



UNIVERSIDADE FEDERAL DO MARANHÃO
Programa de Pós-Graduação em Ciência da Computação

Petterson Sousa Diniz

**Uma Abordagem Utilizando Séries Temporais
para Detecção de Gás em Imagens Sísmicas
com Transformer**

São Luís - MA

2022

Petterson Sousa Diniz

Uma Abordagem Utilizando Séries Temporais para Detecção de Gás em Imagens Sísmicas com Transformer

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Ciência da Computação, ao Programa de Pós-Graduação em Ciência da Computação, da Universidade Federal do Maranhão.

Programa de Pós-Graduação em Ciência da Computação

Universidade Federal do Maranhão

Orientador: Prof. Dr. Anselmo Cardoso de Paiva

Coorientador: Prof. Dr. Aristófanés Correa da Silva

São Luís - MA

2022

Petterson Sousa Diniz

Uma Abordagem Utilizando Séries Temporais para Detecção de Gás em Imagens Sísmicas com Transformer

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Ciência da Computação, ao Programa de Pós-Graduação em Ciência da Computação, da Universidade Federal do Maranhão.

São Luís - MA, 06 de Janeiro de 2023:

Prof. Dr. Anselmo Cardoso de Paiva

Orientador

Universidade Federal do Maranhão

Prof. Dr. Aristóфанes Correa da Silva

Coorientador

Universidade Federal do Maranhão

Prof. Dr. Geraldo Braz Junior

Examinador Interno

Universidade Federal do Maranhão

Prof. Dr. Kelson Romulo Texeira

Aires

Examinador Externo

Universidade Federal do Piauí

São Luís - MA

2022

*A **Deus**, primeiramente e acima de tudo.*

*Ao **NCA**, ao **TecGraf** e a essa instituição (**UFMA**), que eu aprendi a admirar
imensamente, tanto a metodologia como as pessoas que a frequentam.*

*À minha **família** e **amigos**, por toda a fé e apoio.*

*Ao **meu orientador**, **Anselmo**, por toda a determinação e perseverança.*

A estes, dedico este trabalho.

Agradecimentos

Agradeço, primeiramente e acima de tudo, à Deus, por sua imensa bondade comigo, ainda que eu não seja merecedor.

Agradeço a minha família, que me sustentou, me ensinou valores valiosos que levarei para toda a vida e me proporcionaram os alicerces para que eu chegasse até esse momento.

Agradeço ao meu orientador, professor, mentor e maior incentivador, Anselmo Cardoso de Paiva, que acreditou em mim até em momentos em que eu mesmo estava desacreditado. Agradeço por todos os conselhos e a insistência, pois sem estes, eu não estaria concluindo este trabalho.

Agradeço ao meu coorientador, Aristófanés Correa da Silva, por ter me dado a oportunidade de desenvolver este trabalho junto a ele. Sem a sua mentoria este trabalho não teria obtido o sucesso que obtive. Grato por estar sempre presente conosco mesmo nos momentos que parecia que não tínhamos como contornar os problemas, mas que conseguimos passar com muita dedicação e companheirismo.

Agradeço aos meus amigos Johnatan e Thayane, que me foram boas companhias em momentos de descontração e foram de grande ajuda nos momentos mais diversos da minha vida.

Agradeço em especial ao meu amigo e irmão Domingos Alves Dias Júnior, com quem eu dividi grande parte desta jornada. Não apenas acadêmica, mas em todas as áreas. Obrigado, amigo. Você é um amigo!

Agradeço aos meus professores e mestres que acompanharam essa jornada.

Agradeço aos meus amigos João Bandeira e Giovanni Lucca pelo apoio nas escritas de artigos e até mesmo deste trabalho de dissertação.

Agradeço ao Núcleo de Ciência Aplicada e a todos com quem eu tive a felicidade cooperar. Todos vocês foram importantes para o meu crescimento.

Finalmente, agradeço a todos os que contribuíram de forma direta ou indireta para este momento.

"Music is the most advanced form of mathematics."

(GPT-3 Inteligência Artificial)

Resumo

A detecção de reservatórios de gás em imagens sísmicas é complexa, demorada e exige que tenham profissionais especializados para tal tarefa. O grande desafio está na grande quantidade de dados para serem analisados e na dificuldade e recursos necessários para se fazer a comprovação da existência de um reservatório de gás, uma vez que é preciso fazer a perfuração do poço para atestar o fato. Uma opção para contornar a grande quantidade de dados para serem analisados é recorrer ao uso de inteligência artificial através da implementação de redes neurais profundas. Nesse sentido, a escassez de dados comprovadamente corretos impacta no número de amostras anotadas a que temos acesso, o que dificulta de forma significativa o uso de métodos computacionais. Visando contornar as dificuldades citadas, este trabalho propõe o uso de um método de segmentação de reservatórios de gás a partir de uma perspectiva unidimensional com análise de séries temporais. Devido à característica sequencial dos dados sísmicos, é possível se empregar uma metodologia que interpreta cada traço sísmico presente nas imagens de entrada como sendo uma instância isolada da base de dados. Dessa forma, é possível aumentar o número de amostras no conjunto de dados de forma expressiva. Para fazer a segmentação dos reservatórios foi proposta uma nova arquitetura, que consiste em modificar uma arquitetura de rede Transformer para que seja capaz de interpretar séries temporais de dados sísmicos. A rede proposta é do tipo sequência para sequência, o que significa que a rede processa uma sequência de entrada e retorna uma nova sequência contendo a região de gás segmentada. A escolha deste método se deu pelo fato de esta arquitetura ser eficaz em extrair informações contextuais sem a necessidade do uso de recorrência em seu treinamento, o que resulta em uma melhor performance e permite a paralelização do processo, resultando em uma grande economia de tempo, tanto no treinamento quanto na inferência. A base de dados utilizada nos experimentos é privada e foi concedida pela Eneva para o desenvolvimento desta tecnologia. As métricas utilizadas para avaliação dos resultados foram acurácia, sensibilidade, especificidade e AUC, onde o método obteve 97.16%, 79.61%, 97.47% e 88.54%, respectivamente.

Palavras-chave: redes neurais, Transformer, sísmica.

Abstract

The detection of gas reservoirs in seismic images is complex, time consuming and requires specialized professionals for this task. The great challenge lies in the large amount of data to be analyzed and the difficulty and resources needed to prove the existence of a gas reservoir, since it is necessary to drill the well to prove the fact. One option to get around the large amount of data to be analyzed is to resort to the use of artificial intelligence through the implementation of deep neural networks. In this sense, the scarcity of proven correct data impacts the number of annotated samples that we have access to, which significantly hinders the use of computational methods. In order to overcome the aforementioned difficulties, this work proposes the use of a gas reservoir segmentation method from a one-dimensional perspective with time series analysis. Due to the sequential characteristic of seismic data, it is possible to employ a methodology that interprets each seismic trace present in the input images as an isolated instance of the database. In this way, it is possible to increase the number of samples in the data set significantly. To perform the segmentation of the reservoirs, a new architecture was proposed, which consists of modifying a Transformer network architecture so that it is capable of interpreting temporal series of seismic data. The proposed network is of the sequence-to-sequence type, which means that the network processes an input sequence and returns a new sequence containing the segmented gas region. The choice of this method was due to the fact that this architecture is effective in extracting contextual information without the need to use recurrence in its training, which results in a better performance and allows the parallelization of the process, resulting in a great saving of time, both in training and inference. The database used in the experiments is private and was granted by Eneva for the development of this technology. The metrics used to evaluate the results were accuracy, sensitivity, specificity and AUC, where the method obtained 97.16%, 79.61%, 97.47% and 88.54%, respectively.

Keywords: neural networks, Transformer, seismic.

Lista de ilustrações

Figura 1 – Arquitetura da rede Transformer.	22
Figura 2 – Ilustração do mecanismo de auto-atenção.	23
Figura 3 – Atenção e Atenção de Múltiplas Cabeças.	24
Figura 4 – Fluxograma da metodologia proposta.	26
Figura 5 – Parque dos Gaviões	27
Figura 6 – Exemplo de uma linha 2D da base de dados.	27
Figura 7 – Configuração da ROI. a) imagem sísmica; b) primeira etapa: marcação humana; e c) segunda etapa: tamanho adaptativo.	30
Figura 8 – Extração de <i>patches</i> . a) ROI; b) decomposição em sequencia de traços; e c) extração do <i>patch</i> do traço.	31
Figura 9 – Estudo de caso: (a), (b), e (c) são três amostras diferentes. Em vermelho estão os falsos positivos. Em azul estão os verdadeiros positivos. Em verde estão os falsos negativos.	38

Lista de tabelas

Tabela 1 – Distribuição da base de dados.	34
Tabela 2 – Resultado dos experimentos.	34
Tabela 3 – Comparativo - F3-Block	36
Tabela 4 – Comparativo - Gaviões Branco e Vermelho	36
Tabela 5 – Comparativo - Parque dos Gaviões.	37

Lista de abreviaturas e siglas

2D	- <i>Bidimensional</i>
AUC	- <i>Area Under the ROC Curve</i>
CNN	- <i>Convolutional Neural Network</i>
IoU	- <i>Intersection over Union</i>
LSTM	- <i>Long Short-Term Memory</i>
MLP	- <i>Multi Layer Perceptron</i>
NLP	- <i>Natural Language Processing</i>
O&G	- <i>Óleo e Gás</i>
PLN	- <i>Processamento da Linguagem Natural</i>
PSO	- <i>Particle Swarm Optimization</i>
ROC	- <i>Receiver Operating Characteristic</i>
ROI	- <i>Region of Interest</i>
RNA	- <i>Rede Neural Artificial</i>
SVM	- <i>Support Vector Machines</i>

Sumário

1	INTRODUÇÃO	12
1.1	Objetivos	13
1.1.1	Objetivos Específicos	13
1.2	Organização do Trabalho	14
2	TRABALHOS RELACIONADOS	15
3	FUNDAMENTAÇÃO TEÓRICA	18
3.1	Sísmica	18
3.2	Séries Temporais	19
3.3	Redes Neurais Artificiais	19
3.4	Aprendizagem Profunda	20
3.5	Transformer	21
3.5.1	Auto-Atenção	22
3.5.2	Atenção de Múltiplas Cabeças	24
3.5.3	Codificação Posicional	24
3.6	Time2Vec	25
4	METODOLOGIA	26
4.1	Base de Dados	26
4.2	Dimensionalidade	28
4.3	Preprocessamento	29
4.4	Experimentos	30
4.5	Arquitetura Transformer	31
4.6	Inferência e Reconstrução	32
5	EXPERIMENTOS E RESULTADOS	33
5.1	Ambiente de desenvolvimento	33
5.2	Construção das bases de dados	33
5.3	Treinamento	34
5.4	Comparação com a Literatura	36
5.5	Estudo de caso	37
6	CONCLUSÃO	39
	REFERÊNCIAS	40

1 Introdução

A sísmica de reflexão é um dos métodos geofísicos mais utilizados pela indústria de óleo e gás (O&G). Consiste em gerar uma onda de energia massiva na superfície através da detonação controlada de explosivos para capturar o retorno dessa onda após se propagar pelo solo. É possível extrair dados referentes à geometria e estrutura das camadas geológicas, tipos de rocha, litologia e propriedades físicas do solo de uma determinada região em terra ou no mar (MORTON-THOMPSON; WOODS et al., 1993).

Tendo posse dessas informações, pode-se inferir a localização e o volume de prováveis acúmulos de gás (SANTOS, 2019). Para isso, é necessário focar em dois aspectos fundamentais da sísmica: tempo de chegada e forma de onda. O tempo de chegada das ondas refletidas está relacionado com a profundidade das interfaces geológicas, enquanto a forma, a intensidade, o espectro de frequência e a distribuição das frequências estão relacionados ao tipo de rocha de cada camada (litologia) e suas propriedades físicas (MORTON-THOMPSON; WOODS et al., 1993).

No entanto, essa grande quantidade de informação exige muito tempo e esforço de equipes especializadas, que precisam escolher um modelo geológico final dentre a enorme quantidade de possíveis cenários que os dados permitem. Todavia, esse tempo não é compatível com os prazos curtos impostos pela indústria e pelo volume de dados que precisam ser analisados (SANTOS, 2019).

Diversas áreas com contextos similares, com grandes quantidades de dados e curtos prazos, tem feito uso de redes neurais artificiais profundas como solução para esse problema. As redes neurais profundas vem produzindo resultados eficientes e cada vez mais confiáveis, como por exemplo, no processamento de linguagem natural (COLLOBERT; WESTON, 2008), processamento de imagens (KRIZHEVSKY; SUTSKEVER; HINTON, 2012a) e processamento de sinais (ROY; KIRAL-KORNEK; HARRER, 2018).

Em geofísica, o uso de redes neurais enfrenta uma dificuldade que não se encontra nas outras áreas citadas. Existe uma grande dificuldade em se comprovar a existência ou não de reservatórios de gás, pois o processo de perfuração de poços é muito caro, o que faz com que o número de dados confiáveis seja bem reduzido, pois não há certeza da existência de reservatórios até que seja perfurado. Devido a essa limitação na quantidade de dados rotulados, o treinamento se torna bastante difícil, pois exige que muitas amostras estejam rotuladas como região de gás ou não. Entretanto, essa informação, na maioria dos casos, não está disponível (SANTOS, 2019).

Técnicas para geração de dados sintéticos podem ser utilizados para ampliar a quantidade de dados rotulados. O benefício na utilização desse tipo de técnica é o aumento

da robustez do modelo gerado com dados levemente distorcidos, além de impedir que o modelo aprenda possíveis ruídos presentes nos dados. Em sismica, existem trabalhos focados na geração de bases sintéticas para simulação de falhas sísmicas (WU et al., 2020), mas a geração de dados sísmicos que simulem assinaturas de reservatórios de gás é bem mais complicada (SANTOS, 2019).

Outra forma de resolver a falta de dados é processar os dados sísmicos como uma sequência de dados unidimensionais, onde é possível gerar uma maior quantidade de amostras (patches) para treinamento. (SANTOS, 2019). Dessa forma, é possível alimentar uma rede neural com uma quantidade substancialmente maior de dados sem que seja necessário fazer novas anotações de reservatórios de gás.

Este trabalho contribui diretamente para diversas áreas. No campo da geofísica, contribui no desenvolvimento de um novo método automático para detecção de hidrocarbonetos em imagens sísmicas que pode ser usado para o auxílio de geofísicos na análise exploratória de campos com potenciais acumulações de gás. Na área da computação, especificamente no processamento de sinais, contribui na utilização de uma rede neural que obteve resultados promissores em Processamento de Linguagem Natural (PLN) para fazer análise unidimensional de imagens sísmicas e na adaptação a arquitetura da rede proposta para trabalhar de forma otimizada com dados de séries temporais.

1.1 Objetivos

Avaliar um método para detectar potenciais acumulações de gás em imagens sísmicas baseado em uma abordagem unidimensional utilizando a rede neural Transformer com abordagem baseada em séries temporais.

1.1.1 Objetivos Específicos

Para alcançar o objetivo geral deste trabalho, foi necessário atingir os seguintes objetivos específicos:

- Investigar estratégias de pre-processamento dos dados para homogenizar os valores de amplitude, que variam devido a mudança no processo de aquisição;
- Desenvolver ma arquitetura de rede neural Transformer para trabalhar com dados de séries temporais;
- Desenvolver ma camada de saída da rede neural Transformer para que seja capaz de trabalhar na solução de um problema de segmentação;
- Avaliar o desempenho do método proposto através de experimentos feitos em uma base de dados privada.

1.2 Organização do Trabalho

Esta dissertação está organizada em seis capítulos, sendo o primeiro esta introdução.

O Capítulo 2 apresenta os trabalhos relacionados que utilizando redes neurais e aprendizado de máquina para fazer análises e inferências em imagens sísmicas.

O capítulo 3 apresenta a fundamentação teórica e expande termos que são necessários para a compreensão do método proposto.

O capítulo 4 detalha a metodologia proposta, trazendo uma visão geral da arquitetura da rede e uma descrição detalhada de cada uma das etapas do método.

O capítulo 5 mostra os testes realizados e os resultados obtidos.

Por fim, o Capítulo 6 faz as considerações finais deste trabalho, artigos desenvolvidos e apresenta propostas para trabalhos futuros.

2 Trabalhos Relacionados

O uso de aprendizagem profunda tem se difundido bastante nos últimos anos devido a grande evolução na capacidade de processamento que os computadores vem ganhando a cada ano. Métodos utilizando aprendizagem profunda são capazes de aprender de forma dinâmica novos atributos relevantes durante o seu treinamento, diferente dos métodos clássicos, onde é necessário fazer uma extração de características (COLLOBERT; WESTON, 2008; KRIZHEVSKY; SUTSKEVER; HINTON, 2012a; ROY; KIRAL-KORNEK; HARRER, 2018).

Devido a essa capacidade, métodos utilizando aprendizagem profunda tem sido usados em diversas áreas e obtendo bons resultados. Na área de geofísica, encontramos trabalhos que tem diferentes objetivos, por exemplo, detecção de falhas sísmicas, geração de dados sintéticos para treinamento de detecção de anomalias e detecção de gás (WU et al., 2020; SANTOS, 2019).

No trabalho de Araya-Polo et al. (2017) é proposto um método para detecção de falhas sísmicas, que são deslocamentos de camadas de rocha com o objetivo de encontrar hidrocarbonetos que podem estar retidos nessas estruturas. Para este objetivo, o trabalho propõe o uso de redes neurais profundas ao invés de métodos clássicos de processamento sísmico devido à grande capacidade de redes neurais se adaptarem a problemas extremamente não lineares. As métricas de validação utilizadas foram o AUC (*Area Under the ROC Curve*) e o IoU (*Intersection over Union*), onde valores mais próximos de 1 representam melhores resultados. Os maiores índices obtidos nos testes para cada uma dessas métricas foi de 0,902 para o AUC e de 0,311 para o IoU.

Um tipo de rede neural eficiente para fazer processamento de sinais em uma ou mais dimensões são as redes neurais convolucionais (*Convolutional Neural Networks - CNNs*). Sua utilização é bastante versátil, visto que é possível obter resultados a partir de atributos pré calculados, que são processamentos feitos aos dados a fim de gerar características antes de enviar para rede neural, ou utilizar os dados não processados diretamente na rede, deixando a tarefa de extrair características com a própria rede neural. No trabalho de Wang et al. (2018) encontramos um estudo onde o uso de CNNs é feito para interpretar imagens sísmicas a fim de encontrar falhas sísmicas ou domos de sal, que são estruturas que podem indicar a existência de hidrocarbonetos líquidos em alguma região. Os autores concluem que os métodos de interpretação de dados baseados em CNNs foram capazes de capturar com eficiência as características espaciais e estruturais dos dados sísmicos, tendo melhor desempenho do que técnicas convencionais de aprendizado de máquina.

O uso de CNNs associado a atributos sísmicos pré calculados pode ser encontrado

no trabalho de [Huang, Dong e Clee \(2017\)](#), que propõe, além de uma plataforma de Big Data com uma infra estrutura de memória distribuída para processar grandes volumes de dados, o uso de CNNs para obter resultados aprimorados em relação aos métodos utilizados até o momento, como regressão logística. Já [Di, Wang e AlRegib \(2018b\)](#) decide fazer o uso do dado sísmico sem extração de atributos em uma CNN para fazer detecção de falhas sísmicas. Os autores concluem que o uso de CNNs juntamente com os dados originais é capaz de encontrar falhas sísmicas de forma otimizada comparada com métodos que selecionam os atributos manualmente. De forma similar, [Pochet et al. \(2018\)](#) faz uso de CNNs baseada em pacotes (*patches*) em dados sísmicos gerados sinteticamente a fim de detectar falhas sísmicas, porém neste trabalho são analisadas diferentes arquiteturas de CNNs, obtendo resultados de 69% de sensibilidade e 99% de precisão.

A utilização de CNNs também foi feita para classificar camadas no registro sísmico, também chamadas de fácies sísmicas. No trabalho de [Chevitarese et al. \(2018\)](#) é proposto um método que consiste em dividir as imagens sísmicas em imagens menores, chamadas de partições, que alimentam uma CNN. Algumas dessas partições recebem a interpretação de um especialista é utilizada para fazer a classificação das partições não interpretadas. Alguns parâmetros foram variados, como o tamanhos das partições e a quantização dos dados, obtendo 0,98 de f1-score em sua melhor combinação de parâmetros.

A detecção de domos de sal é uma tarefa bastante difícil para os interpretes de dados sísmicos, porém a detecção dessas estruturas é um forte indicativo da presença de um reservatório de gás. No trabalho de [Di, Wang e AlRegib \(2018a\)](#) é apresentado um método para a detecção de sal usando aprendizagem profunda. Neste trabalho, os autores usam uma CNN para fazer a detecção e delineamento automático dos domos de sal e compara os resultados com marcações de especialistas e técnicas de aprendizado de máquina, como o SVM e uma rede MLP. A conclusão do trabalho foi que o uso de CNNs tem vantagem sobre as abordagens clássicas devido a grande quantidade de atributos gerados automaticamente pela rede pelas camadas de convolução, sendo capazes de encontrar padrões que não foram previamente percebidos no dado sísmico.

Uma outra forma de lidar com os dados sísmicos é fazer o uso do sinal independente de cada receptor, chamado de geofone, trabalhando com o traço sísmico. O uso de redes neurais unidimensionais é encontrado no trabalho de [Santos \(2019\)](#), que utiliza uma rede LSTM para fazer a detecção dos reservatórios de gás. Neste trabalho, cada traço sísmico é dividido em pacotes que são a entrada da rede LSTM, junto com a rotulação de cada pacote. A rede gera um mapa de probabilidades de se encontrar gás em cada região da imagem sísmicas. Para o teste foi utilizada a base de dados pública Netherlands F3-Block. Os resultados foram promissores e alcançaram 97% de acurácia.

O trabalho de [Junior et al. \(2021\)](#), além de fazer uso dos dados sísmicos em forma unidimensional, também faz o uso da arquitetura *Transformer* adaptada com camadas de

pré-processamento para introduzir a característica temporal à rede. Neste trabalho, foi feito o uso de uma camada *Time Delay Embedding*, que codifica a característica temporal da sequência antes de passar os dados para a próxima camada. Essa camada é parametrizada de forma estática durante a definição da arquitetura e não é capaz de aprender com os dados que são retro propagados pela rede. No entanto, mesmo com essas limitações, obteve resultados de acurácia superior a 97% na base pública Netherlands F3-Block, com sensibilidade e especificidade também superiores a 97%.

Posteriormente, o trabalho [Junior et al. \(2022\)](#) trouxe mais uma aplicação de uma arquitetura *Transformer* para imagens sísmicas. Desta vez fazendo o uso de pacotes de dados bidimensionais extraídos das imagens sísmicas que alimentam uma arquitetura *Vision Transformer* otimizada com o uso do algoritmo PSO (*Particle Swarm Optimization*). O método desenvolvido foi capaz de alcançar uma acurácia superior a 99% na base pública Netherlands F3-Block, mantendo métricas de sensibilidade e especificidade em 88,60% e 99,56%, respectivamente.

Pode-se observar nos trabalhos apresentados que o uso de redes neurais permite que a análise massiva de dados sísmicos seja feita de forma mais precisa e menos custosa do ponto de vista de recursos humanos. Essas metodologias são capazes de auxiliar geofísicos na análise de levantamentos sísmicos fornecendo informações de possíveis acumulações de hidrocarbonetos ou de estruturas frequentemente associadas a tais eventos. O uso de metodologias unidimensionais possui a vantagem de permitir que o treinamento seja feito com uma quantidade de imagens menor, visto que cada imagem será fragmentada em vários traços sísmicos. Por ter essa característica e ainda ser uma abordagem pouco explorada em comparação com o uso de CNNs, este trabalho se propõe a apresentar uma metodologia unidimensional para detecção de reservatórios de gás utilizando a rede neural Transformer, que tem se mostrado superior a LSTM em outros problemas.

3 Fundamentação Teórica

Este capítulo apresenta a fundamentação teórica que embasa o método apresentado.

3.1 Sísmica

Os métodos sísmicos podem ser divididos em dois grupos: métodos de refração e métodos de reflexão. O método de refração consiste em registrar somente as ondas refratadas com ângulo crítico e tem aplicações na área de sismologia. Esse método tornou possível desvendar a estrutura interna da terra, porém sua aplicação na área de petróleo é bastante restrita, embora tenha tido larga utilização até meados da década de 50 (SHERIFF; GELDART, 1995).

Por outro lado, o método de reflexão é o mais utilizado na indústria de Óleo e Gás (O&G), obtendo mais de 90% dos investimentos em prospecção, por fornecer uma alta definição das feições geológicas em subsuperfícies propícias à acumulação de gás por um custo bem inferior (THOMAS et al., 2004).

O levantamento sísmico se inicia com a geração de ondas elásticas, através de fontes artificiais, que se propagam pelo interior da Terra. Essas ondas são refletidas ou refratadas nas interfaces que separam rochas de diferentes estruturas e retornam à superfície onde são capturadas (ABDULAZIZ; ALSAEGH, 2022).

Para fazer um levantamento sísmico é preciso que existam uma fonte e um conjunto de receptores. A fonte é responsável por gerar um pulso ou uma detonação que irá se propagar pelas camadas geológicas e serão refletidas ou refratadas, até que retornem a superfície com informações valiosas para a pesquisa de O&G. Os receptores são os responsáveis por registrar as reflexões dos pulsos gerados pelas fontes. Podem ser de dois tipos: eletromagnéticos, chamados de geofones e usados para registros em terra, e os de pressão, chamados de hidrofones e usados para levantamentos na água (THOMAS et al., 2004).

A aquisição consiste em gerar uma perturbação mecânica em um ponto da superfície e o registro das reflexões em canais de recepção. A perturbação pode ser gerada por dinamites ou vibradores, quando a captação é feita em terra, ou por canhões de ar em levantamentos marítimos. A duração dos pulsos são de comprimento muito pequeno, da ordem de 200 milissegundos. Os receptores estão dispostos em linha reta e equidistante a uma distância que varia de 20 a 50 metros. A quantidade de receptores é muito grande, podendo ser centenas, o que faz com que o receptor mais distante esteja a quilômetros de distância da fonte (BURSCHIL; BUNESS; SCHMELZBACH, 2022).

3.2 Séries Temporais

Séries temporais são o conjunto de observações feitas sequencialmente ao longo do tempo, tendo como principal característica a dependência das observações vizinhas, a qual temos interesse de analisar e modelar a fim de prever comportamentos futuros. Diferentemente dos modelos de regressão, na análise das séries temporais a ordem dos dados é crucial. É importante ressaltar que o tempo pode ser substituído por outra variável como espaço, profundidade, entre outras (SILVEIRA et al., 2022).

Os dados de séries temporais são comuns em muitas áreas de estudo como Economia (preços de ações, taxa mensal de desemprego), Medicina (eletrocardiograma e eletroencefalograma), Epidemiologia (números diários/mensais de novos casos de uma doença), Meteorologia (precipitação pluviométrica, temperatura), etc (EHLERS, 2007). Em sismica, o uso de séries temporais é encontrado na forma de aquisição, onde cada receptor recebe uma sequência de sinais ao longo do tempo, que representam a estrutura da região que está sendo estudada.

Uma série temporal pode ser contínua ou discreta, sendo contínua quando as observações são feitas continuamente ao longo do tempo ou discreta quando feita em tempos específicos, em grande parte das vezes, equidistantes. Essa terminologia não diz respeito ao tipo de dado, podendo assumir valores discretos ou contínuos independente da forma como as observações forem feitas. Séries contínuas podem ser discretizadas ao se assumir um intervalo de observação a ser considerado (MORETTIN; TOLOI, 2018).

Os dados obtidos pelos receptores são as informações das reflexões ou refrações das ondas ao longo do tempo. Dessa forma, um traço sísmico, que é o resultado da captação dessas reflexões por um único receptor, pode ser tratado como uma série temporal e processadas por redes neurais com foco nesse tipo de dados.

3.3 Redes Neurais Artificiais

Uma rede neural artificial consiste em um modelo matemático que visa imitar o funcionamento do sistema nervoso central humano. São formadas por um conjunto de neurônios artificiais, que são as suas unidades de processamento e estão conectados entre si por sinapses artificiais (SILVA; SPATTI; FLAUZINO, 2010). Suas aplicações variam entre reconhecimento de voz, detecção de objetos, diagnóstico de doenças através da análise de exames, entre outras (ALANIS; ARANA-DANIEL; LOPEZ-FRANCO, 2019; BARONI, 2020).

Através do trabalho de McCulloch e Pitts (1943) tivemos acesso à primeira concepção do que seria um neurônio artificial. O neurônio é uma estrutura simples, porém fundamental para a construção de modelos computacionais complexos. Através da

conexão entre várias dessas estruturas é possível computar diversas funções matemáticas (HAYKIN, 2000).

Um neurônio artificial é descrito pela Equação 3.1,

$$y = g\left(\sum_{i=1}^n x_i w_i + b\right) \quad (3.1)$$

em que x_1, x_2, \dots, x_n representam os sinais de entrada advindos da aplicação; w_1, w_2, \dots, w_n são os pesos sinápticos, b é o termo bias e g é a função de ativação. As ativações comuns são função identidade, sigmóide e tangente hiperbólica.

A função ReLU (*Rectified Linear Units*), descrita na Equação 3.2, foi proposta por Krizhevsky, Sutskever e Hinton (2012b) como uma alternativa capaz de convergir mais rápido do que a tangente hiperbólica em aproximadamente 6 vezes.

$$g(x) = \max(0, x) \quad (3.2)$$

O tipo mais comum de arquitetura mais utilizado é a *Perceptron* de múltiplas camadas, do inglês *Multilayer Perceptron* (MLP), sendo a mais encontrada na literatura (PHAM et al., 2019). Elas possuem pelo menos uma camada oculta de neurônios, que fica entre a camada de entrada e a camada de saída (SILVA; SPATTI; FLAUZINO, 2010). Os neurônios dessas camadas permitem que a rede aprenda funções mais complexas, pois capacita a rede para extrair características mais significativas dos dados fornecidos (HAYKIN, 2000). Após a publicação do livro *Parallel Distributed Processing*, onde foi apresentado o algoritmo de *backpropagation*, as MLPs se popularizaram e foram aplicadas em diversas áreas, pois esse algoritmo permite que as redes sejam treinadas (RUMELHART et al., 1988).

3.4 Aprendizagem Profunda

Segundo Fernandes (2013), animais e humanos aprendem a interpretar o mundo a partir de conceitos simples para depois entenderem conceitos mais complexos, tendo por base os conceitos observados anteriormente. Denomina-se esse processo de aprendizagem como aprendizagem profunda, pois sugere que existem camadas hierárquicas no cérebro com diferentes responsabilidades (HUBEL; WIESEL, 1998).

De forma análoga ao funcionamento cerebral, as técnicas de aprendizagem profunda costumam apresentar várias camadas, repensáveis por fazer o processamento dos dados de forma não-linear a fim de reconhecer padrões (LECUN; BENGIO; HINTON, 2015). Com o uso dessas técnicas é possível extrair características de forma automática a partir de um conjunto de dados, sem a necessidade de um processamento prévio aos dados (HUA et al., 2015). Dessa forma, existe uma abstração feita por parte do próprio modelo que substitui

as etapas de extração, seleção e classificação de características, levando a uma baixa intervenção humana (CHENG et al., 2016). Existem diferentes arquiteturas que fazem o uso de aprendizagem profunda, por exemplo, redes neurais convolucionais, redes neurais recorrentes, redes de crença profunda, redes de memória de longo prazo, auto-codificadores esparsos empilhados, entre outras (HAFEMANN, 2014).

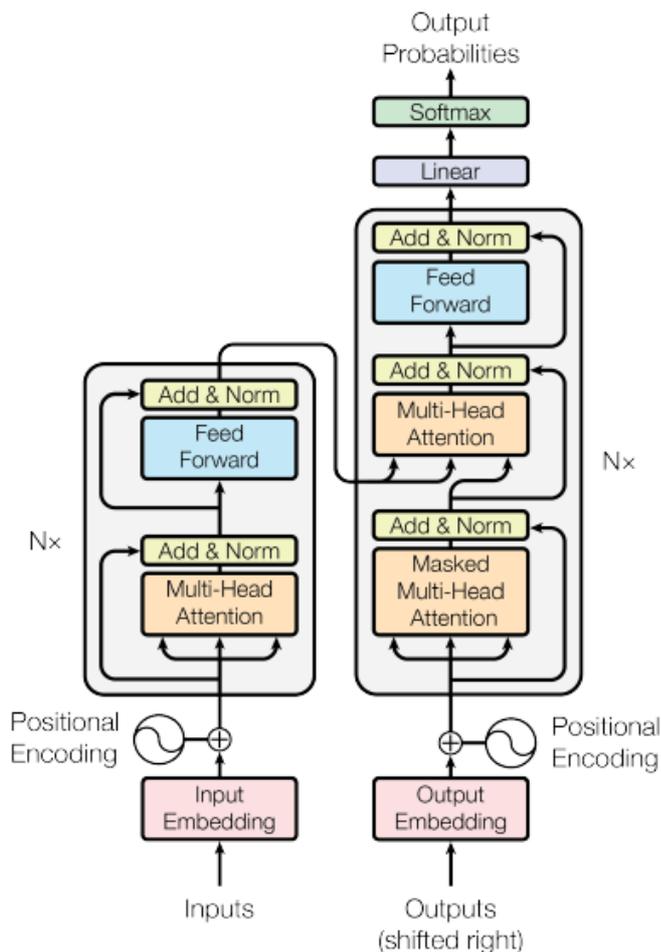
3.5 Transformer

Uma das arquiteturas de aprendizagem profunda que tem apresentado melhores resultados são os denominados Transformer. Esta arquitetura foi proposta para trabalhar com problemas do tipo sequência para sequência de modo a superar a limitação das redes neurais recorrentes (RNN), lidar com sequências muito longas. A principal característica da arquitetura transformer é o mecanismo de atenção, que permite que a rede neural olhe para todos os elementos fornecidos na entrada, não apenas para o mais recente (VASWANI et al., 2017).

Transformer é uma arquitetura de rede neural baseada no mecanismo de auto-atenção (Self-Attention) para processar dados sequenciais sem fazer uso de recorrência. Seu uso mais principal foi no processamento de linguagem natural (PLN), por lidar muito bem com problemas sequência para sequência, mas também tem utilização em outras tarefas, como compreensão de vídeo. O uso do mecanismo de auto-atenção e a falta da necessidade do uso de recorrência faz com que a arquitetura Transformer trabalhe com dados sequenciais sem a necessidade de processar os dados em ordem. Em vez disso, a operação de atenção identifica o contexto para qualquer posição de uma sequência. Essa característica permite que a paralelização seja implementada de forma mais fácil e otimizada em problemas que necessitariam de recorrência para serem resolvidos (VASWANI et al., 2017).

A arquitetura, apresentada na Figura 1, é do tipo codificador-decodificador (encoder-decoder). O codificador é formado pela camada de codificação posicional (*positional encoding*) seguida de um conjunto de camadas que processa os dados de forma iterativa, uma camada processando a saída da camada anterior. Essas camadas são compostas pela camada de atenção de múltiplas cabeças (*multi-head attention*) seguida da camada de *feedforward*. O decodificador trabalha da mesma forma, recebendo como entrada a saída do codificador. Cada camada do codificador processa sua entrada para gerar codificações que contenham informações sobre quais partes da entrada são relevantes entre si. Essas codificações alimentam a próxima camada do codificador. As camadas do decodificador fazem o oposto, processam as informações contextuais para gerar uma sequência de saída (VASWANI et al., 2017).

Figura 1 – Arquitetura da rede Transformer.



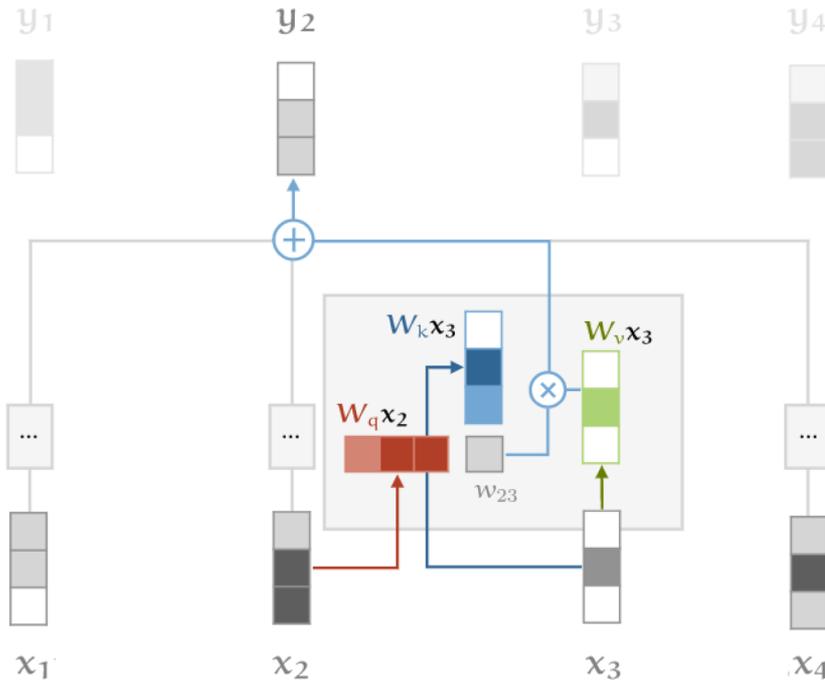
Fonte: (VASWANI et al., 2017).

3.5.1 Auto-Atenção

A auto-atenção é uma operação sequência para sequência, ou seja, uma sequência de valores entra e uma sequência de valores sai. O vetor de entrada é representado por x_1, x_2, \dots, x_t enquanto o vetor de saída correspondente é representado por y_1, y_2, \dots, y_t , ambos de dimensão k .

Para melhor compreensão do mecanismo de auto-atenção é preciso que 3 elementos sejam apresentados: Consultas (*Queries*), Valores (*Values*) e Chaves (*Keys*). Em cada função, como pode ser observado na Figura 2, o vetor de entrada é comparado com os outros vetores para obter sua própria saída y_i (Consulta), para obter a j -ésima saída y_j (Chave) e para calcular cada vetor de saída uma vez que os pesos tenham sido estabelecidos (Valor). Para obter esses dados, precisamos de três matrizes de peso de dimensão $k \times k$ e calcular três transformações lineares para cada posição x_i da entrada (VASWANI et al., 2017).

Figura 2 – Ilustração do mecanismo de auto-atenção.



(BLOEM, 2019).

Essas três matrizes são geralmente conhecidas como K, Q e V. São três camadas de peso treináveis que são aplicadas à mesma entrada codificada. Conseqüentemente, como cada uma dessas três matrizes vem da mesma entrada, podemos aplicar o mecanismo de atenção do vetor de entrada consigo mesmo, uma auto-atenção. Essas matrizes são usadas para calcular os níveis de atenção. Os níveis medem quanto foco colocar em outros lugares ou palavras da sequência de entrada com uma palavra em uma determinada posição (VASWANI et al., 2017).

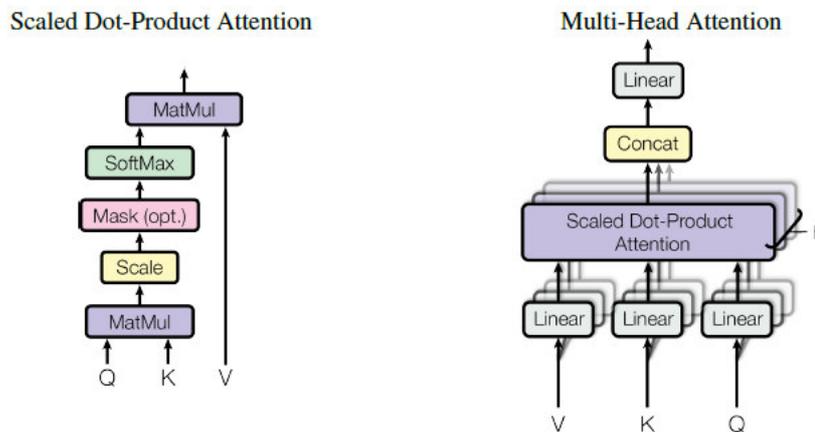
Devido a função *softmax* não funcionar bem para valores grandes, pois elimina os gradientes e retarda o aprendizado, as matrizes Q e K são submetidas à aplicação de um fator de escala, que diminui os valores para que possa se ter gradientes mais estáveis. Após a aplicação da função *softmax*, o resultado é multiplicado pela matriz V para manter o valor das palavras que se deseja focar e minimizar os valores das palavras irrelevantes. Desta forma, a função que calcula a atenção é descrita na Equação 3.3 (VASWANI et al., 2017).

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3.3)$$

3.5.2 Atenção de Múltiplas Cabeças

No mecanismo de atenção explicado anteriormente, os níveis de atenção são calculados para a sequência inteira de uma única vez, o que produz resultados idênticos para sequências que possuem os mesmos elementos, ainda que estes estejam dispostos em uma ordem diferente. É possível dar ao mecanismo de atenção um maior poder de discriminação quando são combinadas múltiplas cabeças de auto-atenção, dividindo a sequência em um número fixo de pedaços e aplicando a auto-atenção a cada um destes (Figura 3). Ao final desse processo, as matrizes de saída são concatenadas e a matriz resultante é associada a um peso adicional antes de ser enviada a camada de *Feed-Forward* (VASWANI et al., 2017).

Figura 3 – Atenção e Atenção de Múltiplas Cabeças.



(VASWANI et al., 2017).

3.5.3 Codificação Posicional

Devido à característica do mecanismo de auto-atenção ser invariante a permutação é preciso criar uma representação de posição e incorporar à representação da palavra. Para isso, é aplicada uma função que mapeia uma posição em uma sequência para um vetor de valores. Essa informação é passada para a rede para que utilize durante o processo de treinamento. A função aplicada está representada na Equação 3.4 (VASWANI et al., 2017).

$$\begin{aligned}
 PE_{(pos,2i)} &= \sin(pos/10000^{2i/d_{model}}) \\
 PE_{(pos,2i+1)} &= \cos(pos/10000^{2i/d_{model}})
 \end{aligned}
 \tag{3.4}$$

3.6 Time2Vec

Problemas envolvendo séries temporais são recorrentes e aparecem em diversas áreas. Alguns exemplos de problemas podem ser prever a quantidade de vezes que um evento acontecerá em um determinado período, prever a hora que um determinado evento voltará a acontecer, entre outros. A entrada para esses problemas pode ser considerada uma sequência onde os valores não estão distribuídos de forma idêntica e independente, mas possuem uma dependência do tempo entre os pontos de dados (KAZEMI et al., 2019).

Para fazer a modelagem de sequências nas quais o tempo é uma característica importante, muitas vezes as redes neurais tem que ser alimentadas com o tempo como uma dimensão independente (LI; DU; BENGIO, 2017; DU et al., 2016; CHOI EDWARD, 2016). As RNNs, que são comumente usadas para problemas desse tipo, falham na tentativa de usar o tempo como um recurso de forma efetiva. Para auxiliar as RNNs a fazer um melhor uso do tempo, os pesquisadores projetam recursos artesanais (*hand-carfted*) que se adequam a problemas específicos (CHOI EDWARD, 2016; BAYTAS et al., 2017; KWON et al., 2018). Tais recursos, no entanto, podem ser caros e exigir um profundo conhecimento e domínio sobre o problema (KAZEMI et al., 2019).

Time2Vec é uma representação vetorial para o tempo que tem seus pesos balanceados durante o treinamento do modelo. Três propriedades são importante para a representação do tempo: capturar padrões periódicos e não periódicos, ser invariante ao reescalonamento do tempo e ser simples para permitir que possa ser combinado com diversos modelos (KAZEMI et al., 2019). Dessa forma, a proposta do Time2vec consiste em representar o tempo de uma forma que mantenha essas três propriedades. Para um dado tempo τ , $t2v(\tau)$ é um vetor de dimensão $k + 1$ definido por:

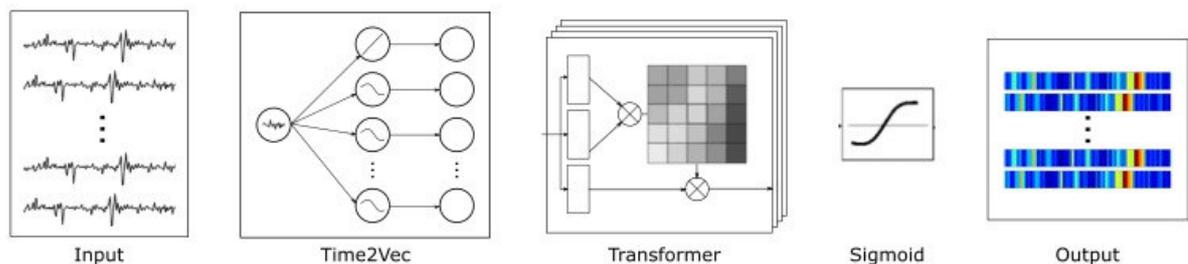
$$t2v(\tau)[i] = \begin{cases} \omega_i\tau + \varphi_i & \text{if } i = 0 \\ \mathcal{F}(\omega_i\tau + \varphi_i) & \text{if } 1 \leq i \leq k \end{cases} \quad (3.5)$$

onde $t2v(\tau)[i]$ é o i -ésimo elemento de $t2v(\tau)$, \mathcal{F} é uma função periódica de ativação e os parâmetros treináveis são ω_i e φ_i . A função escolhida para \mathcal{F} foi a função seno, pois facilita a detecção de eventos periódicos sem a necessidade de fazer engenharia de características (KAZEMI et al., 2019).

4 Metodologia

Este capítulo apresenta a metodologia proposta para a detecção de reservatórios de gás em imagens sísmicas. O diagrama do fluxo do conjunto de etapas aplicado neste estudo está sumarizado na Figura 4. Três etapas principais são necessárias: aquisição dos dados, pré-processamento e detecção de gás. Todos eles são descritos nas seções a seguir.

Figura 4 – Fluxograma da metodologia proposta.



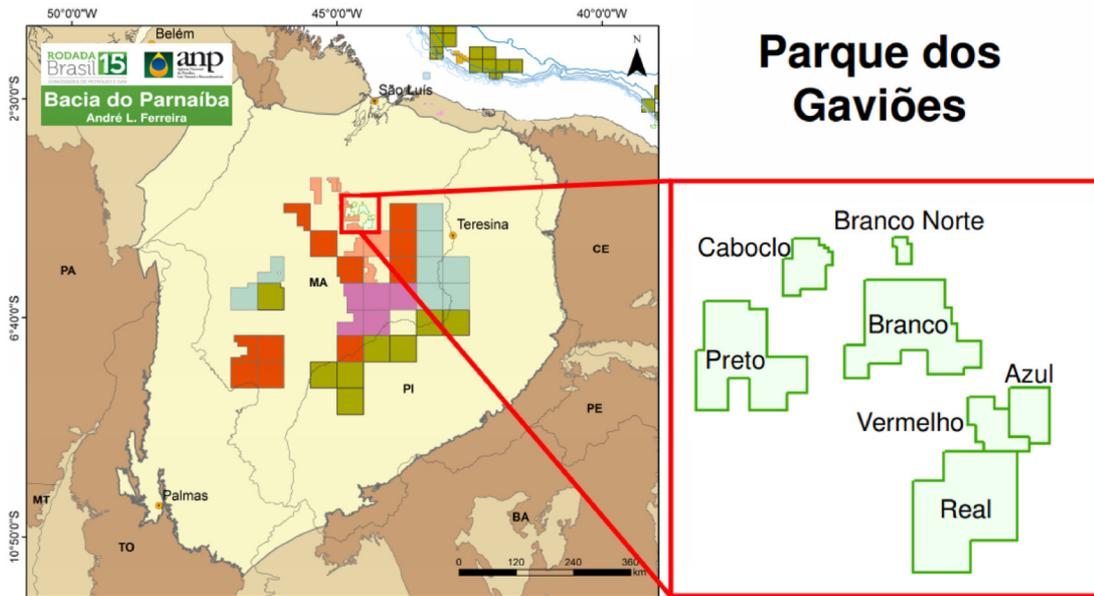
Acervo do autor.

4.1 Base de Dados

A base de dados utilizada para realizar os experimentos é a Bacia Paleozoica do Parnaíba, localizada no Nordeste do Brasil. Parnaíba é uma bacia intra cratônica clássica em formato elíptico e é composta por 600.000 km² e espessura máxima de 3.500 m. O primeiro mapeamento geológico da Bacia do Parnaíba foi feito no início do século XX (1909-10), voltado para recursos de carvão e água, mas sem resultados tangíveis. O antigo Conselho Nacional Brasileiro deu os primeiros passos na exploração de petróleo e gás de 1947 a 1953 (MIRANDA et al., 2018).

A principal área produtora é conhecida como Parque dos Gaviões (Figura 5) em referência às espécies nativas de gaviões que são denominadas como: Gavião Real, Gavião Azul, Gavião Branco, Gavião Branco Norte, Gavião Vermelho, Gavião Caboclo e Gavião Preto (MIRANDA et al., 2018). Todos os dados utilizados neste trabalho são provenientes de um banco de dados privado fornecido pela Eneva, uma empresa brasileira de energia, com negócios complementares em geração, exploração e produção de hidrocarbonetos. É a terceira maior empresa em capacidade térmica do país, responsável por 11% da capacidade térmica a gás instalada no Brasil. Dessa forma, a base de dados utilizada para treinamento e teste é confidencial e privada. Os dados utilizados são de capturas sísmicas bidimensionais (2D) em uma área de 438 km² e possuem demarcação de horizontes, marcações que delimitam o espaço de busca por reservatórios de gás.

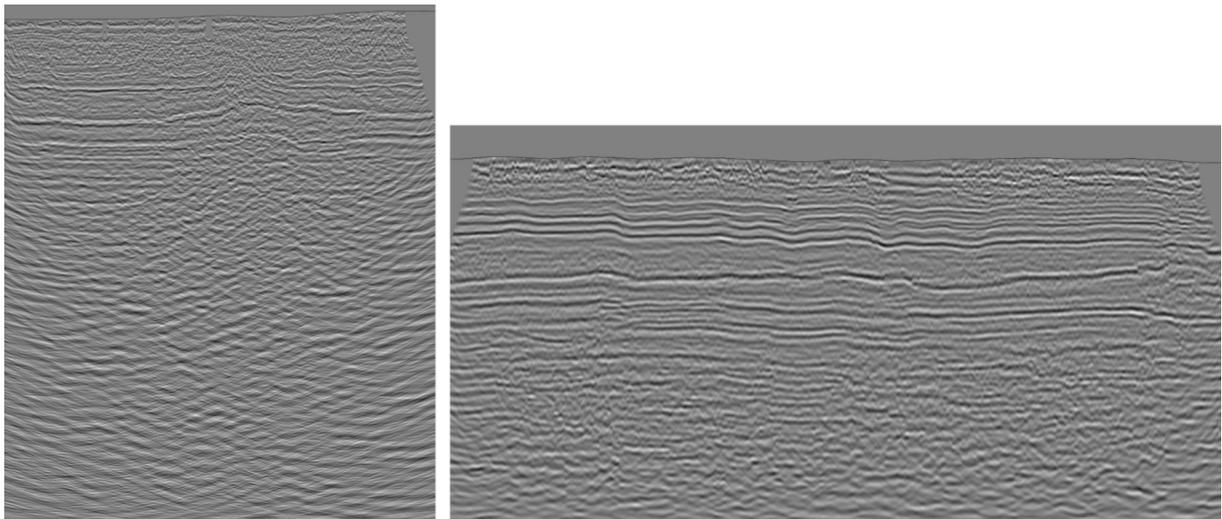
Figura 5 – Parque dos Gaviões



(MIRANDA et al., 2018)

Os dados disponíveis consistem em 380 seções sísmicas na área do Parque dos Gaviões. Esses dados foram obtidos em diferentes intervalos de tempo com diferenças de clima, geologia, processo de aquisição e outros fatores externos. Assim, os dados são diversificados e heterogêneos. A Figura 6 mostra exemplos do banco de dados.

Figura 6 – Exemplo de uma linha 2D da base de dados.



Acervo do autor.

4.2 Dimensionalidade

O processo de reconhecimento de padrões em imagens, em sua grande maioria, faz uso de redes neurais convolucionais profundas. O treinamento supervisionado dessas redes consiste em calibrar os pesos das camadas de convolução utilizando imagens já marcadas e espera-se que após o treinamento a rede seja capaz de rotular as regiões em uma imagem que satisfazem o padrão desejado.

A limitações envolvendo o uso de CNNs é o grande número de imagens necessárias para que a rede generalize de forma satisfatória. Isso acontece devido ao fato de a arquitetura ter alta sensibilidade a ruídos, que é minimizada quando adicionamos mais imagens à base de dados de treinamento.

Entretanto, essa limitação é bastante difícil de ser contornada em sísmica devido ao fato de possuímos um número insuficiente de imagens rotuladas. A rotulação de imagens acaba por não ser trivial, pois mesmo especialistas podem se equivocar facilmente quanto à presença ou ausência de um reservatório de gás em uma determinada região, fazendo com que a única forma de se ter essa certeza é após a exploração, que custa muito tempo e dinheiro.

Além disso, reservatórios de gás encontrados no norte do Brasil, local onde a exploração para esse trabalho foi feita, apresentam pouca extensão horizontal, o que faz com que mesmo recortes pequenos da imagem possua a maior parte da sua borda fora dos reservatórios de gás, além de apresentar uma grande quantidade de ruído (SANTOS, 2019).

É possível obter um maior número de amostras com uma mesma quantidade de imagens quando se diminui a vizinhança ao redor de um ponto desejado da imagem. Por exemplo, em uma imagem 100x100 é possível obter 100 amostras de tamanho 10x10 ou 400 amostras de tamanho 5x5. Isso diminui o poder de associação da rede com uma vizinhança mais distante, mas aumenta a quantidade de amostras que podem ser extraída de um mesmo conjunto de dados.

No processamento de imagens sísmicas é possível reduzir a dimensionalidade do problema e trabalhar com um domínio unidimensional, denominado traço, reduzindo consideravelmente o tamanho das amostras e aumentando sua quantidade. Dessa forma, uma imagem de tamanho 100x100 seria capaz de gerar 1000 amostras de tamanho 10x1 ou 4000 amostras de tamanho 5x1.

A aplicação da redução de dimensionalidade em imagens sísmicas é justificada pelo processo de aquisição dos dados. Antes de serem empilhados e agrupados em forma de imagem (dado bidimensional), o traço sísmico representa a resposta de uma onda unidimensional ao longo do tempo, que é a captação das reflexões da onda viajando pelas camadas do solo com diferentes impedâncias acústicas. esse processo , chamado de modelo

convolucional, é essencialmente unidimensional (SANTOS, 2019).

4.3 Preprocessamento

Esta etapa consiste na limpeza e normalização dos dados e foram feitos os seguintes passos:

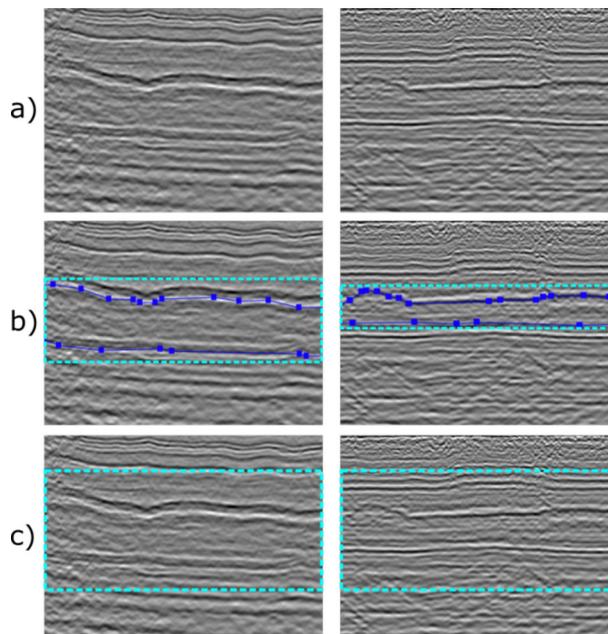
- Normalização dos valores de amplitude do banco de dados.
- Definição de uma região de interesse (ROI).
- Extração dos dados de cada traço individualmente.
- Extração de pacotes (*patches*) do traço para entrada na arquitetura Transformer proposta.

Os valores das amplitudes sísmicas nos dados de entrada variam de forma arbitrária entre um levantamento e outro pois dependem do processo utilizado durante a aquisição e migração dos dados, sendo necessário uniformizar os dados antes do processo de treinamento. A técnica de padronização escolhida foi proposta por (RAJU et al., 2020) que limita os intervalos de valores entre zero (para um valor padrão) e um (para uma derivação padrão). Esta técnica foi aplicada devido à característica dos dados sísmicos se aproximarem de uma distribuição gaussiana.

Os dados são filtrados, através da definição de uma ROI (veja a Figura 7), visando reduzir as informações ruidosas que podem prejudicar o treinamento da rede. Essa ROI é individual para cada imagem e delimita uma área com os padrões sísmicos que o modelo deve aprender a detectar por estruturas de separação em gás ou não-gás. Para esta tarefa há uma preparação manual prévia, que deve analisar os dados e configurar a ROI. Em seguida, são definidos todos as ROIs para o mesmo tamanho considerando o ROI mais significativo no banco de dados, aumentando as coordenadas inferiores no eixo y, se necessário.

Por fim, são extraídos os traços e as *patches* da sísmica (Figura 8). Nesta etapa, a sísmica passa por uma decomposição em uma sequência de traços, no qual os *patches* são extraídos dos traços usando uma janela unidimensional de tamanho 99 e a movemos em um deslocamento de tamanho 1. Após vários experimentos, esse tamanho e movimento apresentaram melhores resultados no banco de dados de validação.

Figura 7 – Configuração da ROI. a) imagem sísmica; b) primeira etapa: marcação humana; e c) segunda etapa: tamanho adaptativo.



Acervo do autor.

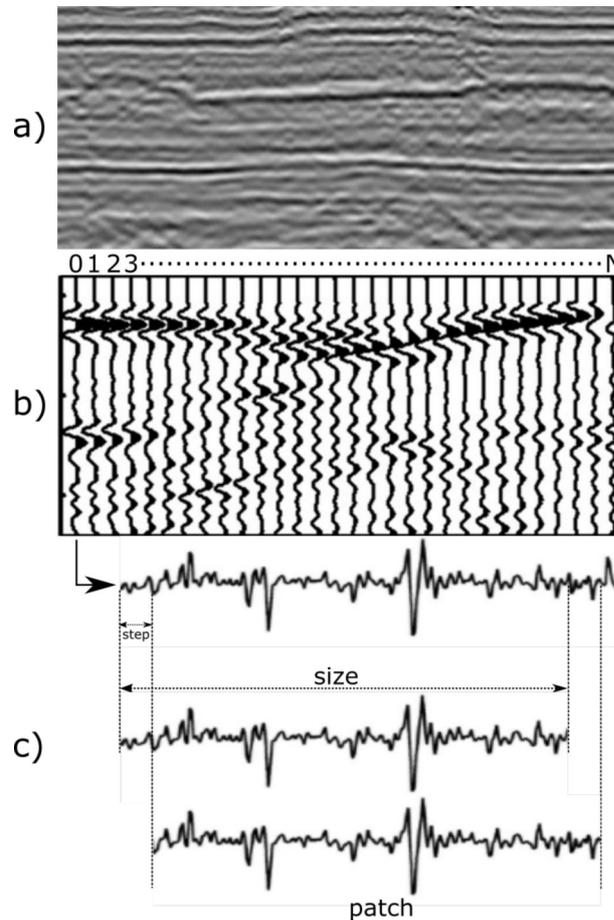
4.4 Experimentos

O processo experimental consiste em definir 2 conjuntos de dados, sendo um deles para treinamento e outro para teste. Dentro do conjunto determinado para treinamento, um percentual dos dados de entrada é separado para compor o conjunto de validação enquanto o restante compõe o conjunto de teste definitivo. Esse percentual é informado como parâmetro antes do treinamento enquanto a seleção dos dados é feita de forma aleatória, respeitando a quantidade predefinida.

Em seguida, toda a base de dados (treino, teste e validação) é submetida ao pré-processamento. A base de teste é reservada para uso posterior enquanto as base de treino e validação são submetidas a um procedimento de balanceamento das classes. Pelo fato de a detecção de gás ser um problema de classe rara, uma vez que temos apenas uma pequena porcentagem dos dados de entrada que possuem alguma interseção com um reservatório de gás (cerca de 2% a 5%), é preciso fazer uso de algumas técnicas para balanceamento das classes, afim de não fazer com que o treinamento fique enviesado em apontar sempre um mesma classe.

O processo de balanceamento de balanceamento de classes é aplicado sobre a base de treino e validação e consiste em separar todos os *patches* com interseção com pelo menos 1 reservatório de gás e escolher, de forma aleatória, um conjunto de *patches* que não fazem interseção, seguindo uma proporção predefinida. Por exemplo, supondo que se tenha 10 *patches* com interseção e a proporção predefinida seja de 1:2, deve-se escolher

Figura 8 – Extração de *patches*. a) ROI; b) decomposição em sequencia de traços; e c) extração do *patch* do traço.



Acervo do autor.

aleatoriamente 20 *patches* sem interseção.

4.5 Arquitetura Transformer

O modelo de traço unidimensional se assemelha a problemas de outros domínios como sinais de som. Ao olharmos para a literatura, vemos que os métodos que apresentam melhor desempenho nesse domínio de dados são as redes neurais recorrentes. As RNNs são capazes de reconhecer uma estrutura de série temporal e processar uma amostra considerando a sequência dos dados. O traço sísmico é uma série temporal que representa a reflexão das ondas através das camadas do solo. Uma arquitetura muito utilizada para esse tipo de problema é a LSTM, que já foi aplicada no contexto de sísmica no trabalho de Santos (2019).

A arquitetura LSTM foi o estado da arte durante muito tempo em problemas envolvendo o processamento de linguagem natural. No trabalho de Vaswani et al. (2017) pudemos observar a arquitetura Transformer superar a LSTM para problema de PNL.

Nesse contexto, uma oportunidade de se aplicar a Transformer em um problema de sísmica surge, uma vez que a arquitetura LSTM já havia mostrado resultados nesse contexto.

O modelo de arquitetura *Trasnformer* proposto consiste em adaptar as camadas de entrada otimizadas para trabalhar com PNL para que possa trabalhar com séries temporais. Para isso, uma arquitetura de classificação de texto foi utilizada como ponto de partida. Em seguida, todo o processo de tokenização juntamente com a camada *Positional Encoding* foram substituídos por uma camada *Time2Vec*. Essa camada possibilitou que a arquitetura *Transformer* fosse capaz de interpretar os dados de forma sequencial. A camada de saída foi expandida para retornar uma sequência de mesmo tamanho da sequência de entrada com cada posição da sequência sinalizando a possibilidade de aquela posição conter ou não gás.

4.6 Inferência e Reconstrução

Para realizar o processo de inferência, é preciso decompor a imagem sísmica em *patches*, processar cada um deles e reconstruir uma imagem de saída combinando o resultado do processamento de cada *patch*.

O processo de reconstrução consiste em posicionar todos os *patches* extraídos da imagem na sua posição de origem. Esse método causa uma sobreposição de *patches* que deve ser resolvida. Para isso, os valores da imagem resultante serão os valores combinados de todos os *patches* que coincidirem sobre uma mesma coordenada da imagem. Para combinar os resultados, foi utilizada uma média aritmética dos valores associados a cada uma das posições.

5 Experimentos e Resultados

A metodologia proposta e apresentada neste trabalho foi o resultado do desenvolvimento de uma técnica para um projeto da empresa de energia Eneva. Os resultados apresentados utilizam os dados que foram disponibilizados pela empresa e que estão descritos com maior detalhes na sessão anterior. O método proposto alcançou resultados coerentes com as expectativas.

5.1 Ambiente de desenvolvimento

O método foi implementado utilizando a linguagem Python. A principal biblioteca utilizada no desenvolvimento do método foi o Keras ([CHOLLET et al., 2015](#)) em conjunto com o TensorFlow-GPU ([ABADI et al., 2015](#)). O computador utilizado para os experimentos possui um processador Intel Core i7-9700K 4.20 GHz, 16 GB de memória RAM e uma placa gráfica Nvidia GeForce GTX 2070. O sistema operacional utilizado foi o Windows 10.

5.2 Construção das bases de dados

A base de dados utilizada foi disponibilizada pela Eneva e corresponde a levantamentos sísmicos obtidos na Bacia Paleozoica do Parnaíba. Não foi possível termos acessos a dados de todos os campos presentes na bacia, então os experimentos foram realizados utilizando os seguintes campos:

- Gavião Branco
- Gavião Vermelho
- Gavião Preto
- Gavião Real

Primeiramente, toda a base de dados é foi submetida ao pré-processamento, o que resultou em uma base mais homogênea e contendo apenas a região de interesse. Para os experimentos, as bases de dados foram divididas em 3 conjuntos: treino, validação e teste. As imagens sísmicas foram escolhidas de forma aleatória obedecendo a proporção de 60% das imagens para o treinamento, 20% para a validação e 20% para o teste. O resultado dessa divisão em números absolutos pode ser encontrada na Tabela 1.

Tabela 1 – Distribuição da base de dados.

Gavião	Treino	Validação	Teste
Real	14	2	5
Preto	15	2	4
Branco	24	4	7
Vermelho	7	1	2
Branco e Preto ->Real e Vermelho	48	8	31
Todos	60	9	18

A divisão das bases para experimentos sugeridas na Tabela 1 foi sugerida com o intuito de avaliar o método proposto sob várias perspectivas. Os campos de exploração sísmica são bastante heterogêneos e, por conta desse fator, decidimos treinar cada um dos campos separadamente e depois fazer um treinamento incluindo todos os campos. Nessa segunda etapa, duas abordagens foram utilizadas: experimento utilizando múltiplos campos sem intersecção e com intersecção.

No experimento sem intersecção, os dois maiores campos (Branco e Preto) são separados para compor os conjuntos de treinamento e validação enquanto os demais campos compõem o conjunto de teste. O objetivo deste experimento é observar a capacidade do método lidar com campos completamente novos. No experimento com intersecção, todas as sísmicas foram escolhidas aleatoriamente sem considerar o campo de origem, criando conjuntos de treino, validação e teste bastante heterogêneos, com o objetivo de observar o comportamento do modelo quando treinado com diferentes origens de dados.

5.3 Treinamento

Com as imagens divididas, o modelo é treinado e os hiper parâmetros são ajustados com base na etapa de validação. O hiper parâmetro que produziu o melhor resultado foi o otimizador Adam (BAE; RYU; SHIN, 2019), com taxa de aprendizado de 0,001, 50 épocas, tamanho do lote igual a 32. Os resultados do treinamento aplicado à base de testes estão descritos na Tabela 2

Tabela 2 – Resultado dos experimentos.

Experimento	ACC (%)	SEN (%)	SPE (%)	AUC (%)
Real	97,35	47,71	98,43	73,07
Preto	98,23	63,33	99,04	81,19
Branco	97,91	77,56	98,11	87,83
Vermelho	98,87	96,12	98,92	97,52
Branco e Preto ->Real e Vermelho	95,59	77,21	95,80	86,51
Todos	97,16	79,61	97,47	88,54

É possível notar um desempenho relevante em todas as métricas, porém é notório

que os experimentos feitos nos campos Real e Preto apresentaram os piores resultados. Pode-se atribuir esses valores ao fato de serem campos relativamente pequenos e com uma significativa diversidade intraclasse. Dentro desses conjuntos de dados foram usados mais de um processo de aquisição e a própria natureza da região é bastante heterogênea, apresentando diferentes características de terreno. Além desses fatores, o campo Real possui aquisições feitas em diferentes períodos, tendo imagens que foram adquiridas ao longo de oito anos (2011 a 2019), o que resulta em equipamentos, aquisição e pós-processamento diferentes.

Por outro lado, o conjunto de dados Vermelho produziu o melhor resultado em todas as métricas, i. e., consegue encontrar efetivamente a classe positiva, gás (96% de sensibilidade), gerando poucos falsos positivos (98% de especificidade). No entanto, vale ressaltar que este é o conjunto de dados com o menor número de amostras e ao observar a Figura 5, essa também é a área do menor campo. Nesta perspectiva, o processo de aquisição de dados é provavelmente mais controlado, apresentando menos ruídos e diferenças geológicas. Portanto, os resultados tendem a apresentar melhores métricas do que os outros experimentos. Vale ressaltar que todas as aquisições do conjunto de dados Vermelho são do mesmo ano e passaram pelo mesmo pós-processamento, produzindo dados mais homogêneos.

Ao observar todas essas variações associadas aos conjuntos de dados, os dois maiores campos (Branco e Preto) utilizados como conjunto de treinamento para gerar um número maior de amostras. Desta forma, obteve-se 48 amostras para treinamento e 8 para validação, além de 31 imagens para o teste (Tabela 1). Este teste também valida a capacidade do modelo proposto em fazer inferências em áreas ainda não exploradas, pois simula um cenário onde o modelo não possui nenhuma amostra dos campos que estão para serem inferidos, algo que ocorre em um cenário real.

É possível perceber que este cenário produziu resultados mais significativos nos valores de sensibilidade, o que significa que o modelo foi mais eficaz em encontrar os traços que possuem gás. Entretanto, a métrica de especificidade teve uma diminuição. Isso ocorre porque mais amostras de traços sem gás são avaliadas no teste, e o modelo tende a gerar mais falsos positivos, pois provavelmente não há amostras com as características do conjunto de dados de teste em treinamento.

Por fim, foi feito um experimento para analisar a capacidade do modelo de aprender com um conjunto de dados de treinamento ainda maior e com dados mais generalizados. Para isso, um teste utilizando todas as imagens foi executado e notou-se que ao inserir amostras diversificadas de todos os campos, o resultado no nível de acerto do traço (a sensibilidade) atinge a melhor métrica. Ainda assim, ao analisar a métrica AUC, também obtivemos os melhores resultados. E, apesar do desequilíbrio entre as classes (gás e sem gás), o método ainda se mostrou eficaz tanto para as métricas de ACC quanto para SPE. Assim,

os experimentos mostram que quanto mais informações generalistas forem inseridas no modelo, mais discriminantes ele poderá ser e, conseqüentemente, gerar melhores resultados independente do campo utilizado.

5.4 Comparação com a Literatura

Nesta seção apresenta uma comparação com os trabalhos relacionados descritos no Capítulo 2. As Tabelas 3 e 4 mostram essa comparação.

Tabela 3 – Comparativo - F3-Block

Método	Base	ACC (%)	SEN (%)	SPE (%)	AUC (%)
(SANTOS, 2019)	F3-inline	97.16	97.83	97.15	98.8
	F3-crossline	96.83	94.77	96.87	98.81
(JUNIOR et al., 2021)	F3-inline	94.51	95.05	94.5	94.78
	F3-crossline	97.33	94.3	97.3	95.85
(JUNIOR et al., 2022)	F3-inline	98.11	75.03	98.88	86.95
	F3-crossline	97.90	77.04	98.65	88.05
Método proposto	F3-inline	98.93	82.87	99.20	91.04
	F3-crossline	98.38	86.24	98.54	92.39

Observa-se que tanto no trabalho de Santos (2019), que propôs uma abordagem usando a rede LSTM, e no trabalho de Junior et al. (2021), com um único modelo Transformer, ambos atingiram resultados superiores nas métricas SEN e AUC em relação ao método proposto. Entretanto, o método proposto atingiu resultados nas métricas ACC e SPE ligeiramente superiores, o que significa que a nova arquitetura Transformer, no conjunto de dados F3-block, perdeu parte da região de gás, mas obteve um menor número de falsos positivos. Em comparação com o trabalho de Junior et al. (2022), o método proposto superou em todas as métricas, com exceção da SPE no conjunto F3-crossline.

Tabela 4 – Comparativo - Gaviões Branco e Vermelho

Método	Base de dados	ACC (%)	SEN (%)	SPE (%)	AUC (%)
(JUNIOR et al., 2021)	Gavião Vermelho	98.93	59.75	99.2	79.48
	Gavião Branco	97.03	68.87	97.25	83.06
Método proposto	Gavião Vermelho	99.12	77.48	99.34	88.41
	Gavião Branco	96.96	75.81	97.53	86.45

Ainda em comparação com o desempenho do método proposto por Junior et al. (2021), podemos comparar o desempenho dos métodos para dois campos presentes no Parque dos Gaviões (Branco e Vermelho). Nota-se que o método proposto obteve melhores números em todas as métricas, demonstrando que é capaz de usar a arquitetura Transformer de forma mais eficiente para a solução do problema.

No trabalho de Junior et al. (2022) foi utilizada a mesma base de dados apresentada neste trabalho, permitindo que uma comparação mais acurada possa ser feita. Na Tabela

5 é possível observar o comportamento de ambos os métodos para a maioria dos testes apresentados anteriormente neste capítulo, à exceção do teste Branco e Preto ->Real e Vermelho que não foi realizado no trabalho citado.

Tabela 5 – Comparativo - Parque dos Gaviões.

Método	Experimento	ACC (%)	SEN (%)	SPE (%)	AUC (%)
(JUNIOR et al., 2022)	Real	93.29	50.13	94.29	72.21
	Preto	88.29	37.63	89.29	63.46
	Branco	88.29	58.02	88.77	73.39
	Vermelho	96.20	67.49	96.85	82.17
	Todos	95.60	75.14	96.14	85.64
Método proposto	Real	97,35	47,71	98,43	73,07
	Preto	98,23	63,33	99,04	81,19
	Branco	97,91	77,56	98,11	87,83
	Vermelho	98,87	96,12	98,92	97,52
	Todos	97,16	79,61	97,47	88,54

Ao observar a Tabela 5, nota-se que o método proposto apresenta valores superiores para todas as métricas. Ambos os métodos apresentam um valor de SPE consideravelmente superior aos valores de SEN. Isso indica que os modelos perdem alguns dados durante a detecção, porém apresenta um menor número de falsos positivos. Isto não significa que o método não seja capaz de detectar algum reservatório, apenas demonstra que o método prioriza apresentar menos falsos positivos em detrimento de uma detecção mais fiel da extensão de cada reservatório, apresentando perdas nas regiões periféricas de cada reservatório enquanto concentra suas indicações em uma região mais central.

5.5 Estudo de caso

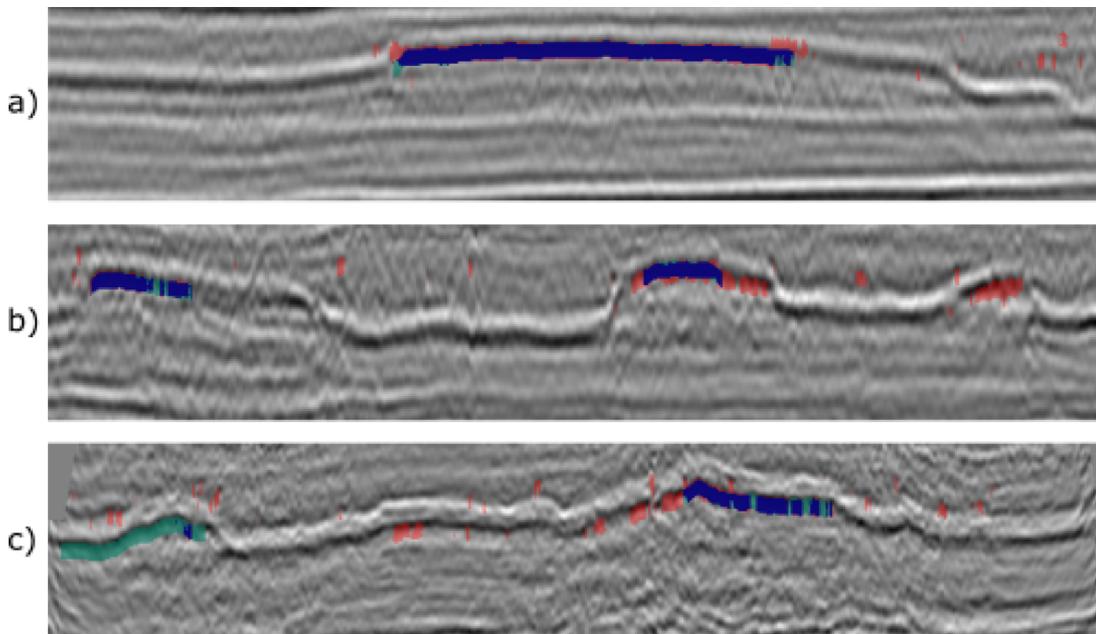
A seguir, é apresentado um estudo de caso com algumas imagens de teste utilizando o modelo treinado com todas as imagens do banco de dados, correspondente ao conjunto de teste Todos na Tabela 1.

É possível notar que o método foi eficaz na tarefa de detecção de gás. Na Figura 9(a), observa-se o melhor caso, onde o método foi capaz de detectar efetivamente o reservatório de gás. Apesar de gerar alguns falsos positivos, o método consegue distinguir a região, facilitando a análise da imagem por um especialista.

Por sua vez, na Figura 9(b), pode-se ver que muitos falsos positivos foram gerados, apesar do método detectar o poço do reservatório de gás. Uma solução seria a aquisição de novas amostras para tornar o modelo mais geral ou novas etapas de pós-processamento para eliminar falsos positivos.

Finalmente, na Figura 9(c), encontra-se o pior caso. Neste, o método não foi eficaz na detecção correta dos reservatórios de gás. Embora alguns sejam encontrados (em

Figura 9 – Estudo de caso: (a), (b), e (c) são três amostras diferentes. Em vermelho estão os falsos positivos. Em azul estão os verdadeiros positivos. Em verde estão os falsos negativos.



Acervo do autor.

azul), o método não consegue encontrar um dos reservatórios (em verde). Isso mostra que, apesar dos resultados promissores, o modelo ainda não consegue detectar alguns dados. Provavelmente porque as amostras extra-classe são muito semelhantes, gerando uma previsão falsa.

Assim, acredita-se que o método proposto é promissor, tanto pelos resultados quantitativos das métricas de validação quanto pelos resultados visuais apresentados pelos estudos de caso. Vale lembrar que a análise de dados não é trivial, requer experiência do especialista e consome muito tempo. Assim, o método proposto aliado à expertise do especialista pode ser fundamental na análise, facilitando a identificação mais rápida desses reservatórios de gás.

6 Conclusão

Neste trabalho foi apresentado um método automático para detecção de gás em sísmica de reflexão. Uma das contribuições é a proposta de uma arquitetura Transformer modificada para fazer a segmentação de reservatórios de hidrocarbonetos, em imagens de sísmica de reflexão. Esta arquitetura proposta consiste em uma melhoria da camada de *embedding* para possibilitar que a rede interprete os traços sísmicos como séries temporais, e o faça de maneira mais precisa. Além disso, a saída da arquitetura proposta é baseada na abordagem sequência para sequência, de modo que produza ao final o traços segmentado.

Foi utilizada uma base de dados privada e disponibilizada pela empresa Eneva. Ao analisarmos as métricas de avaliação e observarmos os resultados visuais comparados com as marcações dos especialistas, pode-se confirmar a qualidade dos resultados obtidos.

O método proposto utiliza dados sísmicos unidimensionais que são obtidos a partir da extração do traços sísmicos a partir de uma imagem sísmica bidimensional. O uso desta estratégia é possível devido a característica temporal presente na captura dos dados e proporciona uma base de dados muito maior utilizando a mesma quantidade de imagens de entrada.

Ao analisar os resultados, pode-se observar que o método se mostrou consistente, com métricas consideravelmente altas e resultados gráficos coerentes. Conclui-se, então, que o método apresentado neste trabalho pode dar suporte para especialistas na tarefa de detectar padrões de assinatura de reservatórios de gás em imagens sísmicas, otimizando o processo de interpretação de imagens sísmicas, o que gera uma economia de tempo e dinheiro.

Como trabalho futuro, pretende-se aplicar outras técnicas de pré-processamento para atenuar os ruídos característicos dos dados sísmicos. Além disso, explorar técnicas de otimização de hiper parâmetros para melhor adequar a rede para este tipo de dados. Por fim, explorar alternativas relacionadas ao pós-processamento de dados ou mesmo técnicas de reconstrução de regiões gasosas para melhor detecção do reservatório de gás.

Referências

- ABADI, M.; AGARWAL, A.; BARHAM, P.; BREVDO, E.; CHEN, Z.; CITRO, C.; CORRADO, G. S.; DAVIS, A.; DEAN, J.; DEVIN, M. et al. Tensorflow: Large-scale machine learning on heterogeneous systems. *Software available from tensorflow.org*, 2015. Citado na página 33.
- ABDULAZIZ, A. M.; ALSAEGH, H. I. Integration of seismic attributes with well logs using artificial neural network for predicting the key pay zones parameters: A case study in the kifl field, south of iraq. *Arabian Journal of Geosciences*, Springer, v. 15, n. 16, p. 1–17, 2022. Citado na página 18.
- ALANIS, A. Y.; ARANA-DANIEL, N.; LOPEZ-FRANCO, C. *Artificial Neural Networks for Engineering Applications*. [S.l.]: Academic Press, 2019. Citado na página 19.
- ARAYA-POLO, M.; DAHLKE, T.; FROGNER, C.; ZHANG, C.; POGGIO, T.; HOHL, D. Automated fault detection without seismic processing. *The Leading Edge*, Society of Exploration Geophysicists, v. 36, n. 3, p. 208–214, 2017. Citado na página 15.
- BAE, K.; RYU, H.; SHIN, H. Does adam optimizer keep close to the optimal point? *arXiv preprint arXiv:1911.00289*, 2019. Citado na página 34.
- BARONI, M. Linguistic generalization and compositionality in modern artificial neural networks. *Philosophical Transactions of the Royal Society B*, The Royal Society, v. 375, n. 1791, p. 20190307, 2020. Citado na página 19.
- BAYTAS, I. M.; XIAO, C.; ZHANG, X.; WANG, F.; JAIN, A. K.; ZHOU, J. Patient subtyping via time-aware lstm networks. In: *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. [S.l.: s.n.], 2017. p. 65–74. Citado na página 25.
- BLOEM, P. *TRANSFORMERS FROM SCRATCH*. 2019. <<http://doi.org/10.7937/K9/TCIA.2015.QJTV5IL5>>. Accessed: 2022-09-17. Citado na página 23.
- BURSCHIL, T.; BUNESS, H.; SCHMELZBACH, C. Near-surface three-dimensional multicomponent source and receiver s-wave survey in the tannwald basin, germany: Acquisition and data processing. *Near Surface Geophysics*, v. 20, n. 4, p. 331–348, 2022. Citado na página 18.
- CHENG, J.-Z.; NI, D.; CHOU, Y.-H.; QIN, J.; TIU, C.-M.; CHANG, Y.-C.; HUANG, C.-S.; SHEN, D.; CHEN, C.-M. Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in us images and pulmonary nodules in ct scans. *Scientific reports*, Nature Publishing Group, v. 6, 2016. Citado na página 21.
- CHEVITARESE, D. S.; SZWARCMAN, D.; SILVA, R. G. e; BRAZIL, E. V. Deep learning applied to seismic facies classification: A methodology for training. In: EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS. *Saint Petersburg 2018*. [S.l.], 2018. v. 2018, n. 1, p. 1–5. Citado na página 16.

- CHOI EDWARD, T. B. M. S. A. F. S. W. S. J. Doctor ai: Predicting clinical events via recurrent neural networks. *Stewart, Jimeng Sun arXiv (2015-11-18) https://arxiv.org/abs/1511.05942 v11*, 2016. Citado na página 25.
- CHOLLET, F. et al. *Keras*. 2015. <<https://keras.io>>. Citado na página 33.
- COLLOBERT, R.; WESTON, J. A unified architecture for natural language processing: Deep neural networks with multitask learning. In: *ACM. Proceedings of the 25th international conference on Machine learning*. [S.l.], 2008. p. 160–167. Citado 2 vezes nas páginas 12 e 15.
- DI, H.; WANG, Z.; ALREGIB, G. Deep convolutional neural networks for seismic salt-body delineation. In: *AAPG Annual Convention and Exhibition*. [S.l.: s.n.], 2018. Citado na página 16.
- DI, H.; WANG, Z.; ALREGIB, G. Seismic fault detection from post-stack amplitude by convolutional neural networks. In: *EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS. 80th EAGE Conference and Exhibition 2018*. [S.l.], 2018. v. 2018, n. 1, p. 1–5. Citado na página 16.
- DU, N.; DAI, H.; TRIVEDI, R.; UPADHYAY, U.; GOMEZ-RODRIGUEZ, M.; SONG, L. Recurrent marked temporal point processes: Embedding event history to vector. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. [S.l.: s.n.], 2016. p. 1555–1564. Citado na página 25.
- EHLERS, R. S. Análise de séries temporais. *Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná*, v. 1, p. 1–118, 2007. Citado na página 19.
- FERNANDES, B. J. T. *Redes neurais com extração implícita de características para reconhecimento de padrões visuais*. Tese (Doutorado) — Universidade Federal de Pernambuco, Recife, 7 2013. Citado na página 20.
- HAFEMANN, L. G. *An analysis of deep neural networks for texture classification*. Dissertação (Mestrado) — Universidade Federal do Paraná, Curitiba, 2014. Citado na página 21.
- HAYKIN, S. S. *Redes neurais: princípios e prática*. 2. ed. Porto Alegre: Bookman, 2000. ISBN 8573077182. Citado na página 20.
- HUA, K.-L.; HSU, C.-H.; HIDAYATI, S. C.; CHENG, W.-H.; CHEN, Y.-J. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *OncoTargets and therapy*, Dove Press, v. 8, 2015. Citado na página 20.
- HUANG, L.; DONG, X.; CLEE, T. E. A scalable deep learning platform for identifying geologic features from seismic attributes. *The Leading Edge*, Society of Exploration Geophysicists, v. 36, n. 3, p. 249–256, 2017. Citado na página 16.
- HUBEL, D. H.; WIESEL, T. N. Early exploration of the visual cortex. *Neuron*, Elsevier, v. 20, n. 3, p. 401–412, 1998. Citado na página 20.
- JUNIOR, D. A. D.; CRUZ, L. B.; DINIZ, J. O. B.; SILVA, A. C.; PAIVA, A. C.; GATTASS, M. Pso-based vit-seismic: A vision transformer approach for gas detection in seismic images. In: *SBC. Anais da X Escola Regional de Computação do Ceará, Maranhão e Piauí*. [S.l.], 2022. p. 11–20. Citado 3 vezes nas páginas 17, 36 e 37.

- JUNIOR, D. A. D.; DINIZ, P.; JUNIOR, L. M.; CIPRIANO, C.; GATTASS, M.; SANTOS, L.; QUEVEDO, R.; AL et. Automatic gas detection using seismic data and transformer neural networks. 2021. Citado 2 vezes nas páginas 16 e 36.
- KAZEMI, S. M.; GOEL, R.; EGHBALI, S.; RAMANAN, J.; SAHOTA, J.; THAKUR, S.; WU, S.; SMYTH, C.; POUPART, P.; BRUBAKER, M. Time2vec: Learning a vector representation of time. *arXiv preprint arXiv:1907.05321*, 2019. Citado na página 25.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2012. p. 1097–1105. Citado 2 vezes nas páginas 12 e 15.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2012. p. 1097–1105. Citado na página 20.
- KWON, B. C.; CHOI, M.-J.; KIM, J. T.; CHOI, E.; KIM, Y. B.; KWON, S.; SUN, J.; CHOO, J. Retainvis: Visual analytics with interpretable and interactive recurrent neural networks on electronic medical records. *IEEE transactions on visualization and computer graphics*, IEEE, v. 25, n. 1, p. 299–309, 2018. Citado na página 25.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015. Citado na página 20.
- LI, Y.; DU, N.; BENGIO, S. Time-dependent representation for neural event sequence prediction. *arXiv preprint arXiv:1708.00065*, 2017. Citado na página 25.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, Springer, v. 5, n. 4, p. 115–133, 1943. Citado na página 19.
- MIRANDA, F. de; VETTORAZZI, A.; CUNHA, P. da C.; ARAGÃO, F.; MICHELON, D.; CALDEIRA, J.; PORSCHE, E.; AL et. Atypical igneous-sedimentary petroleum systems of the parnaíba basin, brazil: seismic, well logs and cores. *Geological Society, London, Special Publications*, Geological Society of London, v. 472, n. 1, p. 341–360, 2018. Citado 2 vezes nas páginas 26 e 27.
- MORETTIN, P. A.; TOLOI, C. M. *Análise de séries temporais: modelos lineares univariados*. [S.l.]: Editora Blucher, 2018. Citado na página 19.
- MORTON-THOMPSON, D.; WOODS, A. M. et al. *Development geology reference manual: AAPG methods in exploration series, no. 10*. [S.l.]: AAPG, 1993. Citado na página 12.
- PHAM, B. T.; NGUYEN, M. D.; BUI, K.-T. T.; PRAKASH, I.; CHAPI, K.; BUI, D. T. A novel artificial intelligence approach based on multi-layer perceptron neural network and biogeography-based optimization for predicting coefficient of consolidation of soil. *Catena*, Elsevier, v. 173, p. 302–311, 2019. Citado na página 20.
- POCHET, A.; DINIZ, P. H.; LOPES, H.; GATTASS, M. Seismic fault detection using convolutional neural networks trained on synthetic poststacked amplitude maps. *IEEE Geoscience and Remote Sensing Letters*, IEEE, v. 16, n. 3, p. 352–356, 2018. Citado na página 16.

- RAJU, V. G.; LAKSHMI, K. P.; JAIN, V. M.; KALIDINDI, A.; PADMA, V. Study the influence of normalization/transformation process on the accuracy of supervised classification. In: IEEE. *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*. [S.l.], 2020. p. 729–735. Citado na página 29.
- ROY, S.; KIRAL-KORNEK, I.; HARRER, S. Deep learning enabled automatic abnormal eeg identification. In: IEEE. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.], 2018. p. 2756–2759. Citado 2 vezes nas páginas 12 e 15.
- RUMELHART, D. E.; MCCLELLAND, J. L.; GROUP, P. R. et al. *Parallel distributed processing: Foundations*. United States: Bradford Book, 1988. v. 1. ISBN 9780262680530. Citado na página 20.
- SANTOS, L. F. T. *Detector de assinaturas de gás em levantamentos sísmicos utilizando LSTM*. Dissertação (Mestrado) — Pontifícia Universidade Católica do Rio de Janeiro, 2019. Citado 8 vezes nas páginas 12, 13, 15, 16, 28, 29, 31 e 36.
- SHERIFF, R. E.; GELDART, L. P. *Exploration seismology*. [S.l.]: Cambridge university press, 1995. Citado na página 18.
- SILVA, I. d.; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas. *São Paulo: Artliber*, p. 33–111, 2010. Citado 2 vezes nas páginas 19 e 20.
- SILVEIRA, A.; MATTOS, V.; NAKAMURA, L.; AMARAL, M.; KONRATH, A.; BORNIA, A. Análise do valor p determinado pela estatística τ na aplicação do teste de dickey-fuller aumentado. *Trends in Computational and Applied Mathematics*, 2022. Citado na página 19.
- THOMAS, J. E. et al. Fundamentos de engenharia de petróleo. 2ª edição. *Editores Interciência*, 2004. Citado na página 18.
- VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. Attention is all you need. *arXiv preprint arXiv:1706.03762*, 2017. Citado 5 vezes nas páginas 21, 22, 23, 24 e 31.
- WANG, Z.; DI, H.; SHAFIQ, M. A.; ALAUDAH, Y.; ALREGIB, G. Successful leveraging of image processing and machine learning in seismic structural interpretation: A review. *The Leading Edge*, Society of Exploration Geophysicists, v. 37, n. 6, p. 451–461, 2018. Citado na página 15.
- WU, X.; GENG, Z.; SHI, Y.; PHAM, N.; FOMEL, S.; CAUMON, G. Building realistic structure models to train convolutional neural networks for seismic structural interpretation. *Geophysics*, Society of Exploration Geophysicists, v. 85, n. 4, p. WA27–WA39, 2020. Citado 2 vezes nas páginas 13 e 15.