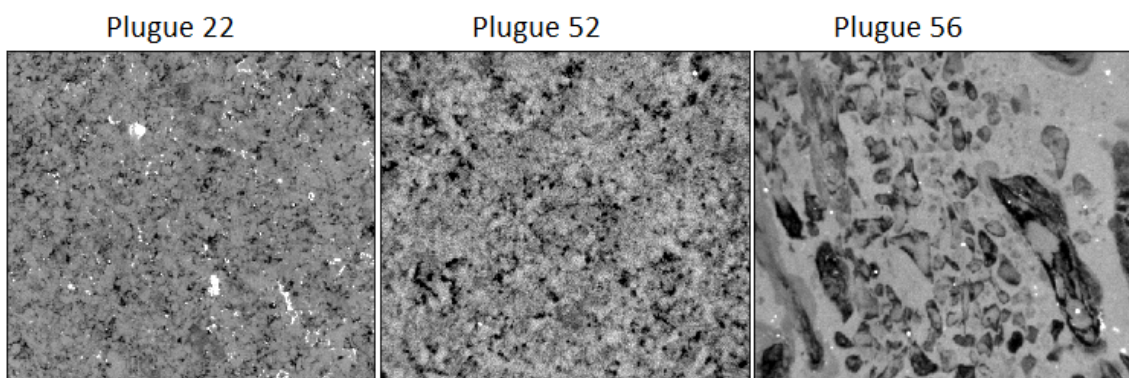


Figura C.6: Exemplos da classe grainstone do fold 2

C.3 Fold 3

A Figura C.7 apresenta um exemplo de cada plugue da classe esferulitito, a Figura C.8 apresenta um exemplo de cada plugue da classe estromatólito e a Figura C.9

apresenta um exemplo de cada plugue da classe grainstone.

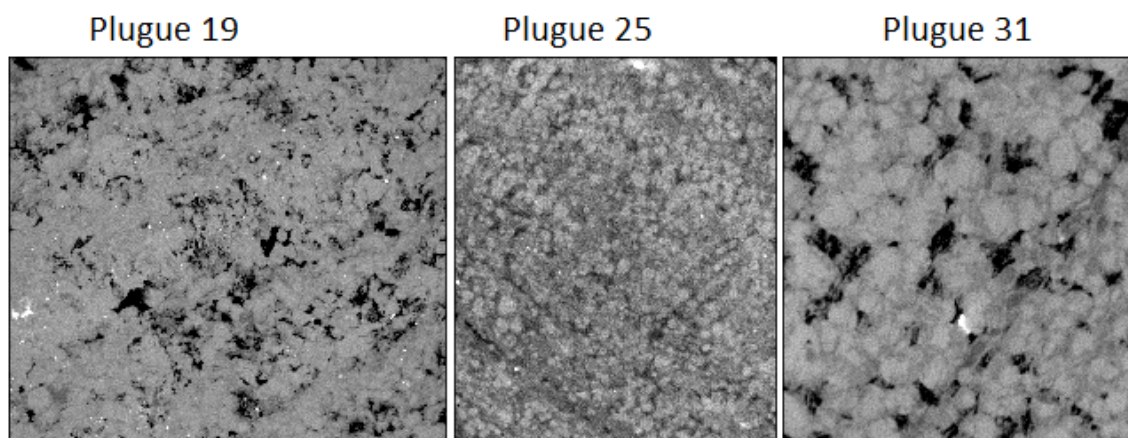


Figura C.7: Exemplos da classe esferulítico do fold 3

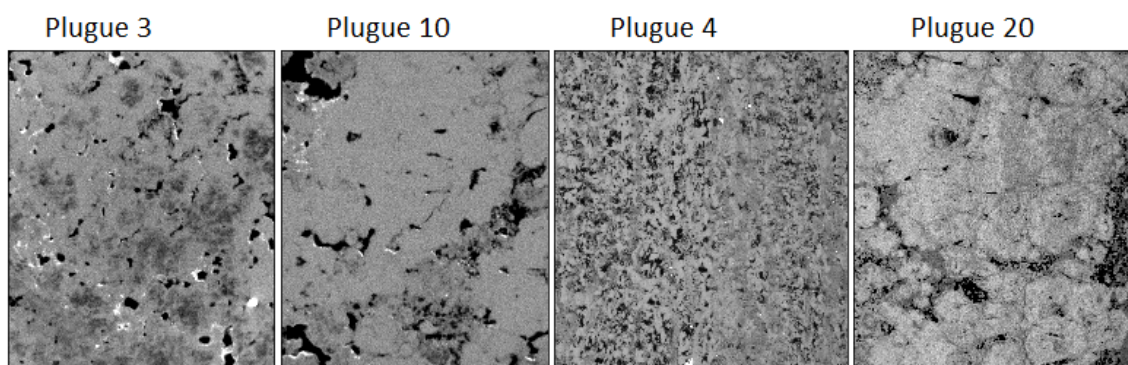


Figura C.8: Exemplos da classe estromatólito do fold 3

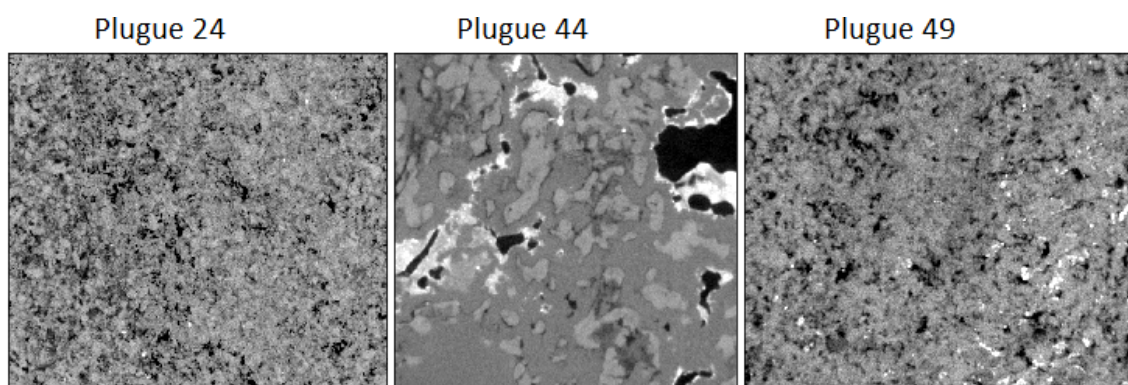


Figura C.9: Exemplos da classe grainstone do fold 3

C.4 Fold 4

A Figura C.10 apresenta um exemplo de cada plugue da classe esferulítico, a Figura C.11 apresenta um exemplo de cada plugue da classe estromatólito e a Figura C.12 apresenta um exemplo de cada plugue da classe grainstone.

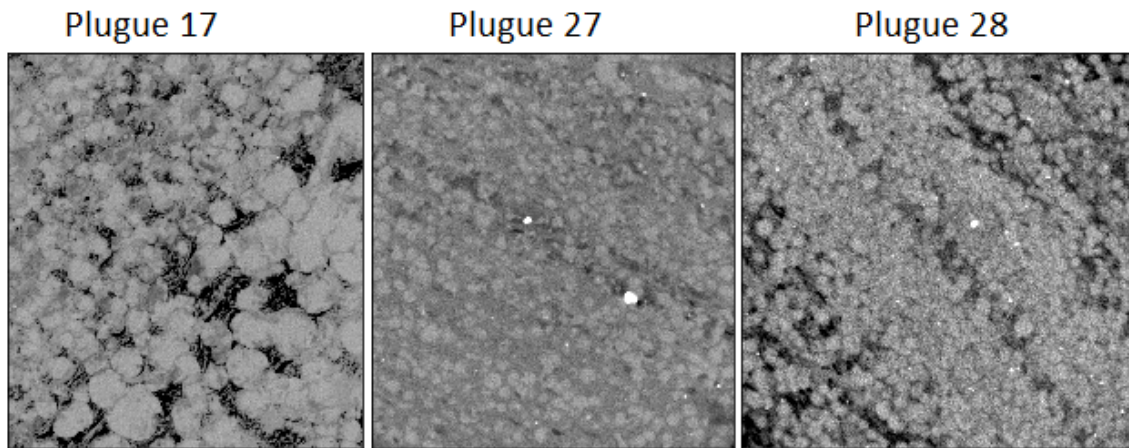


Figura C.10: Exemplos da classe esferulítico do fold 4

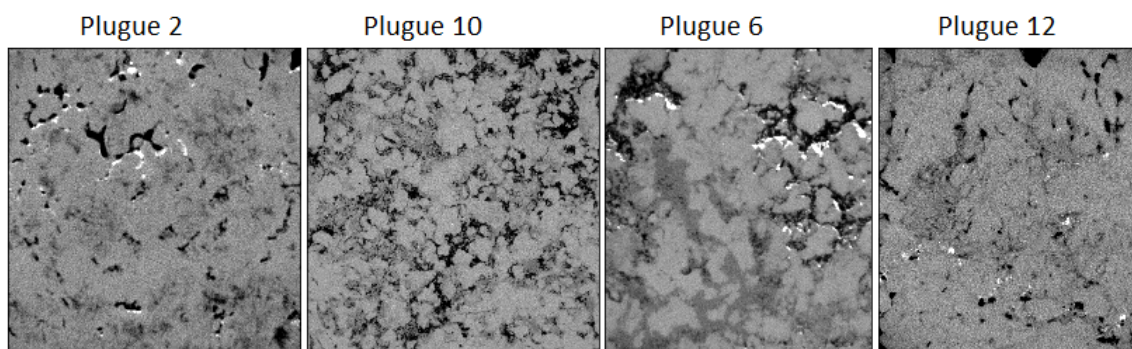


Figura C.11: Exemplos da classe estromatólito do fold 4

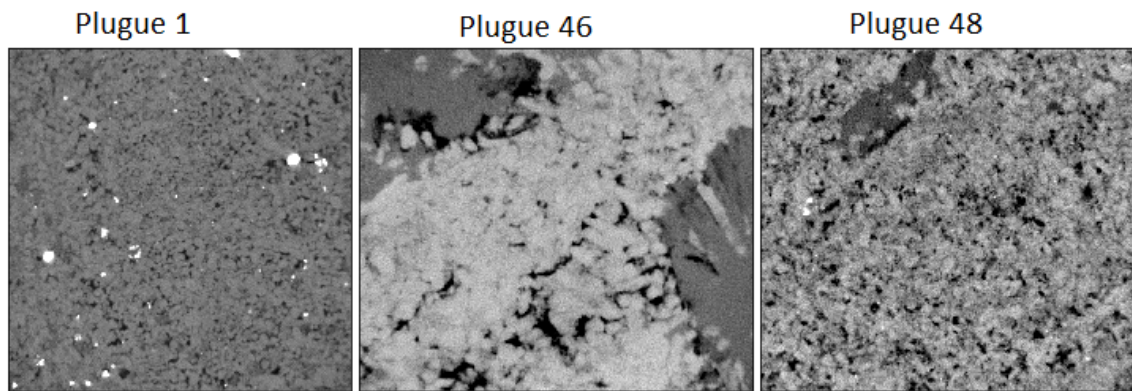


Figura C.12: Exemplos da classe grainstone do fold 4

C.5 Fold 5

A Figura C.13 apresenta um exemplo de cada plugue da classe esferulítico, a Figura C.14 apresenta um exemplo de cada plugue da classe estromatólito e a Figura C.15 apresenta um exemplo de cada plugue da classe grainstone.

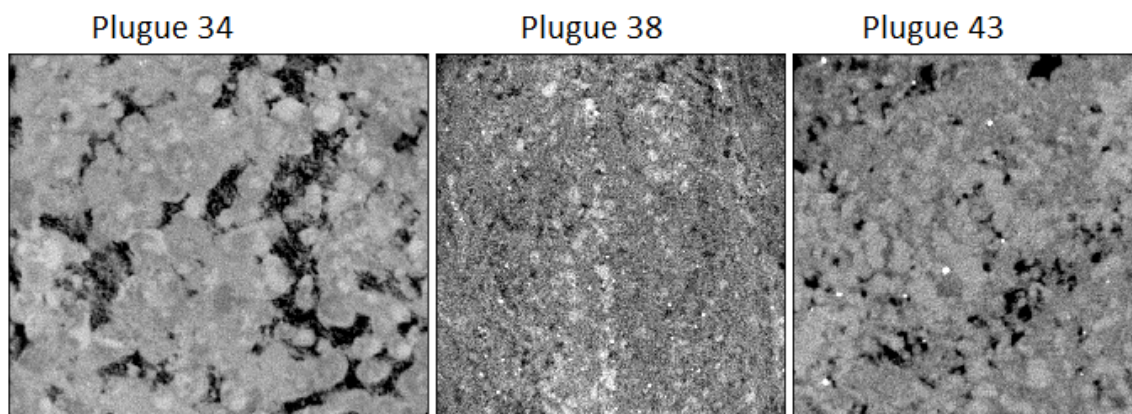


Figura C.13: Exemplos da classe esferulítico do fold 5

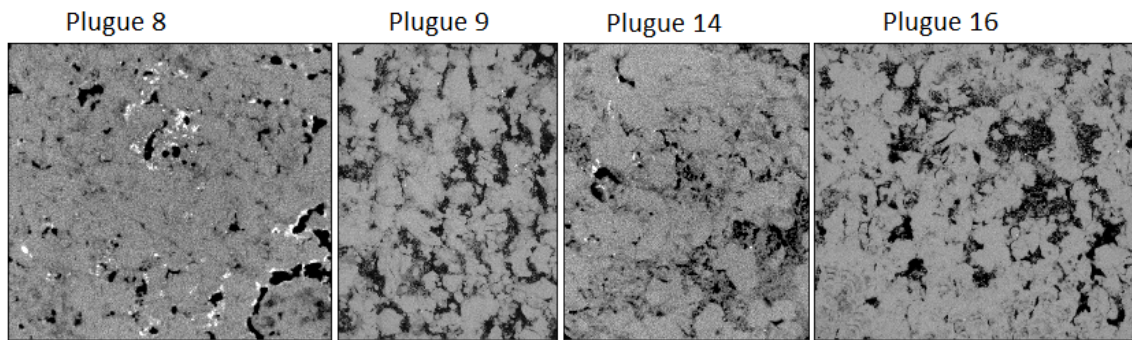


Figura C.14: Exemplos da classe estromatólito do fold 5

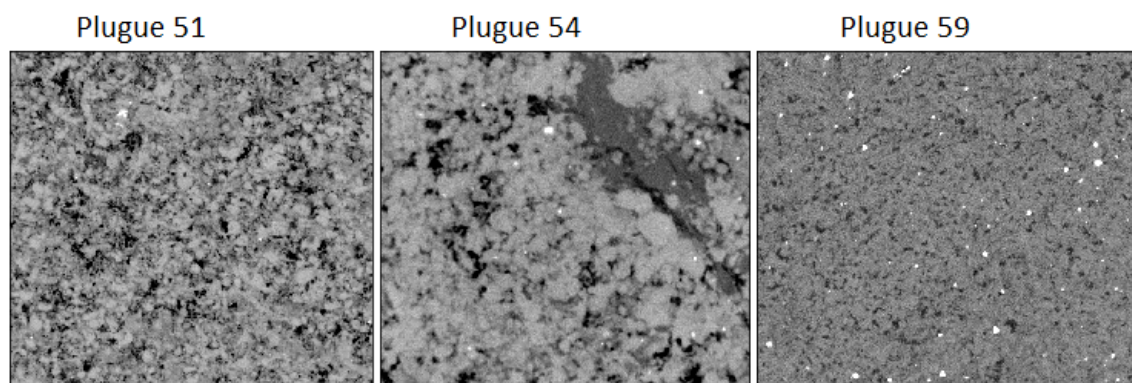


Figura C.15: Exemplos da classe grainstone do fold 5

C.6 Fold 6

A Figura C.16 apresenta um exemplo de cada plugue da classe esferulítico, a Figura C.17 apresenta um exemplo de cada plugue da classe estromatólito e a Figura C.18 apresenta um exemplo de cada plugue da classe grainstone.

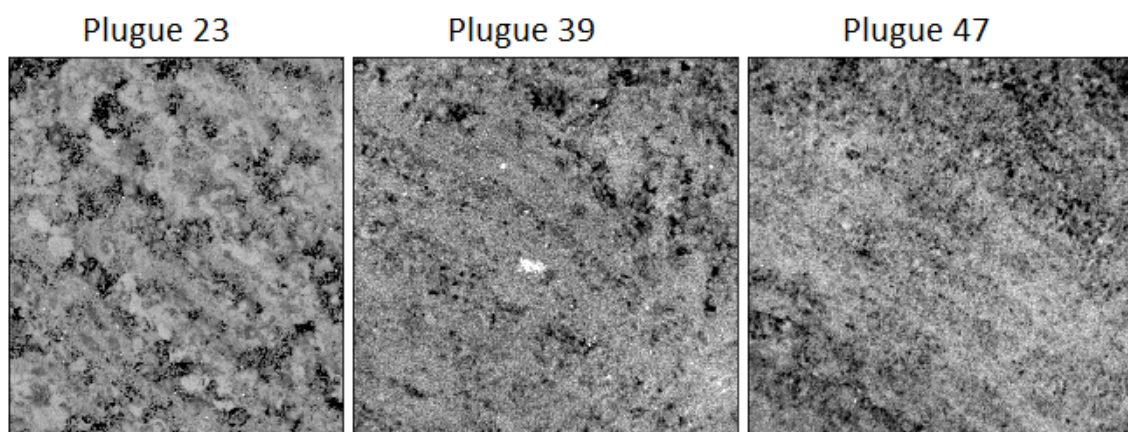


Figura C.16: Exemplos da classe esferulitito do fold 6

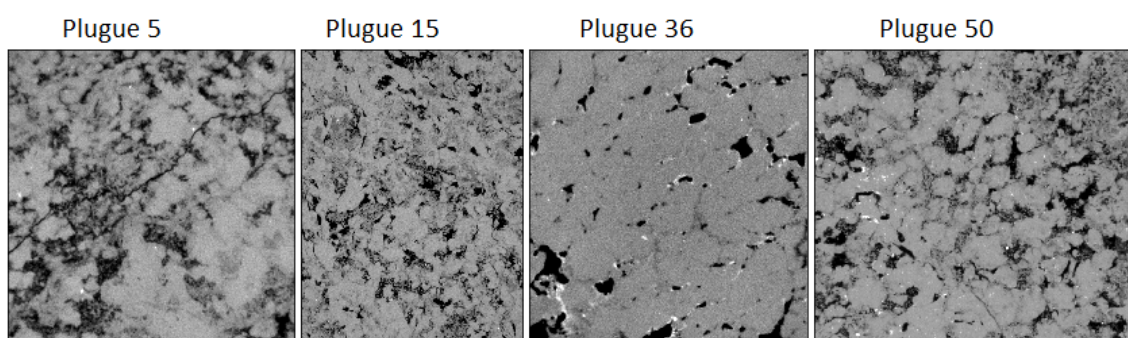


Figura C.17: Exemplos da classe estromatólito do fold 6

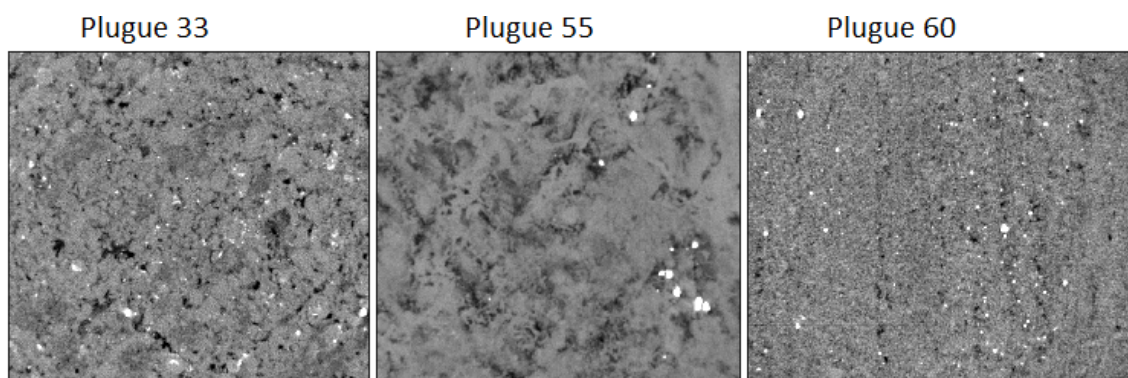


Figura C.18: Exemplos da classe grainstone do fold 6

Apêndice D

Resultados

Neste apêndice, listam-se todas as tabelas com resultados dos modelos apresentados neste trabalho, divididos em seções de treino, validação e teste.

D.1 Treino

Nessa seção, apresentam-se os resultados obtidos no conjunto de treino, iniciando-se com a avaliação por imagens individuais, seguida da avaliação por plugue, a qual usa a classe mais frequente das imagens daquele plugue, gerando, assim, uma classificação única por plugue.

D.1.1 Avaliação por Imagem

Tabela D.1: Tabela com acurácia e índice SP dos modelos ao avaliar por imagem individual no conjunto de treino

Tipo de Entrada	Modelo	Acurácia (%)	Índice SP (%)
Métodos Clássicos	RNA	$73,4 \pm 11,1$	$72,8 \pm 11,0$
Imagem Original	CNN+SPP	$96,5 \pm 06,8$	$96,2 \pm 07,6$
Imagem Redimensionada	CNN+SPP	$99,7 \pm 00,5$	$99,7 \pm 00,6$
Imagem Redimensionada	CNN	$99,8 \pm 00,8$	$99,8 \pm 00,9$
Imagem Redimensionada	Inception	$99,2 \pm 01,5$	$99,2 \pm 01,5$

Tabela D.2: Tabela com recall da classe Esferulítico avaliando por imagem individual no conjunto de treino

Tipo de Entrada	Modelo	Esferulítico (%)
Métodos Clássicos	RNA	78 ± 08
Original	CNN+SPP	97 ± 06
Redimensionada	CNN+SPP	100 ± 00
Redimensionada	CNN	100 ± 02
Redimensionada	Inception	100 ± 01

Tabela D.3: Tabela com recall da classe Estromatólito avaliando por imagem individual no conjunto de treino

Tipo de Entrada	Modelo	Estromatólito (%)
Métodos Clássicos	RNA	75 ± 17
Original	CNN+SPP	98 ± 05
Redimensionada	CNN+SPP	100 ± 01
Redimensionada	CNN	100 ± 00
Redimensionada	Inception	99 ± 02

Tabela D.4: Tabela com recall da classe Grainstone avaliando por imagem individual no conjunto de treino

Tipo de Entrada	Modelo	Grainstone (%)
Métodos Clássicos	RNA	67 ± 16
Original	CNN+SPP	94 ± 14
Redimensionada	CNN+SPP	99 ± 01
Redimensionada	CNN	100 ± 01
Redimensionada	Inception	99 ± 03

D.1.2 Avaliação por Plugue

Tabela D.5: Tabela com acurácia e índice SP dos modelos ao avaliar por plugue no conjunto de treino

Tipo de Entrada	Modelo	Acurácia (%)	Índice SP (%)
Métodos Clássicos	RNA	77 ± 12	76 ± 12
Imagem Original	CNN+SPP	98 ± 05	98 ± 06
Imagem Redimensionada	CNN+SPP	100 ± 00	100 ± 00
Imagem Redimensionada	CNN	100 ± 00	100 ± 01
Imagem Redimensionada	Inception	100 ± 01	100 ± 01

Tabela D.6: Tabela com recall da classe Esferulitito avaliando por plugue no conjunto de treino

Tipo de Entrada	Modelo	Esferulitito (%)
Métodos Clássicos	RNA	80 ± 08
Original	CNN+SPP	99 ± 04
Redimensionada	CNN+SPP	100 ± 00
Redimensionada	CNN	100 ± 02
Redimensionada	Inception	100 ± 00

Tabela D.7: Tabela com recall da classe Estromatólito avaliando por plugue no conjunto de treino

Tipo de Entrada	Modelo	Estromatólito (%)
Métodos Clássicos	RNA	79 ± 19
Original	CNN+SPP	100 ± 02
Redimensionada	CNN+SPP	100 ± 00
Redimensionada	CNN	100 ± 00
Redimensionada	Inception	100 ± 00

Tabela D.8: Tabela com recall da classe Grainstone avaliando por plugue no conjunto de treino

Tipo de Entrada	Modelo	Grainstone (%)
Métodos Clássicos	RNA	71 ± 18
Original	CNN+SPP	95 ± 13
Redimensionada	CNN+SPP	100 ± 00
Redimensionada	CNN	100 ± 00
Redimensionada	Inception	99 ± 03

D.2 Validação

Nessa seção, apresentam-se os resultados obtidos no conjunto de validação.

D.2.1 Avaliação por Imagem

Tabela D.9: Tabela com acurácia e índice SP dos modelos ao avaliar por imagem individual no conjunto de validação

Tipo de Entrada	Modelo	Acurácia (%)	Índice SP (%)
Métodos Clássicos	RNA	$60,9 \pm 10,4$	$57,7 \pm 11,8$
Imagem Original	CNN+SPP	$69,8 \pm 10,8$	$68,5 \pm 11,4$
Imagem Redimensionada	CNN+SPP	$80,1 \pm 07,6$	$79,7 \pm 07,3$
Imagem Redimensionada	CNN	$65,1 \pm 06,8$	$62,6 \pm 07,3$
Imagem Redimensionada	Inception	$87,6 \pm 09,6$	$87,5 \pm 09,4$

Tabela D.10: Tabela com recall da classe Esferulítico avaliando por imagem individual no conjunto de validação

Tipo de Entrada	Modelo	Esferulítico (%)
Métodos Clássicos	RNA	70 ± 22
Original	CNN+SPP	67 ± 17
Redimensionada	CNN+SPP	78 ± 14
Redimensionada	CNN	45 ± 12
Redimensionada	Inception	84 ± 15

Tabela D.11: Tabela com recall da classe Estromatólito avaliando por imagem individual no conjunto de validação

Tipo de Entrada	Modelo	Estromatólito (%)
Métodos Clássicos	RNA	64 ± 19
Original	CNN+SPP	73 ± 19
Redimensionada	CNN+SPP	77 ± 19
Redimensionada	CNN	74 ± 16
Redimensionada	Inception	85 ± 19

Tabela D.12: Tabela com recall da classe Grainstone avaliando por imagem individual no conjunto de validação

Tipo de Entrada	Modelo	Grainstone (%)
Métodos Clássicos	RNA	47 ± 24
Original	CNN+SPP	68 ± 19
Redimensionada	CNN+SPP	86 ± 19
Redimensionada	CNN	73 ± 15
Redimensionada	Inception	95 ± 06

D.2.2 Avaliação por Plugue

Tabela D.13: Tabela com acurácia e índice SP dos modelos ao avaliar por plugue no conjunto de validação

Tipo de Entrada	Modelo	Acurácia (%)	Índice SP (%)
Métodos Clássicos	RNA	66 ± 13	60 ± 20
Imagem Original	CNN+SPP	77 ± 13	76 ± 14
Imagem Redimensionada	CNN+SPP	90 ± 09	90 ± 09
Imagem Redimensionada	CNN	77 ± 12	70 ± 22
Imagem Redimensionada	Inception	92 ± 11	92 ± 12

Tabela D.14: Tabela com recall da classe Esferulítico avaliando por plugue no conjunto de validação

Tipo de Entrada	Modelo	Esferulítico (%)
Métodos Clássicos	RNA	73 ± 22
Original	CNN+SPP	78 ± 24
Redimensionada	CNN+SPP	$90 \pm 15, 54$
Redimensionada	CNN	50 ± 26
Redimensionada	Inception	91 ± 19

Tabela D.15: Tabela com recall da classe Estromatólito avaliando por plugue no conjunto de validação

Tipo de Entrada	Modelo	Estromatólito (%)
Métodos Clássicos	RNA	69 ± 23
Original	CNN+SPP	78 ± 23
Redimensionada	CNN+SPP	93 ± 21
Redimensionada	CNN	87 ± 18
Redimensionada	Inception	86 ± 20

Tabela D.16: Tabela com recall da classe Grainstone avaliando por plugue no conjunto de validação

Tipo de Entrada	Modelo	Grainstone (%)
Métodos Clássicos	RNA	53 ± 27
Original	CNN+SPP	76 ± 26
Redimensionada	CNN+SPP	98 ± 08
Redimensionada	CNN	89 ± 22
Redimensionada	Inception	100 ± 00

D.3 Teste

Nessa seção, apresentam-se os resultados obtidos no conjunto de teste.

D.3.1 Avaliação por Imagem

Tabela D.17: Tabela com acurácia e índice SP dos modelos ao avaliar por imagem individual no conjunto de teste

Tipo de Entrada	Modelo	Acurácia (%)	Índice SP (%)
Métodos Clássicos	RNA	$54,4 \pm 10,1$	$50,0 \pm 11,8$
Imagem Original	CNN+SPP	$64,0 \pm 10,8$	$62,5 \pm 11,7$
Imagem Redimensionada	CNN+SPP	$75,5 \pm 08,78$	$75,1 \pm 08,7$
Imagem Redimensionada	CNN	$61,7 \pm 07,3$	$59,2 \pm 07,5$
Imagem Redimensionada	Inception	$81,0 \pm 11,9$	$79,9 \pm 11,9$

Tabela D.18: Tabela com recall avaliando por imagem da classe Esferulitito no conjunto de teste

Tipo de Entrada	Modelo	Esferulitito (%)
Métodos Clássicos	RNA	$64,94 \pm 26,30$
Original	CNN+SPP	$58,47 \pm 20,63$
Redimensionada	CNN+SPP	$71,79 \pm 14,39$
Redimensionada	CNN	$41,74 \pm 11,45$
Redimensionada	Inception	$72,84 \pm 20,80$

Tabela D.19: Tabela com recall avaliando por imagem da classe Estromatólito no conjunto de teste

Tipo de Entrada	Modelo	Estromatólito (%)
Métodos Clássicos	RNA	57 ± 19
Original	CNN+SPP	68 ± 14
Redimensionada	CNN+SPP	73 ± 16
Redimensionada	CNN	71 ± 16
Redimensionada	Inception	82 ± 21

Tabela D.20: Tabela com recall avaliando por imagem da classe Grainstone no conjunto de teste

Tipo de Entrada	Modelo	Grainstone (%)
Métodos Clássicos	RNA	41 ± 27
Original	CNN+SPP	64 ± 17
Redimensionada	CNN+SPP	83 ± 18
Redimensionada	CNN	70 ± 14
Redimensionada	Inception	87 ± 13

D.3.2 Avaliação por Plugue

Tabela D.21: Tabela com acurácia e índice SP dos modelos ao avaliar por plugue no conjunto de teste

Tipo de Entrada	Modelo	Acurácia (%)	Índice SP (%)
Métodos Clássicos	RNA	57 ± 13	45 ± 27
Imagem Original	CNN+SPP	67 ± 14	62 ± 21
Imagem Redimensionada	CNN+SPP	81 ± 11	81 ± 12
Imagem Redimensionada	CNN	70 ± 15	57 ± 31
Imagem Redimensionada	Inception	84 ± 13	83 ± 14

Tabela D.22: Tabela com recall avaliando por plugue da classe Esferulitito no conjunto de teste

Tipo de Entrada	Modelo	Esferulitito (%)
Métodos Clássicos	RNA	68 ± 27
Original	CNN+SPP	63 ± 32
Redimensionada	CNN+SPP	81 ± 21
Redimensionada	CNN	38 ± 26
Redimensionada	Inception	77 ± 25

Tabela D.23: Tabela com recall avaliando por plugue da classe Estromatólito no conjunto de teste

Tipo de Entrada	Modelo	Estromatólito (%)
Métodos Clássicos	RNA	59 ± 24
Original	CNN+SPP	69 ± 19
Redimensionada	CNN+SPP	77 ± 20
Redimensionada	CNN	85 ± 20
Redimensionada	Inception	84 ± 21

Tabela D.24: Tabela com recall avaliando por plugue da classe Grainstone no conjunto de teste

Tipo de Entrada	Modelo	Grainstone (%)
Métodos Clássicos	RNA	$44,44 \pm 31,96$
Original	CNN+SPP	$68,89 \pm 26,16$
Redimensionada	CNN+SPP	$87,78 \pm 22,29$
Redimensionada	CNN	$83,33 \pm 19,08$
Redimensionada	Inception	$92,22 \pm 14,34$

Apêndice E

Produção Científica

Neste apêndice, lista-se o resumo publicado em um congresso internacional sobre o tema deste projeto.

E.1 Congressos

E.1.1 Artigo - Sociedade Brasileira de Geofísica - 2019

Artigo aprovado para 16th Congresso Internacional da Sociedade Brasileira de Geofísica (SBGF) no ano de 2019

Título: Convolutional Neural Network for micro-CT image classification of carbonate rocks samples

Autores: ANJOS, C. E. M.; VARGAS, M. R.; MARTINS NETA, A. P.; MARTINS, J. S.; MEDEIROS, L. C.; EVSUKOFF, A. G.

Resumo: Developing the Brazilian pre-salt reservoirs remains a challenge, mainly because of its geological peculiarities, as the heterogeneity of carbonate rocks. These difficulties stimulate the application of several methods that help in the characterization of reservoirs, such as X-ray computed tomography (CT). It is a useful non-destructive technique for imaging features within rocks, based on variations in material compositions. Another technique that has been studied is artificial intelligence, its implementation grants to automate some activities related to oil exploration. Given the above, this work aims to propose a convolutional neural network for lithology classification using microtomography images of carbonate rock samples from Brazilian pre-salt. For that, a set of images of 60 rock samples with distinct characteristics was used. The results obtained achieved 52.46% of accuracy on the validation set, which were within expectation, given the level of uncertainty associated with the lithological classification in this type of reservoir. More complex models and different image preprocessing are a future research line and are already being investigated.