Universidade Federal da Paraíba

Centro de Tecnologia

Programa de Pós-Graduação em Engenharia Mecânica

Mestrado - Doutorado

CLASSIFICAÇÃO DE PADRÕES EM IMAGENS SÍSMICAS UTILIZANDO INTELIGÊNCIA ARTIFICIAL

por

José Fabrício Lima de Souza

Tese de Doutorado apresentada à Universidade Federal da Paraíba para obtenção do grau de Doutor.

João Pessoa – Paraíba **Dezembro, 2019**

JOSÉ FABRÍCIO LIMA DE SOUZA

CLASSIFICAÇÃO DE PADRÕES EM IMAGENS SÍSMICAS UTILIZANDO INTELIGÊNCIA ARTIFICIAL

Tese apresentada ao curso de Pós – Graduação em Engenharia Mecânica da Universidade Federal da Paraíba, em cumprimento às exigências para obtenção do Grau de Doutor.

Orientador: Dr. Moisés Dantas dos Santos

João Pessoa – Paraíba **Dezembro, 2019**

Catalogação na publicação
Seção de Catalogação e Classificação

S729c Souza, Jose Fabricio Lima de.
Classificação de padrões em imagens sísmicas utilizando inteligência artificial / Jose Fabricio Lima de Souza. - João Pessoa, 2020.
80 f. : il.

Orientação: Moisés Dantas dos Santos.
Tese (Doutorado) - UFPB/CT.

1. Imagem sísmica. 2. Reconhecimento de padrões. 3. Aprendizado de máquina. I. Santos, Moisés Dantas dos. II. Título.

UFPB/BC

CLASSIFICAÇÃO DE PADRÕES EM IMAGENS SÍSMICAS UTILIZANDO INTELIGENCIA ARTIFICIAL

por

JOSE FABRICIO LIMA DE SOUZA

Tese aprovada em 20 de dezembro de 2019

Prof. Dr. MOISES DANTAS DOS SANTOS
Orientador - UFPB

Prof. Dr. ABEL CAVALCANTE LIMA FILHO
Examinador Interno - UFPB

Prof. Dr. GUSTAVO CHARLES PEIXOTO DE OLIVEIRA
Examinador Interno - UFPB

Prof. Dr. LEONARDO VIDAL BATISTA
Examinador Externo – UFPB

Prof. Dr. JOSÉ ANTONIO BARBOSA
Examinador Externo - UFPE

João Pessoa – PB

2019

AGRADECIMENTOS

A Deus, que em sua infinita misericórdia, me fortaleceu em todo o processo.

A minha esposa Luiza Paula e a minha filha Késia Anacleto, pela paciência de ambas, que foi fundamental durante os momentos difíceis.

Aos meus familiares que acreditaram que eu chegaria lá.

Ao meu orientador Moisés Dantas dos Santos pela sua competência e praticidade durante todo o processo.

A Edvaldo Neto e a Gabriel pelo apoio durante as etapas de programação, devo muito a vocês por isso.

Ao Gustavo Peixoto pelas críticas construtivas.

Aos demais professores do curso de pós-graduação pelo empenho na ministração de suas disciplinas.

A todos os amigos que me apoiaram para que esse dia acontecesse.

Parafraseando 1Co 3.6

“Eu ralei, Moisés me orientou, muitos me ajudaram, mas foi Deus que deu todo o conhecimento”

CLASSIFICAÇÃO DE PADRÕES EM IMAGENS SÍSMICAS UTILIZANDO INTELIGÊNCIA ARTIFICIAL

RESUMO

A classificação de regiões com maior probabilidade de acúmulo de hidrocarbonetos é um procedimento que envolve uma análise especializada de dados geofísicos e geológicos das bacias sedimentares. Parte da análise desses dados é realizada através da interpretação de dados obtidos com o método de sísmica de reflexão, sendo uma etapa que requer uma quantidade de tempo considerável, além de ser uma tarefa trabalhosa, mesmo para um intérprete experiente. Detectar zonas propícias a acumulação de hidrocarbonetos ("plays e leads") sob o ponto de vista da visão computacional é um tema emergente que demanda enormes desafios. O objetivo deste trabalho foi avaliar abordagens alternativas para a classificação automática de regiões que apresentem a possibilidade de acúmulo de hidrocarbonetos, a partir do uso de técnicas de aprendizado de máquina para a identificação de padrões em imagens sísmicas. Nesse sentido, foram utilizados Redes Neurais Artificiais (RNA), Redes Neurais Convolucionais (CNN) e segmentação semântica com uma arquitetura do tipo U-Net. Foi empregado um banco de imagens sísmicas provenientes da Bacia de Sergipe-Alagoas (nordeste do Brasil) como imagens de entrada para treinamento, validação e teste. Indicadores de desempenho tais como acurácia, precisão, recall, F_1-*Score*, erro e *IoU* foram utilizados para avaliar a rede durante a fase de treinamento e validação. Os resultado se mostraram bastante satisfatórios, principalmente envolvendo a CNN e a U-Net, e esta última apresentou um resultado mais significativo.

Palavras-chave: imagem sísmica, reconhecimento de padrões, aprendizado de máquina.

CLASSIFICATION OF PATTERNS IN SEISMIC IMAGES USING ARTIFICIAL INTELLIGENCE

Abstract

The classification of regions most likely to accumulate hydrocarbons is a procedure that involves a specialized analysis of geophysical and geological data from sedimentary basins. Part of the analysis of these data is performed through the interpretation of data obtained with the seismic reflection method, being a step that requires a considerable amount of time, in addition to being a laborious task, even for an experienced interpreter. Detecting areas conducive to the accumulation of hydrocarbons ("plays and leads") from the point of view of computer vision is an emerging theme that demands enormous challenges. The objective of this work was to evaluate alternative approaches for the automatic classification of regions that present the possibility of accumulation of hydrocarbons, using machine learning techniques to identify patterns in seismic images. In this sense, Artificial Neural Networks (RNA), Convolutional Neural Networks (CNN) and semantic segmentation with a U-Net architecture. A database of seismic images from the Sergipe-Alagoas Basin (northeastern Brazil) was used as input images for training, validation and testing. Performance indicators such as accuracy, precision, recall, F_1-$Score$, error and IoU were used to assess the network during the training and validation phase. The results were quite satisfactory, mainly involving CNN and U-Net, and the latter showed a more significant result.

Keywords: seismic imaging, pattern recognition, machine learning.

S U M Á R I O

LISTA DE FIGURAS

LISTA DE TABELAS

ABREVIAÇÕES

$\boldsymbol{acc}$ - acurácia

ANP: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis

CNN: *Convolutional Network Neural*

ELU: *Exponencial Linear Unit*

$\boldsymbol{F_1}$ - F_1-$Score$

IA: Inteligência Artificial

$\boldsymbol{prec}$ - precisão

RELU: *Rectifier Linear Unit*

$\boldsymbol{rec}$**:** revocação

RNAs: Redes Neurais Artificiais

LISTA DE SÍMBOLOS

$\boldsymbol{A}, \boldsymbol{B}$ – conjunto quaisquer

$\boldsymbol{A} \cap \boldsymbol{B}$ – interseção entre dois conjuntos

$\boldsymbol{A} \cup \boldsymbol{B}$ – união entre dois conjuntos

$\boldsymbol{E}$ – erro médio quadrático

$\boldsymbol{d}$ - constante

$\boldsymbol{f(x)}$ – função

$\boldsymbol{FP}$ – falso positivo

$\boldsymbol{FN}$ – falso negativo

$\boldsymbol{I_c}$ – índice de confiança

$\boldsymbol{IoU}$: Interseção sobre a união

$\boldsymbol{l_r}$ – taxa de aprendizagem

$\boldsymbol{\mathcal{L}}$ - função de limiarização

$\boldsymbol{m_c}$ – termo do momento

$\boldsymbol{\mathcal{P}}$ – número de pixels

$\boldsymbol{\mathcal{P}_w}$ – número de pixels brancos

$\boldsymbol{P_d^\theta}$ – matriz de co-ocorrência

$\boldsymbol{\widehat{P}_d^\theta}$ – matriz de co-ocorrência normalizada

$\boldsymbol{p(i, j, \theta, d)}$ – elementos da matriz de co-ocorrência

$\boldsymbol{N_g}$ – número total de níveis do atributo

$\boldsymbol{S}$ – somatório dos termos de entrada de um nó

$\boldsymbol{SQ}$ – soma de quadrados

$\boldsymbol{VP}$ – verdadeiro positivo

$\boldsymbol{VN}$ – verdadeiro negativo

$\boldsymbol{\alpha}$ – alfa

$\boldsymbol{\beta}$ – parâmetro de controle

$\boldsymbol{\lambda}$ – lambda

$\boldsymbol{\theta}$ – teta

CAPÍTULO I

1. Introdução

No momento em que você estiver lendo esse texto já seremos mais 7,5 bilhões de pessoas no mundo. Segundo a Wikipédia, a população mundial cresce a uma taxa de 1,2% ao ano, e com isso, maior a demanda de energia necessária para manter tudo em pleno funcionamento, o que impulsiona cada vez mais pesquisas por fontes renováveis. Embora as pesquisas nesse campo tenham avançado bastante, os combustíveis fósseis continuarão sendo a maior fonte de energia para atender a essa demanda nas próximas três décadas.

A grande dificuldade se encontra no fato de que boa parte das reservas restantes se encontram em regiões de grande profundidade (THOMAS, 2004). A determinação da localização dessas regiões é realizada através de interpretação de imagens obtidas pelo método sísmico de reflexão, "método este que auxilia a modelar as condições de formação e acumulação de hidrocarbonetos em subsuperfície" (MATOS, 2004). A redução do risco na determinação dessas acumulações é o grande desafio do setor de petróleo e gás. Essa é uma tarefa que requer tempo e um estudo minucioso de dados geofísicos e geológicos (SONG *et. al*., 2017). Somente após exaustivo prognóstico do comportamento de diversas camadas em subsuperfície é que os geólogos e geofísicos indicam potenciais regiões para a realização de perfuração de um poço, etapa esta que envolve um alto custo de realização. Por exemplo, a taxa diária paga pelo aluguel de sondas de perfuração *offshore* tem impacto cada vez maior no custo total do projeto de desenvolvimento de um campo: sondas mais modernas, capazes de perfurar em profundidades de água superiores a 2.000 metros são alugadas, segundo (GRIFFTHS, 2015) citado por (SIMÕES, 2017), por quase US$ 500.000,00 por dia.

Com o advento da tecnologia, a área da Inteligência Artificial, se torna um processo fundamental para a fase de interpretação dessas regiões de interesse no tocante a tomada de decisões do intérprete. Cada vez mais o campo de reconhecimento de imagens através de

inteligência artificial vem ganhando espaço no campo da pesquisa científica, como por exemplo, na identificação de tumores, de digitais, de objetos e até mesmo no desenvolvimento de veículos autônomos.

Logo, o desenvolvimento de técnicas direcionadas a análise e interpretação de imagens através do reconhecimento de padrões vêm se tornando um recurso importante para a indústria petrolífera, pois permite que especialistas elaborem interpretações refinadas a partir de dados sísmicos (BARNES et al., 2002; DU et al., 2015). Portanto, o desenvolvimento de técnicas que auxiliem os intérpretes na tomada de decisão em relação a novas regiões de exploração que tornem mais rápidas e eficientes à interpretação é de grande valia para a engenharia de petróleo.

Assim, o objetivo principal deste trabalho é apresentar técnicas de aprendizado de máquina para o reconhecimento de padrões em imagens sísmicas, e com isso, auxiliar o intérprete na tomada de decisão para identificar regiões favoráveis à formação de hidrocarbonetos.

1.1 Objetivo Geral

O objetivo geral desta tese é identificar de forma automática regiões favoráveis ao acúmulo de hidrocarbonetos.

1.2 Objetivo Específico

Utilizar ferramentas de Inteligência Artificial, a saber, Redes Neurais Artificias, Redes Neurais Convolucionais e ferramentas *autoenconders* que auxiliem na tomada de decisão quanto a identificação de regiões favoráveis ao acúmulo de hidrocarbonetos.

1.3 Estado da Arte

Detectar com precisão regiões de acúmulo de hidrocarbonetos é uma atividade que envolve muito tempo, mesmo para um intérprete experiente (ROY *et al.,* 2014; SONG, *et al.*, 2017). Outra dificuldade é o fato de que na maioria das vezes essas interpretações recaem na subjetividade do intérprete (ZHAO *et al.*, 2015) o que provoca a necessidade de melhorar cada vez mais a precisão quanto a identificação das regiões de acúmulo de hidrocarbonetos, tendo em vista o alto custo de perfuração.

Embora as técnicas automáticas de reconhecimento de padrões ainda não sejam consideradas a "última palavra" nas decisões tomadas por especialistas, elas vem se tornando ferramentas importantes no apoio à tomada de decisões (COLÉOU et al., 2003). Para um conhecimento mais aprofundado dessas técnicas de reconhecimento de padrões ZHAO *et al.* (2015) apresenta as vantagens e desvantagens das técnicas mais utilizadas na análise de imagens sísmicas.

Entre as técnicas utilizadas destaca-se o uso de ferramentas de aprendizado de máquina, tais como: Redes Neurais Artificiais supervisionadas, como por exemplo, as redes *Perceptron* de Multiplas Camadas (PMC*)* e as redes não supervisionadas, dentre elas temos os Mapas Auto-Organizáveis (da sigla, SOM, *Self-organized map*, em inglês). Com o surgimento da Rede Neural de Aprendizado Profundo (*Deep Learning*), em particular, as Redes Neurais Convolucionais – CNN (do inglês *Convolutional Neural Network*) se tornara a técnica mais utilizada como ferramenta de análise sísmica, tendo em vista a vantagem de não necessitar de descritores para alimentar a rede.

De acordo com (MA e GOMES, 2015), as redes neurais artificiais tem ganhado muita popularidade na classificação de imagens sísmicas tendo em vista o seu poder de identificação e sua capacidade de usar múltiplas entradas.

Os Mapas Auto Organizáveis como análise e interpretação de imagens sísmicas são ferramentas bastante utilizada no auxílio de clusterização de regiões que apresentam características semelhantes (STRECKER e HUDEN, 2002; MATOS et al., 2007, 2010; ROY et al., 2010; ROY, 2013). Por se tratarem de uma rede não supervisionada, a classificação dos dados sísmicos é baseada inteiramente nas caracteríticas da resposta, sem exigir o uso de qualquer informação de poços. Nos Mapas Auto Organizáveis os dados de entrada da rede são obtidos a partir de atributos sísmicos (CHOPRA E MARFURT, 2006; TANER, 2001). Os dados de entrada podem ser amplitudes sísmicas (MATOS, 2010; PRIEZZHEV e

MANRAL, 2012), atributos de textura em matrizes de co-ocorrência em níveis de cinza (ÂNGELO *et al.*, 2009; MATOS *et al.*, 2011) e formas de ondas (ROY et al., 2010; SARASWAT et al., 2012; SONG et al., 2017). Segundo SAGGAF (2003) a ferramenta SOM é bastante útil para prospecção de petróleo em regiões onde existe pouca ou nenhuma informação da região a ser estudada. Outra vantagem da ferramenta SOM é a capacidade de clusterização dos dados, que podem ser armazenados de forma "amigável ao intérprete" (ZHAO et al., 2015), os clusters ordenados podem ser mapeados para uma barra de cores gradacional (COLÉOU *et al.*, 2003). Porém, segundo DU *et al.* (2015) esse processo de clusterização deve ser feito com cuidado, pois a escolha do número de clusters pode influenciar nos resultados obtidos. Um número muito baixo de clusters pode fornecer resultados grosseiros na classificação, enquanto que um número elevado pode levar a resultados redundantes.

Embora existam diversos trabalhos utilizando redes neurais artificiais, a grande maioria deles se concentram na clusterização de imagens sísmicas através da técnica SOM, deixando a cargo do especialista a identificação das regiões mais propícias a concentração de hidrocarbonetos.

O trabalho de ZHANG *et al.* (2001) faz uso de RNA para classificar traços sísmicos pela análise de formas de onda. Por sua vez, WEST *et al.* (2002) combina análise de textura com uma RNA probalística para quantitativamente mapear fácies sísmicas. Enquanto que CHERAZZI *et al.* (2013) realizaram a identificação de padrões de falhas em dados sísmicos através de RNAs usando como descritores informações de mergulho e azimute obtidos a partir de um cubo de *steering*.

Até meados de 2017 não foram observados trabalhos na obtenção de características que facilitem a obtenção de um padrão de aprendizagem para classificação automática de imagens sísmicas. Porém, com o avanço da tecnologia de aprendizado de máquina através das CNNs, a área da inteligência artificial vem ganhando cada vez mais destaque na análise de dados sísmicos (DI *et al.*, 2018).

A principal diferença entre as CNNs e as redes neurais tradicionais é o fato de que as RNAs necessitam de descritores para realizar a análise da imagem sísmica, diferentemente das CNNs cuja única necessidade de entrada é a imagem a ser analisada, o processo de obtenção das características/atributos é realizado pela própria rede (GOODFELLOW *et al.*, 2016).

A publicação de trabalhos que utilizam as CNNs como ferramenta de análise de imagens sísmicas teve um impulso relativamente grande a partir de 2018. O grande destaque das redes convolucionais têm sido a sua capacidade de generalização (WANG *et al.,* 2018), além de que, diferentemente das redes tradicionais, as CNN podem obter seus próprios atributos para o processo de treinamento da rede (GOODFELLOW *et al.,* 2016), eliminando assim a tarefa de procurar os melhores atributos que possibilitem a generalização da rede.

Quanto a produção científica relativa ao uso de redes profundas, a maioria dos trabalhos se divide em dois tipos de procedimentos: classificação baseada em patches ou segmentação de imagens. A análise das vantagens e desvantagens desses procedimentos podem ser analisadas em (ZHAO *et al.,* 2018).

Com relação à classificação baseada em patches podemos destacar inicialmente o trabalho de (ARAYA-POLO *et al.,* 2017) que faz uso de CNN na identificação de falhas em dados sísmicos em dados sintéticos 3D. Segundo o autor, a justificativa para o uso de dados sintéticos é o fato de que em dados reais a rede neural seria vinculada ao desempenho humano e pela qualidade dos dados. O trabalho de HUANG, *et al.* (2017) por sua vez tratou da identificação de falhas geológicas em imagens 3D a partir de uma combinação de CNN e redes neurais tradicionais. Podemos destacar ainda os trabalhos de WU *et al.* (2018) e POCHET *et al.* (2018) que identificaram falhas a partir da aplicação de CNN. No primeiro caso o treinamento da rede foi realizado a partir de dados sintéticos gerados a partir de funções senoidais e aplicação de ruídos, enquanto que no segundo foram utilizados *patches* de sinal acústico. Os trabalhos de (ZENG *et al.,* 2019; ALAUDAH *et al.,* 2018; LEWIS e VIGH, 2017) fazem uso de CNN para classificação automática de corpos de sal em imagens sísmicas 2D. Enquanto que (SHI *et al.,* 2019; WALDELAND *et al.,* 2018) classificam corpos de sal numa estrutura 3D.

No que diz respeito a segmentação de imagens, destacamos o trabalho de (CHEVITARESE *et al.,* 2018) que fez uso da CNN para um mapeamento estratigráfico de imagens sísmicas. Por sua vez (PETERS *et al.,* 2019) faz uso da CNN para uma segmentação semântica de imagens sísmicas a partir de uma função de perda que possibilita o treinamento da rede com imagens parcialmente rotuladas. Podemos destacar ainda o trabalho de ZENG *et al*. (2019) e SHI *et al.* (2018) que fizeram a segmentação de corpos de sal a partir de CNN. Enquanto que SILVA *et al.* (2019) utilizou CNN para geração de banco de dados segmentados para treinamento de redes neurais.

O nosso trabalho empregou como abordagem deste problema a utilização de ferramentas de reconhecimento de padrões: Redes Neurais Artificiais, Redes Neurais Convolucionais e técnicas de segmentação de imagens a partir de *autoenconders* com uma arquitetura do tipo U-Net para realizar identificação de regiões reconhecidas como portadoras de alto potencial para a acumulação de hidrocarbonetos.

CAPÍTULO II

2. Fundamentação Teórica

O objetivo dessa seção é apresentar as noções teóricas envolvidas nesta pesquisa no que diz respeito a noções de geologia de petróleo, método sísmico de reflexão, redes neurais e alguns classificadores de desempenho da rede.

2.1 Noções de Geologia do Petróleo

De acordo com (THOMAS, 2001) para que ocorra o acúmulo de hidrocarbonetos é necessário alguns pré-requisitos, a saber:

(a) Presença de rocha geradora
(b) Presença de rocha reservatório
(c) Presença de rocha capeadora ou selante
(d) Migração
(e) Trapas ou armadilhas
(f) Relações temporais adequadas

O processo de geração do petróleo é o resultado da transformação de matéria orgânica com a contribuição do fluxo de calor oriundo do interior da terra. Inicialmente é necessária a existência de uma rocha, rica em matéria orgânica capaz de gerar hidrocarbonetos sólidos, líquidos e gasosos (rocha geradora). Também é necessário que exista uma rocha porosa (rocha reservatório) de boa permeabilidade, tais como arenitos e carbonatos. Após o processo de geração do petróleo, é necessário que ocorra a migração do mesmo através da rocha reservatório. Sendo assim, para que ocorra o acúmulo, é necessário que a migração do

mesmo seja interrompida por uma barreira de baixa permeabilidade (rocha capeadora). A Figura 1 é uma representação gráfica desse processo.

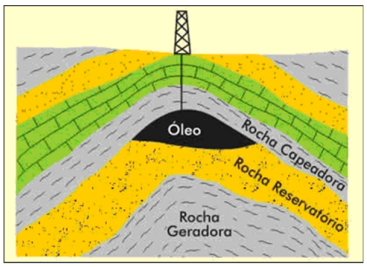


Figura 1: Relações espaciais entre rochas geradora, reservatório e capeadora. (TEIXEIRA *et al.*, 2001)

Para que tenhamos a formação de uma jazida de petróleo se faz necessária a existência de trapas ou armadilhas, que podem ter diferentes origens, características e dimensões (Figura 2). Essas armadilhas podem ser classificadas em estratigráficas, estruturais e mistas, e podem exibir arranjos que combinam tanto características estratigráficas como estruturais.

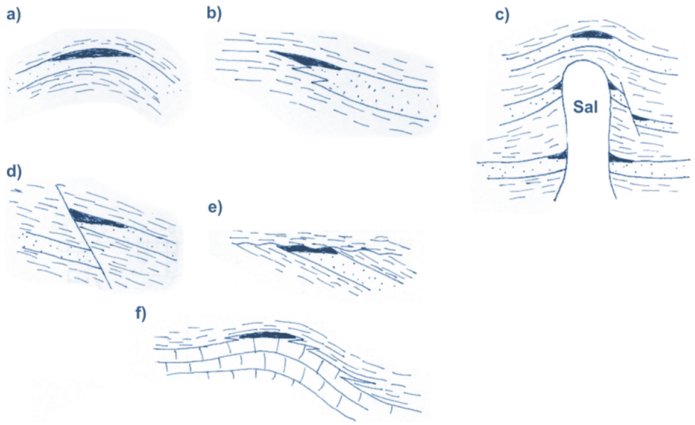


Figura 2: Exemplos de trapas comuns para o petróleo e gás natural. **(a)** Trapa estrutural em anticlinal, **(b)** Trapa estratigráfica associada à variação lateral de fácies e de espessura, **(c)** Trapas associadas a domos de sal, **(d)** Trapa associada à falha, **(e)** Trapa estratigráfica associada à discordância e **(f)** Trapa estratigráfica associada a construções recifais. Modificado de (RAILSBACK, 2015).

Infelizmente, nem sempre a estrutura geológica garante a existência de um acúmulo de hidrocarbonetos. Essas estruturas de acordo com a certeza ou não da presença de hidrocarbonetos recebem alguns nomes específicos, a saber, *play*, *lead* ou prospecto.

O play é um modelo de uma estrutura que poderia reunir os elementos de um reservatório (existe rocha porosa e existe camada selante), e esses elementos estão arranjados em uma geometria que de uma forma geral apresenta alto potencial para o acúmulo de hidrocarbonetos.

O *lead* é um *play* que encontramos em uma bacia que possui um alto potencial de acúmulo de hidrocarbonetos. Ele ainda não foi testado a partir da perfuração de um poço.

O prospecto é um lead que foi perfurado e constatado o acúmulo de hidrocarbonetos.

2.2 Método Sísmico de Reflexão

A base para utilização do Método Sísmico de Reflexão se deve ao fato de que o subsolo é geralmente composto por diferentes camadas de sedimentos. Essas diferentes camadas geológicas são caracterizadas por possuírem diferentes propriedades físicas, entre elas destacamos a impedância acústica; é através dela que o método é aplicado.

Segundo (THOMAS, 2004) entre os métodos de prospecção de hidrocarbonetos, o método sísmico de reflexão é amplamente utilizado na indústria de petróleo. Ele é considerado uma das ferramentas essenciais na descoberta, quantificação e qualificação de depósitos de petróleo e gás, além de serem também utilizados para exploração de aquíferos, jazidas minerais, estudos de meio ambiente, dentre outros.

O objetivo principal do método sísmico é a criação de um modelo de dados que, após processado, permite acessar informações importantes a respeito da geologia da região de exploração. Entre os modelos existentes para a exploração de petróleo o método sísmico de reflexão tem apresentado relativo destaque, por ser um método indireto de exploração da superfície e ser mais barato em comparação aos métodos direto, como a perfuração de poços. Outra vantagem é que ele permite a cobertura de uma vasta área de aquisição.

A seguir vamos tecer um breve comentário sobre como funciona o método sísmico no que diz respeito à aquisição, processamento e interpretação dos dados obtidos. Para maiores detalhes sobre o método sísmico ver (THOMAS, 2004) e (ROBINSON e TREITEL, 1980).

2.2.1 Aquisição

Tanto em terra (Figura 3) quanto no mar (Figura 4), o processo de aquisição consiste na geração de uma perturbação mecânica, em terra esta perturbação pode ser gerada a partir de dinamite ou por um caminhão de vibração, no mar são usados canhões de ar comprimido. Essas ondas se propagam através das camadas rochosas até encontrarem interfaces entre duas camadas de rocha com impedâncias acústicas diferentes. Parte delas então é refratada, continuando a viagem para baixo, e outra parte é refletida, retornando à superfície onde estarão dispostas linhas de receptores (geofones no caso terrestre ou hidrofones no caso marítimo).

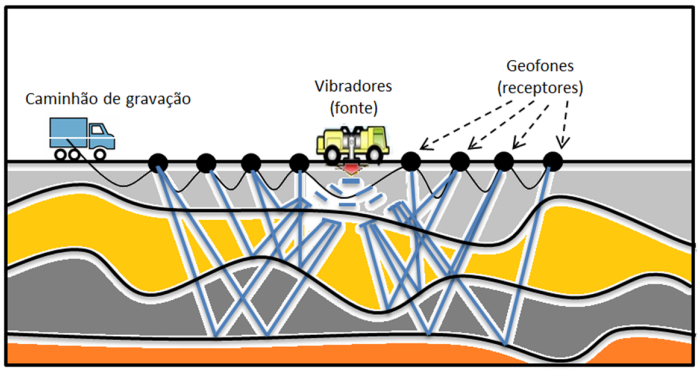


Figura 3: Método sísmico de reflexão terrestre (THOMAS, 2004)

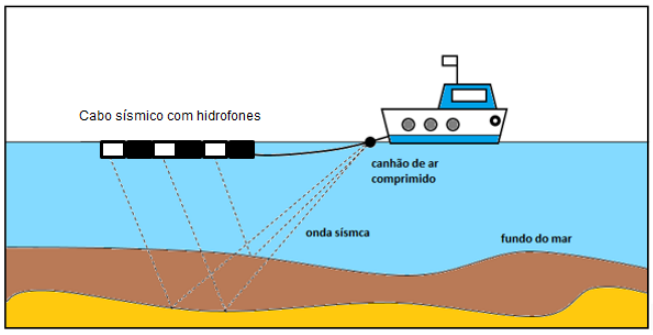


Figura 4: Método sísmico marítimo (THOMAS, 2004)

As ondas refletidas são captadas por receptores que ficam situados em posições específicas. Estes podem ser eletromagnéticos, utilizados para captação em terra, também

conhecidos como geofones. E para captação de dados em regiões oceânicas, são utilizados receptores de pressão, também chamados de hidrofones. As ondas, após captadas pelos sensores, dão origem a traços sísmicos e a composição desses traços são organizados em sismogramas (Figura 5).

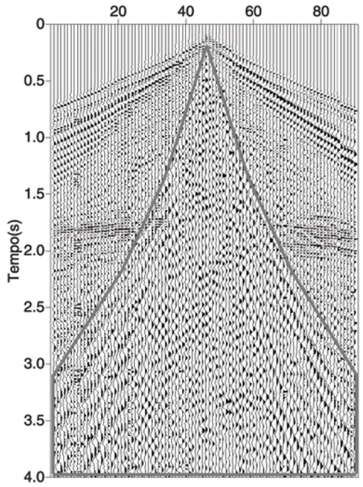


Figura 5: Sismograma fruto de uma família de traços
(SILVA e PORSANI, 2006)

2.2.2 Processamento

De acordo com (THOMAS, 2004) o processamento sísmico tem como objetivo a obtenção de imagens da subsuperfície com a maior precisão possível, atenuando os ruídos orientes do processo. Essa atenuação, segundo (SILVA e PORSANI, 2006) é realizada a partir da utilização de operações matemáticas, entre elas temos correção estática, correção de amplitude e Filtragem F-K. Na correção estática o objetivo é corrigir atrasos ou antecipações nas chegadas das reflexões sísmicas. A correção de amplitude por sua vez é necessária devido a fato de que as ondas sísmicas ao se propagarem sofrerem atenuação devido a absorção do meio, esses efeitos são corrigidos através da aplicação de ganhos de amplitudes dos traços sísmicos. A Filtragem F-K permite a remoção de ruídos com base nas velocidades de propagação. Como exemplo, temos o ruído do tipo *Ground Roll* bastante comum em sismogramas terrestres. Eles estão associados a zonas de baixa velocidade e são

caracterizados nos sismogramas por um padrão linear com elevada amplitude, baixa frequência e baixa velocidade (YILMAZ, 2001).

A Figura 6 mostra uma sequência convencional com as principais etapas do processamento sísmico. Para maiores detalhes dessas etapas ver (THOMAS, 2004 e (GOMES *et al.*, 2011).

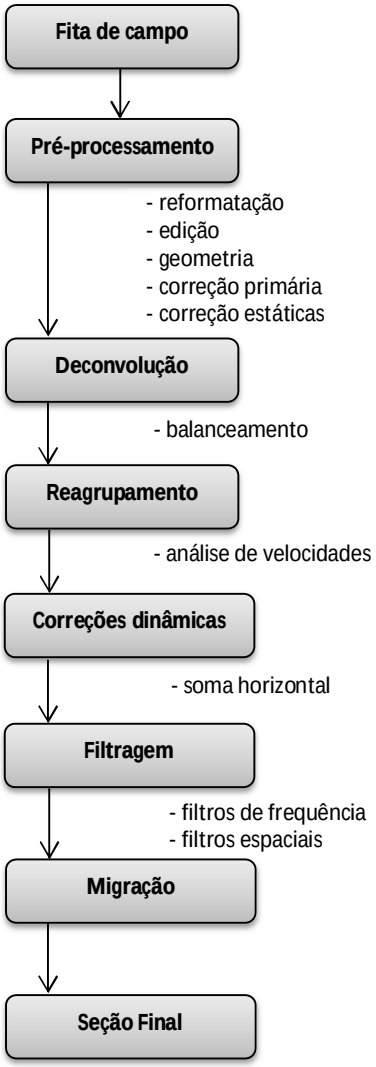


Figura 6: Sequência convencional de Processamento Sísmico
Adaptado de (THOMAS, 2004)

A conclusão do processamento permite obter volumes de imagens 2D e 3D onde geólogos e geofísicos interpretam as imagens obtidas no processo sísmico na busca de

situações favoráveis à acumulação de hidrocarbonetos. A Figura 7 ilustra um exemplo de uma imagem ou seção sísmica.

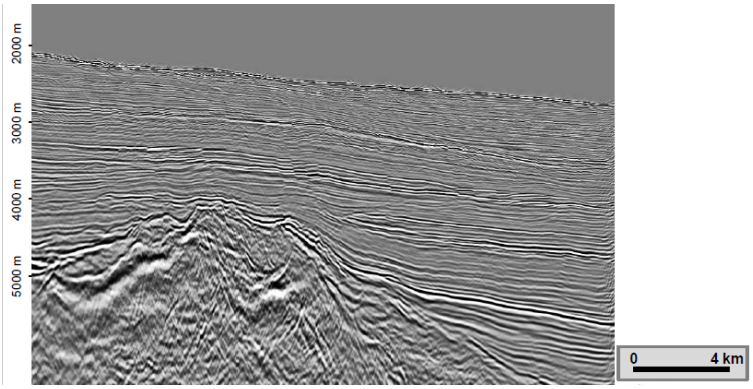


Figura 7: Imagem sísmica após a etapa de processamento.
FONTE: Bacia Sergipe-Alagoas

2.2.3 Interpretação

Após processadas, as seções sísmicas migradas na escala do tempo ou em profundidade são enviadas aos intérpretes, para que estes possam inferir sobre as possíveis estruturas, tais como domos ou diápiros de sal, presença de falhas, dobramento entre outras. A localização dessas estruturas é de fundamental importância, pois estas permitem a criação de um modelo geológico da área em estudo. E, consequentemente, essas feições geológicas são importantíssimas no processo de geração dos hidrocarbonetos. A Figura 8 mostra a interpretação de feições geológicas predefinidas a partir de uma seção sísmica. De acordo com DE CASTRO e HOLTZ (2004), a interpretação sísmica pode ser classificada, de acordo com o foco, em dois tipos: estrutural e estratigráfica. A interpretação estrutural basicamente tenta identificar as camadas geológicas ou, de forma equivalente, as interfaces entre as camadas, bem como as falhas geológicas que recortam as camadas. Na interpretação estratigráfica o foco do trabalho está em entender a maneira como as camadas se formaram ao longo do tempo.

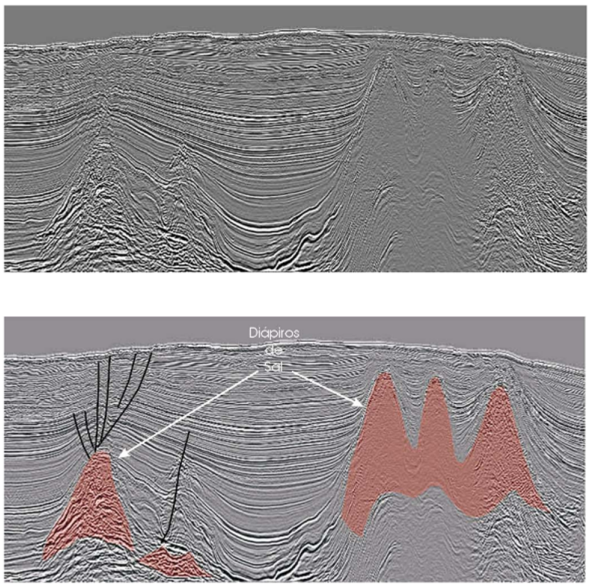


Figura 8: Interpretação sísmica da diápiros de sal: Na parte superior temos uma imagem sísmica não interpretada, enquanto que na parte inferior temos a interpretação da imagem identificando diápiros de sal. (DE CASTRO e HOLTZ, 2004)

2.3 Inteligência Artificial

Fornecer o conceito correto de Inteligência Artificial (IA) não é algo simples, de acordo com PATTERSON e GIBSON (2017) podemos entender a IA como sendo da área computação que procura reproduzir, por meios computacionais, atividades humanas como raciocinar, planejar, aprender e resolver problemas. De forma bem resumida, a IA é a tentativa de reproduzir em uma máquina a capacidade humana de ser inteligente. Entre as formas de IA podemos destacar as Redes Neurais Artificiais e as Redes Neurais Convolucionais.

2.3.1 Redes Neurais Artifiais

Segundo GOODFELLOW *et al.* (2016) o desenvolvimento da teoria das Redes Neurais Artificiais (RNAs) teve início através do estudo do funcionamento do cérebro humano, Figura 9(a). O pontapé inicial foi dado pelo neurofisiologista Warren McCulloch e o matemático Walter Pitts quando escreveram um artigo sobre como os neurônios poderiam funcionar e para isso, eles modelaram uma rede neural simples usando circuitos elétricos. Para criação do modelo, eles utilizaram conceitos matemáticos e algoritmos lógicos denominados lógica de limiar (em inglês, *threshold logic*). A partir deste modelo foram desenvolvidas duas linhas de pesquisa, uma focada em processos biológicos do cérebro e a outra na aplicação de redes neurais artificiais.

De acordo com (PATTERSON e GIBSON, 2017) as RNAs são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de seres humanos e que adquirem conhecimento através da experiência. Elas são formadas por um conjunto de neurônios artificiais que têm a capacidade de coletar, utilizar e armazenar informações baseadas em aprendizagem. Estes neurônios estão organizados em três camadas: camada de entrada, camada oculta e camada de saída, conforme Figura 9(b).

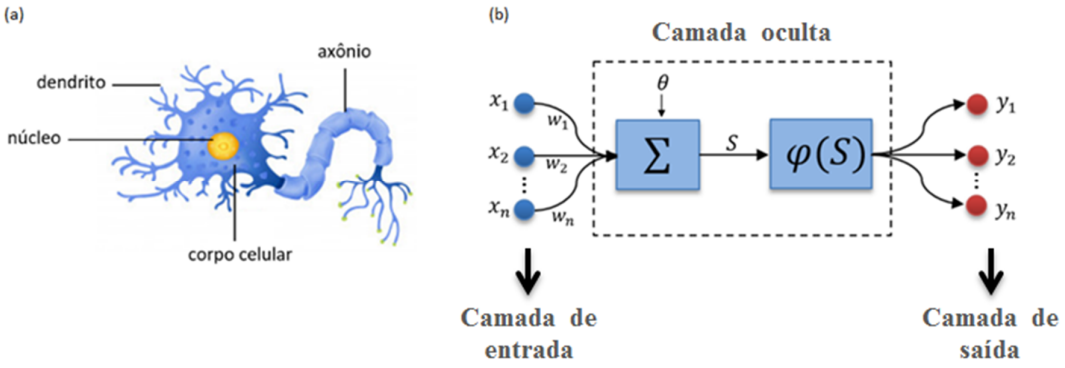


Figura 9: Estrutura de neurônios: (a) biológico e (b) artificial.
(Adaptado de GOODFELLOW *et al.* (2016))

A aprendizagem da rede é realizada a partir de ajustes dos pesos sinápticos da rede de forma a alcançar o resultado desejado. Basicamente, a arquitetura de uma rede neural é composta por neurônios (nós) conectados por meio de links direcionados, como ilustra a Figura 10.

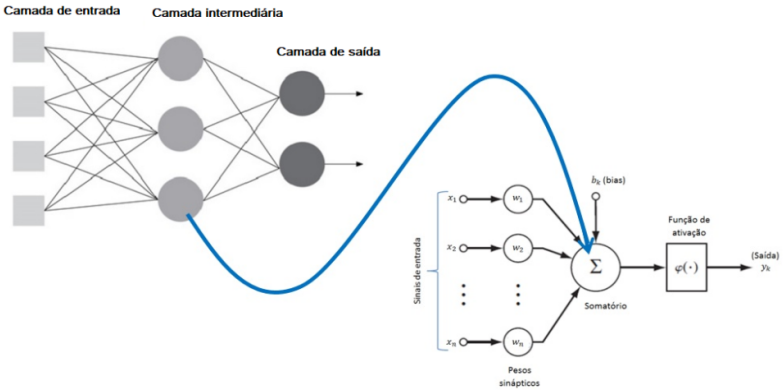


Figura 10: Arquitetura de uma RNA

Adaptado de PATTERSON e GIBSON (2017)

Em cada nó da camada intermediária são agregados todos os sinais de entrada que foram ponderados pelos pesos sinápticos em conjunto com o limiar de ativação (bias). Uma função de ativação, cujo objetivo é limitar a saída dentro de um intervalo de valores razoáveis a ser assumido, é aplicada ao somatório S indicado em (1), e assim é obtido um valor de saída.

$$\boldsymbol{S = \sum_{i=1}^{n} x_i . w_i + \theta} \qquad \textbf{(1)}$$

As redes neurais aprendem por comparação entre os valores de saída da rede e o valor desejado para a resposta. A comparação é feita pela minimização da função erro médio quadrático (GOODFELLOW *et al.* 2016).

$$\boldsymbol{E = \frac{1}{2}(y - \hat{y})} \qquad (2)$$

Onde E é o erro, y é o valor de saída da rede e $\hat{y}$ é o valor desejado.

Entre os processos de aprendizagem um dos mais utilizados é o *backpropagation* onde a minimização do erro é realizada a partir de ajustes, em cada iteração, nas matrizes de pesos.

Como descrito no conceito de redes neurais, eles aprendem através da experiência, logo, para isso é necessário uma etapa de treinamento da rede, o que se faz necessário um número razoável de amostras, chamadas de conjunto de treinamento, conjunto este que ainda pode ser subdividido em subconjunto de treinamento e subconjunto de validação, cujo objetivo é verificar o desempenho da rede durante a fase de treinamento. Depois de treinada, a capacidade de generalização da rede é avaliada a partir de dados não observados por ela, este conjunto é chamado de conjunto de teste. A avaliação do desempenho de uma rede é realizada a partir de métricas, entre elas, podemos destacar a acurácia, a especificidade, a sensibilidade, a precisão, índice de Jaccar entre outras. Para se ter mais detalhes sobre essas métricas, verificar a Seção 2.3.4.

2.3.2 Redes Neurais Convolucionais

Os avanços das tecnologias digitais da informação e, mais especificamente, das Unidades Gráficas de Processamento, bem como o aumento exponencial da quantidade de dados disponível em formato digital levaram ao desenvolvimento recente das Redes Neurais Artificiais Profundas, desenvolvidas pelo grupo de pesquisa de LeCun (LECUN *et al.,* 2015) no final da década de 90 para realizar o reconhecimento de dígitos em cheques. Isso causou uma mudança progressiva dos métodos clássicos de reconhecimento de padrões, sendo substituídos gradativamente por técnicas de Deep Learning, que tem alcançado resultados superiores (KRIZHEVSKY *et al.,* 2012). Quanto ao termo aprendizado profundo, não há um consenso sobre a distinção precisa entre redes profundas e rasas. De modo relativamente ambíguo, afirma-se que redes profundas são redes que possuem muitas camadas de treinamento sobre enormes massas de dados disponíveis (LECUN *et al.,* 2015).

Nas técnicas clássicas de aprendizagem de máquina, as imagens a serem analisadas são primeiramente convertidas em vetores de atributos ou características, e estes vetores alimentam um sistema de reconhecimento de padrões (DUDA *et al.,1999*). Em uma Rede Neural Convolucional (CNN) (em inglês: *Convolutional Neural Network*), a própria imagem (pré-processada ou não) alimenta diretamente a primeira camada da rede (Figura 11), que extrai automaticamente, na etapa de treinamento, um conjunto de atributos de baixo nível, codificados na forma de filtros convolucionais.

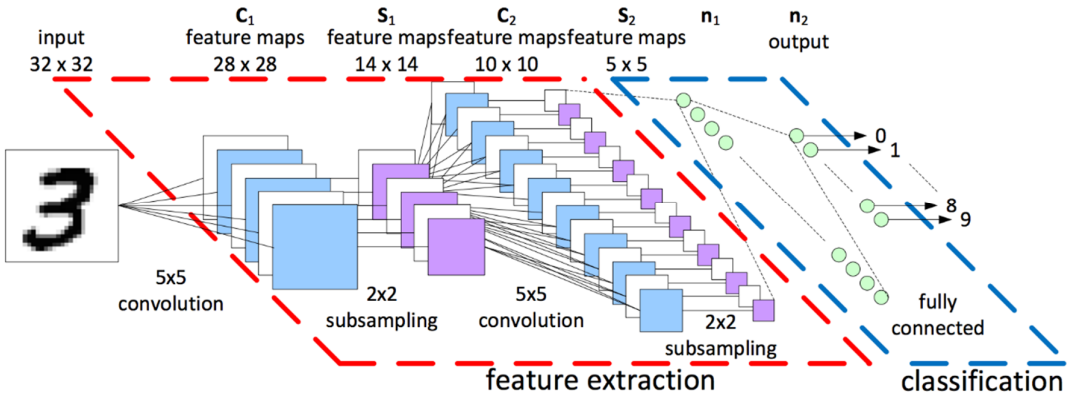
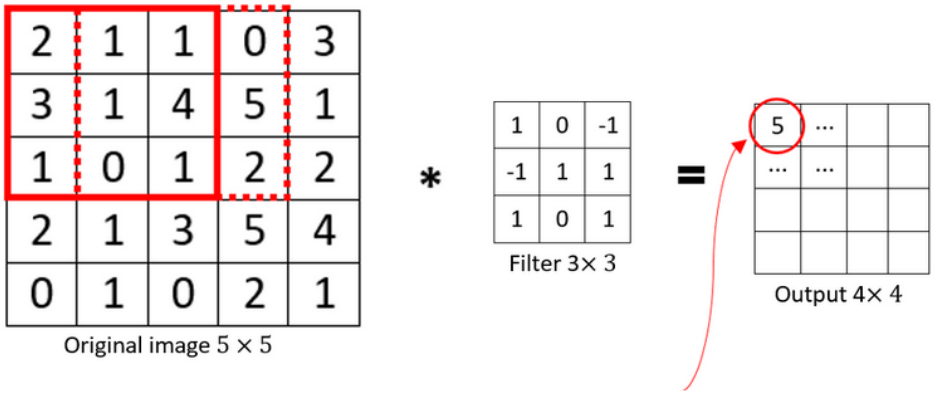


Figura 11: Estrutura de uma CNN.

Disponível em: https://github.com/tavgreen/landuse_classification (Acesso em set. 2019)

Com respeito aos filtros convolucionais, os aspectos matemáticos podem ser entendidos como um produto de matrizes, conforme mostra a Figura 12. Uma operação de convolução possui três componentes principais: a entrada, o detector de características também conhecido como kernel de convolução, e o mapa de características que é o resultado da operação.



$$2 \times 1 + 1 \times 0 + 1 \times (-1) + 3 \times (-1) + 1 \times 1 + \cdots + 0 \times 1 + 1 \times 1$$

Figura 12: Operação de convolução 3×3.

Adapatado de ZHAO *et al.* (2015)

Em geral, os filtros convolucionais da primeira camada representam objetos visuais simples, como pontos, retas e fronteiras entre tons de cinza ou cores. A convolução da imagem de entrada com esses múltiplos filtros produz um conjunto de mapas de ativação,

que correspondem a imagens filtradas. A próxima camada convolucional processa os mapas de ativação da camada precedente, utilizando outros filtros também definidos automaticamente na etapa de treinamento, produzindo novos mapas de ativação, que serão então filtrados pela próxima camada convolucional, e assim sucessivamente (Figura 13).

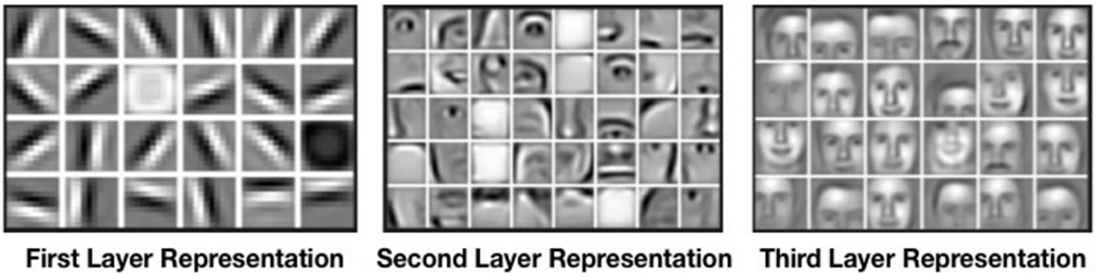


Figura 13: Extração de características de uma CNN
Disponível em:
http://www.lapix.ufsc.br/ensino/visao/visao-computacionaldeep-learning/deep-learningglossario/
Acesso em set. 2019

Uma outra operação importante realizada nas redes convolucionais é a operação de *max pooling* que visa realçar as características mais importantes da imagem. Esta operação varre a imagem a partir de uma janela, ver Figura 14, destacando o maior valor contido na mesma.

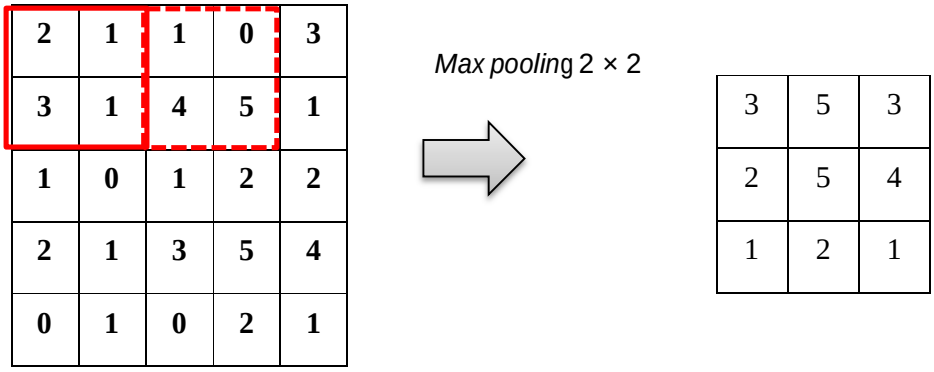


Figura 14: Operação de *max pooling*.
Adapatado de ZHAO *et al.* (2015)

Como algoritmo de aprendizado da rede, o algoritmo de *backpropagation* vem sendo o mais utilizado em Redes Neurais Convolucionais Profundas (GOODFELLOW, 2016).

Um dos aspectos relevantes de uma Rede Neural Convolucional Profunda é a aprendizagem hierárquica: quanto mais se avança pelas camadas da rede, mais complexas as características aprendidas, que constituem agregados das características das camadas anteriores. Cria-se assim uma hierarquia de atributos, em que a sucessão de camadas cria níveis de abstração cada vez mais altos.

2.3.3 *Autoenconders* e U-Net

Autoencoder é um tipo de rede neural artificial não supervisionada que compacta e codifica os dados de maneira bastante eficiente, e em seguida, aprende a reconstruir os dados de entrada da maneira mais próxima possível, mantendo as características principais dos dados que alimentaram a rede (GOODFELLOW *et al.*, 2016). Este processo é realizado com o intuito de representar dos dados de forma a reduzir a sua dimensionalidade ignorando os ruídos. Para isto a rede realiza dois tipos de transformações sobre os dados: a primeira denominada de função de codificação, que transforma os dados de entrada em uma forma de representação codificada dos mesmos; e a segunda, chamada de função decodificadora que realiza o processo inverso reconstruindo os dados da representação de maneira a ficar o mais próximo possível dos dados de entrada originais. A arquitetura de um *autoenconder*, é

ilustrada na Figura 15, nela podemos observar as unidades de codificação (lado esquerdo) e decodificação (lado direito). Na codificação, são aplicados blocos convolucionais seguidos de max pooling, uma técnica de downsampling, para codificar a imagem de entrada e representá-la em várias dimensões. A decodificação consiste nas técnicas de upsampling e concatenação, seguidas de operações convolucionais.

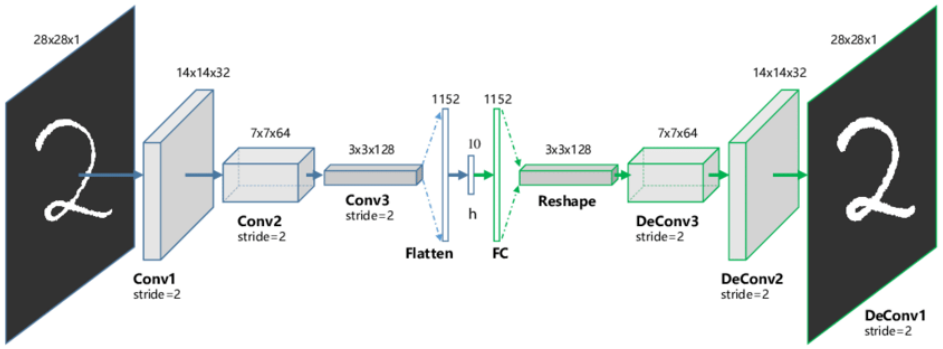


Figura 15: Modelo de estrutura *autoenconder* (GUO *et. al.*, 2017)

Entre as arquiteturas de *Autoenconders* que são utilizadas na segmentação de imagens, podemos destacar a do tipo U-net, proposta inicialmente para a segmentação semântica de imagens biomédicas (RONNEBERG *et al.*, 2015). A segmentação da imagem é realizada classificando cada pixel a partir do contexto de toda a imagem. Sua arquitetura é descrita na Figura 16 e é composta por camadas de contração (lado esquerdo) e camadas de expansão (lado direito), apresentando uma simetria entre suas camadas no formato de “U”.

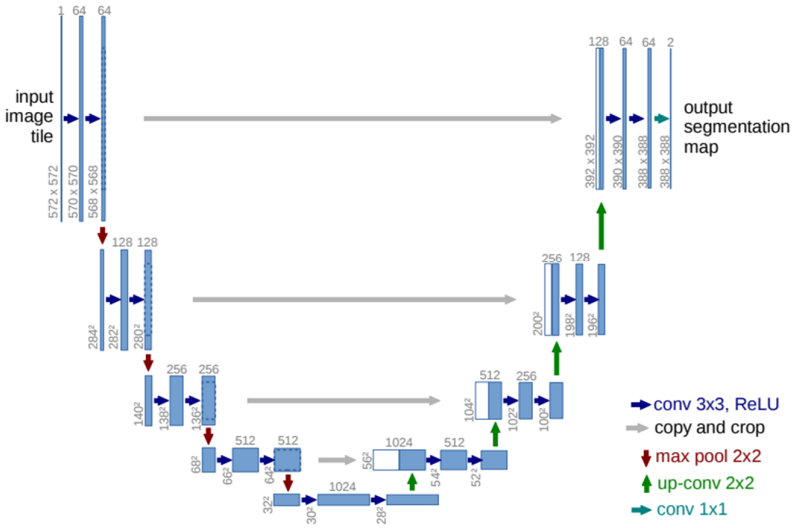


Figura 16: Arquitetura da U-Net
Disponível em:
https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/
Acesso em set. 2019

Uma das vantagens de se trabalhar com a U-Net é a sua capacidade de generalização dos resultados mesmo em domínios de baixa variância, isto é, bancos onde as imagens possuem muitas características em comum. Nessas condições a U-Net é capaz de capturar detalhes finos das imagens rapidamente, possibilitando um treinamento efetivo com quantidades limitadas de dados.

2.3.4 Matriz de Confusão

Com respeito ao desempenho de um classificador, existem diversas medidas de avaliação. A métrica de avaliação mais adequada depende do tipo de dado em questão e da natureza da classificação realizada. Todas elas são baseadas na chamada matriz de confusão, que nada mais é do que uma tabela que permite a visualização do desempenho de um algoritmo de aprendizado supervisionado (MONARD e BARANAUSKAS, 2003). Cada coluna da matriz representa as instâncias de uma classe prevista, enquanto as linhas representam os casos de uma classe real. A matriz de confusão é ilustrada na Figura 17.

		Valor Previsto	
		Positivo	Negativo
Valor Verdadeiro	Positivo	VP verdadeiro positivo	FN falso negativo
	Negativo	FP falso positivo	VN verdadeiro negativo

Figura 17: Matriz de confusão
Adaptado de MONARD e BARANAUSKAS (2003)

Essa matriz é uma forma intuitiva de saber como seu classificador está se comportando. Como exemplo, vejamos um classificador cuja tarefa é classificar cachorro e gato a partir de imagens. A Figura 18 mostra o resultado obtido pela matriz de confusão.

		Valor Previsto	
		Gato	Cachorro
Valor Real	Gato	45	10
	Cachorro	5	40

Figura 18: Classificação de gatos e cachorros.
Adaptado de MONARD e BARANAUSKAS (2003)

Do lado esquerdo temos o valor real e no topo temos o valor previsto pelo classificador. Com base nesses valores temos que:

- O modelo classificou 45 instâncias como gato o que eram realmente gatos.
- O modelo classificou 10 instâncias como cachorro o que na verdade eram gatos.
- O modelo classificou 5 instâncias como gato que na verdade era cachorro.
- O modelo classificou 40 instâncias como cachorro que eram realmente cachorros.

Como saber se o classificador está classificando bem as instâncias que são realmente cachorros? Para isso são utilizadas algumas métricas de avaliação de desempenho, entre elas temos a acurácia, *Recall*, precisão e $F_1 Score$ (LEVER *et al.*, 2016).

A taxa de acurácia representada em (3) é a proporção *de predições corretas*, sejam eles verdadeiros positivos ou verdadeiros negativos. Ou seja, ela mede o total de acertos em relação ao total de testes realizados. Esta medida é altamente suscetível a desbalanceamentos do conjunto de dados e pode facilmente induzir a uma conclusão errada sobre o desempenho do sistema, pois podemos ter uma acurácia alta em uma referida classe que possui um número relativamente maior em comparação a outra classe.

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \tag{3}$$

O Recall também como revocação, sensibilidade ou taxa de verdadeiros positivos, dada por (4), mede a proporção de classificações positivas dentre todos os objetos realmente positivos. Ou seja, a métrica *Recall* mede o quão frequente o seu classificador classifica como X os que são realmente da classe X.

$$Recall = \frac{VP}{VP + FN} \tag{4}$$

A precisão, dada por (5), é uma medida do quão exato está a classificação para as amostras positivas. Ou seja, a precisão mede a proporção dos que foram classificados como positivos, dos que realmente eram positivos.

$$Precisão = \frac{VP}{VP + FP} \tag{5}$$

Para explicar melhor a diferença entre as métricas Recall e precisão considere que um programa de computador para o reconhecimento de cães em cenas de um vídeo identifica 7 cães em uma cena contendo 9 cães e alguns gatos. Se 4 das identificações estão corretas, mas 3 são, na verdade, gatos, a precisão do programa é $\frac{4}{7}$ enquanto a sua revocação é $\frac{4}{9}$.

Outra métrica de avaliação de desempenho é a chamada F_1Score, ver (6). Ela é uma média harmônica entre precisão e *recall* de modo a trazer um número único que indique a qualidade geral do seu modelo e trabalha bem até com conjuntos de dados que possuem classes desproporcionais com respeito a quantidade de dados entre as classes.

$$F_1Score = \frac{2 \times precisão \times Recall}{precisão + Recall} \quad (6)$$

Estas métricas têm seus valores em um intervalo que varia de 0 a 1, valores próximos de 1 representam os melhores resultados.

Logo, com respeito a classificação para cachorros e gatos obtida na Figura 17 temos uma acurácia de 85%, um *Recall* de 82%, uma precisão de 90% e um F_1Score de 86%. O que revela que o classificador está identificando relativamente bem gatos e cachorros.

2.3.5 Índice de Similaridade de Jaccard

O Índice de Similaridade de Jaccard, também chamado de interseção sobre a união (IoU), é uma medida estatística utilizada para medir a similaridade entre amostras finitas de um conjunto e é calculado a partir da razão entre o número de elementos da interseção e o número de elementos da união entre os conjuntos. O (IoU) é uma métrica bastante popular entre os especialistas de segmentação e imagem (IGLOVIKOV e SHVETS, 2018; NARAYANA *et al.*, 2012).

O IoU mede a taxa de sobreposição entre a imagem prevista e a verdade básica. Em suma, o índice de similaridade de Jaccard compara elementos entre dois conjuntos e observa entre eles quais são semelhantes e quais são distintos. Quanto maior for a semelhança, maior será o índice de Jaccard, que varia entre 0 e 1. Uma combinação perfeita implicaria $IoU = 1{,}0$. A Figura 19 ilustra uma representação do conceito do Índice de Jaccard.

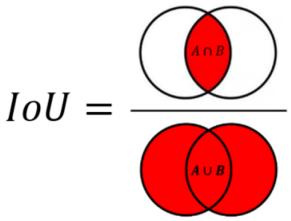


Figura 19: Representação geométrica da IoU

Por exemplo, sendo A e B dois conjuntos finitos quaisquer de modo que $A \cup B$ representa a união desses conjuntos e $A \cap B$ a interseção desses conjuntos. Então, o índice de similaridade de Jaccard é calculado a partir da expressão abaixo:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{7}$$

O símbolo | . | representa a cardinalidade de um conjunto.

CAPÍTULO III

3. Materiais e Métodos

A pesquisa foi desenvolvida a partir de 3 etapas. Nas duas primeiras, tínhamos disponíveis apenas 4 imagens, com as quais fizemos uso de técnicas de identificação de padrões a partir da classificação do tipo "*lead"* ou "*no lead"*. Na primeira etapa foi utilizada uma RNA a partir de descritores de Haralick como entrada. Na segunda etapa, fez-se uso de uma CNN cujos descritores são obtidos pela própria rede. E por fim, na terceira etapa, o banco de imagens foi ampliado a partir da subdivisão das mesmas de forma que todas as imagens foram fixadas no tamanho 450 x 342 pixels, o que fez o número de imagens subir de 4 para 8. A partir disso foi realizado o uso de *autoenconders* para uma segmentação semântica a partir de uma arquitetura do tipo U-Net.

3.1 Aquisição das imagens

As imagens sísmicas utilizadas neste trabalho foram obtidas a partir do método sísmico de reflexão. Elas foram adquiridas na bacia Sergipe-Alagoas no nordeste do Brasil, a partir da Agência Nacional de Petróleo, Gás Natural e Biocombustíveis (ANP) (HEASER, 2015). A Figura 20 corresponde a uma das imagens sísmicas utilizadas neste trabalho rotuladas por um especialista. A região em amarelo destacada corresponde ao *lead*.

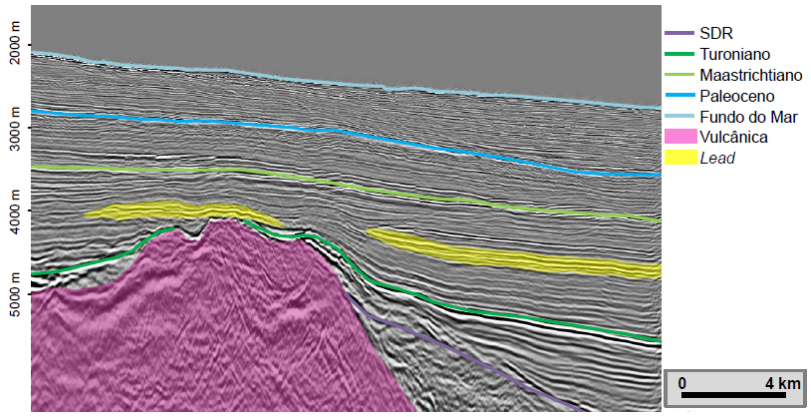


Figura 20: Amostra de uma seção sísmica coletado do bloco *offshore* da Bacia Sergipe-Alagoas segmentada por um especialista. Em amarelo está a região de interesse.
FONTE: Bacia Sergipe-Alagoas

Sendo assim, o nosso conjunto de dados é formado por 4 imagens rotuladas, sendo duas de dimensões 1052 x 505 pixels e duas de 1028 x 695 pixels. Foram a partir dessas quatro imagens que as duas primeiras etapas do experimento da tese foram realizadas. A Figura 21 ilustra uma dessas imagens com seu respectivo rótulo.

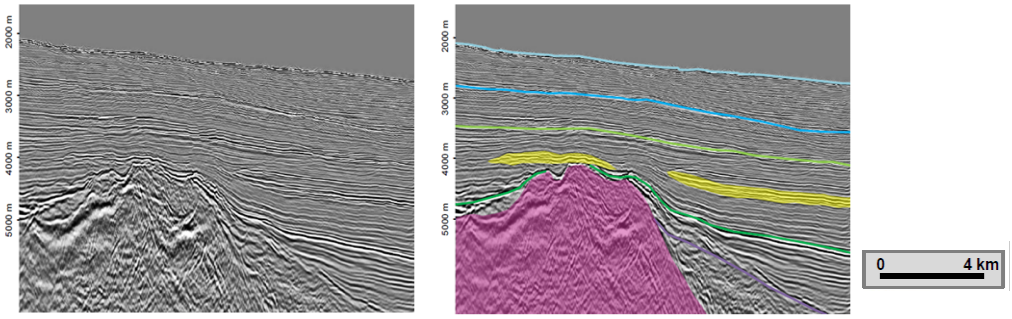


Figura 21: Uma das imagens utilizadas nas etapas de RNA e CNN. A identificação dos elementos foi realizada por um especialista.

Com respeito a etapa a qual foi utilizada uma arquitetura do tipo U-Net, foram utilizadas 8 imagens, obtidas da subdivisão das quatro imagens anteriores. Tal procedimento foi realizado no sentido de aumentar o conjunto de dados para treinamento da rede. A Figura 22 ilustra as oito imagens utilizadas.

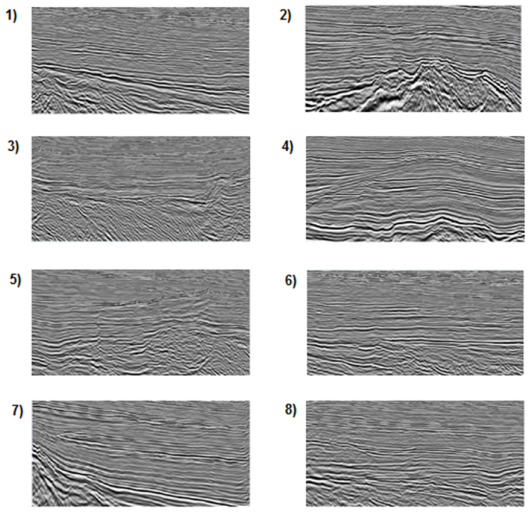


Figura 22: Imagens do *Dataset* utilizado para a U-Net

3.2 Pré-Processamento dos dados

Esta etapa consiste na preparação dos dados antes que os mesmos sejam lançados na rede para fins de treinamento. Inicialmente, retiramos da imagem original os elementos que não fazem parte do estudo (turoniano, maastrichtiano, paleoceno, fundo do mar e vulcânico) deixando apenas a região de interesse os *leads*. Em seguida, através do processo de binarização, separamos a região em "*leads*" e "*no leads"*. Os pixels pertencentes à região do *lead* são coloridos de branco, e os pixels fora dessa região, denominados *"no lead"*, foram coloridos de preto. O resultado desse processo pode ser observado na Figura 23.

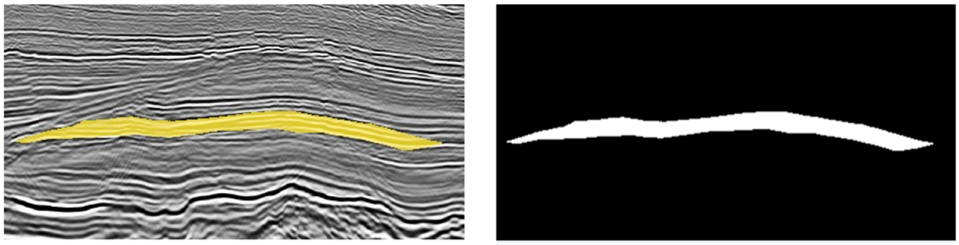


Figura 23: Imagem sísmica original e binarizada. Do lado esquerdo temos uma imagem sísmica original indicando somente a região de interesse, e do lado direito temos a respectiva máscara, indicando os pixels pertencentes a região de interesse pintados de branco e caso contrário, pintados de preto.

O processo de geração dos *patches* da imagem foi realizado a partir de um processo que denominamos de janela deslizante. Essa janela, com um tamanho pré-definido, percorre toda imagem através de deslocamentos (horizontais e verticais) que denominamos de *strides*. Os *patches* são gerados a partir da razão entre o número de pixels brancos $(\mathcal{P}_w)$ sobre o número total de pixels contidos na janela $(\mathcal{P})$. No que diz respeito às etapas em que foram utilizadas RNA e CNN o objetivo foi obter um conjunto de treino e validação formado por duas classes de *patches* (*lead* ou *no lead*). O que determina a que classe determinado *patch* pertence é o resultado da razão $\frac{\mathcal{P}_w}{\mathcal{P}}$. Se $\frac{\mathcal{P}_w}{\mathcal{P}} \geq \lambda$, então o *patch* foi caracterizado como *lead*, caso contrário, foi considerado *no lead*. Onde $0 \leq \lambda \leq 1$ faz o papel de limiar de separação das classes. Para o nosso trabalho foi considerado um limiar de 0,5. Com isso, foi observado que o número de *leads* foi bem inferior em comparação ao número de *no leads*, então, com o objetivo de aumentar o número de subimagens com região classificada como *lead* realizamos o procedimento de janela deslizante utilizando dois *strides* diferentes $(s_1 > s_2)$. De modo que, a janela deslizante inicia o processo a partir de um *stride* s_1 e a partir do momento que for detectado um *patch* da classe *lead*, ela passa a se deslocar com um *stride* s_2 gerando assim superposições nos *patches*. A medida que a janela gera um *patch* que pertence a classe *no lead* ela retorna a deslizar com um *stride* s_1. A Figura 24 ilustra as etapas do processo a partir de uma imagem 8×8 a qual será "varrida" por uma janela deslizante 3×3, utilizando *strides* de valores 3 e 2.

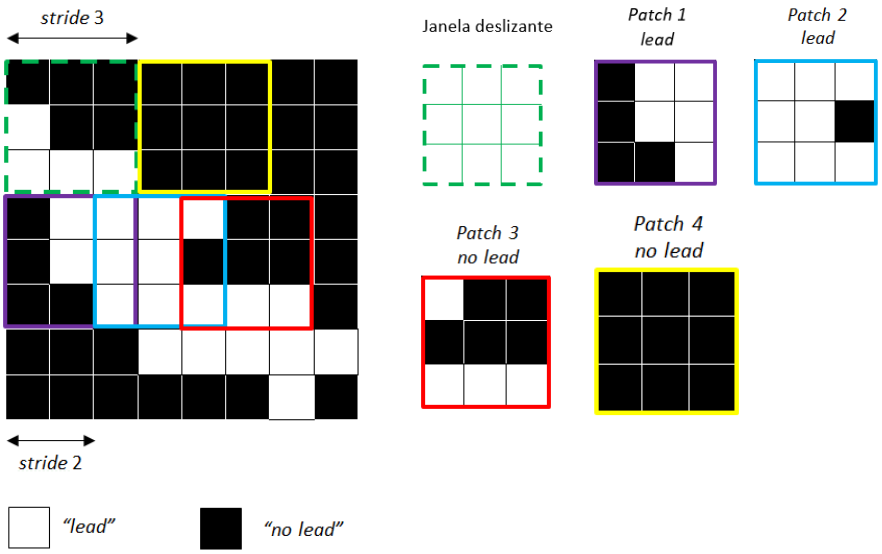


Figura 24: Processo de geração de *patches* nas etapas de RNA e CNN: Uma janela deslizante 3×3 varre a imagem a paritr de dois *strides* a partir de um limiar $\lambda = 0{,}5$. Neste exemplo, os *patches* 1 e 2 (azul e roxo) representam região de interesse, enquanto que os *patches* 3 e 4 não representa a região de interesse (vermelho e amarelo).

O processo tem início no lado superior esquerdo da imagem e obtém a razão $\frac{\mathcal{P}_W}{\mathcal{P}}$, se esse valor for menor que 0,5 o *patch* é classificado como *no lead* (quadrado tracejado verde) e a janela passa para a próxima subimagem com um *stride* de 3, e realiza novamente o processo de obtenção da razão $\frac{\mathcal{P}_W}{\mathcal{P}}$. Caso a janela obtenha uma razão maior ou igual a 0,5 (quadrado roxo), a janela desliza até o próximo *patch* com um *stride* de 2, gerando assim sobreposição de *patches* (quadrado azul). A janela continua realizando o processo de gerar *patches* até obter um *patch* cuja razão $\frac{\mathcal{P}_W}{\mathcal{P}} \geq 0{,}5$ retornando assim ao *stride* de valor 3 sem sobreposições. O processo continua até que a janela tenha "varrido" toda a linha horizontal, reiniciando o processo a partir do ponto de partida, porém com um deslocamento vertical da janela. Após percorrer toda a imagem o processo é encerrado.

No que diz respeito a geração de *patches* utilizando a U-Net o objetivo não é obter um conjunto com dois tipos de classes, e sim, obter um conjunto de *patches* de treino, validação e teste de modo que a rede possa aprender com base nas informações contidas em cada *patch*. Com o intuito de obter *patches* que contenham partes da região de interesse, só serão selecionados *patches* cuja razão $\frac{\mathcal{P}_W}{\mathcal{P}} \geq \lambda$. A Figura 25 ilustra o processo de geração dos

patches a partir de uma imagem 8×8 a qual será "varrida" por uma janela deslizante 3×3, utilizando um *stride* de 2 com um limiar $\lambda = 0{,}5$.

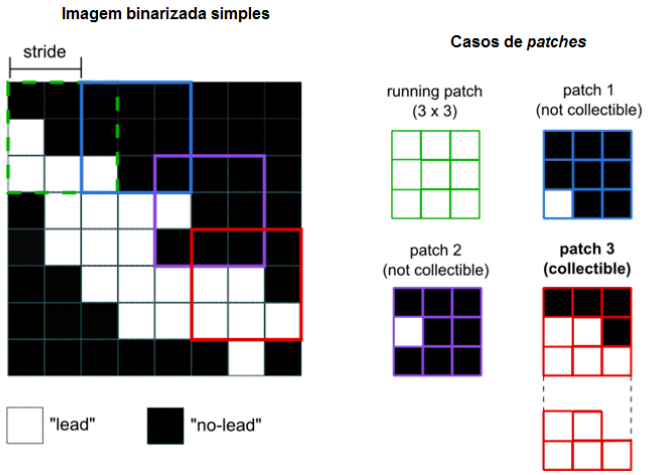


Figura 25: Processo de geração de *patches*: é utilizada uma janela deslizante 3×3 sobre uma imagem binária com um *stride* igual a 2 e limiar $\lambda = 0{,}2$. Neste exemplo, os *patches* 1 e 2 (azul e roxo) não representam região de interesse, enquanto que o *patch* 3 representa a região de interesse (vermelho).

O processo é semelhante ao que foi realizado nas etapas anteriores. Um *patch* só será selecionado se $\frac{\mathcal{P}_w}{\mathcal{P}} \geq 0{,}5$. Por exemplo, o *patch* 3 é um caso de *patch* selecionável.

Nas etapas envolvendo RNA e CNN foi utilizada uma janela deslizante de dimensão 20×20, com um *strides* de 10 e de 20 com um limiar $\lambda = 0{,}5$. A partir disso foram gerados cerca de 4.000 *patches* de treinamento, 1.300 de validação e 1.000 de teste. Na U-Net a janela possuía uma dimensão de 80×80 com um *stride* igual a 2 e um limiar $\lambda = 0{,}2$. Com isso o total de *patches* obtidos foi cerca de 24.000 de treinamento, 5.000 de validação e 600 de teste.

Vale salientar ainda que, como ferramenta de aumento da variabilidade dos dados de treinamento e validação, foi aplicada técnicas de *data augmentation* sobre os *patches*, que incluem rotações aleatórias, mudanças de escala e espelhamentos (*flipping* horizontal). O objetivo de realizarmos esse procedimento é evitar que a rede se torne especialista em um

conjunto de dados específico (*overfitting*), provocando assim erros de classificação em novas amostras (PETTERSON e GIBSON, 2017).

3.3 Arquitetura da rede

3.3.1 Arquitetura da RNA

Para realizar o processo de classificação de padrões utilizamos uma RNA *feed-forward* com algoritmo de aprendizagem *back-propagation* com minimização do erro através do gradiente descendente estocástico. Como função de ativação foi utilizada a função tangente sigmoide $f(x) = \frac{2}{1+e^{-\beta x}} - 1$, onde β é um parâmetro de controle. Os pesos iniciais para todos os links foram definidos de forma aleatória. Com respeito à camada de entrada, representa vetores obtidos a partir dos descritores de Haralick. Utilizamos os descritores de textura para identificação de características para o treinamento da rede tendo em vista que a análise de imagens através de texturas vem sendo utilizada para a diferenciação de diversos tipos de estruturas encontrados na natureza, como exemplo em imagens geológicas (HARALICK e SHANMUGAN, 1973). Os descritores são obtidos a partir de cálculos estatísticos de segunda ordem, considerando as ocorrências de cada nível de cinza em pixels diferentes ao longo de diferentes direções. Para obtenção de características que auxiliem no treinamento da rede foram utilizados os seguintes parãmetros: segundo momento angular, entropia, contraste, variância, homogeneidade, probabilidade máxima, momento de diferença de 3ª ordem, momento de diferença inverso de 3ª ordem, variância inversa, dissimilaridade e média da soma. Os descritores foram obtidos para distância $d = 1$ e ângulos de variação nos valores de 0°, 45°, 90° e 135°. Para maiores detalhes sobre os descritores de Haralick ver Apêndice.

Assim, os vetores de entrada da rede serão formados por 44 atributos previsores. Com respeito à camada oculta, foram utilizados 128 neurônios e em cada um deles foi aplicado à função tangente sigmoide. Na camada de saída foram utilizados 2 neurônios, em cada um deles foi aplicada a função sigmoidal (Figura 26).

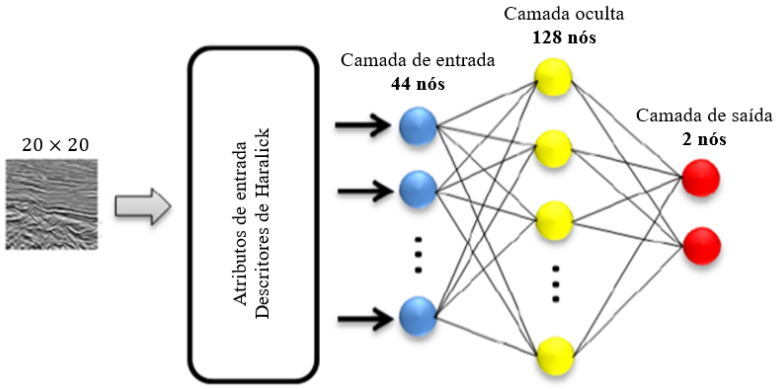


Figura 26: Arquitetura da RNA para as etapas de treino, validação e teste

Na busca pela melhor configuração da rede neural foi realizado um planejamento fatorial do tipo 2^k (ver APÊNDICE B), onde foram analisados além dos tamanhos das subimagens ($10 \times 10, 15 \times 15$ *e* 20×20), outros fatores tais como, taxa de aprendizagem da rede (l_r) que é um parâmetro que interfere na convergência da rede. Os valores avaliados foram $l_r = 0{,}001; 0{,}006; 0{,}01; 0{,}03$. Um outro fator avaliado foi o termo do momento (m_c), que ajuda a melhorar a taxa de aprendizado evitando a oscilação do mesmo. Os valores avaliados foram $m_c = 0{,}01; 0{,}06; 0{,}1; 0{,}3$. A métrica utilizada como fator de decisão foi a $F_1 Score$, essa escolha é justificada por ser um indicador mais robusto (POWERS, 2011; SASAKI, 2007). Os melhores resultados foram obtidos para subimagens de tamanho 20×20, $l_r = 0{,}03$ e $m_c = 0{,}01$.

3.3.2 Arquitetura da CNN

Neste trabalho utilizamos um modelo de CNN composta por sete camadas conforme ilustrado na Figura 27. Ela é composta de três camadas de convolução com 32, 64 e 128 filtros respectivamente, todos de tamanho 3×3. Após cada camada de convolução foram aplicadas camadas de *max-pooling* de tamanho 2×2. Ao final da CNN os resultados foram aplicados a uma RNA densa com 128 neurônios na camada oculta e 2 neurônios na camada de saída.

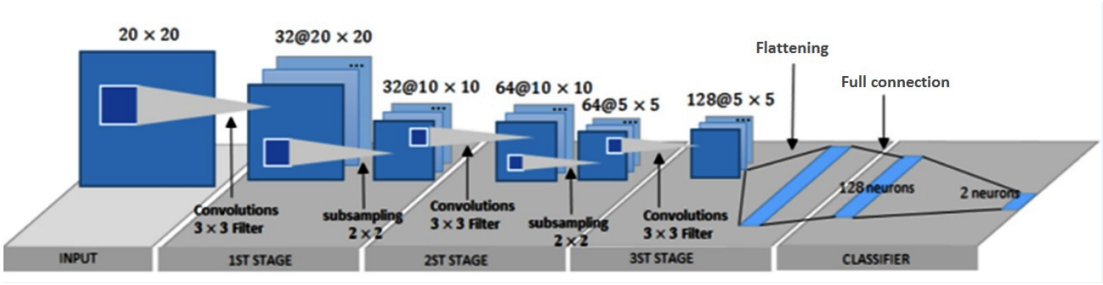


Figura 27: Arquitetura da CNN para as etapas de treino, validação e teste.

A rede foi treinada com um batch de tamanho 64 com 16 amostras por classe (159.490 parâmetros no total). Em todas as camadas de convolução foi utilizada a função de ativação Relu assim como na rede neural densa. Foi utilizado o Adam como algoritmo de otimização e os pesos foram inicializados com valores aleatórios. Nos neurônios de saída foi utilizada a função softmax.

3.3.3 Arquitetura U-Net

Dentre as arquiteturas de CNN já existentes voltadas para segmentação, optamos por utilizar a U-Net devido à sua capacidade de generalizar em domínios de baixa variância, isto é, bancos onde as imagens não são muito diferentes entre si. Nessas condições a U-Net é capaz de capturar detalhes finos das imagens rapidamente, possibilitando um treinamento efetivo com quantidades limitadas de dados. Apesar de desenvolvida para interpretação de imagens médicas (RONNEBERGER, 2015), a U-Net pode ser também utilizada para interpretações sísmicas, como por exemplo detecção de corpos de sal (ZENG, 2019).

Devido à baixa dimensionalidade e quantidade das imagens de entrada deste trabalho, a U-Net padrão sofre superajuste de imediato se fazendo necessário reduzir a quantidade de filtros por camada além do uso adicional de *dropout* entre camadas convolucionais para prevenir este superajuste (SRIVASTAVA, 2014). Por fim, o modelo resultante está descrito na Tabela 01.

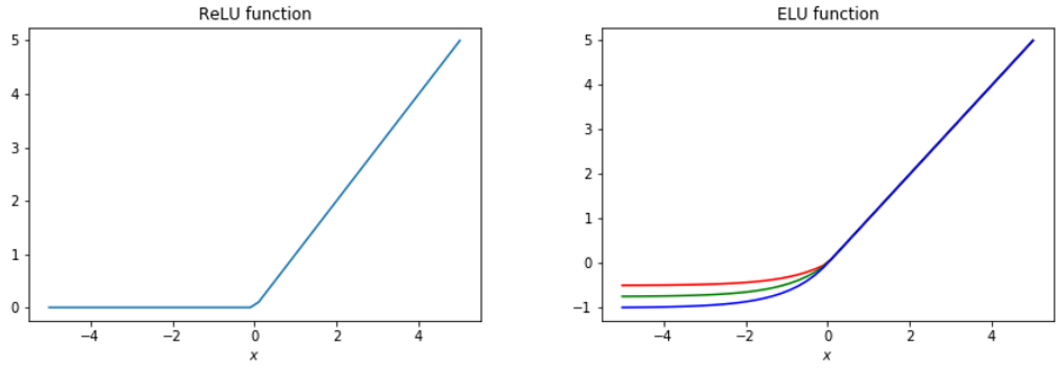
Tabela 01: Descrição da arquitetura utilizada, totalizando 1,940,817 parâmetros.

Camada	**Tamanho do núcleo**	**Camada de *maxpolling***	**Tamanho da saída**
entrada	-	-	80 × 80 × 1
flat-conv1 + dropout=0.3	3 × 3	-	80 × 80 × 16
down-conv1	3 × 3	2 × 2	40 × 40 × 16
flat-conv2 + dropout=0.3	3 × 3	-	40 × 40 × 32
down-conv2	3 × 3	2 × 2	20 × 20 × 32
flat-conv3 + dropout=0.4	3 × 3	-	20 × 20 × 64
down-conv3	3 × 3	2 × 2	10 × 10 × 64
flat-conv4 + dropout=0.4	3 × 3	-	10 × 10 × 128
down-conv4	3 × 3	2 × 2	5 × 5 × 128
flat-conv5 + dropout=0.5	3 × 3	-	5 × 5 × 256
up-conv1 + concatenate w/ flat-conv4	2 × 2	$\frac{1}{2} \times \frac{1}{2}$	10 × 10 × 256
flat-conv6 + dropout=0.4	3 × 3	-	10 × 10 × 128
up-conv2 + concatenate w/ flat-conv3	2 × 2	$\frac{1}{2} \times \frac{1}{2}$	20 × 20 × 128
flat-conv7 + dropout=0.4	3 × 3	-	20 × 20 × 64
up-conv3 + concatenate w/ flat-conv2	2 × 2	$\frac{1}{2} \times \frac{1}{2}$	40 × 40 × 64
flat-conv8 + dropout=0.3	3 × 3	-	40 × 40 × 32
up-conv4 + concatenate w/ flat-conv1	2 × 2	$\frac{1}{2} \times \frac{1}{2}$	80 × 80 × 32
flat-conv9 + dropout=0.3	3 × 3	-	80 × 80 × 16
Output	1 × 1	-	80 × 80 × 1

Segundo (PEDAMONTI, 2018), apesar da função de ativação ReLU (Rectifier Linear Unit) apresentar resultados satisfatórios em praticamente todas as aplicações de deep learning, suas variantes não-lineares podem apresentar resultados melhores a depender do domínio. Dentre tais variantes, a função ELU (Exponential Linear Unit) pode apresentar melhorias ao se trabalhar com imagens sísmicas, evitando problemas pertinentes a ReLU como nós mortos (nós da rede cuja resposta será sempre 0) e por estar apresentando convergência mais rápida, como observado por (ZENG *et al.*, 2019). Por esses motivos, optou-se por utilizar ELU como função de ativação em todas as camadas, exceto a camada de saída que possui uma função de ativação sigmóide.

As funções ReLU e ELU são definidas a partir das equações citadas em (8) e na Figura 28 vemos a representação gráfica de ambas.

$$ReLU(x) = \max(0; x) \qquad ELU(x) = \begin{cases} \alpha(\exp(x) - 1) & se \quad x \leq 0 \\ s & se \quad x > 0 \end{cases} \tag{8}$$



(a) (b) $\alpha = 0.5$; $\alpha = 1$; $\alpha = 1.5$.

Figura 28: Representação gráfica das funções: (a) ReLU; (b) ELU.

O treinamento foi realizado durante 100 épocas utilizando o algoritmo de otimização Adam (KINGMA, 2015) iniciando com um *learning rate* de $5e^{-4}$. A cada época, o IoU foi utilizado para medir a performance da rede no conjunto de validação, em caso de melhorias em relação à época anterior, a rede na época atual é salva para fins de adotar a rede em sua melhor época como rede final. No final de cada nova época, a IoU era calculado para medir o desempenho da rede no conjunto de dados de validação. Após verificar se o

desempenho da rede na época mais atualizada foi melhor do que desempenho entre todos a épocas anteriores, a época atual era armazenada como a rede mais recente.

3.4 Pós-processamento

Essa etapa do processo, realizada apenas na arquitetura U-Net, foi subdividida em etapas menores, a saber, reconstrução, limiarização e retirada de regiões desconexas. Elas serão explicadas a seguir.

3.4.1 Reconstrução

Para a etapa de reconstrução, fazemos uso de uma imagem de teste a qual é percorrida por uma janela deslizante de tamanho 80 x 80. O primeiro estágio do processo corresponde ao envio para a rede do primeiro patch localizado no lado superior esquerdo da figura, em seguida o mesmo foi segmentado pela rede e a imagem de saída é salva. O segundo estágio foi realizado a partir de um *stride* (deslocamento) de 10 pixels para direita a partir do primeiro patch, conforme Figura 29. Esse segundo patch foi enviado para a rede o qual foi segmentado de acordo com o aprendizado da mesma. Após segmentada, o processo de reconstrução é realizado a partir da junção do patch segmentado no estágio anterior com a segmentação do patch atual, de maneira que as regiões idênticas fiquem sobrepostas. O processo continua até o final da margem direita superior. Em seguida o processo foi reiniciado a partir do patch inicial porém com um deslocamento de 10 pixels para baixo, para logo depois percorrer horizontalmente a figura a partir de *strides* de tamanho 10 pixels. A tarefa é realizada até percorrer toda imagem e assim obter a reconstrução final. Dessa forma, obtemos a reconstrução da imagem como uma montagem incremental de *patches* adjacentes de deslocamento de 10 *pixels.*

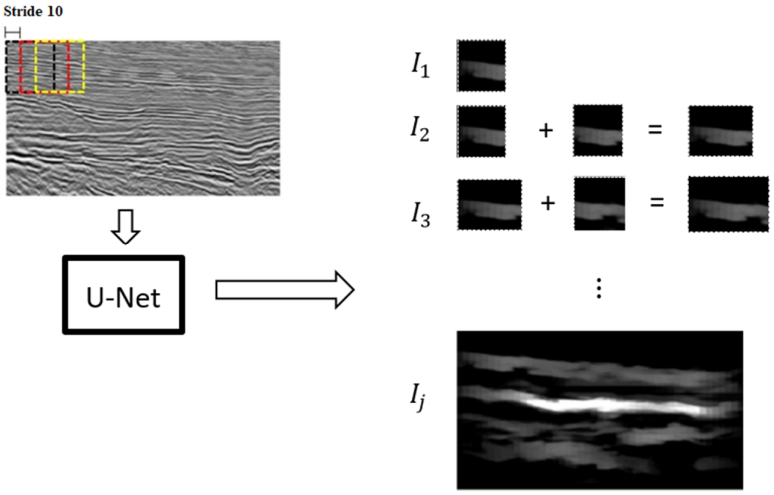


Figura 29: Etapa de segmentação e reconstrução da imagem: Uma imagem original é analisada pela U-Net a partir de *patches* de tamanho 80×80. Em seguida, depois de segmentados, os *patches* são unidos por uma sucessão e composições a partir de um *stride* de tamanho 10.

Esse processo, porém, gera muitas sobreposições na imagem e com isso faz com que os pixels aumentem de valor nas sobreposições, alterando assim o nível de cinza da imagem, provocando uma espécie de imagem desfocada, o que leva a ocorrência de falsos positivos como pode ser observado na etapa I_j da Figura 29. Para resolver este problema, foi aplicada uma técnica de limiarização.

3.4.2 Limiarização

O processo de segmentação da imagem realizada pela U-Net, associa a cada pixel p do objeto reconstruído um valor normalizado, que podemos chamar de *Índice de Confiança* ($I_c(p)$). Esse valor é variável de acordo com o nível de cinza da imagem reconstruída. A limiarização proposta irá mapear o índice I_c em toda a imagem de forma a obter como saída uma nova máscara binária da imagem.

Sejam $\mathcal{G}$ a imagem reconstruída em nível de cinza e $\mathcal{B}$ a representação binária dessa imagem. O processo de binarização é feito a partir da função $\mathcal{L}: \mathcal{G} \rightarrow \mathcal{B}$ indicada em (9) de modo que para cada $p \in \mathcal{G}$ temos:

$$\mathcal{L}(p) = \begin{cases} 0 & se \quad I_c(p) < \tau \\ 1 & se \quad I_c(p) \geq \tau \end{cases} \qquad (9)$$

Onde τ representa o índice de limiarização, para o nosso caso definimos $\tau = 0{,}5$.

3.4.3 Retirada de regiões desconexas

O próximo passo quanto ao pós-processamento é a retirada de valores discrepantes que permaneceram após o processo de limiarização. Esses valores geram regiões desconexas (ou ilhas isoladas) que chamamos de *outliers* que podem causar a falsa impressão de regiões de interesse. Para realizar a tarefa de retirada dos *outliers* utilizamos um algoritmo de identificação de componentes conectados disponíveis no módulo *scikit-image* do *Python* para separar regiões.

Considerando a característica de continuidade local esperada para as regiões principais observadas nas imagens do nosso banco, assumimos que apenas o componente com a maior área de pixels deve ser mantido. Embora a maior área possa não ser o melhor critério para escolher, foi verificado que o número de outliers encontrados possui uma área muito menor em comparação com a área principal. De modo mais formal, podemos entender a retirada dos *outliers* como uma função $\mathcal{R}$ que trasnforma uma imagem binarizada $\mathcal{B}$ em uma imagem de fundo mais limpa $\mathcal{B}^*$, conforme ilustra a Figura 30.

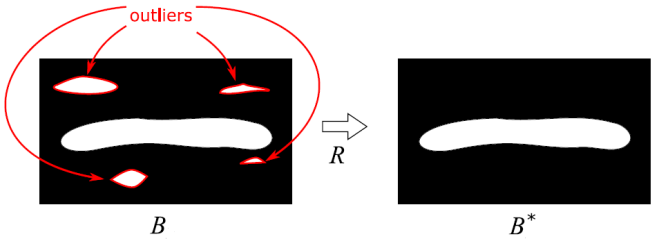


Figura 30: Processo de remoção dos *outliers* de uma imagem binarizada após a limiarização. As áreas menores marcadas (lado esquerdo) são eliminadas, restando apenas a área maior, a região de "*lead* " (lado direito).

Os resultados obtidos no pós-processamento serão apresentados na seção posterior. Sendo assim, o fluxo de trabalho realizado em cada etapa seguiu a seguinte ordem: Aquisição das imagens, pré-processamento, definição da arquitetura da rede e treinamento da rede, classificação (RNA e CNN) ou segmentação (U-Net) e no caso do uso de *autoenconders*, foi utilizado o pós-processamento. A Figura 31 ilustra as etapas seguidas em cada arquitetura.

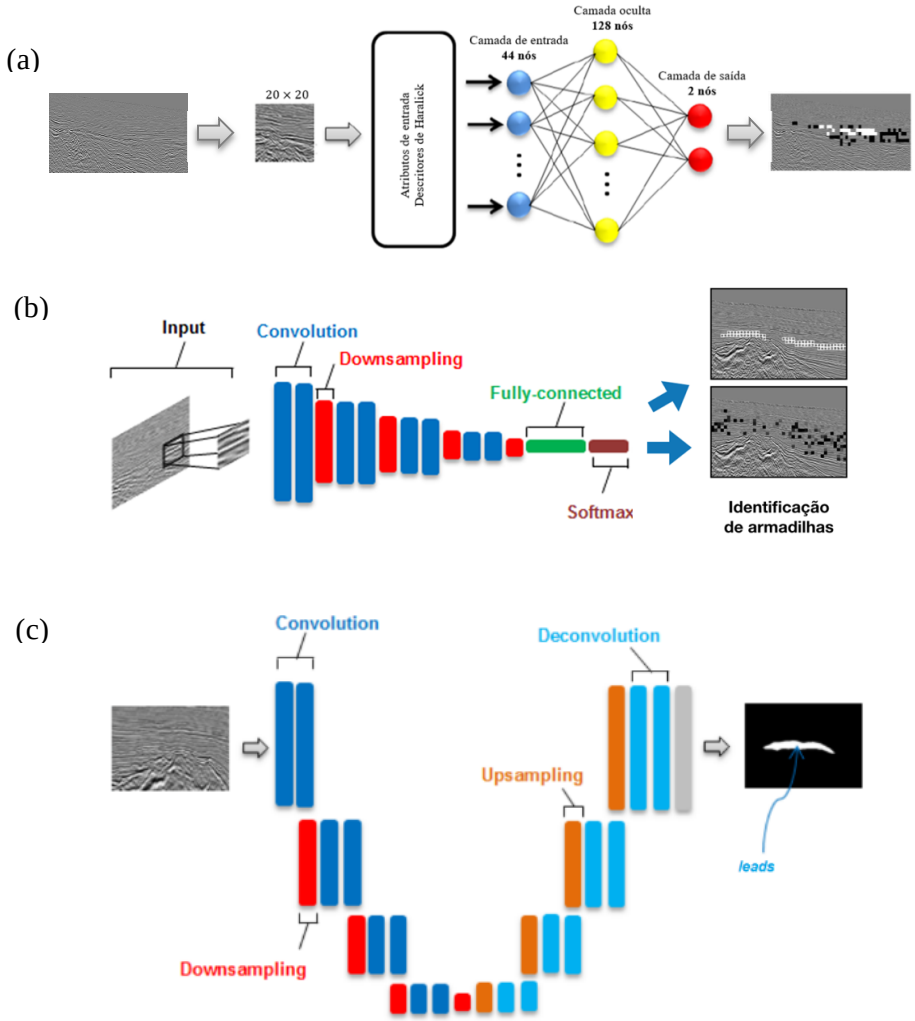


Figura 31: Ilustração das etapas de cada arquitetura: (a) Arquitetura de uma RNA utilizando os descritores de Haralick como entrada. (b) Arquitetura de uma CNN, nela, os atributos são obtidos a partir de filtros convolucionais que extraem características que auxiliam no aprendizado da rede. (c) *Autoenconders* com arquitetura tipo U-Net

CAPÍTULO IV

4. Resultados e discussões

Nesta etapa da pesquisa iremos averiguar os resultados obtidos utilizando as ferramentas de aprendizado de máquina citadas nas seções anteriores.

4.1 Classificação binária a partir do uso de RNA e CNN

A principal diferença entre a utilização de RNA em relação a CNN se deve ao fato de que na CNN a extração de características que irão auxiliar no aprendizado da rede é definida pela própria rede, e não de forma manual como na RNA. Em ambas as técnicas foi utilizado o mesmo conjunto de treinamento, validação e teste. Também foi realizado técnicas de *data augmentation* tais como: rotação ($\pm 20°$), escala ($\pm 20\%$) e *flipping* horizontal. No caso da CNN, com o intuito de evitar o *overfitting* foi aplicada a técnica de *dropout* eliminando aleatoriamente 25% dos nós de uma das camadas que compõe a rede. A saída de ambas as redes foram obtidas a partir de um filtro de probabilidade de 0,6.

A Figura 32 ilustra o resultado da classificação realizada pela CNN, no lado esquerdo temos a imagem rotulada de forma manual a partir de um especialista, em contraste com a classificação realizada pela CNN no teste às cegas. Os quadrados pretos representam falsos positivos enquanto que os quadrados brancos representam verdadeiros positivos. A capacidade de classificação da CNN é evidente, tendo em vista que embora existam falsos positivos, isso ocorre em subimagens que possuem características morfológicas semelhantes com as subimagens da região de interesse. Em paralelo, a Figura 33 mostra os resultados semelhantes obtidos pela RNA também no teste às cegas das mesmas regiões classificadas pela CNN. Embora tenhamos observado um número maior de falsos positivos na classificação realizada pela RNA, devemos salientar que ambas foram eficazes no reconhecimento de camadas de formação de rochas, fundo do mar e água.

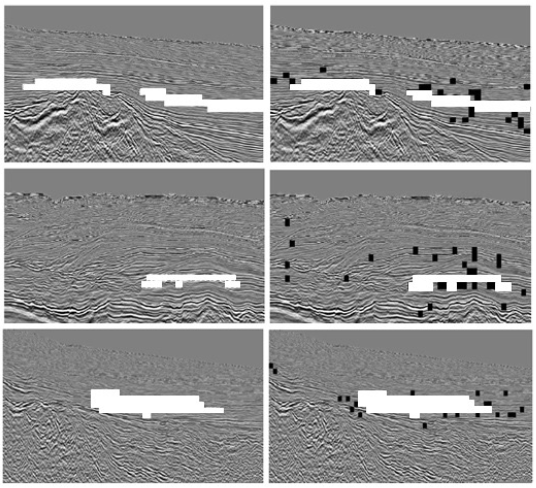


Figura 32: Teste às cegas usando CNN: Classificação de imagens através de testes às cegas para detecção de *leads*, na coluna da esquerda temos a rotulagem manual e na coluna da direita temos a rotulagem artificial através da CNN. Quadrados brancos correspondem a sub-imagens da região de interesse; os quadrados pretos, à direita, são falsos positivos erroneamente classificados.

A Tabela 02 informa alguns valores quantitativos do teste às cegas realizados pelas redes. Nela podemos enxergar a superioridade da CNN em relação a RNA utilizando os descritores de Haralick. Em termos de acurácia por exemplo, a RNA oscilou entre 76% e 89%, enquanto que na CNN o valor mínimo atingido foi de 91%.

Tabela 02: Medidas estatísticas (acurácia, precisão, *recall* e $F_1 - Score$) para o processo de classificação da região de interesse obtida por uma CNN e por uma RNA para o conjunto de teste.

Imagem	**Classificador**	**Medida estatística**			
		acc	$prec$	rec	F_1
1 (cima)	CNN	0,94	0,95	0,94	0,94
	RNA	0,89	0,88	0,91	0,90
2 (meio)	CNN	0,91	0,97	0,92	0,94
	RNA	0,76	0,81	0,77	0,79
3 (baixo)	CNN	0,91	0,93	0,91	0,92
	RNA	0,80	0,85	0,80	0,82

Vale ressaltar que foram realizados diversos testes quanto ao número de camadas ocultas da RNA, como por exemplo: 50/50, 50/128, 128/128, 128/256 e 256/256, mas a

utilização de uma única camada com 128 neurônios foi satisfatória para os resultados obtidos.

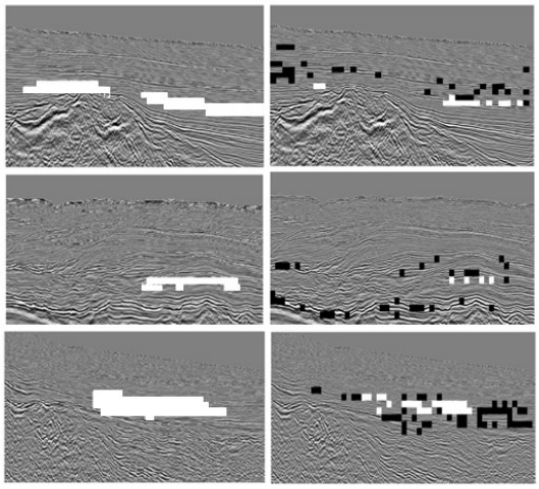


Figura 33: teste às cegas usando RNA: Classificação de imagens através de testes às cegas para detecção de *leads*, na coluna da esquerda temos a rotulagem manual e na coluna da direita temos a rotulagem artificial através da RNA a partir de descritores de Haralick. Quadrados brancos correspondem a sub-imagens da região de interesse; os quadrados pretos, à direita, são falsos positivos erroneamente classificados.

A Figura 34 mostra os gráficos da acurácia e do erro nas etapas de treino (linha vermelha) e validação (linha azul) obtidos com a RNA para um número de 50 épocas.

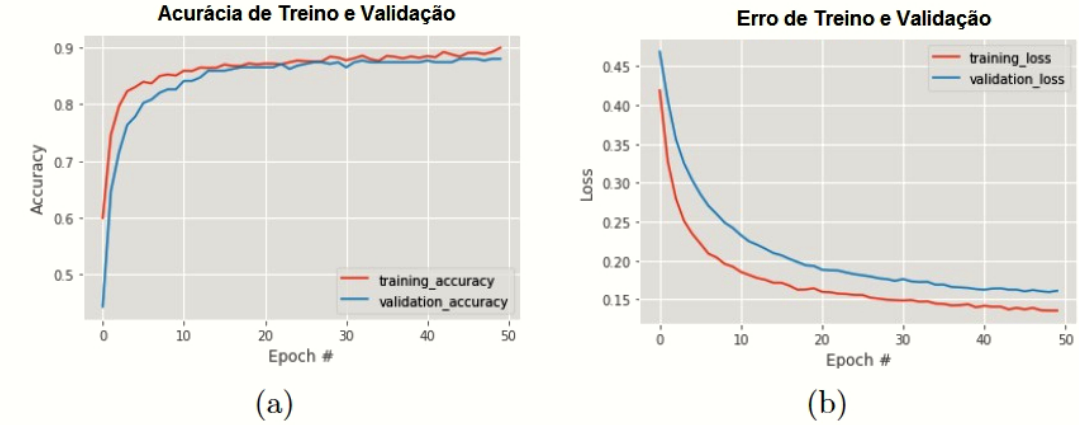


Figura 34: Curvas de acurácia (à esquerda) e perda (à direita) nas etapas de treino e validação da rede neural artificial.

De forma semelhante a Figura 35 mostra os gráficos de acurácia e do erro obtidos utilizando a CNN durante a fase de treino (linha vermelha) e validação (linha azul). Vale salientar a boa capacidade de aprendizado de ambas as redes, em destaque a CNN que atingiu um valor de 100% de acurácia na fase de treino.

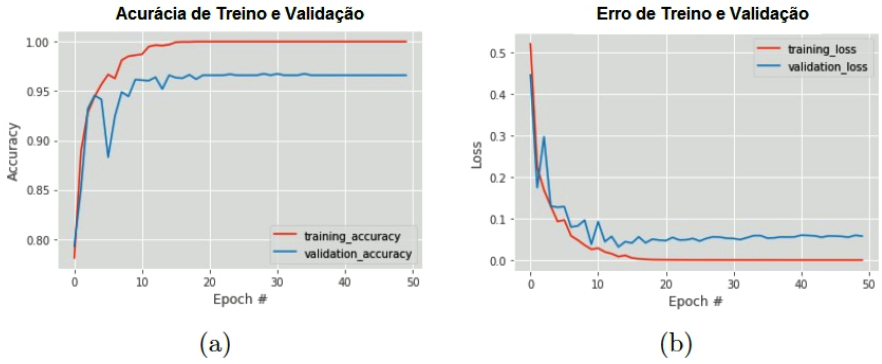


Figura 35: Curvas de acurácia (à esquerda) e perda (à direita) nas etapas de treino e validação da CNN.

4.2 Segmentação semântica a partir de *autoenconders*

Após a fase de treino da rede as imagens brutas são inseridas na rede a fim de segmentar a região de interesse. Até a obtenção do resultado final, a imagem passa por um processo de pós-processamento, conforme explicado na seção 3.4, a qual envolve reconstrução, limiarização e remoção de *outliers*. O resultado para uma das imagens de teste pode ser observado na Figura 36. Após a reconstrução da imagem segmentada conforme Figura 36(a), a etapa de limiarização foi realizada conforme seção 3.4.2 com um limiar $\tau = 0{,}5$. O resultado obtido pode ser observado na Figura 36(b), o que revela a presença de *outliers* (Figura 36(c)). Tais regiões podem ser frutos de características de textura na imagem muito semelhante a região de interesse o que pode ter acarretado a classificação de falsos positivos ou até mesmo a inclusão de regiões que não foram detectadas pelo especialista. Após a retirada dos *outliers* temos um resultado final de acordo com a Figura 36(d).

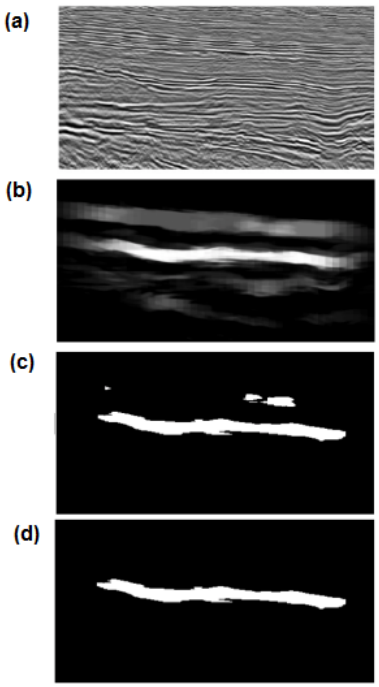


Figura 36: Etapas do pós-processamento: (a) Imagem de teste é enviada para a U-Net; (b) Imagem depois da reconstrução; (c) Imagem limiarizada; (d) Imagem após a retirada dos *outliers.*

4.3 Validação Cruzada

No sentido de garantir a capacidade de generalização da U-net, realizamos três rodadas de validação cruzada, um processo que consiste em repetir as etapas do treinamento com a troca dos conjuntos de treinamento, validação e teste. Confirmando assim a capacidade de generalização da rede, anulando a possibilidade de viés da base de dados, ou seja, que os resultados dependem da disposição dos conjuntos de treino, validação e teste.

A Tabela 03 mostra os resultados obtidos para acurácia e erro de validação em cada rodada de validação cruzada.

Tabela 03: Valores de acurácia e do erro nas três rodadas da validação cruzada da U-Net

Rodada	Acurácia	Erro
1	0,8425	0,0739
2	0,8848	0,0538
3	0,8370	0,0771

Além disso, a Figura 37 mostra os valores da perda obtida após a execução do código para cada treinamento. Como pode ser observado, a perda de validação se estabiliza em valores entre 5% e 10%, mostrando assim que os resultados da rede U-Net são razoavelmente precisos para a aplicação pretendida.

Figura 37: Perfis de perda de dados normalizados para treinamento e validação computados após 250 épocas por rodada. Em todos os casos, a perda de validação se estabiliza em valores entre 5% e 10%.

4.4 Resíduos de Segmetação

Outra perspectiva que traz enriquecimento ao método é exibida por resíduos de segmentação, conforme ilustrado na Figura 38. Os resíduos são as diferenças absolutas em pixels existentes entre a imagem da verdade do solo e o pós-processado. Como visto, os respectivos erros estão consistentemente concentrados em torno da região de interesse, como esperado, garantindo assim um interpretação dada a margem de tolerância.

Figura 38: Resíduos de segmentação: Erros absolutos em pixels, obtidos ao comparar a imagem real do solo e a imagem final após ser analisada pela rede U-net. A estreita área marginal ao redor da parte principal mostra uma evidência de que a previsão da rede é consistente com a interpretação original.

4.5 Resultados da Segmentação da U-Net

A seguir, mostramos um resumo da capacidade de segmentação da U-net, fornecendo assim uma visão qualitativada metodologia. A Figura 39 é uma sinopse do trabalho implementado: Na primeira coluna temos a imagem original rotulada pelo especialista, na coluna central temos a respectiva versão binária da imagem e por fim, na terceira coluna temos a versão final pós-processada de três imagens de teste.

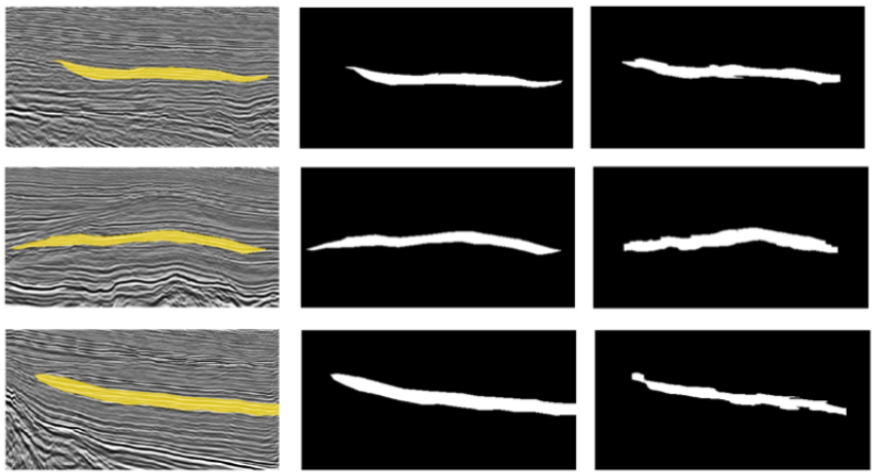


Figura 39: Sinopse da implementação: Quadro geral do trabalho implementado nesta tese destacando três imagens distintas da bacia SEAL: imagens rotuladas originais (1ª coluna), versões binárias das imagens originais (segunda coluna) e versões pós-processadas das imagens analisado pela U-Net (terceira coluna). A última área de pixel branco concorda fortemente com o região previamente interpretada (em amarelo).

CAPÍTULO V

5.1 Conclusões

Propusemos aqui a utilização de técnicas de aprendizagem de máquina para identificar padrões em imagens sísmicas. Realizamos três abordagens distintas, as duas primeiras realizaram classificações do tipo *lead* ou *no lead*. Para isso utilizamos Redes Neurais Artificiais, tendo como entrada os Descritores de *Haralick* e Redes Convolucionais. A terceira técnica envolveu o uso de segmentação semântica a partir de uma arquitetura tipo U-Net. Com base nos resultados obtidos, concluímos que:

- No que diz respeito à classificação das imagens sísmicas a partir da RNA e CNN os resultados mostraram ser promissores, porém, a abordagem feita com a CNN revelou uma qualidade superior em relação a RNA. As redes convolucionais mostraram ser robustas o suficiente para extrair as características da imagem necessárias para o aprendizado mais substancial da rede em comparação com a RNA, tal resultado é relevante, tendo em vista que a procura por descritores que alimentem a RNA pode ser um trabalho dispendioso.

- Em relação ao fato da CNN ter obtido resultados superiores em comparação a RNA a partir do uso de descritores de Haralick, isso não significa que os referidos descritores não sejam uteis na tarefa de reconhecimento de estruturas potencialmente portadoras de hidrocarbonetos.

- A terceira técnica que envolveu o processo de segmentação da imagem, apresentou duas vantagens importantes em comparação com as outras técnicas utilizadas. Uma delas é uma maior precisão dos resultados, principalmente após a etapa de pós-processamento, e a outra, é a sua capacidade de desenvolver a tarefa com um número

baixo de dados de treinamento. Se tornando assim uma ferramenta fundamental como auxílio na tomada de decisões realizadas por um especialista.

- O fato de que as técnicas de aprendizado de máquina tenham mostrado bons resultados para o desafio apresentado, isso não significa que elas possam substituir os mesmos, tendo em vista que a rede necessita ser alimentada por imagens rotuladas por alguém que possua os conhecimentos necessários para realizar a caraterização das imagens de treinamento. Porém tais ferramentas apresentam enorme potencial de aplicação para este tipo de atividade na indústria.

5.2 Recomendações de Trabalhos Futuros

A partir desta tese, podemos recomendar os seguintes trabalhos futuros:

- Construir um banco de dados de imagens sísmicas a partir de técnicas de *autoenconders* e *morphing* fornecendo assim uma base de dados significativa para o treinamento de uma rede, tendo em vista que a obtenção de um banco de dados como esse é de difícil acesso.

- Utilização de ferramentas da inteligência artificial como a U-Net na detecção de microssismos a partir de imagens 4D. Microssismos são eventos de baixa amplitude induzido por atividades de faturamento hidráulico na exploração de reservatórios.

- Fazer uso de inteligência artificial para automatizar as etapas preliminares a classificação de *plays* e *leads*.

5.3 Trabalhos aceitos

Trabalho publicado em revista internacional

SOUZA, J. F. L., SANTOS, M. D., MAGALHAES, R. M., NETO, E. M., OLIVEIRA, G. P. e ROQUE, W. L., 2019. *Automatic classification of hydrocarbon "leads" in seismic images through artificial and convolutional neural networks.* Computers & Geosciences, v. 132, p. 23-32.

Trabalho a ser submetido em revista internacional

SOUZA, J. F. L., SANTANA, G. L., BATISTA, L. V., OLIVEIRA, G. P., ROEMERS, E. Ee SANTOS, M. D., 2019. *U-Net architecture for segmentation of hydrocarbon "leads"in seismic images.* Computers & Geosciences.

Referências

ALAUDAH, Y., GAO, S. e ALREGIB, G., 2018. *Learning to label seismic structures with deconvolution networks and weak labels*. In: SEG Technical Program Expanded Abstracts 2018. Society of Exploration Geophysicists, 2018. p. 2121-2125.

ANP. Agência Nacional de Petróleo, Gás Natural e Biocombustíveis. Disponível em: http://rodadas.anp.gov.br/pt/concessao-de-blocos-exploratorios-1/13-rodada-de-licitacao-de-blocos Acesso em: 05 out. 2017.

ANGELO S. M., MATOS M. e MARFURT K. J., 2009. *Integrated seismic texture segmentation and clustering analysis to improved delineation of reservoir geometry.* In: 79th Annual International Meeting of the SEG. Expanded Abstracts, 1107-1111.

ARAYA-POLO, Mauricio et al., 2017. *Automated fault detection without seismic processing.* The Leading Edge, v. 36, n. 3, p. 208-214.

BARNES AE & LAUGHLIN K., 2002. *Investigation of methods for unsupervised classification of seismic* data. In: 72thAnnual International Meeting, SEG. Expanded Abstracts, 2221-2224.

CHEHRAZI, A., RAHIMPOUR-BONAB, H. e REZAEE, M. R., 2013. *Seismic data conditioning and neural network-based atribute selection for enhanced fault detection. Petroleum Geoscience,* Vol. **19**, pp. 169–183.

CHEVITARESE, D. *et al.*, 2018. *Seismic facies segmentation using deep learning*. AAPG Annual and Exhibition, 2018.

CHOPRA, S. e MARFURT, K., 2006. *Seismic Attributes – a promising aid for geologic prediction.* Canadian Society of Exploration Geophysicists Recorder Special Edition 31, 110–121.

COLEOU, T., POUPON, M., and AZBEL, K., 2003. *Unsupervised seismic facies classification: A review and comparison of techniques and implementation*: The Leading Edge, 22, 942–953, doi: 10.1190/1.1623635.

DE CASTRO, A. S., HOLZ, M., 2004. A tectônica de sal e a deposição de sedimentos em águas profundas na região sul da Bacia de Santos.

Di, H., WANG, Z., e ALREGIB, G., 2018. *Why using CNN for seismic interpretation? An investigation.* In *SEG Technical Program Expanded Abstracts 2018* (pp. 2216-2220). Society of Exploration Geophysicists.

DU, H., CAO, J., XUE, Y., WANG, X., 2015. *Seismic facies analysis based on self-organizing map and empirical mode decomposition*. Journal of Applied Geophysics 112: 52-61.

GOODFELLOW, I., BENGIO, Y. e COURVILE, A., 2016. *Deep Learning* MIT Press.

GOMES, M. P., VITAL, H., MACEDO, J. W. P., 2011. Fluxo de processamento aplicado a dados de sísmica de alta resolução em ambiente de Plataforma Continental: exemplo: Macau-RN. *Revista Brasileira de Geofísica*, 2011, 29.1: 173-186.

GUO, Xifeng, *et al.*, 2017. *Deep clustering with convolutional autoencoders.* In: *International Conference on Neural Information Processing.* Springer, p. 373-382.

HARALICK, R. M. e SHANMUGAN, K., 1973. *Textural Features for Image Classification.* IEEE Transactions on Systems, Man, and Cybernetics. Volume: SMC-3, Issue: 6, Nov.

HEASER, B., 2015. *National agency of petroleum, natural gas and biofuels, superintendence of de_nition of blocks*: Geologic summary and block bidding - 13th round.

HINES, W. W., MONTGOMERY, D. C. e GOLDSMAN, D. M., 2006. Probabilidade e Estatística na Engenharia - 4ª Ed. LTC.

HUANG, L., DONG, X. e CLEE, T. E., 2017. *A scalable deep learning platform for identifying geologic features from seismic attributes.* The Leading Edge, v. 36, n. 3, p. 249-256.

IGLOVIKOV, Vladimir e SHVETS, Alexey., 2018. *Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation*. arXiv preprint arXiv:1801.05746.

KINGMA, Diederik P.; BA, Jimmy., 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

LEVER, J. KRZYWINSKI, M. e ALTMAN, N., 2016. Points of significance: classification evaluation.

LEWIS, W. e VIGH, D., 2017. *Deep learning prior models from seismic images for full-waveform inversion.* In: SEG Technical Program Expanded Abstracts. Society of Exploration Geophysicists, 2017. p. 1512-1517.

MA, Y. Z. e GOMEZ, E., 2015. Uses and abuses in applying neural networks for predictions in hydrocarbon resource evaluation. Journal of Petroleum Science and Engineering, v. 133, p. 66-75.

MATOS, M. C., 2004. *Reconhecimento de Padrões Sísmicos utilizando Análise Tempo-Frequência.* Tese de Doutorado apresentada pelo Programa de Pós-Graduação em Engenharia Elétrica da PUC-Rio.

MATOS, M. C., OSÓRIO, P. L. M. & JOHANN, P. R. S., 2007. *Unsupervised seismic facies analysis using wavelet transform and self-organizing maps*. Geophysics, 72: P9-P21.

MATOS, M. C., MARFURT, K. J., e JOHANN, P. R. S., 2010. *Seismic Interpretation of self-organizing maps using 2D colors*. Revista Brasileira de Geofísica. Vol. 28. São Paulo.

MATOS, M. C., YENUGU, M. M., ANGELO, S. M. e MARFURT, K. J., 2011. *Integrated seismic texture segmentation and cluster analysis applied to channel delineation and chert reservoir characterization.* Geophysics, vol. 76. n° 5. P: P11-P21.

MONARD, M. C. e BARANAUSKAS, J. A., 2003. Conceitos sobre aprendizado de máquina. Sistemas Inteligentes: fundamentos e aplicações, v.1, n. 1, p.32.

NARAYANA, V., REDDY, E. S. e PRASAD, M., 2012. *Seetharama. Automatic image segmentation using ultra fuzziness. International Journal of Computer Applications*, 2012, 49.12.

PATTERSON, Josh; GIBSON, Adam., 2017. *Deep learning: A practitioner's approach.*" O'Reilly Media, Inc.".

PEDAMONTI, Dabal., 2018. Comparison of non-linear activation functions for deep neural networks on MNIST classification task. *arXiv preprint arXiv:1804.02763.*

PETERS, B., GRANEK, J. e HABER, E., 2019. *Automatic classification of geologic units in seismic images using partially interpreted examples.* In: 81st EAGE Conference and Exhibition 2019. 2019.

POCHET, A. *et al.*, 2018. Seismic Fault Detection Using Convolutional Neural Networks Trained on Synthetic Poststacked Amplitude Maps. IEEE Geoscience and Remote Sensing Letters, v. 16, n. 3, p. 352-356.

POWERS, D. M. W., 2011. "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation" *Journal of Machine Learning Technologies*. 2(1): 37–63.

PRIEZZHEV, I. e MANRAL, S., 2012. *3D Seismic waveform classification. International Geophysical Conference and Oil & amp*; Gas Exhibition, Istanbul, Turkey, 17-19 September 2012: 1-4.

RAILSBACK, L. B., 2015. *Petroleum Geoscience and Subsurface Geology*. Prepared for GEOL 4320/6320 Petroleum Geology Course.

ROBINSON, E. A. e TREITEL, S., 1980. *Gheophysical Signal Analysis.* Englewood Cliffs: Prentice-Hall.

RONNEBERGER, O., FISCHER, P. e BROX, 2015. T. *U-net: Convolutional networks for biomedical image segmentation*. In: International Conference on Medical image computing and computer-assisted intervention. Springer, p. 234-241.

ROY, A., ARACELI, S. R., KWIATKOWSKI, T. e MARFURT, K. J., 2014. *Generative topographic mapping for seismic facies estimation of a carbonate wash, Veracruz Basin, Southern Mexico*: Interpretation, 2, no. 1, SA31–SA47.

ROY, A., 2013. *Latent space classification of seismic facies,* PhD Thesis, University of Oklahoma, Oklahoma.
ROY, A., MATOS, M., MARFURT. J., 2010. *Automatic Seismic Facies Classification with Kohonen Self Organizing Maps* - a Tutorial. Geohorizons Journal of Society of Petroleum Geophysicists: 6-14.

SAGGAF, M. M., TOKSÔZ, M. N., MARHOON, M. I., 2003. *Seismic facies classification and identification by competitive neural networks.* Geophysics 68(6): 1984-1999.

SARASWAT, P. & SEN, M. K., 2012. *Artificial immune-based self-organizing maps for seismic-facies analysis.* Geophysics 77(4): O45-O53.

SASAKI, Y., 2007. *The truth of the F-measure.*

SHI, Y., WU, X. e FOMEL, S., 2019. *SaltSeg: Automatic 3D salt segmentation using a deep convolutional neural network.* Interpretation, v. 7, n. 3, p. SE113-SE122.

SHI, Y., WU, X. e FOMEL, S., 2018. Automatic salt-body classification using a deep convolutional neural network. In: SEG Technical Program Expanded Abstracts 2018. Society of Exploration Geophysicists, p. 1971-1975.

SILVA, R. M. *et. al*., 2019. *Netherlands Dataset: A New Public Dataset for Machine Learning in Seismic Interpretation.* arXiv preprint arXiv:1904.00770.

SILVA, M. G., PORSANI, M. J., 2006. *Aplicação de balanceamento espectral e DMO no processamento sísmico da Bacia do Tacutu.* Revista Brasileira de Geofísica, v. 24, n. 2, p. 273-290.

SIMÕES, T. A., 2017. *Identificação de zona de Produção e Recuperação de Óleo baseadas em unidades de Fluxo Hidráulico e Simulações computacionais.* Tese de doutorado do Programa de Pós Graduação em Engenharia Mecânica da UFPB.

SRIVASTAVA, Nitish, et al., 2014. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15.1: 1929-1958.

SONG, C., LIU, Z., WANG, Y., LI, X., e HU, G., 2017. *Multi-waveform classification for seismic facies analysis*. Computers and Geosciences, vol. 101.

SOUZA, J. F. L., SANTOS, M. D., MAGALHÃES, R. M., NETO, E. M., OLIVEIRA, G. P. e ROQUE, W. L., 2019. *Automatic classification of hydrocarbon "leads" in seismic images through artificial and convolutional neural networks.* **Computers & Geosciences**, v. 132, p. 23-32, 2019.

STRECKER, U., e HUDEN, R., 2002. *Datamining of 3Dpost-stack attribute volumes using Kohonen self-organizing* maps: The Leading Edge, 21, 1032–1037, doi: 10.1190/1.1518442.

TANER, M. T., 2001. *Seismic attributes.* CSEG recorder, v. 26, n. 7, p. 49-56.

TEIXEIRA, W., TOLEDO, M. C. M., FAIRCHILD, T. R., TAIOLI, F., 2001. Decifrando a terra.

THOMAS, J. E. et al. (Org.), 2004. Fundamentos de Engenharia de Petróleo. 2. ed. Rio de Janeiro: Interciência.

VIEIRA, S. Estatística experimental., 1999. 2 ed. São Paulo, Atlas.

WALDELAND, A. U. *et al.*, 2018. *Convolutional neural networks for automated seismic interpretation.* The Leading Edge, v. 37, n. 7, p. 529-537.

WEST, B. P., MAY, S. R., EASTWOOD, J. E., e ROSSEN, C. 2002. *Interactive seismic facies classification using textural attributes and neural networks.* The Leading Edge, 21(10), 1042–1049.

WU, X. et al., 2018. *Convolutional neural networks for fault interpretation in seismic images.* In: SEG Technical Program Expanded Abstracts. Society of Exploration Geophysicists, 2018. p. 1946-1950.

YILMAZ, Öz., 2001. *Seismic data analysis: Processing, inversion, and interpretation of seismic data*. Society of exploration geophysicists.

WANG, Z., DI, H., SHAFIQ, ALAUDAH, Y., AL REGIB, G., 2018. *Successful lever aging of image processing and machine learning in seismic structural interpretation: A review,* The Leading Edge 37, 451–461.

WEST, B. P., MAY, S. R., EASTWOOD, J. E., e ROSSEN, C. 2002. *Interactive seismic facies classification using textural attributes and neural networks.* The Leading Edge, 21(10), 1042–1049.

YILMAZ, Öz., 2001. *Seismic data analysis: Processing, inversion, and interpretation of seismic data*. Society of exploration geophysicists.

ZENG, Y., JIANG, K., CHEN, J., 2019. *Automatic seismic salt interpretation with deep convolutional neural networks*. In: Proceedings of the 2019 3rd International Conference on Information System and Data Mining. ACM, p. 16-20.

ZHANG, L., QUIEREN, J., SCHUELKE, J., 2001. *Self-Organizing Map (SOM) Network for Tracking and Classifying Seismic Traces.* Computational Neural Networks for Geophysical Data Processing: Handbook of Geophysical Exploration, vol. 30. chapter 10.

ZHAO, T., JAYARAM, V., ROY, A., MARFURT, K. J., 2015. *A comparison of classification techniques for seismic facies recognition.* Interpretation 3(4): SAE29-SAE58.

ZHAO, T., 2018. *Seismic facies classification using different deep convolutional neural networks*. In: SEG Technical Program Expanded Abstracts 2018. Society of Exploration Geophysicists, 2018. p. 2046-2050.

APÊNDICES

A. Descritores de Haralick

Este apêndice estabelece um formalismo matemático para explicar a ideia por trás do cálculo dos descritores de Haralick e seu significado. Os descritores de textura de Haralick descrevem uma metodologia de classificação de imagens a partir de estatística de segunda ordem, ou seja, leva em conta o posicionamento espacial relativo da ocorrência dos níveis de cinza em uma imagem. Isso é feito a partir de matrizes de co-ocorrência a partir da determinação de frequência de ocorrência de um determinado nível de cinza da imagem.

Matrizes de coocorrência (GLCM - *gray level coocurrence matrix*) são matrizes que mostram a organização espacial da ocorrência de níveis de cinza em uma imagem. São dispostas em uma organização bi-dimensionais de níveis de cinza, onde pares de pixels são separados por uma relação espacial fixa. Esta relação define a distância e a direção (d, θ), que um pixel de referência possui em relação ao pixel vizinho. Normalmente são utilizados quatro direcionamentos: 0°, 45°, 90° e 135°, conforme ilustra a Figura 40.

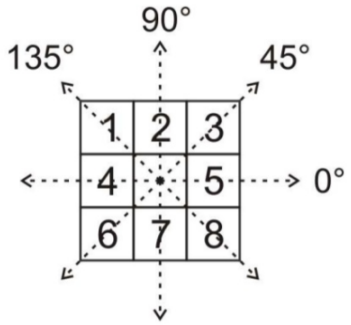


Figura 40: Variações angulares de θ utilizadas no cálculo das matrizes de co-ocorrência, considerando $d = 1$.

Dada uma matriz A obtem-se a matriz de co-ocorrência $P_d^{(\theta)} = [p(i,j,\theta,d)]$, onde $p(i,j,\theta,d)$ representa a frequência dos pontos na imagem de intensidade i que são vizinhos, a uma distância d, de pontos de intensidade j, em direção do ângulo θ. A versão normalizada dessa matriz é dada por:

$$\hat{P}_d^{(\theta)} = \frac{p(i,j,\theta,d)}{\sum_{i=0}^{N_g}\sum_{j=0}^{N_g} p(i,j,\theta,d)} \tag{10}$$

Onde:

$\hat{P}_d^{(\theta)}$: matriz de co-ocorrência normalizada

$p(i,j,\theta,d)$: elementos da matriz de co-ocorrência

N_g : número total de níveis do atributo

Para ilustrar o cálculo da matriz de co-ocorrência, considere uma matriz A de níveis de cinza dada por:

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 2 & 2 \\ 2 & 2 & 2 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 \\ 1 & 2 & 0 & 1 & 1 \end{bmatrix}$$

Logo a seguir temos o cálculo das matrizes de co-ocorrência para $\theta = 0°, 45°, 90°, 135°$ e $d = 1$.

$$P_1^{(0°)} = \begin{bmatrix} 4 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix} \qquad \hat{P}_1^{(0°)} = \begin{bmatrix} 0{,}2 & 0{,}15 & 0{,}05 \\ 0{,}15 & 0{,}15 & 0{,}05 \\ 0{,}05 & 0{,}05 & 0{,}15 \end{bmatrix}$$

$$P_1^{(45°)} = \begin{bmatrix} 2 & 2 & 1 \\ 3 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} \qquad \hat{P}_1^{(45°)} = \begin{bmatrix} 0{,}125 & 0{,}125 & 0{,}0625 \\ 0{,}1875 & 0{,}0625 & 0{,}125 \\ 0{,}125 & 0{,}125 & 0{,}0625 \end{bmatrix}$$

$$P_1^{(90°)} = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 1 & 4 \\ 3 & 3 & 0 \end{bmatrix} \qquad \hat{P}_1^{(90°)} = \begin{bmatrix} 0{,}1 & 0{,}15 & 0{,}05 \\ 0{,}15 & 0{,}05 & 0{,}2 \\ 0{,}15 & 0{,}15 & 0 \end{bmatrix}$$

$$P_1^{(135°)} = \begin{bmatrix} 3 & 1 & 2 \\ 2 & 1 & 2 \\ 2 & 3 & 0 \end{bmatrix} \qquad \hat{P}_1^{(135°)} = \begin{bmatrix} 0{,}1875 & 0{,}0625 & 0{,}125 \\ 0{,}125 & 0{,}0625 & 0{,}125 \\ 0{,}125 & 0{,}1875 & 0 \end{bmatrix}$$

Com base nas matrizes de probabilidade obtidas a partir das matrizes de co-ocorrência são calculados os descritores de Haralick. A Tabela 04 abaixo é listada os descritores utilizados nesta tese obtidos a partir da matriz de co-ocorrência.

Tabela 04: descritores de Haralick utilizados no trabalho

Descritor	**Expressão matemática**	**Significado**
Segundo Momento Angular	$\sum_i \sum_j \hat{p}^2(i,j)$	Mede a energia da imagem
Entropia	$\sum_i \sum_j \hat{p}(i,j) \log(\hat{p}(i,j))$	Mede a aleatoriedade do pixel
Contraste	$\sum_i \sum_j (i-j)^2 . \hat{p}(i,j)$	Mede a intensidade entre tons de cinza entre um pixel e seus vizinho
Variância	$\sum_i \sum_j (i-\mu)^2 . \hat{p}(i,j)$	Mede a heterogeneidade do pixel
Homogeneidade	$\sum_i \sum_j \frac{1}{1+(i-j)^2} \hat{p}(i,j)$	Mede a auto-correlação espacial
Probabilidade máxima	$\max\{\hat{p}(i,j)\}$	Mede a direção predominante da textura
Diferença de momento de 3ª ordem	$\sum_i \sum_j (i-j)^3 . \hat{p}(i,j)$	Mede o nível de distorção da imagem
Momento inverso de 2ª ordem	$\sum_i \sum_j \frac{\hat{p}(i,j)}{(i-j)^2}$	Mede o momento inverso
Momento inverso de 3ª ordem	$\sum_i \sum_j \frac{\hat{p}(i,j)}{(i-j)^3}$	Mede o momento inverso
Dissimilaridade	$\sum_i \sum_j \lvert i-j \rvert \hat{p}(i,j)$	Mede o desvio de valores entre pares de pixels
Média da soma	$\sum_i \sum_j i . \hat{p}(i,j)$	Mede a intensidade da imagem

B. **Planejamento Fatorial 2^k**

O planejamento fatorial tem sido muito aplicado em pesquisas básicas e tecnológicas e é classificado como um método do tipo simultâneo, onde as variáveis de interesse que realmente apresentam influências significativas na resposta são avaliadas ao mesmo tempo. Para realizar um planejamento fatorial, escolhem-se as variáveis a serem estudadas e efetuam-se experimentos em diferentes níveis desses fatores. A seguir são realizados experimentos para todas as combinações possíveis dos níveis selecionados. De um modo geral, o planejamento fatorial pode ser representado por b^k, onde "k" é o número de fatores "b" é o número de níveis escolhidos. Em geral, os planejamentos fatoriais do tipo 2^k são os mais comuns. Um dos aspectos favoráveis deste tipo de planejamento é a realização de poucos experimentos. Torna-se óbvio que com um número reduzido de níveis não é possível explorar de maneira completa uma grande região no espaço das variáveis. Entretanto, pode-se observar tendências importantes para a realização de investigações posteriores, (VIEIRA, 1999).

Segundo (HINES, 2006) o planejamento 2^2 é o mais simples planejamento do tipo 2^k, pois somente dois fatores *A* e *B* são envolvidos, cada um deles em dois níveis. Em geral, consideramos esses níveis como os níveis “alto” e “baixo” do fator. A Figura 41 mostra o planejamento 2^2. Note que esse planejamento pode ser representado geometricamente por um quadrado. Usa-se uma notação especial para representar as combinações de tratamento. Em geral, uma combinação de tratamento é representada por uma série de letras minúsculas. Se uma letra está presente, então o fator correspondente é rodado no nível mais alto na combinação de tratamento; se ela está ausente, o fator é rodado em seu nível baixo.

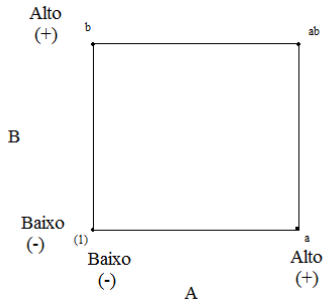


Figura 41: Planejamento fatorial 2^2.
Adaptado de (HINES, 2006)

Por exemplo, a combinação de tratamento $\boldsymbol{a}$ indica que o fator A está em seu nível mais alto e o fator B em seu nível mais baixo. A combinação de tratamento com ambos os fatores em nível baixo é denotada por (1).

Os efeitos de interesse no planejamento 2^k são os efeitos principais A e B e a interação de dois fatores AB. Sejam, também, $(1), a, b$ e ab os totais de todas as n observações tomadas nesses pontos do planejamento. O valor dos efeitos principais de um planejamento 2^k pode ser determinado pelas equações abaixo:

$$A = \frac{1}{2n}[a + ab - b - (1)] \tag{11}$$

$$B = \frac{1}{2n}[b + ab - a - (1)] \tag{12}$$

$$AB = \frac{1}{2n}[ab + (1) - a - b] \tag{13}$$

As quantidades entre colchetes nas equações 11 a 13 são chamas de *contrastes*. Para obter mais detalhes sobre as equações acima ver (HINES, 2006).

Nessas equações, os coeficientes do contraste são sempre $+1$ ou -1. Uma tabela de sinais mais e menos, como a Tabela 05, pode ser usada para se determinar o sinal de cada combinação de tratamento para um contraste particular.

Tabela 05: Sinais dos Efeitos no planejamento 2^k

Combinações de Tratamentos	Efeito Fatorial			
	I	A	B	AB
(1)	+	−	−	+
a	+	+	−	−
b	+	−	+	−
ab	+	+	+	+

A magnitude desses efeitos principais pode ser obtida a partir do valor do contraste de cada efeito. Isso é realizado a partir da Equação 26 que expressa a relação entre um contraste e sua soma de quadrados:

$$SQ = \frac{(Contraste)^2}{n.\sum(coeficientes\ do\ contraste)^2} \quad (14)$$

Portanto, as somas de quadrados para A, B e AB são:

$$SQ_A = \frac{[a + ab - b - (1)]^2}{4n} \quad (15)$$

$$SQ_B = \frac{[b + ab - a - (1)]^2}{4n} \quad (16)$$

$$SQ_{AB} = \frac{[ab + (1) - a - b]^2}{4n} \quad (17)$$

Por exemplo, deseja-se analisar qual a influência dos fatores taxa de aprendizado e constante de momento nos resultados da rede quanto ao resultado da acurácia da rede. A Figura 41 indica os níveis de cada fator.

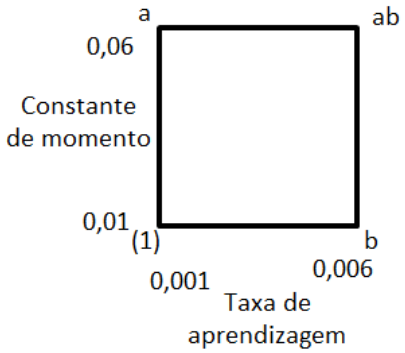


Figura 42: Aplicação do planejamento fatorial para a acurácia.

Para o resultado foram realizados quatro testes cujos resultados encontram-se indicados na Tabela 06 a seguir.

Tabela 06: Matriz de planejamento fatorial 2^2

Combinações de Tratamentos	Fatores		Respostas				Total	Média
	A	B						
(1)	−	−	0,815	0,817	0,815	0,798	3,244	0,811
a	+	−	0,839	0,851	0,853	0,847	3,400	0,850
b	−	+	0,820	0,741	0,768	0,765	3,084	0,774
ab	+	+	0,847	0,848	0,847	0,848	3,392	0,848

Com base na Tabela 06 determinam-se os valores dos efeitos principais conforme é mostrado a seguir.

$$A = \frac{1}{2n}[a + ab - b - (1)] = \frac{1}{2 \times 4}[3{,}400 + 3{,}392 - 3{,}084 - 3{,}244] = 0{,}058$$

$$B = \frac{1}{2n}[b + ab - a - (1)] = \frac{1}{2 \times 4}[3{,}084 + 3{,}392 - 3{,}400 - 3{,}244] = -0{,}021$$

$$AB = \frac{1}{2n}[ab + (1) - a - b] = \frac{1}{2 \times 4}[3{,}392 + 3{,}244 - 3{,}400 - 3{,}084] = 0{,}152$$

Com base nas estimativas numéricas dos efeitos pode-se observar que a iteração entre os efeitos margem de treinamento e constante de momento são grandes e tem uma direção positiva, ou seja, aumentando-se a margem de treinamento e a constante de momento, aumenta-se o percentual de acertos na rede.

A magnitude desses efeitos pode ser determinada a partir do valor do contraste, conforme é mostrado a seguir.

$$SQ_A = \frac{[a + ab - b - (1)]^2}{4n} = \frac{[0{,}464]^2}{4 \times 4} = 0{,}013$$

$$SQ_B = \frac{[b + ab - a - (1)]^2}{4n} = \frac{[-0{,}168]^2}{4 \times 4} = 0{,}002$$

$$SQ_{AB} = \frac{[ab + (1) - a - b]^2}{4n} = \frac{[1{,}216]^2}{4 \times 4} = 0{,}092$$

Para o caso de 3 fatores cada um com dois níveis, esse planejamento é chamado de planejamento fatorial 2^3 e tem oito combinações de tratamento. Geometricamente, o planejamento pode ser mostrado como na Figura 20, com as oito rodadas formando os vértices do cubo. Esse planejamento permite que sejam estimados três efeitos principais (A, B e C) juntamente com três interações de dois fatores (AB, AC e BC) e uma interação de três fatores (ABC).

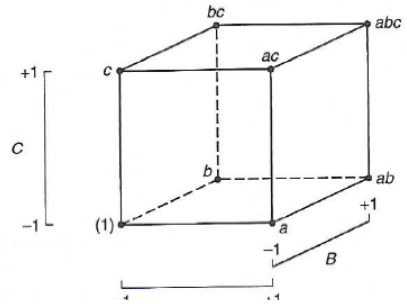


Figura 43: Planejamento 2^3. (HINES, 2006)

Os valores dos efeitos principais são determinados pelas equações a seguir;

$$A = \frac{1}{4n}[a + ab + ac + abc - b - c - bc - (1)] \quad (18)$$

$$B = \frac{1}{4n}[b + ab + bc + abc - a - c - ac - (1)] \quad (19)$$

$$C = \frac{1}{4n}[c + ac + bc + abc - a - b - ac - (1)] \quad (20)$$

$$AB = \frac{1}{4n}[ab + (1) + abc + c - b - a - bc - ac] \quad (21)$$

$$AC = \frac{1}{4n}[ac + (1) + abc + b - a - c - ab - bc] \quad (22)$$

$$BC = \frac{1}{4n}[bc + (1) + abc + a - b - c - ab - ac] \quad (23)$$

$$ABC = \frac{1}{4n}[abc - bc - ac + c - ab + b + a - (1)] \quad (24)$$

C. Artigo aceito em revista internacional

Computers & Geosciences
Volume 132, November 2019, Pages 23-32

Automatic classification of hydrocarbon "leads" in seismic images through artificial and convolutional neural networks

J.F.L. Souza [a, b], M.D. Santos [a, b], R.M. Magalhães [a, b], E.M. Neto [c], G.P. Oliveira [a, b] ✉, W.L. Roque [a]

[a] Petroleum Engineering Modelling Laboratory, Federal University of Paraíba, João Pessoa, Brazil
[b] Postgraduate Program in Mechanical Engineering, Federal University of Paraíba, João Pessoa, Brazil
[c] Postgraduate Program in Informatics, Federal University of Paraíba, João Pessoa, Brazil

Received 10 January 2019, Revised 21 June 2019, Accepted 1 July 2019, Available online 11 July 2019.