



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Marcus Vinicius de Oliveira Saraiva

**Uso de Redes Adversárias Geradoras
Conditionais para Construção de Modelos de
Velocidades Sísmicas**

Vitória, ES

2022

Marcus Vinicius de Oliveira Saraiva

Uso de Redes Adversárias Geradoras Condicionais para Construção de Modelos de Velocidades Sísmicas

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

Universidade Federal do Espírito Santo – UFES

Centro Tecnológico

Programa de Pós-Graduação em Informática

Orientador: Prof. Dr. Thomas Walter Rauber

Coorientador: Dr. Alexandre Rodrigo Maul

Vitória, ES

2022

Ficha catalográfica disponibilizada pelo Sistema Integrado de
Bibliotecas - SIBI/UFES e elaborada pelo autor

S243u Saraiva, Marcus Vinicius de Oliveira, 1983-
Uso de Redes Adversárias Geradoras Condicionais para
Construção de Modelos de Velocidades Sísmicas / Marcus
Vinicius de Oliveira Saraiva. - 2022.
103 f. : il.

Orientador: Thomas Walter Rauber.
Coorientador: Alexandre Rodrigo Maul.
Dissertação (Mestrado em Informática) - Universidade
Federal do Espírito Santo, Centro Tecnológico.

1. Inteligência Artificial. 2. Geofísica. 3. Redes Neurais. 4.
Modelo de Velocidades Sísmicas. 5. Redes Neurais Geradoras. 6.
Sísmica. I. Rauber, Thomas Walter. II. Maul, Alexandre
Rodrigo. III. Universidade Federal do Espírito Santo. Centro
Tecnológico. IV. Título.

CDU: 004

Marcus Vinicius de Oliveira Saraiva

Uso de Redes Adversárias Geradoras Condicionais para Construção de Modelos de Velocidades Sísmicas

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

Trabalho aprovado. Vitória, ES, 27 de outubro de 2022:

Prof. Dr. Thomas Walter Rauber

Orientador



Documento assinado digitalmente

ALEXANDRE RODRIGO MAUL

Data: 31/10/2022 14:03:41-0300

Verifique em <https://verificador.iti.br>

Dr. Alexandre Rodrigo Maul

Coorientador

Prof. Dr. Renato Antônio Krohling

Membro interno



Documento assinado digitalmente

MARCOS DE CARVALHO MACHADO

Data: 01/11/2022 16:23:36-0300

Verifique em <https://verificador.iti.br>

Dr. Marcos de Carvalho Machado

Membro externo



Documento assinado digitalmente

ANTONIO COSME DEL REY

Data: 21/11/2022 14:26:26-0300

Verifique em <https://verificador.iti.br>

Dr. Antônio Cosme Del Rey

Membro externo

Prof. Dr. Avelino Forechi Silva

Membro externo

Vitória, ES

2022



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

PROTOCOLO DE ASSINATURA



O documento acima foi assinado digitalmente com senha eletrônica através do Protocolo Web, conforme Portaria UFES nº 1.269 de 30/08/2018, por
THOMAS WALTER RAUBER - SIAPE 2201072
Departamento de Informática - DI/CT
Em 28/10/2022 às 14:23

Para verificar as assinaturas e visualizar o documento original acesse o link:
<https://api.lepisma.ufes.br/arquivos-assinados/593351?tipoArquivo=O>



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

PROTOCOLO DE ASSINATURA



O documento acima foi assinado digitalmente com senha eletrônica através do Protocolo Web, conforme Portaria UFES nº 1.269 de 30/08/2018, por
RENATO ANTONIO KROHLING - SIAPE 1921734
Departamento de Engenharia de Produção - DEP/CT
Em 31/10/2022 às 09:18

Para verificar as assinaturas e visualizar o documento original acesse o link:
<https://api.lepisma.ufes.br/arquivos-assinados/593863?tipoArquivo=O>



Emitido em 27/10/2022

ATA DE DEFESA Nº 9/2022 - ARA-CCEM (11.02.16.01.03.02.08)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 31/10/2022 14:23)

AVELINO FORECHI SILVA

PROFESSOR DO ENSINO BASICO TECNICO E TECNOLÓGICO

ARA-CCEM (11.02.16.01.03.02.08)

Matricula: 2296127

Para verificar a autenticidade deste documento entre em <https://sipac.ifes.edu.br/documentos/> informando seu número: **9**, ano: **2022**, tipo: **ATA DE DEFESA**, data de emissão: **31/10/2022** e o código de verificação: **d11ee77319**

Aos amores da minha vida: Maura, Pedro e Luísa!

Agradecimentos

Agradeço aos meus pais, Izabel e Antônio Marcos, por me ensinarem a valorizar a educação, exercitar a criatividade e exercer o pensamento livre. À minha esposa Maura pela parceria, compreensão, incentivo e leitura crítica dessa dissertação. Às minhas irmãs Karine e Joice por compartilharem a guarda das mais preciosas memórias de nossa infância. Aos meus tios Maria Lúcia, Paulo César e José Francisco por terem possibilitado meu primeiro contato com um computador.

Agradeço à todos os professores que tive nessa jornada, em especial a Prof^a. Claudine Badue pela recepção ao PPGI e aos meus orientadores Prof. Thomas Rauber, Alexandre Maul e Avelino Forechi pelos ensinamentos, incentivo e tempo empregado à minha orientação.

Agradeço ao ensino público, gratuito e, em sua maior parte, de qualidade do qual tive a oportunidade de usufruir desde os meus primeiros rabiscos até o dia de hoje.

Agradeço à Petrobras pela disponibilização do dado sísmico e da infraestrutura de hardwares utilizados para realização desse trabalho, bem como aos gestores Ana Paula Diniz e Mathias Erdtmann pelo incentivo e flexibilização do horário de trabalho, de forma que fosse possível assistir as aulas e elaborar esta dissertação. Aos meus colegas de trabalho com quem compartilhei minhas dúvidas, incertezas e alegrias relacionadas ao desenvolvimento desta dissertação, em especial ao Antônio Cosme Del Rey e Jorcy Neto, que participaram ativamente na concepção das ideias que deram origem a este trabalho.

*“Repetir repetir — até ficar diferente.
Repetir é um dom do estilo.”
(Manoel de Barros)*

Resumo

Na indústria de petróleo e gás um modelo de velocidades sísmicas acurado é essencial no decorrer das etapas de processamento sísmico, caracterização de reservatórios e cálculo de volume de hidrocarbonetos. A modelagem de velocidade por inversão do campo de onda completo, do inglês *Full Waveform Inversion* (FWI) é uma técnica iterativa avançada que fornece um modelo de velocidades sísmicas intervalares de alta resolução, embora com um custo computacional elevado devido às simulações numéricas baseadas na equação da onda necessárias em cada iteração. Este trabalho propõe um método de geração de modelos de velocidades sísmicas intervalares tão detalhados quanto os obtidos por meio da FWI, porém utilizando uma fração do tempo e dos recursos computacionais gastos com a modelagem tradicional, utilizando a FWI. Para tanto, foram empregadas Redes Adversárias Geradoras Condicionais, do inglês *conditional Generative Adversarial Networks* (cGAN), do tipo *pix2pix*, utilizando dados tridimensionais. O treinamento desta rede foi realizado com três entradas condicionantes, Amplitude Sísmica, Velocidade Média Tomográfica e o *grid* de Tempo Duplo de Trânsito, a saída desejada é o dado de velocidades sísmicas de alta resolução oriundo da FWI. Dados sísmicos reais foram utilizados para treinar e validar a arquitetura de rede proposta e três métricas de avaliação foram adotadas como critérios de qualidade (Análise Visual, o Erro Percentual Absoluto Médio e a Medida do Índice de Similaridade Estrutural). Com base nessas métricas os resultados do conjunto de validação sugerem que a cGAN foi capaz de gerar modelos de velocidades sísmicas de alta resolução, equiparáveis aos modelos gerados pela FWI, permitindo extrair dos dados de entrada as principais estruturas geológicas e variações laterais de velocidade. Os resultados experimentais indicam que o método proposto tem potencial para aumentar a velocidade dos processos de caracterização geofísica de reservatórios, economizando tempo e recursos computacionais.

Palavras-chaves: Redes Adversárias Geradoras Condicionais. Inversão do Campo de Onda Completo. Modelo de Velocidades Sísmicas.

Abstract

In the oil and gas industry, estimating an accurate seismic velocity model is an essential step in seismic processing, reservoir characterization, and hydrocarbon volume calculation. Full Waveform Inversion (FWI) velocity modeling is an iterative advanced technique that provides a high-resolution seismic interval velocity model, although at a very high computational cost due to the wave-equation-based numerical simulations required at each FWI iteration. This study proposes a method of generating seismic velocity models, as detailed as those obtained through FWI, however using a fraction of the time and computational resources spent in traditional FWI velocity modeling. To do so, a 3D conditional Generative Adversarial Network (cGAN), known as pix2pix was used. The cGAN training was performed using three conditional inputs, Seismic Amplitude, Tomographic Average Velocity, and Two Way Time Grid, the desired output was the high-resolution seismic interval velocity model from the FWI modeling. Real-world data were used to train and validate the proposed network architecture, and three evaluation metrics (percent error, structural similarity index measure, and visual analysis) were adopted as quality criteria. Based on these metrics, the results evaluated upon the validation set suggest that the cGAN was able to accurately match real FWI generated outputs, enabling it to extract from input data the main geological structures and lateral velocity variations. Experimental results indicate that the proposed method, when deployed, has the potential to increase the speed of geophysical reservoir characterization processes, saving time and computational resources.

Keywords: Conditional Generative Adversarial Network (cGAN). Full Waveform Inversion (FWI). Seismic Velocity Modeling.

Lista de ilustrações

Figura 1 – Janelas de geração e processo de migração do óleo e gás das rochas geradoras até as armadilhas nas rochas reservatórios.	23
Figura 2 – Tipos de ondas sísmicas, a) Onda P, b) Onda S, c) Onda <i>Love</i> e d) Onda <i>Rayleigh</i>	24
Figura 3 – Ondas P, Ondas S e ondas de superficiais (<i>Love</i> e <i>Rayleigh</i>) registradas por um sismômetro.	24
Figura 4 – Parte da energia da Onda P é refletida e parte é transmitida pela interface de rochas com impedâncias acústicas contrastantes.	26
Figura 5 – Figura esquemática de uma aquisição sísmica <i>streamer</i>	27
Figura 6 – Modelo convolucional.	28
Figura 7 – Incidência oblíqua à superfície de reflexão. Além das ondas P, transmitida e refletida, são geradas ondas S, transmitida e refletida, a partir de uma Onda P incidente. Nessa situação, a partição da energia é governada pelas Equações de Zoeppritz 2.9.	29
Figura 8 – Tipos mais usuais de ordenação do dado sísmico durante o processamento: Tiro comum, Receptor comum, Afastamento comum e CMP.	30
Figura 9 – Processo de estimativas de velocidade RMS utilizada para correção de NMO. Notar pontos interpretados e interpolados linearmente no espectro de velocidades na figura central.	31
Figura 10 – a) Representação esquemática da correção de NMO b) Empilhamento de um traço sintético após a correção de NMO.	32
Figura 11 – Modelagem FWI: a) Sismograma real comparado com sismograma sintético gerado a partir do modelo de velocidades intervalares inicial b) Sismograma real comparado com sismograma sintético gerado a partir do modelo de velocidades intervalares final.	33
Figura 12 – Representação simplificada da lacuna de informação entre o modelo de velocidades oriundo da tomografia e o volume sísmico de amplitudes. O objetivo do FWI é preencher esta lacuna.	33
Figura 13 – a) Seção sísmica arbitrária passando pelo poço 1-RJS-540-RJ da Bacia de Campos b) Mesma seção interpretada. Em azul os carbonatos albianos, em rosa a seção evaporítica (Sal) em preto o reservatório preenchido com óleo. Profundidade do reservatório em torno e 5 <i>km</i> a partir do nível do mar.	34

Figura 14 – Representação 3D do volume sísmico e suas principais direções de fatiamento: <i>in-line</i> , <i>cross-line</i> e <i>time-slice</i> . O traço sísmico é representado pela linha vermelha.	35
Figura 15 – a) Seção vertical pelo <i>pes Hippocampi major</i> de um coelho desenhada por Camillo Golgi b) Células de Purkinje e Células Granulosas de um cerebelo de um pombo desenhadas por Santiago Ramón y Cajal. . . .	37
Figura 16 – Ilustração da sinapse: processo de comunicação entre neurônios mediado por moléculas de neurotransmissores.	38
Figura 17 – Modelo de McCulloch e Pitts (1943) com duas entradas e uma saída, análogo a um neurônio com dois <i>dendritos</i> e um <i>axônio</i>	39
Figura 18 – Modelo de Rosenblatt (1958), <i>perceptron</i> , com duas entradas e uma saída.	40
Figura 19 – Frank Rosenblatt operando o <i>perceptron</i> MARK I.	41
Figura 20 – Rede neural com uma camada oculta e um neurônio em cada camada, $a^{(1)}$ é a camada de entrada, $z^{(2)} = w^{(2)}a^{(1)} + b^{(2)}$, $a^{(2)} = \sigma(z^{(2)})$, $z^{(3)} = w^{(3)}a^{(2)} + b^{(3)}$, $a^{(3)} = \sigma(z^{(3)})$, $\hat{y}$ é a saída da rede, y a saída desejada e C a função de perda.	42
Figura 21 – Exemplos de indexação dos parâmetros em uma rede com camada oculta.	43
Figura 22 – Regularização L_1 e L_2 . Curvas de nível em vermelho representa a função de custo e as regiões em azul as restrições de valores para os parâmetros impostas pela regularização.	45
Figura 23 – <i>Dropout</i> aplicado em um <i>perceptron</i> . Os neurônios em tracejado foram escolhidos aleatoriamente para serem ignorados durante uma época do treinamento.	45
Figura 24 – Imagem de um número escrito à mão com dimensões de 28×28 pixels reorganizado na forma de um vetor de tamanho $28 \times 28 = 784$	46
Figura 25 – Lenet (LECUN et al., 1998), trabalho pioneiro na utilização de Redes Neurais Convolucionais para classificação de números escritos a mão.	47
Figura 26 – Operação de correlação cruzada de um impulso unitário com um filtro à esquerda e convolução com o mesmo filtro à direita. A correlação cruzada retorna o mesmo resultado da convolução, contudo refletido em torno dos eixos x e y	48
Figura 27 – Processo de convolução 2D e <i>Max Pooling</i> no contexto das redes neurais convolucionais.	48
Figura 28 – Tipos de normalizações no contexto das redes neurais convolucionais. N é o eixo do <i>batches</i> , C dos canais, e W, H os eixos espaciais.	49
Figura 29 – Vizualização de alguns filtros da Rede Neural Convolucional VGG16.	50
Figura 30 – Estrutura básica de uma rede auto-codificadora: <i>encoder</i> , espaço latente e <i>decoder</i>	51

Figura 31 – Estrutura de uma U-Net com atalhos nas conexões entre <i>encoder</i> e <i>decoder</i>	51
Figura 32 – Microscopia de efeito Kerr mostrando a formação de domínios magnéticos no grão central de uma seção polida de aço.	53
Figura 33 – Rede de Hopfield com cinco entradas, seu vetor de entrada $\mathbf{x}$ e sua matriz de pesos $\mathbf{W}$	54
Figura 34 – Rede de Hopfield treinada para um único padrão de entrada $\mathbf{x}$	55
Figura 35 – Rede de Hopfield utilizada para reconstrução de uma imagem degradada, acima, e incompleta, abaixo.	56
Figura 36 – Rede de Hopfield à esquerda (o número de neurônios é igual ao tamanho da dimensão do dado de entrada), Máquina de Boltzmann ao centro (existem conexões entre todos os neurônios, os de entrada e os ocultos) e Máquina de Boltzmann Restrita à direita (apenas conexões entre os neurônios ocultos e os de entrada).	56
Figura 37 – Visualização de amostras produzidas por redes adversárias geradoras adversárias. À direita de cada imagem, com borda em amarelo, imagens reais, utilizadas para treinamento.	58
Figura 38 – Visualização de amostras produzidas por redes adversárias geradoras condicionais. A amostra de cada linha foi condicionada por sua respectiva categoria, no caso, os números de 0 a 9 (MIRZA; OSINDERO, 2014).	58
Figura 39 – Processo de treinamento de uma GAN condicional treinada para mapear bordas para fotos. O discriminador aprende a classificar entre imagens sintetizadas pelo gerador, $G(x)$, e imagens reais y . O objetivo do gerador é produzir imagens que não possam ser distinguidas de imagens reais. Tanto o gerador quanto o discriminador são condicionados pela imagem de entrada x	59
Figura 40 – Exemplos de imagens produzidas pela rede <i>pix2pix</i> tendo como entrada uma imagem condicionante, imagem de satélite para mapas, bordas para fotos, dia para noite e foto em tons de cinza para foto colorida.	60
Figura 41 – Visualização 3D do volume sísmico de amplitudes disponibilizado para estudo. Todo o dado possui uma modelagem de velocidades por tomografia, contudo a modelagem FWI fica restrita a porção norte.	62
Figura 42 – <i>Xline</i> central dos dados disponíveis após normalização: Amplitude Sísmica, Velocidade Média (Tomografia), Gride de Tempo Duplo (TWT), Velocidade Intervalar (FWI).	64
Figura 43 – Exemplo de sub-volumes de treinamento, acima a entrada da rede, com três canais e abaixo a saída desejada, com apenas um canal.	65
Figura 44 – Arquitetura da Rede Geradora.	68

Figura 45 – Arquitetura de Rede Discriminadora PatchGAN, originalmente proposta por Isola et al. (2017) e posteriormente implementada utilizando convoluções 3D por Cirillo, Abramian e Eklund (2020). O dado de entrada é subdividido em volumes de dimensão $32 \times 32 \times 32 \times 1$ e classificado em dado real, oriundo do conjunto de treinamento, ou falso, produzido pela Rede Geradora.	69
Figura 46 – Arquitetura de GAN condicional proposta: a Rede Geradora produz um sub-volume de velocidade intervalar tendo como entrada o dado condicionante, a Rede Discriminadora classifica se o sub-volume é oriundo do conjunto de dados reais ou dos dados gerados.	69
Figura 47 – Marcas provocadas pela junção dos sub-volumes sem sobreposição . . .	71
Figura 48 – Soma de duas janelas Hann com sobreposição de 50%.	71
Figura 49 – Janela Hann 3D aplicada a um sub-volume produzido pela rede Geradora (corte central na direção <i>inline</i>)	71
Figura 50 – SSIM e MAPE do dado de validação por época de treinamento.	74
Figura 51 – Comparação entre as distribuições da velocidade intervalar FWI e a velocidade intervalar produzida pela GAN.	75
Figura 52 – <i>Crossline</i> 1500: Velocidade estimada pelo Modelo Gerador, Velocidade FWI real e Erro Percentual Absoluto entre os volumes reais e gerados pela cGAN.	75
Figura 53 – <i>Inline</i> 670: Velocidade estimada pelo Modelo Gerador, Velocidade FWI real e Erro Percentual Absoluto entre os dois volumes.	76
Figura 54 – Amostras verticais retiradas na região do volume de validação.	77
Figura 55 – <i>Crossline</i> 1500: Velocidade estimada pelo Modelo Gerador, Amplitude Sísmica e Velocidade estimada pelo Modelo Gerador em transparência com a Amplitude Sísmica. Nota-se uma coerência das principais estruturas geológicas presentes nos dois volumes.	77
Figura 56 – <i>Inline</i> 218: Velocidade estimada pelo Modelo Gerador, Amplitude Sísmica e Velocidade estimada pelo Modelo Gerador em transparência com a Amplitude Sísmica. Nota-se uma coerência das principais estruturas geológicas presentes nos dois volumes.	78

Lista de abreviaturas e siglas

cGAN	conditional Generative Adversarial Networks
CMP	Common Midpoint
FID	Fréchet Inception Distance
FWI	Full Waveform Inversion
GAN	Generative Adversarial Networks
GPU	Graphics Processing Units
IRENA	International Renewable Energy Agency
MAPE	Mean Absolute Percentage Error
NMO	Normal Move-Out
OBC	Ocean Bottom Cable
OBN	Ocean Botton Nodes
RMS	Root Mean Square
RNN	Recurrent Neural Networks
RTM	Reverse Time Migration
SSIM	Structural Similarity Index Measure
TWT	Two Way Time

Sumário

1	INTRODUÇÃO	19
2	REFERENCIAL TEÓRICO	22
2.1	Geologia do petróleo e método sísmico	22
2.1.1	Reservatórios de hidrocarbonetos	22
2.1.2	Ondas sísmicas	23
2.1.3	Dado sísmico	26
2.2	Redes Neurais	35
2.2.1	Redes neurais biológicas	36
2.2.2	Redes neurais artificiais	37
2.2.3	Redes Neurais Convolucionais	45
2.2.4	Redes Neurais Baseadas em Energia	52
2.2.5	Redes Adversárias Geradoras	56
2.3	Trabalhos Relacionados	60
3	MATERIAIS E MÉTODOS	62
3.1	Dados disponíveis	62
3.1.1	Dados Brutos	62
3.1.2	Base de dados	63
3.2	Bibliotecas Numéricas	66
3.3	Equipamentos	66
4	METODOLOGIA EXPERIMENTAL	67
4.1	Arquitetura da Rede	67
4.1.1	Rede Geradora	67
4.1.2	Rede Discriminadora	67
4.1.3	GAN condicional	68
4.2	Treinamento	68
4.3	Inferência e reconstrução do volume	70
4.4	Métricas	72
5	RESULTADOS E DISCUSSÕES	74
5.1	Resultados	74
5.2	Discussões	79
6	CONCLUSÕES E TRABALHOS FUTUROS	81

REFERÊNCIAS	82
APÊNDICES	91
APÊNDICE A – PUBLICAÇÃO 1	92
APÊNDICE B – PUBLICAÇÃO 2	101

1 Introdução

Diante da ameaça real do aquecimento global, causado pelo excesso de emissões de gases que contribuem para o efeito estufa e consequente mudança climática, a substituição dos combustíveis fósseis por uma matriz energética renovável é um desafio e uma prioridade para o Brasil e para o mundo (ARIAS et al., 2021). Contudo, a sociedade ainda é extremamente dependente de combustíveis fósseis, e o segmento de óleo e gás provê mais da metade da energia primária mundial (PETROBRAS, 2022).

O Acordo de Paris, um tratado internacional que dispõe sobre medidas de redução de emissão de gases estufa, do qual o Brasil é signatário, objetiva um aumento máximo na temperatura média global de menos de 2°C, preferencialmente limitando esse aumento em 1,5°C até o final do século XXI com a finalidade de reduzir substancialmente os efeitos das mudanças climáticas em curso (UNFCCC, 2015).

Para tanto, os países signatários se comprometeram a aplicar medidas para reduzir as emissões dos gases de efeito estufa a partir do ano de 2020. A Agência Internacional para as Energias Renováveis, do inglês: *International Renewable Energy Agency* (IRENA) apresenta um plano de ação para que essa transição energética aconteça até 2050, quando seria atingido um equilíbrio entre emissões e retenção de carbono, resultando em uma emissão líquida igual a zero (IRENA, 2022).

Além dos esforços de médio e longo prazo para atingir as metas do Acordo de Paris, no curto e médio prazo, durante a transição, se faz necessário que as reservas existentes de óleo e gás, assim como as novas reservas a serem descobertas sejam exploradas de maneira eficiente e racional, de forma a garantir a segurança energética mundial e propiciar espaço para o surgimento de novas tecnologias que possam acelerar a substituição dos combustíveis fósseis por uma matriz energética renovável (RAIKAR; ADAMSON, 2019; COVERT; GREENSTONE; KNITTEL, 2016; DROEGE, 2011).

A utilização de novas ferramentas no campo do aprendizado de máquinas tem se revelado promissora nos mais diversos campos da ciência (ABBASI; GOLDENHOLZ, 2019; GHODDUSI; CREAMER; RAFIZADEH, 2019; SHARMA et al., 2020; BARON, 2019) e pode contribuir para o enfrentamento deste contexto desafiador, ajudando incorporar mais informação aos modelos geofísicos seja a partir dos novos dados a serem adquiridos e ou a partir dos dados já existentes, porém subutilizados.

A estimativa correta do campo de velocidades é uma etapa fundamental do processamento sísmico (SCHULTZ, 1998). Um modelo de velocidades bem estimado favorecerá todas as outras etapas do fluxo de processamento sísmico, principalmente nas etapas de migração e de posicionamento em profundidade (MAUL, 2020).

A inversão do campo de ondas completo, do inglês *Full Waveform Inversion* (FWI), é um método avançado para estimativa do campo de velocidades sísmicas. Trata-se de um método inverso que utiliza o dado sísmico bruto como entrada de um processo iterativo que objetiva estimar um modelo de velocidades intervalares de alta resolução. Contudo a FWI é computacionalmente custosa, devido às simulações numéricas baseadas na equação da onda exigidas em cada iteração (TARANTOLA, 1984).

O uso de *data-driven surrogates* (aproximação por modelos baseados em dados) tem apresentado resultados promissores nos últimos anos (ZHANG; LIU; SUN, 2020; RAISSI; PERDIKARIS; KARNIADAKIS, 2019; ZHU et al., 2019). Esta abordagem emprega ferramentas de aprendizado de máquinas para extrair dos dados disponíveis um modelo que gera resultados muito próximos dos resultados obtidos por modelos numéricos, baseados em equações que descrevem fenômenos físicos.

Esses modelos baseados em dados não precisam ter, necessariamente, um significado físico. O foco são resultados coerentes com a realidade, sendo estes alcançados utilizando uma fração dos recursos computacionais normalmente demandados pelos métodos tradicionais (RAISSI, 2018).

Diante deste contexto, esta dissertação analisa a viabilidade de se utilizar redes neurais artificiais, do tipo Redes Adversárias Geradoras Condicionais, do inglês *conditional Generative Adversarial Networks* cGANs, para gerar modelos de velocidades sísmicas tão detalhados quanto os gerados pelo método FWI, com o objetivo de economizar tempo e recursos computacionais.

Para tanto, será empregado a arquitetura de cGAN pix2pix (ISOLA et al., 2017), adaptada para utilização de dados sísmicos volumétricos. O condicionamento desta rede será realizado utilizando três canais de entrada, Amplitude Sísmica, Velocidade Média Tomográfica e *Grid* de Tempo Duplo de Trânsito, a saída desejada será o dado de velocidades sísmicas intervalares de alta resolução, oriundo da FWI.

Dados sísmicos reais serão utilizados para treinar e validar a arquitetura de rede proposta e três métricas de avaliação serão adotadas como critérios de qualidade (Análise Visual, o Erro Percentual Absoluto Médio e a Medida do Índice de Similaridade Estrutural).

No Capítulo 1 é realizada uma contextualização e introdução ao tema da dissertação, onde são apresentados a motivação do trabalho, o problema de pesquisa, objetivos e estrutura da dissertação.

O Capítulo 2 apresenta a fundamentação teórica geral, necessária para o entendimento do problema, abordando os aspectos básicos de geologia do petróleo, aquisição e processamento sísmico, modelos de velocidades sísmicas, redes neurais artificiais e trabalhos relacionados.

O Capítulo 3 aborda os materiais e métodos utilizados, discorrendo sobre os dados

disponíveis, pré-processamento, construção dos *datasets* de treino, validação e inferência, bibliotecas numéricas e equipamentos utilizados.

O Capítulo 4 apresenta a metodologia aplicada para a solução do problema, desde o processo de construção das redes Geradora e Discriminadora, passando pelos processos de treinamento, inferência, reconstrução do volume até a escolha das métricas de avaliação.

O capítulo 5 apresenta os resultados e discussões sobre os experimentos realizados, onde serão abordados os pontos positivos e contribuições deste trabalho, bem como as limitações da metodologia proposta.

O Capítulo 6 apresenta as conclusões e trabalhos futuros.

2 Referencial Teórico

2.1 Geologia do petróleo e método sísmico

2.1.1 Reservatórios de hidrocarbonetos

Os hidrocarbonetos de origem orgânica (petróleo e gás natural) se formaram a partir da morte e deposição de algas e outros micro-organismos juntamente com sedimentos de granulometria fina, em bacias sedimentares, sob condições anóxicas ([THOMAS et al., 2001](#)).

A deposição sucessiva desses sedimentos juntamente com a matéria orgânica, associada a subsidência, cria condições de temperatura e pressão propícias às transformações físico-químicas necessárias para formação da rocha geradora de petróleo e gás natural.

À profundidades em torno de 2km e temperaturas entre 60° e 120°C, a maior parte da matéria orgânica é convertida em óleo, este intervalo de profundidade e temperatura é conhecido como janela do óleo. Acima de 120°C a maior parte da matéria orgânica é convertida em metano, esse intervalo é conhecido como janela do gás ([EARLE, 2015](#)).

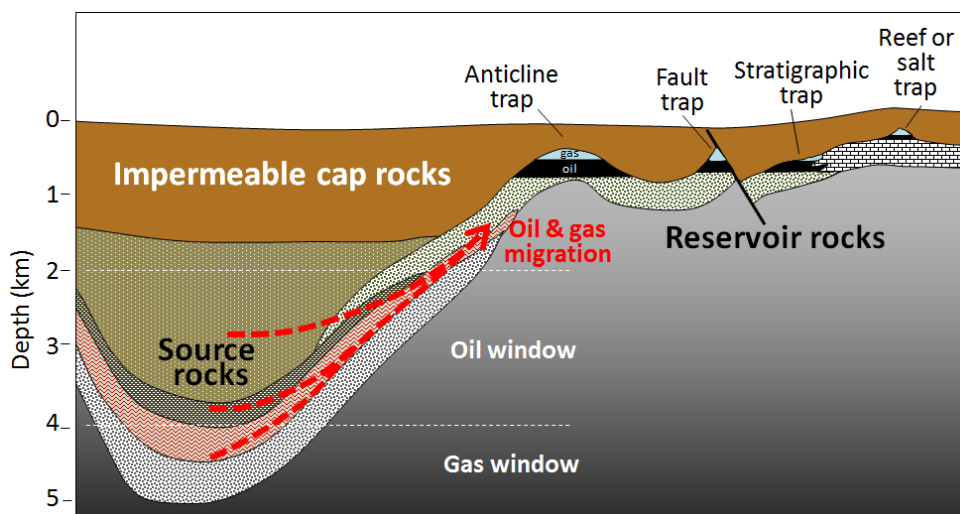
Devido à sua baixa densidade e a outros efeitos como compactação das rochas, expansão termal ou em consequência da mudança de fase, após a geração, os hidrocarbonetos de origem orgânica migram para porções mais rasas, ainda em subsuperfície, e ocasionalmente se acumulam em um tipo de rocha porosa e permeável, chamada de rocha reservatório. Essa acumulação só acontecerá caso exista alguma barreira à migração do fluido, essa barreira é formada por uma rocha de baixíssima permeabilidade, a rocha selante ([AAPG, 2022](#)). Nas bacias sedimentares brasileiras, as principais rochas reservatórios são arenitos turbidíticos e carbonatos microbiais e as principais rochas selantes são folhelhos e evaporitos ([BRUHN et al., 2017](#)).

Além da rocha geradora, rocha reservatório e rocha selante, para que volumes significativos de hidrocarbonetos sejam acumulados, é necessário a existência de uma armadilha. Esta pode ser estrutural (deformação da rocha gera uma geometria que aprisiona o fluido), estratigráfica (variação lateral de litologia impede a migração do fluido da rocha reservatório) ou ainda uma combinação das duas ([EARLE, 2015](#)) (Figura 1).

Esses quatro elementos (rocha geradora, rocha reservatório, rocha selante e armadilha) juntamente com o processo de migração do óleo, de forma síncrona, formam o sistema petrolífero ([MAGOON; DOW, 1994](#)).

O petróleo e gás natural são constituídos, em sua maior parte, por uma mistura de hidrocarbonetos (compostos químicos orgânicos formados por carbono e hidrogênio),

Figura 1 – Janelas de geração e processo de migração do óleo e gás das rochas geradoras até as armadilhas nas rochas reservatórios.



Fonte: (EARLE, 2015).

mas podem conter pequenas quantidades de enxofre, nitrogênio e oxigênio. As moléculas de hidrocarbonetos que constituem o petróleo e o gás natural podem variar entre cadeias curtas (de um a quatro átomos de carbono) até cadeias longas (acima de 38 átomos de carbono) (THOMAS et al., 2001).

2.1.2 Ondas sísmicas

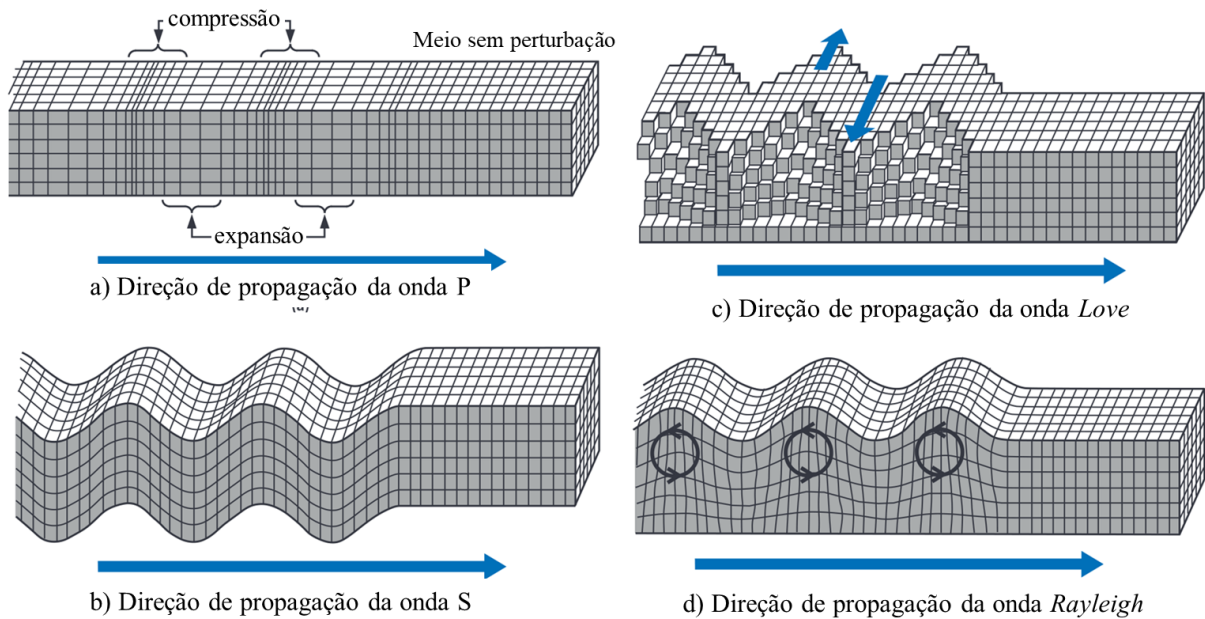
As ondas sísmicas são ondas mecânicas que se propagam pelo interior da terra e podem ser geradas por perturbações naturais como terremotos, vulcões e grandes movimentos de massa ou geradas por fontes artificiais, equipamentos concebidos especificamente para este fim, como explosivos, canhões de ar e *vibroscis* (SHERIFF, 2002).

Em sólidos infinitos e homogêneos, somente dois tipos de ondas mecânicas se propagam, as ondas compressoriais e as ondas cisalhantes. No caso de sólidos finitos, ou seja, com superfícies livres, surgem outros tipos de ondas, as superficiais, como as ondas *Rayleigh* e *Love* (THOMAS, 2003) (Figura 2.)

Dentre os tipos de ondas, as compressoriais possuem maior velocidade de propagação, por isso são conhecidas como ondas primárias ou Ondas P, pois são as primeiras ondas a serem registradas por um sismômetro¹. As ondas cisalhantes possuem velocidade menor que as ondas compressoriais, normalmente são responsáveis pelo segundo evento no registro do sismômetro, por isso são conhecidas como ondas secundárias ou Ondas S (SANTOS;

¹ Sismômetros são instrumentos de detecção e registro dos movimentos do solo. Normalmente são compostos por uma massa suspensa por uma mola e um sistema de amortecimento, de forma que, por inércia, o sistema fica em repouso, enquanto o solo se movimenta devido a passagem de uma onda sísmica (HAVSKOV; ALGUACIL, 2005).

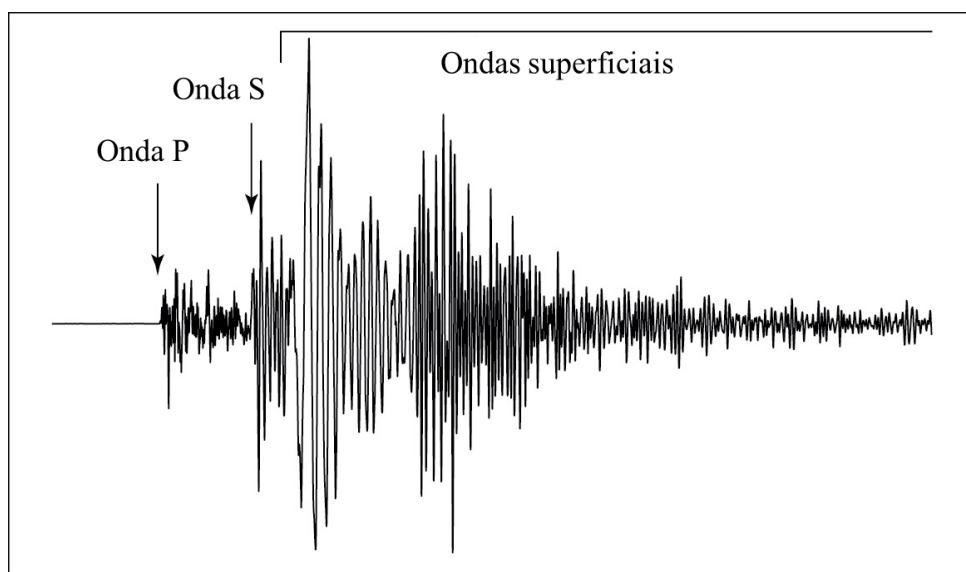
Figura 2 – Tipos de ondas sísmicas, a) Onda P, b) Onda S, c) Onda *Love* e d) Onda *Rayleigh*.



Fonte: Modificado de: (SANTOS; CATAPANG; REYTA, 2019).

CATAPANG; REYTA, 2019). As ondas superficiais (*Love* e *Rayleigh*) são registradas em seguida (Figura 3). No contexto da geofísica de petróleo, as ondas superficiais são tratadas como ruído (*ground-roll*) e eliminadas do dado de registro durante a etapa de processamento sísmico.

Figura 3 – Ondas P, Ondas S e ondas de superficiais (*Love* e *Rayleigh*) registradas por um sismômetro.



Fonte: Modificado de (UPSEIS, 2022).

O comportamento das Ondas P e Ondas S (sua propagação no espaço em função do tempo) pode ser modelado pela equação tridimensional da onda em meios elásticos,

dada pela Eq. 2.1.

$$\frac{\partial^2 \mathbf{u}}{\partial t^2} - \mathbf{V}_P^2 \nabla(\nabla \cdot \mathbf{u}) - \mathbf{V}_S^2 \Delta \mathbf{u} = \mathbf{f} \quad (2.1)$$

Onde $\mathbf{u} = \mathbf{u}(x, y, z, t)$ é a função que descreve o deslocamento provocado pela onda sísmica, $\mathbf{V}_P$ é a velocidade da onda P, $\mathbf{V}_S$ a velocidade da onda S, dadas respectivamente pelas Equações 2.2 e 2.3 e $\mathbf{f}$ a função que modela a fontes das ondas sísmicas.

$$\mathbf{V}_P = \sqrt{\frac{\lambda + 2\mu}{\rho}} \quad (2.2)$$

$$\mathbf{V}_S = \sqrt{\frac{\mu}{\rho}} \quad (2.3)$$

Onde λ e μ são os parâmetros elásticos de *Lamé* e ρ é a densidade do meio. O segundo parâmetro de *Lamé* possui uma interpretação física direta, μ é o módulo de cisalhamento do material. Já o primeiro parâmetro de *Lamé*, λ , não possui um significado físico direto, contudo, para materiais muito incompressíveis e pouco cisalhantes, ou para fluidos, onde $\mu = 0$, podemos aproximar λ do módulo *bulk* (κ), uma vez que $\kappa = \lambda + \frac{2\mu}{3}$.

Rochas, sedimentos e fluidos em subsuperfície possuem $\lambda > 0$, dessa forma, pelas Equações 2.2 e 2.3, temos que $\mathbf{V}_P$ sempre será maior que $\mathbf{V}_S$. Outra conclusão possível de se estabelecer analisando a Eq. 2.3 é que Ondas S não se propagam em fluidos, uma vez que o módulo de cisalhamento, μ , é zero nestes meios.

Uma forma alternativa de expressar os parâmetros elásticos é através das impedâncias acústica e elástica, dadas respectivamente pelas Equações 2.4 e 2.5.

$$I_P = \rho V_P \quad (2.4)$$

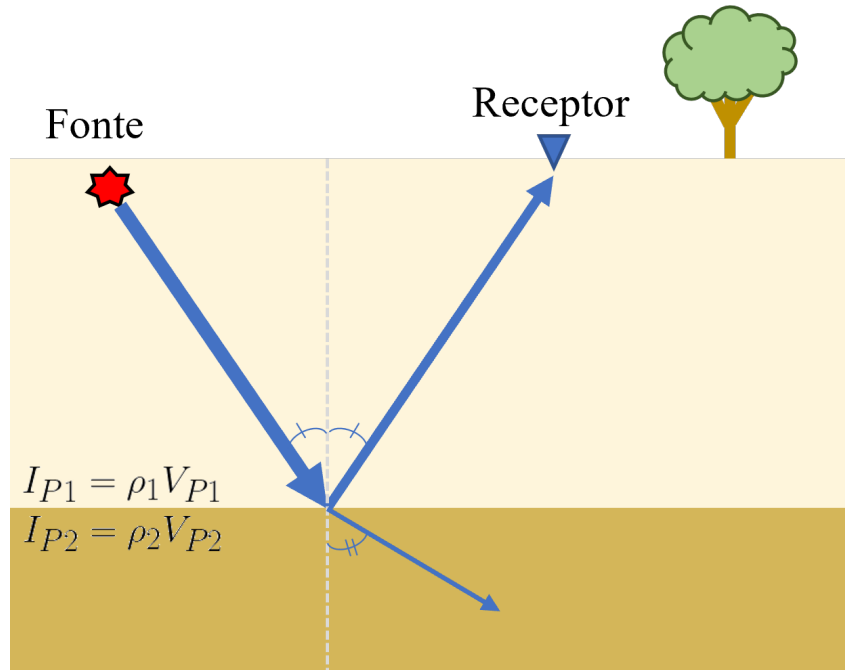
$$I_S = \rho V_S \quad (2.5)$$

Caso exista um contraste de impedância na interface das rochas por onde as ondas se propagam, parte da energia dessas ondas é refletida e parte é transmitida. A proporção da energia transmitida e da energia refletida varia em função do contraste de impedâncias e do ângulo da onda incidente (Figura 4).

Simulações numéricas que utilizam a equação da onda elástica (2.1) conseguem descrever por completo a propagação das ondas sísmicas, porém, são computacionalmente custosas. Contudo, em várias situações, a utilização da equação da onda acústica com densidade constante (Eq. 2.6) é suficiente e possuem um custo computacional menor.

$$\frac{\partial^2 \mathbf{u}}{\partial t^2} - \mathbf{V}_P^2 \nabla^2 \mathbf{u} = \mathbf{f} \quad (2.6)$$

Figura 4 – Parte da energia da Onda P é refletida e parte é transmitida pela interface de rochas com impedâncias acústicas contrastantes.



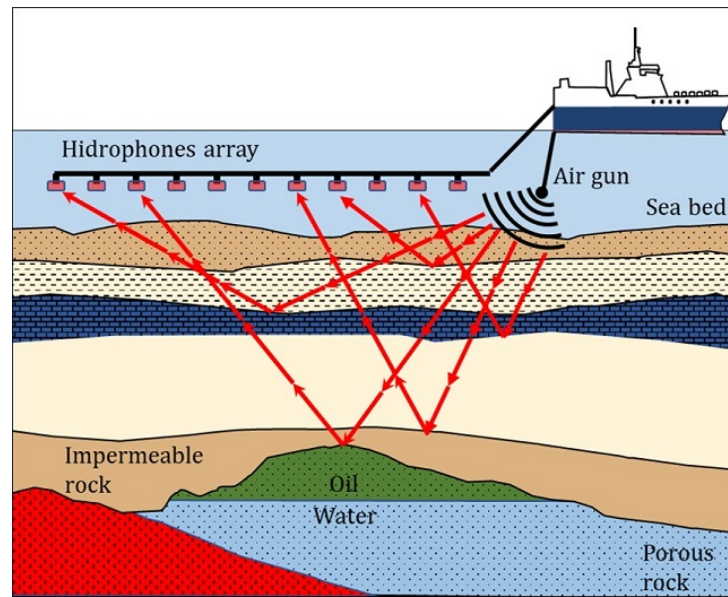
2.1.3 Dado sísmico

A indústria do óleo e gás natural emprega a sísmica de reflexão como o principal método indireto de investigação da subsuperfície. Este método utiliza fontes controladas de ondas sísmicas e arranjos de receptores para medir o tempo de trânsito e amplitude dessas ondas refletidas na interface de rochas com contraste de impedância sísmica (BIONDI, 2006) (Figura 5). Além das reflexões, outros eventos como difrações e refrações também são capturados.

A depender do ambiente, a aquisição sísmica pode ser do tipo terrestre ou marítima. A aquisição sísmica terrestre convencional utiliza cargas explosivas, enterradas a uma profundidade entre 4 m e 10 m, como fonte do sinal sísmico. Este tipo de fonte explosiva pode ser substituída por um vibrador mecânico acoplado a um veículo de grande porte conhecido como *vibroscis*. Nestas duas situações são utilizados arranjos com centenas ou mesmo milhares de geofones como receptores (YILMAZ, 2001). Os geofones são compostos por uma massa magnética suspensa por uma mola que move-se livremente dentro de uma bobina, induzindo um sinal elétrico quando excitado por ondas sísmicas (SHERIFF, 2002).

As aquisições sísmicas marítimas podem ser do tipo *Streamer*, *Ocean Bottom Cable* (OBC) ou *Ocean Bottom Nodes* (OBN), nos três casos a fonte de ondas sísmicas é um canhão de ar acoplado a um navio (BIONDI, 2006).

A aquisição sísmica do tipo *Streamer* utiliza um arranjo de sensores de pressão (hidrofones) presos à cabos, que são arrastados por um navio projetado especialmente para

Figura 5 – Figura esquemática de uma aquisição sísmica *streamer*.

Fonte: (SARAIVA et al., 2021).

realizar esse tipo de tarefa. Os cabos ficam a uma profundidade constante, abaixo do nível do mar, normalmente entre 5 m e 15 m (CIESLEWICZ; LAWTON, 1998), com o intuito de minimizar efeitos de interferência destrutiva provocadas pela reverberação de ondas sísmicas entre o fundo do mar e a superfície livre²

A aquisição sísmica *Streamer* é de difícil operação em campos com muitas obstruções, como plataformas e infraestrutura de escoamento de óleo, pois, por questão de segurança, o navio não pode aproximar os cabos dessas instalações. Uma forma de contornar o problema das obstruções são aquisições sísmicas do OBC e OBN.

Nas aquisições do tipo OBC, cabos com milhares de sensores são dispostos no assoalho oceânico. Esta técnica permite um acoplamento dos sensores diretamente com os sedimentos do fundo do mar possibilitando a utilização de acelerômetros ou geofones em conjunto com os sensores de pressão, que permite o registro das ondas sísmicas cisalhantes (BIONDI, 2006).

A aquisição OBN utiliza sensores individuais, *nodes*, posicionados em uma malha regular no assoalho oceânico. Cada *node* é uma unidade autônoma e assim como no caso do OBC, também permite o uso de acelerômetros ou geofones para o registro das Ondas S, além do sensor de pressão, pois estão acoplados aos sedimentos marinhos.

O objetivo final de uma aquisição sísmica, seja ela terrestre ou marítima, é obter

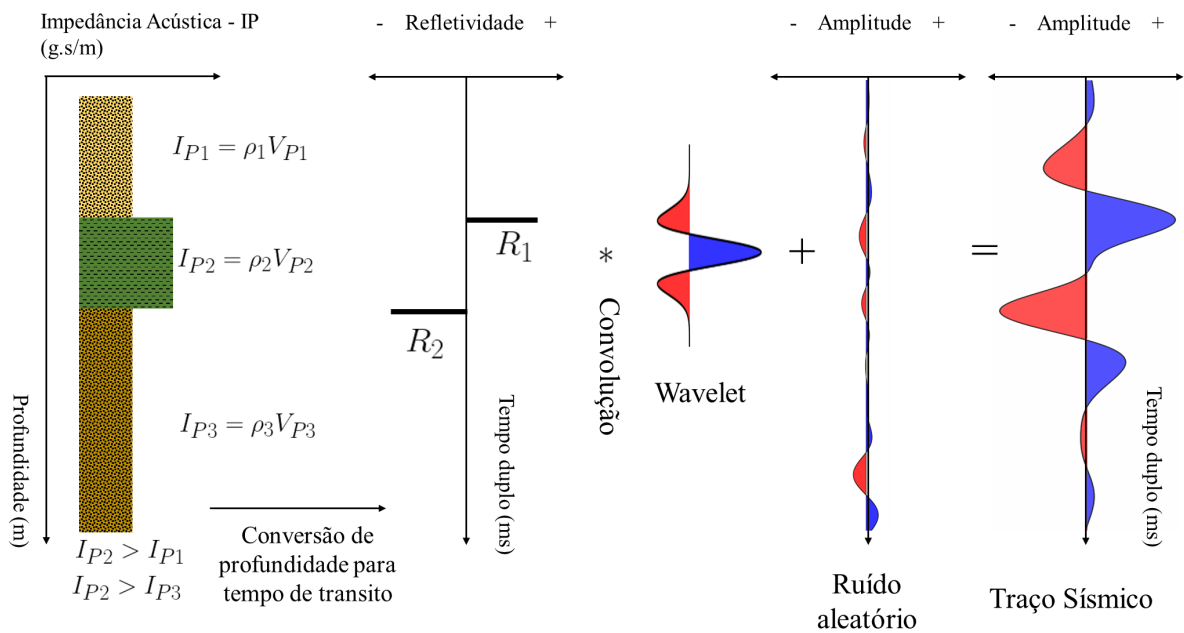
² Um líquido, sendo composto de moléculas relativamente compactas com fortes forças coesivas, tende a reter seu volume e formará uma superfície livre em um campo gravitacional se não for confinado por cima (CENGEL; CIMBALA, 2013). Como as moléculas de gás são amplamente espaçadas, com forças coesivas desprezíveis, a interface entre um líquido e um gás, como no caso da superfície marinha e a atmosfera, pode ser considerada uma superfície livre.

informações sobre a geometrias e propriedades das rochas em subsuperfície.

O modelo mais básico para modelagem direta do dado sísmico é o modelo convolucional (RUSSELL, 1988). Este modelo considera o registro sísmico como sendo o resultado de uma convolução entre a refletividade das rochas em subsuperfície com uma *wavelet* representativa da fonte sísmica e a adição de um ruído aleatório. Considera-se também que o afastamento entre fonte e receptor é zero (Figura 6 e Equação 2.7).

$$s(t) = r(t) * w(t) + \epsilon \quad (2.7)$$

Figura 6 – Modelo convolucional.



Na Equação 2.7, $s(t)$ é o traço sísmico ou sinal sísmico, $*$ o operador de convolução, $r(t)$ é a refletividade sísmica, $w(t)$ é a *wavelet* representativa da fonte sísmica e ϵ o ruído aleatório, inerente à aquisição sísmica. A convolução é uma operação matemática que pondera duas funções, neste caso a função refletividade com a *wavelet*, resultando no traço sísmico.

A refletividade sísmica $r(t)$, também conhecida como coeficiente de reflexão, fornece a porcentagem da energia da onda incidente que será refletida. Por exemplo, na Figura 6 a refletividade da primeira interface, será dada pela Equação 2.8.

$$R_1 = \frac{\rho_2 V_{P2} - \rho_1 V_{P1}}{\rho_2 V_{P2} + \rho_1 V_{P1}} = \frac{I_{P2} - I_{P1}}{I_{P2} + I_{P1}} \quad (2.8)$$

Note que, quando a refletividade for positiva, a onda refletida terá a mesma fase da onda incidente e quando a refletividade for negativa, a onda refletida terá uma fase

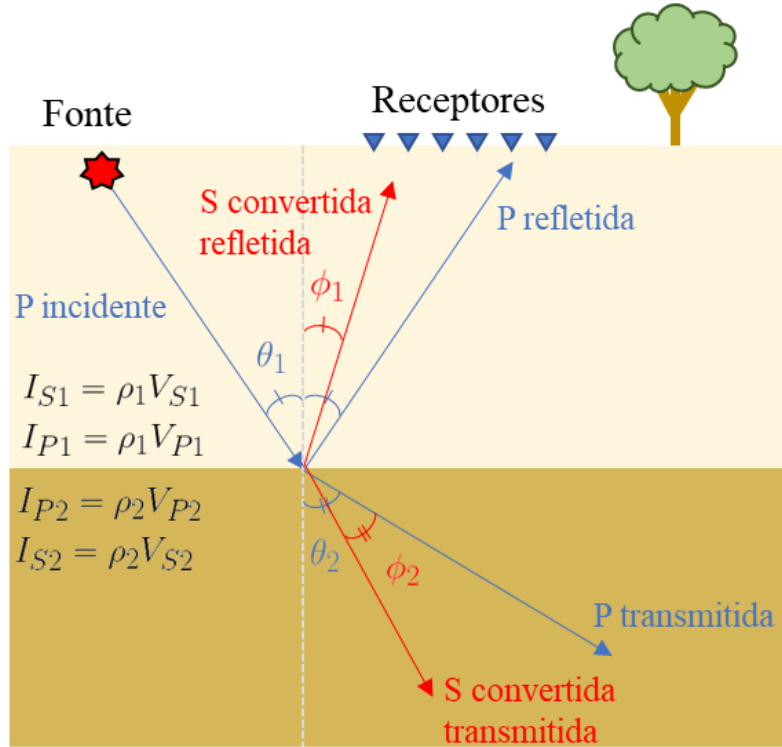
oposta à onda incidente.

Quando a incidência da onda sísmica não for normal à superfície de reflexão, existirá conversão de Ondas P em Ondas S (Figura 7). Neste caso será necessário utilizar as Equações de Zoeppritz (Equação 2.9), ou suas simplificações, para descrever a partição de energia em uma interface geológica.

$$\begin{bmatrix} R_P \\ R_S \\ T_P \\ T_S \end{bmatrix} = \begin{bmatrix} -\sin \theta_1 & -\cos \phi_1 & \sin \theta_2 & \cos \phi_2 \\ \cos \theta_1 & -\sin \phi_1 & \cos \theta_2 & -\sin \phi_2 \\ \sin 2\theta_1 & \frac{V_{P1}}{V_{S1}} \cos 2\phi_1 & \frac{\rho_2 V_{S2}^2 V_{P1}}{\rho_1 V_{S1}^2 V_{P2}} \sin 2\theta_2 & \frac{\rho_2 V_{S2} V_{P1}}{\rho_1 V_{S1}^2} \cos 2\phi_2 \\ -\cos 2\phi_1 & \frac{V_{S1}}{V_{P1}} \sin 2\phi_1 & \frac{\rho_2 V_{P2}}{\rho_1 V_{P1}} \cos 2\phi_2 & -\frac{\rho_2 V_{S2}}{\rho_1 V_{P1}} \sin 2\phi_2 \end{bmatrix}^{-1} \begin{bmatrix} \sin \theta_1 \\ \cos \theta_1 \\ \sin 2\theta_1 \\ \cos 2\phi_1 \end{bmatrix} \quad (2.9)$$

Onde R_P , R_S , T_P , e T_S são, respectivamente, os coeficientes das ondas P refletida, S refletida, P transmitida e S transmitida, e θ_1 , θ_2 , ϕ_1 e ϕ_2 são os ângulos das ondas P refletida, P transmitida, S refletida e S transmitida.

Figura 7 – Incidência oblíqua à superfície de reflexão. Além das ondas P, transmitida e refletida, são geradas ondas S, transmitida e refletida, a partir de uma Onda P incidente. Nessa situação, a partição da energia é governada pelas Equações de Zoeppritz 2.9.

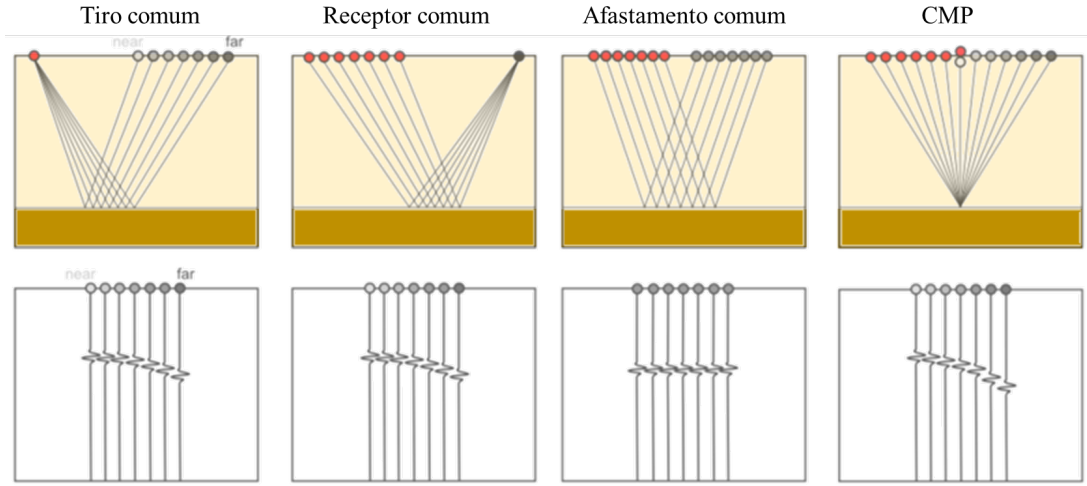


Durante a aquisição, a energia emitida por uma única fonte é registrada por milhares de receptores e a mesma região em subsuperfície é amostrada várias vezes, por diferentes arranjos de fontes e receptores. Quando plotamos o sinal de uma linha de receptores para

um determinado tiro, temos um sismograma no domínio do tiro, conhecido no jargão da indústria como *shot gather* (BIONDI, 2006). Esse é o domínio original da aquisição sísmica, no qual o dado é armazenado e enviado para o processamento.

Durante o processamento sísmico se torna conveniente ordenar o dado em outros domínios. Um domínio muito utilizado é o CMP (*common midpoint*), onde o dado é ordenado por pares de fontes e receptores equidistantes de forma a amostrar várias vezes o mesmo ponto em subsuperfície. A projeção vertical desse ponto na superfície é o CMP. Ao ordenar o dado pelos CMPs obtém-se os *CMP gathers* (BIONDI, 2006). A depender do objetivo ainda é possível ordenar o dado sísmico pelo receptor comum (*Common receiver gather*) e afastamento comum (*Common offset gather*) (Figura 8).

Figura 8 – Tipos mais usuais de ordenação do dado sísmico durante o processamento: Tiro comum, Receptor comum, Afastamento comum e CMP.



Fonte: Modificado de (SUBSURFWIKI, 2022).

As reflexões são as feições de maior interesse, pois evidenciam a presença de uma interface geológica. Em um *CMP gather* as reflexões se apresentam como hipérboles, cuja a excentricidade depende da velocidade média quadrática (RMS) da onda sísmica no meio, quanto maior a velocidade menor a excentricidade, e mais fechada será a hipérbole de reflexão (YILMAZ, 2001).

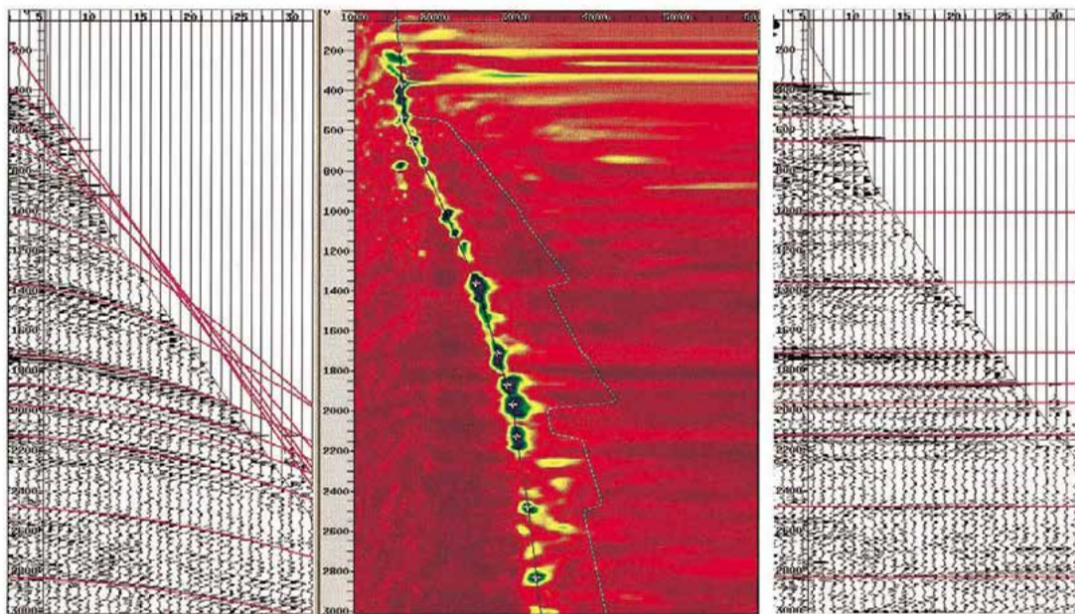
Para um modelo com camadas horizontais e afastamento curtos, a hipérbole de reflexão será dada pela Equação 2.10 (YILMAZ, 2001).

$$t^2 = t_0^2 + \frac{x^2}{v_{RMS}^2} \quad (2.10)$$

Onde x é o afastamento entre fonte e receptor, t_0 é o tempo duplo de trânsito para afastamento igual a zero (*zero offset*). Para situações onde existam refletos inclinados, ainda é necessário aplicar uma correção, conhecida como correção de DMO (*dip move-out*). Apesar de existirem algumas propostas para automatizar o processo de estimativa

da função de velocidade RMS correspondente as hipérboles de reflexão, a exemplo de (CALDERÓN-MACÍAS; SEN; STOFFA, 1998), (BISWAS et al., 2019) e (FERREIRA et al., 2021), tradicionalmente este trabalho é interpretativo e manual. Consiste na análise de um painel com o espectro de velocidades e na escolha de um valor de velocidade que melhor descreve a hipérbole de reflexão, tendo como referência a energia do painel do espectro de velocidades (CHAURIS; NOBLE, 1998). Essa energia é calculada comparando uma série de hipérboles teóricas com as hipérboles de reflexão no CMP *gather* (Figura 9).

Figura 9 – Processo de estimativas de velocidade RMS utilizada para correção de NMO. Notar pontos interpretados e interpolados linearmente no espectro de velocidades na figura central.



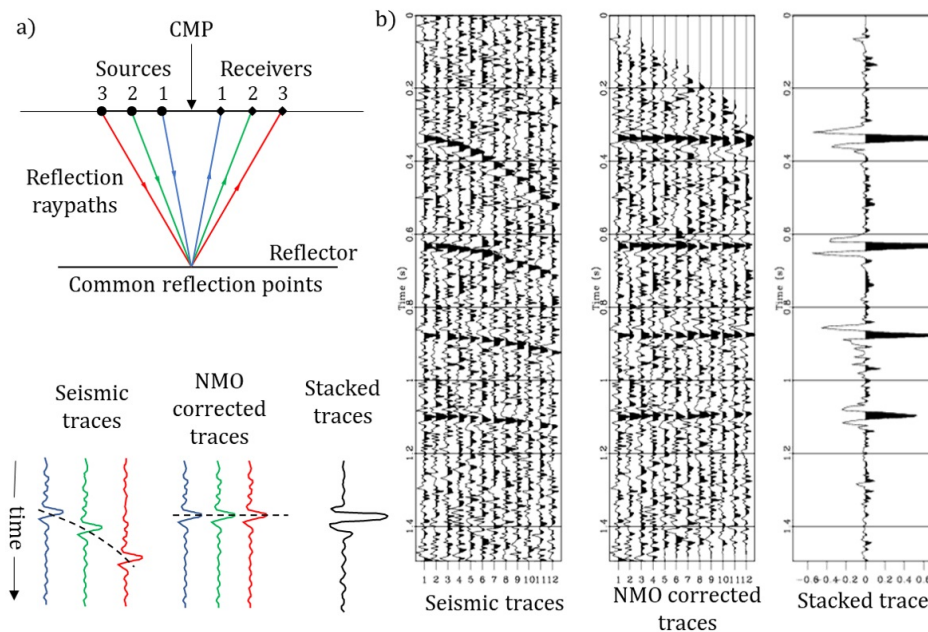
Fonte: (YILMAZ, 2001).

Uma vez estimada a função de velocidade RMS com o auxílio do espectro de velocidades, é possível realizar a correção de NMO (*Normal Move-Out*), onde as hipérboles são horizontalizadas, ou seja, os pares fontes/receptores são corrigidos para o afastamento igual a zero. Essa horizontalização permite o empilhamento do dado sísmico, uma soma sobre o eixo horizontal do CMP *gather*, onde os ruídos aleatórios são atenuados e os dados coerentes reforçados, resultando em uma amostra vertical da subsuperfície, conhecida como traço sísmico (SHERIFF, 2002) (Figura 10).

Além da estimativa de velocidade RMS, existem outros processos, mais sofisticados, que permitem uma estimativa mais acurada do campo de velocidades sísmicas, como a tomografia e a inversão do campo de ondas completo do inglês *Full Waveform Inversion* (FWI).

Tomografia sísmica é o procedimento que investiga o tempo de trânsito das ondas sísmica entre fonte e receptor com o objetivo de estimar a distribuição de velocidades da subsuperfície (SCHUSTER, 2017). A modelagem direta pode ser realizada utilizando

Figura 10 – a) Representação esquemática da correção de NMO b) Empilhamento de um traço sintético após a correção de NMO.



Fonte: Modificado de: (LIU et al., 2009).

traçado de raios em um modelo de velocidades intervalares inicial. Como o mesmo ponto em subsuperfície é amostrado por vários pares fonte/receptor é possível calcular os resíduos dos tempos de trânsito e atualizar o modelo de velocidades, que será utilizado como modelo inicial na próxima iteração tomográfica. Este processo é repetido até que os resíduos atinjam valores de corte previamente estabelecidos.

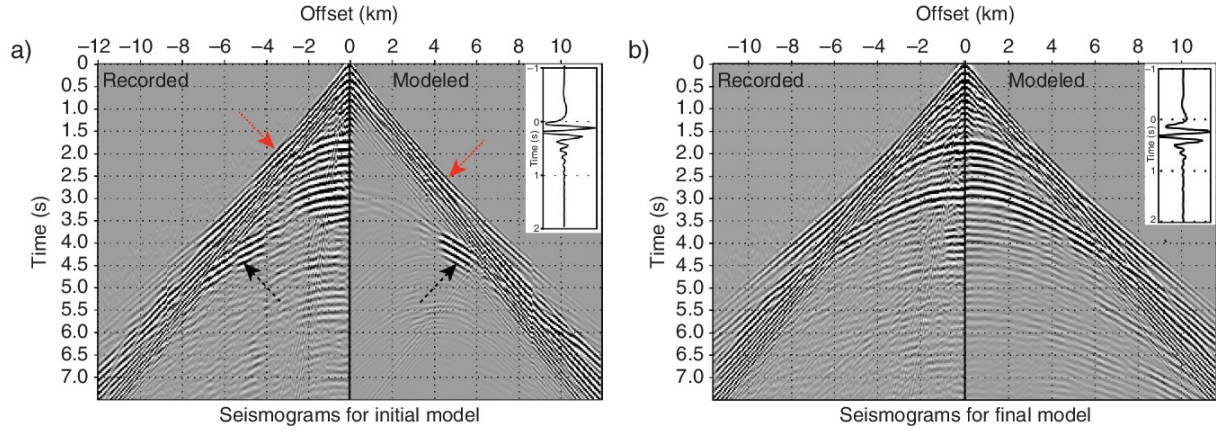
A FWI é uma técnica de inversão geofísica inicialmente proposta por Lailly e Bednar (1983) e Tarantola (1984) que parte de um modelo de velocidades intervalares inicial suave para gerar modelos de velocidades de alta resolução.

Este modelo de velocidades inicial é utilizado, juntamente com a modelagem por diferenças finitas da equação da onda, para gerar sismogramas sintéticos, similares aos dados obtidos em campo. O modelo de velocidades inicial é atualizado iterativamente até que a diferença entre os sismogramas sintéticos e o dado sísmico real seja minimizada. A Figura 11 exemplifica este processo.

O tempo para a construção de um modelo de velocidades intervalares utilizando a FWI é significativo, pode variar de três meses a um ano, a depender da complexidade da área e do tamanho da aquisição sísmica.

Em teoria, a modelagem utilizando FWI é capaz de gerar modelos de velocidade com alta resolução a partir de um modelo suave. Contudo, na prática, para garantir a convergência e reduzir o número de interações, é conveniente utilizar como entrada o modelo de velocidades oriundo da tomografia (VIGH et al., 2016). De acordo com

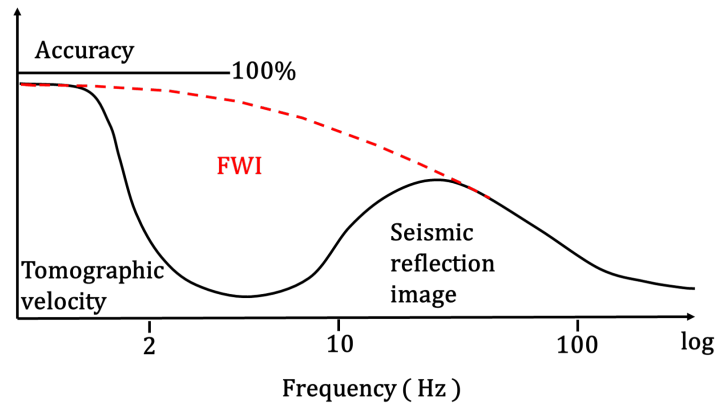
Figura 11 – Modelagem FWI: a) Sismograma real comparado com sismograma sintético gerado a partir do modelo de velocidades intervalares inicial b) Sismograma real comparado com sismograma sintético gerado a partir do modelo de velocidades intervalares final.



Fonte: (VIRIEUX et al., 2017).

Claerbout (1985), a FWI preenche a lacuna de informação entre o modelo de velocidades tomográfico, com baixo conteúdo de frequências, e o volume de amplitudes sísmicas, com alto conteúdo de frequências, de forma a aumentar a largura de banda útil do dado (Figura 12).

Figura 12 – Representação simplificada da lacuna de informação entre o modelo de velocidades oriundo da tomografia e o volume sísmico de amplitudes. O objetivo do FWI é preencher esta lacuna.



Fonte: Modificado de (CLAERBOUT, 1985).

O sinal oriundo da aquisição sísmica é registrado em tempo duplo de trânsito (*Two Way Time* - TWT), dessa forma, existe um raio de possibilidades para origem do sinal sísmico. Ele pode ter origem em diversos pontos em subsuperfície correspondente ao mesmo TWT. Para contornar esse problema e posicionar os eventos sísmicos em suas posições devidas, de modo a formar uma imagem da subsuperfície, aplica-se o processo de migração sísmica ao dado (YILMAZ, 2001).

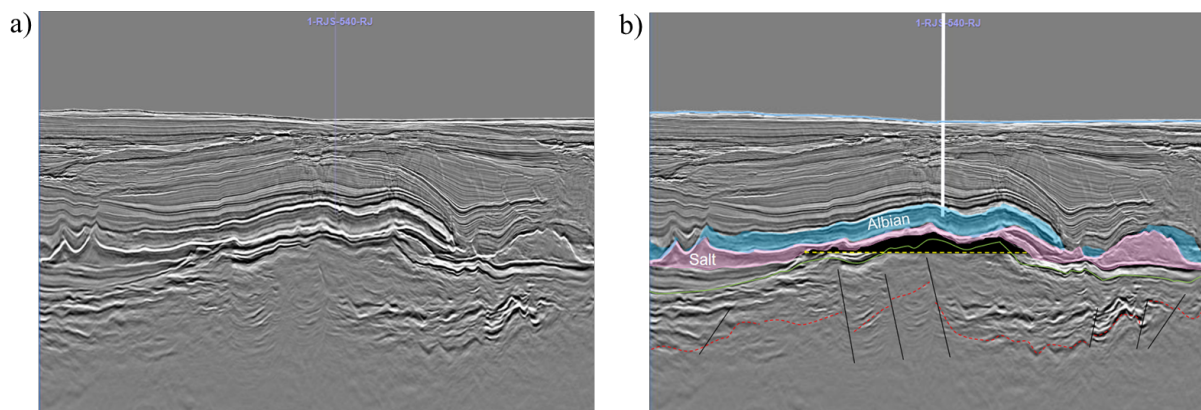
A migração sísmica pode ocorrer antes ou depois do empilhamento do dado sísmico, no domínio do tempo ou da profundidade (LINER, 1999). Ao classificar os algoritmos de migração por sua complexidade e qualidade dos resultados, temos que a mais simples é a Migração Kirchhoff pós-empilhamento no domínio do tempo, seguida pelas Migração Kirchhoff pré-empilhamento no domínio do tempo, Migração Kirchhoff pré-empilhamento no domínio da profundidade e por fim a Migração RTM (*Reverse Time Migration*) (ZHOU et al., 2018; YILMAZ, 2001).

Atualmente o padrão da indústria são as migrações Kirchhoff pré-empilhamento no domínio da profundidade e a Migração RTM (PEREIRA-DIAS et al., 2018; PENNA et al., 2013). Contudo esses algoritmos de migração são extremamente dependentes de um modelo de velocidades acurado, quanto melhor o modelo de velocidades, mais bem posicionados os eventos sísmicos vão estar e melhor vai ser a imagem sísmica (MAUL, 2020; MAUL et al., 2021).

Ao final do processo de migração e empilhamento sísmico, o resultado é o volume sísmico de amplitudes. Esse dado, juntamente com os poços presentes na região de estudo, permite o mapeamento sísmico das principais estruturas em subsuperfície, como interfaces entre camadas (horizontes sísmicos) e falhas geológicas no domínio da profundidade em um espaço tridimensional (SIMM; BACON; BACON, 2014). A Figura 13 apresenta a interpretação de uma seção sísmica.

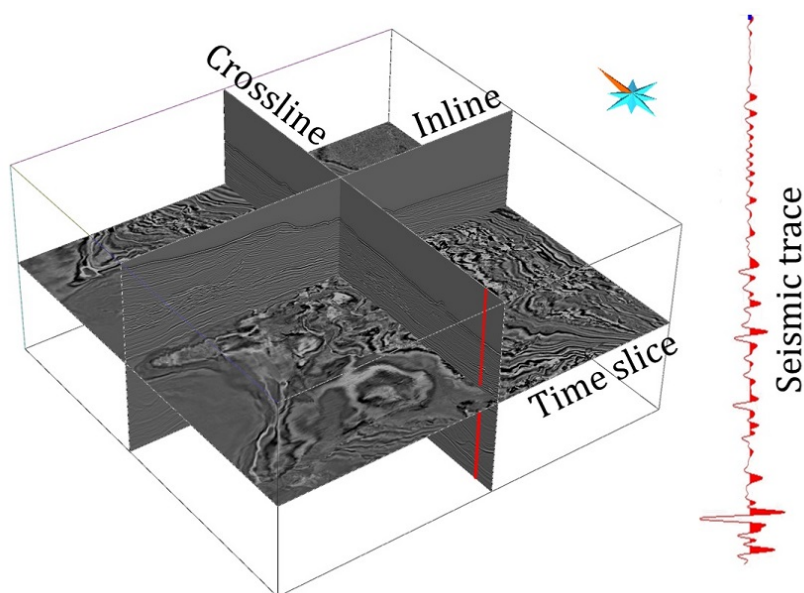
Para realizar o mapeamento sísmico, o interprete pode fatiar o dado de diversas formas. Um fatiamento na direção de aquisição é conhecido como seção *in-line*, perpendicular a direção de aquisição como *cross-line*, perpendicular ao eixo vertical como *depth-slice* se no domínio da profundidade ou *time-slice* se no domínio do tempo duplo de trânsito. Uma amostra vertical do volume de amplitudes é conhecida como traço sísmico (Figura 14).

Figura 13 – a) Seção sísmica arbitrária passando pelo poço 1-RJS-540-RJ da Bacia de Campos b) Mesma seção interpretada. Em azul os carbonatos albianos, em rosa a seção evaporítica (Sal) em preto o reservatório preenchido com óleo. Profundidade do reservatório em torno e 5 km a partir do nível do mar.



Fonte: Modificado de (ANP, 2020).

Figura 14 – Representação 3D do volume sísmico e suas principais direções de fatiamento: *in-line*, *cross-line* e *time-slice*. O traço sísmico é representado pela linha vermelha.



Os processos descritos aqui formam um resumo simplificado das principais etapas do processamento sísmico. Várias etapas foram omitidas de forma a permitir o foco na ideia central: obter uma imagem da subsuperfície que permita a interpretação das feições geométricas e a inferência dos parâmetros elásticos das rochas que compõem um sistema petrolífero.

2.2 Redes Neurais

A iniciativa de modelar matematicamente as redes neurais biológicas deu início a um novo campo de estudos, as redes neurais artificiais (MCCULLOCH; PITTS, 1943). Apesar da inspiração biológica inicial, as redes neurais artificiais se tornaram uma ferramenta com aplicações nas mais diversas áreas, desde processamento e reconhecimento de imagens (EGMONT-PETERSEN; RIDDER; HANDELS, 2002; KRIZHEVSKY; SUTSKEVER; HINTON, 2012), processamento de linguagem natural (GOLDBERG, 2017; DEVLIN et al., 2018), reconhecimento de voz (NASSIF et al., 2019), detecção de falhas em processos industriais (EREN, 2017) até como *surrogates* em simulações numéricas de fenômenos físicos complexos (ZHANG; LIU; SUN, 2020; RAISSI; PERDIKARIS; KARNIADAKIS, 2019; ZHU et al., 2019).

Será apresentada uma breve evolução histórica do estudo das redes neurais biológicas seguido de uma introdução aos modelos mais simples de redes neurais artificiais (modelos de McCulloch e Pitts (1943) e Rosenblatt (1958)), passando pelas redes neurais convolucionais (LECUN et al., 1998). Na sequência, será dada ênfase aos modelos gera-

dores (*generative models*), capazes de aprender a distribuição do conjunto de dados de treinamento e, a partir desta, gerar novos exemplos. Dentre as arquiteturas de modelos geradores, serão apresentadas as redes de Hopfield (HOPFIELD, 1982), Máquinas de Boltzmann, Máquinas de Boltzmann Restritas (HUEMBELI et al., 2022) e Redes Adversárias Geradoras (GOODFELLOW et al., 2014), do inglês Generative Adversarial Networks (GANs).

2.2.1 Redes neurais biológicas

O prêmio nobel da medicina de 1906 foi concedido à Camillo Golgi e Santiago Ramón y Cajal em reconhecimento ao estudo das estruturas do sistema nervoso realizados no final do século XIX (NOBEL PRIZE OUTREACH, 2022). Golgi desenvolveu o método de tingimento histológico utilizando nitrato de prata que permitiu a caracterização das células nervosas com o uso do microscópio óptico, inaugurando uma nova era da anatomia microscópica do sistema nervoso (DROUIN; PILOQUET; PÉREON, 2015).

Ramón y Cajal aprimorou a técnica de tingimento criada por Golgi e, por meio de suas observações, propôs que o sistema nervoso é composto por bilhões de células anatomicamente e funcionalmente distintas, chamadas de neurônios. Ele chegou a conclusão de que o fluxo de informação entre os neurônios se dava por meio de sinais elétricos modulados por descontinuidades existentes entre essas células (HORTSCH, 2009).

A descontinuidade entre os neurônios só foi verificada na década de 1950, com a possibilidade de utilização da microscopia eletrônica, na qual se constatou que a lacuna entre dois neurônios consecutivos é da ordem de 20 *nm*, bem abaixo da resolução máxima teórica do método de microscopia óptica, da ordem de 200 *nm*. Essa pequena descontinuidade entre neurônios, que permite a passagem de informação na forma de um sinal elétrico ou químico, é conhecida como sinapse (GRAY, 1959).

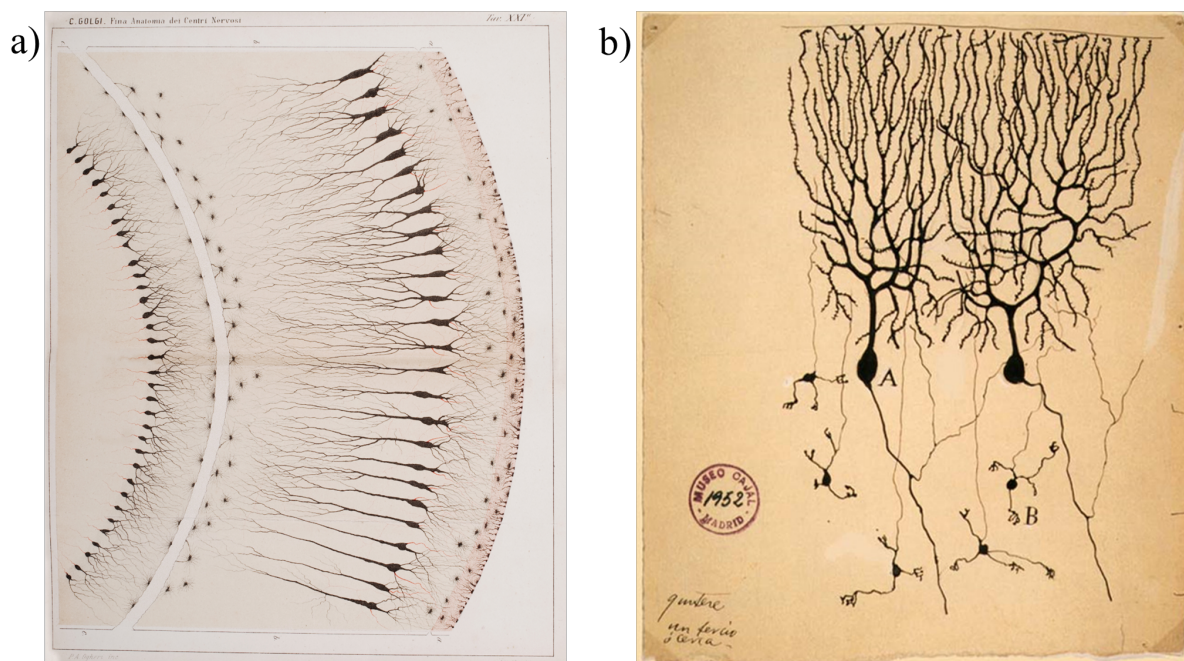
A Figura 15 apresenta exemplos das ilustrações científicas realizadas por Camillo Golgi e Santiago Ramón y Cajal no final do século XIX após observações microscópicas de estruturas do sistema nervoso central.

A diversidade morfológica dos neurônios é grande, eles podem variar em forma, tamanho e complexidade, contudo existem elementos comuns à todos os tipos de neurônios. Um neurônio típico é composto pelo *soma*, *dendritos* e *axônios* (MASLAND, 2004).

O *soma* é o corpo celular do neurônio, onde se localizam o núcleo e o citoplasma. *Dendritos* são prolongamentos celulares que atuam na recepção de estímulos de outros neurônios e na transmissão desse estímulo para o *soma* por meio de impulsos elétricos. O *axônio* é a parte do neurônio responsável pela condução dos impulsos elétricos que partem do *soma* em direção ao próximo neurônio (WNUK et al., 2018).

A transmissão de informação entre os neurônios é mediada por uma complexa

Figura 15 – a) Seção vertical pelo *pes Hippocampi major* de um coelho desenhada por Camillo Golgi b) Células de Purkinje e Células Granulosas de um cerebello de um pombo desenhadas por Santiago Ramón y Cajal.



Fonte: a) (DROUIN; PILOQUET; PÉREON, 2015) b) (RAMÓN Y CAJAL, 1899).

cadeia de processos bioquímicos, na qual neurônios eletricamente estimulados liberam moléculas de neurotransmissores nas fendas sinápticas. Os neurotransmissores liberados se ligam aos receptores de um neurônio diferente, desencadeando um novo estímulo elétrico após atingirem um certo patamar de diferença de potencial causado pela movimentação de íons induzida pelos neurotransmissores, viabilizando a transmissão de informação entre essas células (LUČÍČ, 2020), como ilustra a Figura 16.

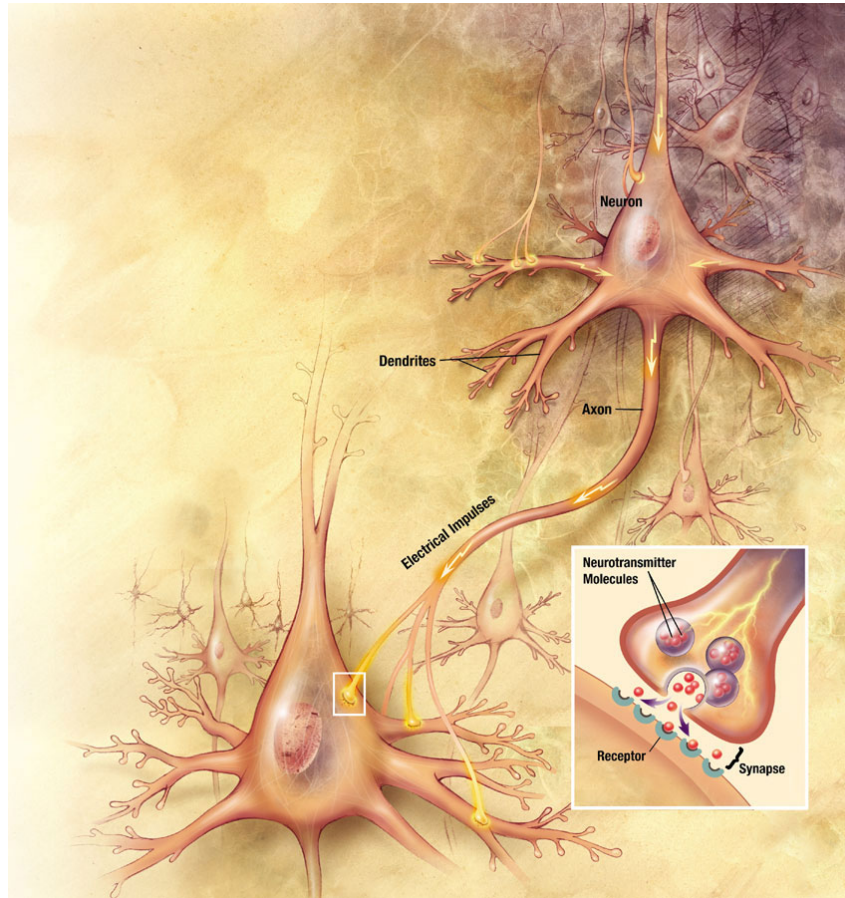
O processo evolutivo permitiu que certas regiões do cérebro se diferenciasssem de forma a se especializarem em funções específicas, como função motora, visão, olfato, cognição, etc., criando redes complexas de neurônios (WNUK et al., 2018).

Entender como o cérebro funciona é um desafio para a biologia moderna. Apesar da grande quantidade de dados coletados nos últimos anos ainda há muito a ser investigado para entender as regras que governam o processamento de informação e as relações entre estruturas e funções do cérebro. A modelagem matemática tem se mostrado uma ferramenta útil nesse sentido (SIDIROPOULOU; PISSADAKI; POIRAZI, 2006).

2.2.2 Redes neurais artificiais

Um dos trabalhos pioneiros a modelar matematicamente uma rede neural foi realizado por McCulloch e Pitts (1943). Este modelo considera o neurônio como um circuito elétrico composto por entradas $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \{0, 1\}^n$, um conjunto de pesos

Figura 16 – Ilustração da sinapse: processo de comunicação entre neurônios mediado por moléculas de neurotransmissores.



Fonte: (KRAMES, 2022).

$\mathbf{w} = (w_1, w_2, \dots, w_n) \in \{-1, 1\}^n$ responsáveis por classificar como excitatório ou inibitório os sinais da entrada, um limiar θ e uma saída y , dados pela Equação 2.11.

$$y = f(g(\mathbf{x})) = \begin{cases} 1, & \text{se } g(\mathbf{x}) \geq \theta \\ 0, & \text{se } g(\mathbf{x}) < \theta \end{cases} \quad (2.11)$$

Onde $g(\mathbf{x}) = \sum_{i=1}^n w_i x_i$ para um número n de entradas, $f(\cdot)$ é uma função degrau e θ é o limiar a partir do qual o sinal é ativado.

A Figura 17 apresenta um diagrama esquemático da modelagem matemática de McCulloch e Pitts (1943) com duas entradas e uma saída, análogo a um neurônio com dois *dendritos* e um *axônio*.

O neurônio de McCulloch e Pitts (1943) é capaz de replicar algumas funções lógicas linearmente separáveis como *E*, *OU* e *NÃO*, a depender do valor conferido a θ . As tabelas 1 e 2 apresentam o resultado obtido utilizando o neurônio artificial, com os pesos $w_n = 1$ (apenas sinais excitatórios), para as funções *E* e *OU*.

Apesar de suas limitações, como por exemplo, só trabalhar com entradas e saídas

Figura 17 – Modelo de McCulloch e Pitts (1943) com duas entradas e uma saída, análogo a um neurônio com dois *dendritos* e um *axônio*.

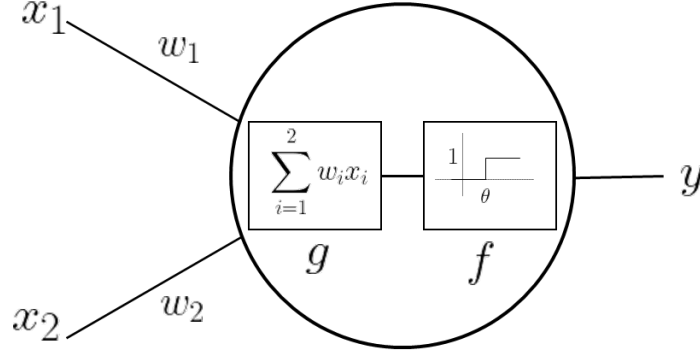


Tabela 1 – Neurônio modelando a função lógica E .

x_1	x_2	$x_1 \wedge x_2$	θ	$\sum_{i=1}^n w_i x_i$	y
0	0	0	2	0	0
0	1	0	2	1	0
1	0	0	2	1	0
1	1	1	2	2	1

Tabela 2 – Neurônio modelando a função lógica OU .

x_1	x_2	$x_1 \vee x_2$	θ	$\sum_{i=1}^n w_i x_i$	y
0	0	0	1	0	0
0	1	1	1	1	1
1	0	1	1	1	1
1	1	1	1	2	1

binárias, possuir somente soluções lineares, necessitar de parametrização manual e carecer de um algoritmo de aprendizagem, o modelo de McCulloch e Pitts (1943) foi uma importante contribuição para o campo das redes neurais artificiais. Este modelo já apresentava alguns dos elementos básicos das redes neurais artificiais mais modernas, como o conceito de pesos e de função de ativação.

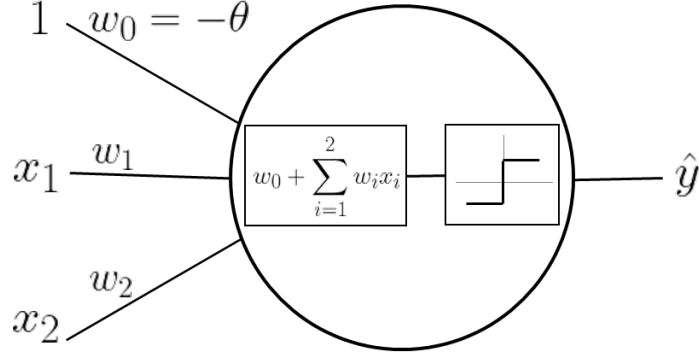
A primeira rede neural artificial a utilizar um algoritmo de aprendizagem recebeu o nome de *perceptron* e foi proposta por Rosenblatt (1958).

A rede *perceptron* não exige uma programação explícita para realizar uma tarefa. Os parâmetros mais adequados são encontrados durante o processo de treinamento, que utiliza pares de entradas ($\mathbf{x}$) e saídas desejadas (y) de forma a encontrar uma regra geral para mapear $\mathbf{x} \rightarrow y$. Este tipo de treinamento, onde existem pares de entradas e saídas desejadas, é conhecido como aprendizagem supervisionada.

A rede *perceptron* é composta por uma camada com entradas $\mathbf{x} \in \mathbb{R}^n$, um conjunto de pesos $\mathbf{w} \in \mathbb{R}^n$, um viés $\theta \in \mathbb{R}$, uma função de ativação do tipo sinal $\sigma = \text{sign}(\cdot)$ e uma camada de saída $\hat{y} \in \{-1, 1\}$, onde n é o número de variáveis do vetor de entrada $\mathbf{x}$. É possível adicionar o viés ao vetor de pesos, para isto basta criar uma nova entrada unitária $x_0 = 1$ e um termo $w_0 = -\theta$. A Figura 18 apresenta o diagrama esquemático de

um *perceptron*. A saída é dada por $\hat{y} = \sigma(\mathbf{w}^T \cdot \mathbf{x})$ (BISHOP; NASRABADI, 2007).

Figura 18 – Modelo de Rosenblatt (1958), *perceptron*, com duas entradas e uma saída.



O algoritmo de aprendizagem do *perceptron* é iterativo, de início todos os pesos recebem valores iguais a zero. Para um conjunto de dados $X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i), \mathbf{y} = (y_1, y_2, \dots, y_i)$ de tamanho i , uma amostra $(\mathbf{x}, y)$ é escolhida aleatoriamente e classificadas pelo *perceptron*, retornando um valor de $\hat{y}$. Caso $\hat{y} \neq y$ os valores dos parâmetros são atualizados de acordo com a seguinte regra: $\mathbf{w} \leftarrow \mathbf{w} + \eta(y - \hat{y})\mathbf{x}$, onde η é uma constante de ponderação conhecida como taxa de aprendizagem. Esse processo é repetido até que não exista nenhuma amostra classificada incorretamente ou o número máximo de iterações previamente definido seja atingido (Algoritmo 1).

O treinamento do *perceptron* converge para problemas de classificação com separação linear entre as classes. Caso o problema não seja linearmente separável, não haverá convergência (MINSKY; PAPERT, 1969).

Algoritmo 1: Regra de aprendizagem do *perceptron* para problemas linearmente separáveis

Entrada: $\mathbf{x}, y$
Saída: $\hat{y}$

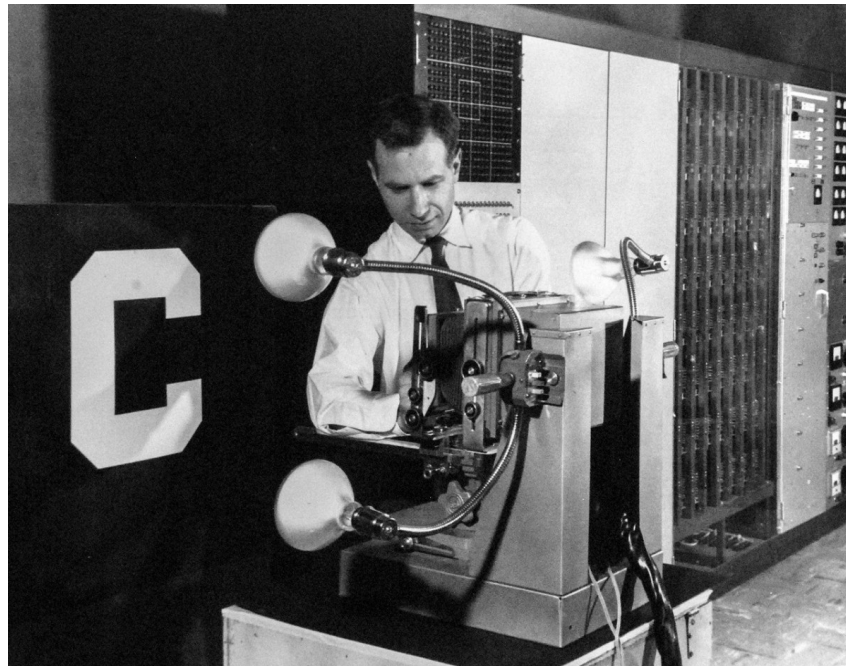
```

1 início
2    $\mathbf{w} = \{w_0, w_1, \dots, w_n\}$ 
3   enquanto  $\sum_{k=1}^i (y_k - \hat{y}_k) \neq 0$  faça
4     sorteie um par  $\mathbf{x}_k, y_k$ 
5      $\hat{y} = \text{sign}(\mathbf{w}^T \cdot \mathbf{x}_k)$ 
6      $\mathbf{w} \leftarrow \mathbf{w} + \eta(y_k - \hat{y}_k)\mathbf{x}$ 
7   fim
8 fim
9 retorna  $\mathbf{w}$ 
```

O *perceptron* foi originalmente implementado na forma de *software*, em um IBM 704, um computador do tamanho de uma sala com peso em torno de cinco toneladas.

Posteriormente foi implementado na forma de *hardware*, o MARK I, que utilizava como sensores de entrada uma matriz de 20×20 foto-sensores de sulfeto de cádmio, equivalente a uma imagem de 400 pixels (Figura 19). Os pesos foram implementados por um conjunto de potenciômetros, ajustados por motores elétricos de forma a permitir o ajuste durante o processo de treinamento (BISHOP; NASRABADI, 2007).

Figura 19 – Frank Rosenblatt operando o *perceptron* MARK I.



Fonte: (NATIONAL MUSEUM OF THE U.S. NAVY, 1960).

Apesar de revolucionário, por ser a primeira rede neural a utilizar um mecanismo de aprendizado, o *perceptron* fica limitado a resolver problemas com separação linear entre classes, pois trabalha somente com combinações lineares de pesos e entradas, $\mathbf{w}^T \cdot \mathbf{x}$. As limitações do *perceptron* foram detalhadamente descritas no trabalho de Minsky e Papert (1969) onde é demonstrado que um *perceptron* não é capaz de resolver uma função lógica básica, o OU exclusivo (XOU).

O trabalho de Rosenblatt (1958) já explicitava esta limitação do *perceptron* e propôs uma solução utilizando *perceptrons* multicamadas para problemas com separação não lineares, contudo não existia à época algoritmo de treinamento para esse tipo de rede multicamadas.

Na segunda metade da década de 80, Rumelhart, Hinton e Williams (1986) deram visibilidade ao algoritmo de retro-propagação do erro ao utilizá-lo para treinar um *perceptron* com camada oculta. Esse tipo de algoritmo possibilita o cálculo do gradiente de uma função de custo $C(y, \hat{y})$ em relação a cada parâmetro de uma rede neural, permitindo a minimização da função de custo utilizando métodos de otimização por gradiente descendente.

Os requisitos para aplicação do algoritmo de retro-propagação do erro são os

seguintes: as função de perda $C(\cdot)$ e as funções de ativação $\sigma(\cdot)$ utilizadas na rede neural devem ser numericamente diferenciáveis; a função de perda total possa ser escrita como uma média da função de perda individual dos exemplos de treinamento $C = \frac{1}{n} \sum_x C_x$ (NIELSEN, 2019).

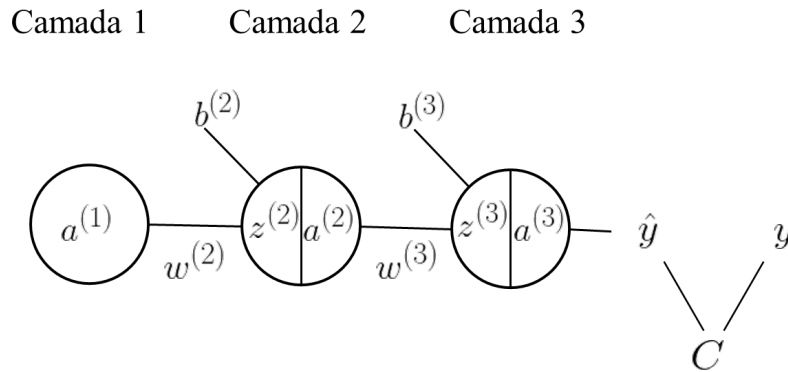
No trabalho de Rumelhart, Hinton e Williams (1986) foi utilizado como função de perda a Equação 2.12 (Erro Quadrático Médio) e como função de ativação a Equação 2.13 (Função Sigmóide), onde $\hat{y}$ é saída da rede e y o resultado desejado.

$$C = \frac{1}{2} \sum_{k=1}^n (\mathbf{y} - \hat{\mathbf{y}})^2 \quad (2.12)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.13)$$

A Figura 20 apresenta uma rede neural com apenas uma camada oculta e um neurônio em cada camada, cuja finalidade é exemplificar o mecanismo de retro-propagação dos erros. Têm-se que $a^{(1)}$ é a camada de entrada, $z^{(2)} = w^{(2)}a^{(1)} + b^{(2)}$, $a^{(2)} = \sigma(z^{(2)})$, $z^{(3)} = w^{(3)}a^{(2)} + b^{(3)}$, $a^{(3)} = \sigma(z^{(3)})$, $\hat{y}$ é a saída da rede, y a saída desejada e C a função de perda.

Figura 20 – Rede neural com uma camada oculta e um neurônio em cada camada, $a^{(1)}$ é a camada de entrada, $z^{(2)} = w^{(2)}a^{(1)} + b^{(2)}$, $a^{(2)} = \sigma(z^{(2)})$, $z^{(3)} = w^{(3)}a^{(2)} + b^{(3)}$, $a^{(3)} = \sigma(z^{(3)})$, $\hat{y}$ é a saída da rede, y a saída desejada e C a função de perda.



Para encontrar o quanto C varia com $w^{(3)}$ e $b^{(3)}$ é preciso derivar C em relação a $w^{(3)}$ e em relação a $b^{(3)}$. O caminho na rede de $w^{(3)}$ até C e de $b^{(3)}$ até C é uma função composta, $C(a_2(z_2(w^{(3)}, b^{(3)})))$, então, pela regra da cadeia temos:

$$\frac{\partial C}{\partial w^{(3)}} = \frac{\partial z^{(3)}}{\partial w^{(3)}} \frac{\partial a^{(3)}}{\partial z^{(3)}} \frac{\partial C}{\partial a^{(3)}} \quad (2.14)$$

$$\frac{\partial C}{\partial b^{(3)}} = \frac{\partial z^{(3)}}{\partial b^{(3)}} \frac{\partial a^{(3)}}{\partial z^{(3)}} \frac{\partial C}{\partial a^{(3)}} \quad (2.15)$$

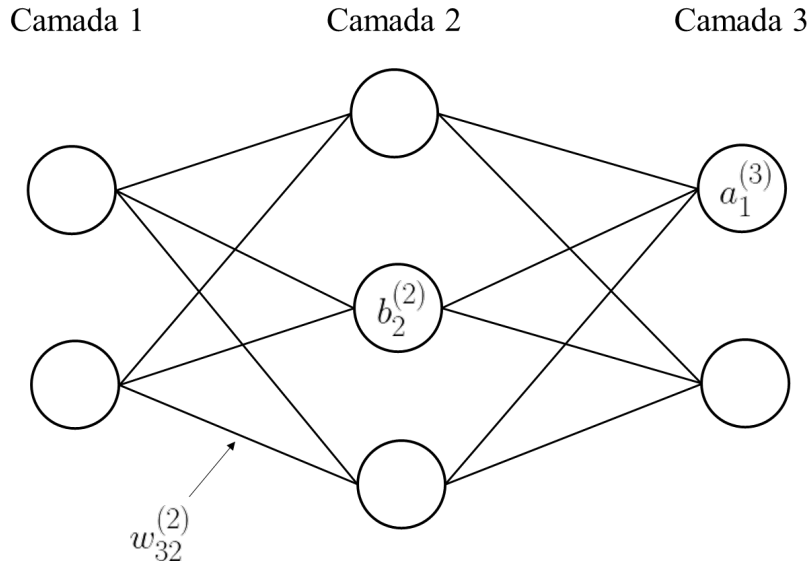
Analogamente, para encontrar o gradiente de C em relação a $w^{(2)}$ e $b^{(2)}$, temos:

$$\frac{\partial C}{\partial w^{(2)}} = \frac{\partial z^{(2)}}{\partial w^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial w^{(3)}}{\partial a^{(2)}} \frac{\partial z^{(3)}}{\partial w^{(3)}} \frac{\partial a^{(3)}}{\partial z^{(3)}} \frac{\partial C}{\partial a^{(3)}} \quad (2.16)$$

$$\frac{\partial C}{\partial b^{(2)}} = \frac{\partial z^{(2)}}{\partial b^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial w^{(3)}}{\partial a^{(2)}} \frac{\partial z^{(3)}}{\partial w^{(3)}} \frac{\partial a^{(3)}}{\partial z^{(3)}} \frac{\partial C}{\partial a^{(3)}} \quad (2.17)$$

Estendendo o raciocínio para uma rede neural com mais de um neurônio por camada, temos a seguinte indexação, de acordo com Nielsen (2019): $w_{jk}^{(l)}$, $b_j^{(l)}$ e $a_j^{(l)}$, onde l é o índice da camada, k a posição do neurônio na camada $(l-1)$, j a posição do neurônio na camada l . Dessa forma, como exemplos de indexação, temos que $w_{32}^{(2)}$ é o peso que conecta o segundo neurônio da primeira camada ao terceiro neurônio da segunda camada, $b_2^{(2)}$ é o viés do segundo neurônio da segunda camada e $a_1^{(3)}$ é a ativação referente ao primeiro neurônio da terceira camada (Figura 21).

Figura 21 – Exemplos de indexação dos parâmetros em uma rede com camada oculta.



Fonte: Modificado de Nielsen (2019).

Os valores intermediários, z , e as ativações, a , serão dados por:

$$z_j^{(l)} = \sum_k w_{jk}^{(l)} a_k^{(l-1)} + b_j^{(l)} \quad (2.18)$$

$$a_j^{(l)} = \sigma(\sum_k w_{jk}^{(l)} a_k^{(l-1)} + b_j^{(l)}) \quad (2.19)$$

Com o objetivo de escrever os gradientes $\frac{\partial C}{\partial w_{jk}^{(l)}}$ e $\frac{\partial C}{\partial b_j^{(l)}}$ na forma vetorial, se torna conveniente definir a grandeza $\delta_j^{(l)}$, dada pela Equação 2.20.

$$\delta_j^{(l)} \equiv \frac{\partial C}{\partial z_j^{(l)}} \quad (2.20)$$

Pode-se interpretar $\delta_j^{(l)}$ como a parcela da função de custo no neurônio j da camada l . O vetor $\boldsymbol{\delta}^{(l)}$ representa o conjunto dos δ de todos os neurônios dessa camada. Para a última camada, L , o valor $\delta_j^{(L)}$ será dado pela Equação 2.21.

$$\delta_j^{(L)} = \frac{\partial C}{\partial a_j^{(L)}} \sigma'(z_j^{(L)}) \quad (2.21)$$

Na forma matricial, temos que $\boldsymbol{\delta}^{(L)}$ é dado pela Equação 2.22.

$$\boldsymbol{\delta}^{(L)} = \nabla_a C \odot \sigma'(\mathbf{z}^{(L)}) \quad (2.22)$$

Ao escolher como função de custo o Erro Quadrático Médio, dado pela Equação 2.12, temos que $\frac{\partial C}{\partial a_j^{(L)}} = (a_j^{(L)} - y_j)$ e seu equivalente na forma matricial $\nabla_a C = (\mathbf{a}^{(L)} - \mathbf{y})$, o que permite a reescrita da Equação 2.22 como:

$$\boldsymbol{\delta}^{(L)} = (\mathbf{a}^{(L)} - \mathbf{y}) \odot \sigma'(\mathbf{z}^{(L)}) \quad (2.23)$$

Para as outras camadas da rede neural, a Equação 2.20 pode ser escrita na forma matricial como:

$$\boldsymbol{\delta}^{(l)} = ((\mathbf{w}^{(l+1)})^T \boldsymbol{\delta}^{(l+1)}) \odot \sigma'(\mathbf{z}^{(l)}) \quad (2.24)$$

Logo, conhecendo $\boldsymbol{\delta}^{(L)}$ na última camada L , é possível calcular $\boldsymbol{\delta}^{(L-1)}$ da camada anterior por meio da Equação 2.24 e recursivamente os $\boldsymbol{\delta}^{(L-2)}$, $\boldsymbol{\delta}^{(L-3)}$, ..., $\boldsymbol{\delta}^{(2)}$, das camadas anteriores, retro-propagando esse cálculo por todas as camadas da rede neural.

Uma vez calculado δ é possível encontrar o gradiente de C para qualquer peso w ou viés b da rede neural por meio das Equações 2.25 e 2.26.

Ao contrário do *perceptron* simples, sem camadas intermediárias, onde existia prova de convergência para problemas linearmente separáveis, o *perceptron* multicamadas possui uma função de perda não convexa, ou seja, não há garantia de que o algoritmo de gradiente descendente vá encontrar um mínimo global (RUMELHART; HINTON; WILLIAMS, 1986).

Rumelhart, Hinton e Williams (1986) propuseram o método de descida de gradiente no qual o gradiente é acumulado para todo o conjunto de treinamento (x, y) . Posteriormente os parâmetros da rede são atualizados proporcionalmente ao gradiente acumulado. $\Delta w = -\eta \frac{\partial C}{\partial w}$.

$$\frac{\partial C}{\partial w_{jk}^{(l)}} = a_k^{(l-1)} \delta_j^{(l)} \quad (2.25)$$

$$\frac{\partial C}{\partial b_j^{(l)}} = \delta_j^{(l)} \quad (2.26)$$

Utilizando as Equações 2.22, 2.24, 2.25 e 2.26 é possível implementar um algoritmo para o cálculo da retro-propagação dos erros e treinamento por descida de gradiente.

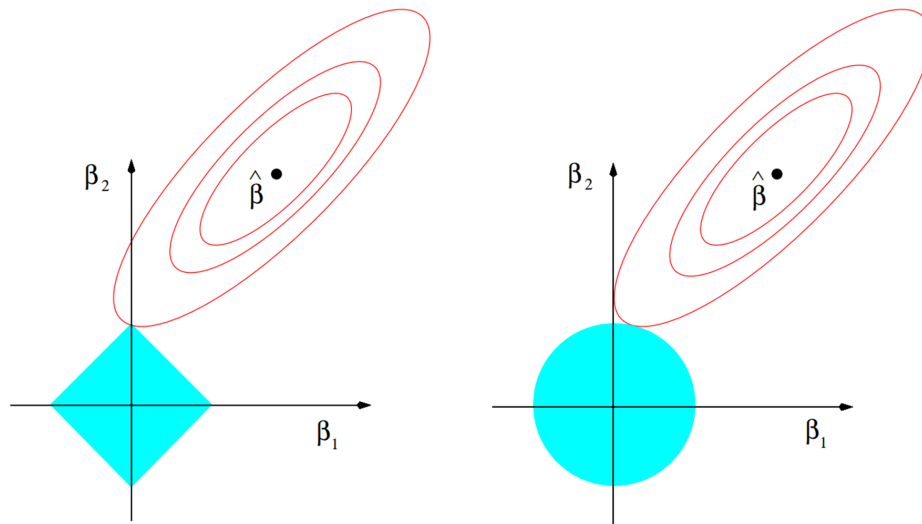
De acordo com Goodfellow, Bengio e Courville (2016) um problema central no campo do aprendizado de máquinas é fazer com que os algoritmos propostos tenham um bom desempenho tanto para o dado de treino quanto para novos dados de entrada. A capacidade de um modelo produzir bons resultados para dados não utilizados durante o treinamento é chamada de generalização.

Quanto mais parâmetros uma rede neural possuir, maior a sua flexibilidade para realizar aproximações de funções complexas. Contudo, quanto maior o número de parâmetros, maior a demanda por complexidade nos dados de treinamento. Caso contrário, a rede

tenderá a memorizar o dado de entrada em vez de encontrar uma regra representativa do mesmo, resultando em uma baixa performance ao lidar com novos dados (*overfitting*).

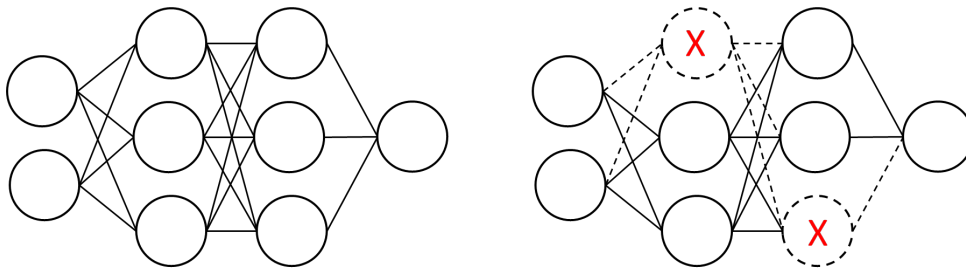
A regularização é um procedimento que permite o treinamento de uma rede neural com muitos parâmetros de forma a evitar o *overfitting*. A regularização pode ser realizada de forma explícita, como por exemplo adicionando uma penalidade à função de perda proporcional aos parâmetros da rede (ex: Regularização L_1 e L_2 (Figura 22)) ou de forma implícita, como a exemplo do *dropout* (Figura 23), quando uma porcentagem dos parâmetros da rede é ignorada durante o treinamento.

Figura 22 – Regularização L_1 e L_2 . Curvas de nível em vermelho representa a função de custo e as regiões em azul as restrições de valores para os parâmetros impostas pela regularização.



Fonte: Modificado de [Hastie, Tibshirani e Friedman \(2001\)](#).

Figura 23 – *Dropout* aplicado em um *perceptron*. Os neurônios em tracejado foram escolhidos aleatoriamente para serem ignorados durante uma época do treinamento.



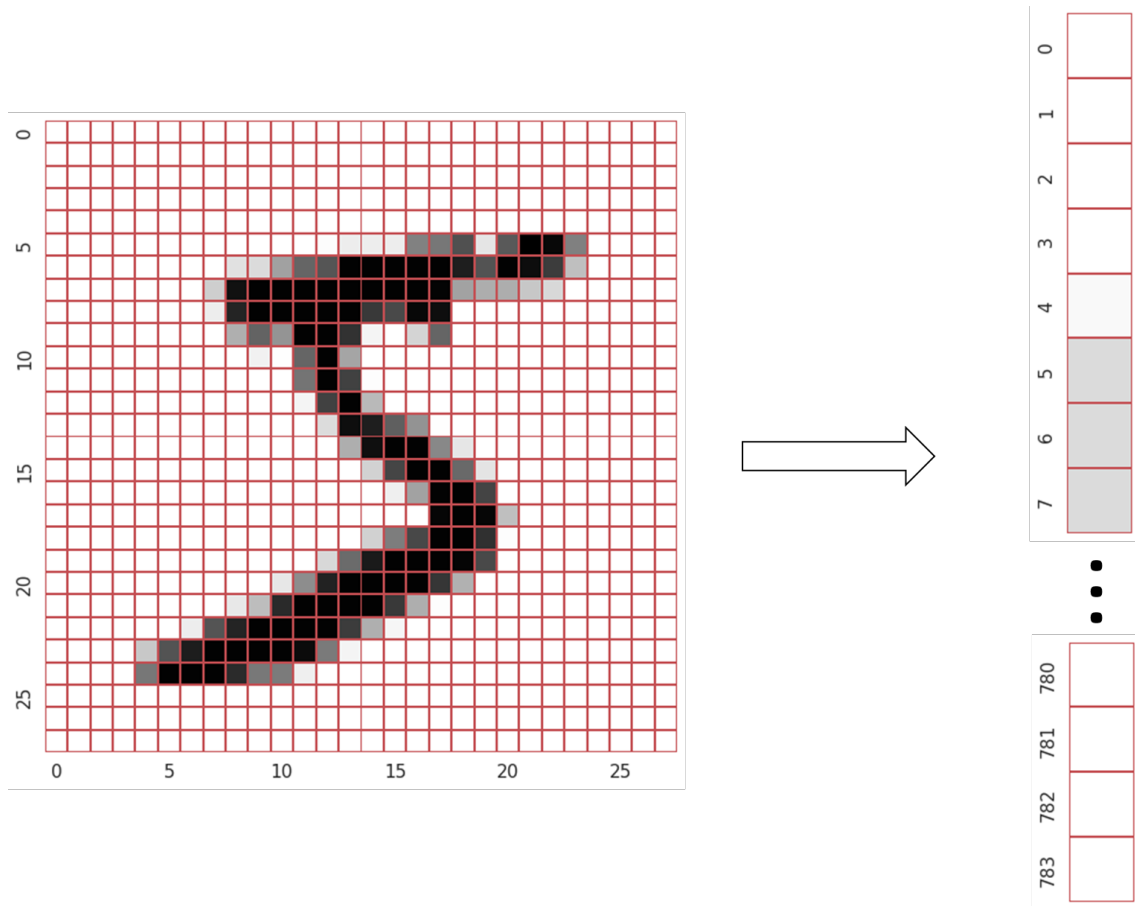
2.2.3 Redes Neurais Convolucionais

O uso do *perceptron* Multicamadas como uma ferramenta para classificação ou regressão de dados tabulares é consolidado na literatura científica ([KADRA et al., 2021](#)), todavia é possível utilizá-lo para realizar classificação de imagens, bastando reorganizar o

dado bi-dimensional em um vetor (Figura 24). Entretanto, os *perceptrons* multicamadas são sensíveis à translação, ou seja, um mesmo objeto em posição diferente no *grid* da imagem, é percebido de forma diferente pela rede (LECUN et al., 1998).

Outra implicação da reorganização de dados bi-dimensionais em vetores é a explosão do tamanho do vetor de entrada com o aumento da dimensionalidade da imagem, e por consequência do número de parâmetros do *perceptron* multicamadas. A Figura 24 apresenta uma imagem com apenas um canal e dimensões de 28×28 pixels, o que equivale a um vetor de tamanho 784. Caso a primeira camada oculta do *perceptron* tenha, por exemplo, 64 neurônios, existiriam 47.936 parâmetros, 47.872 pesos e 64 vieses, uma quantidade elevada, mas ainda viável computacionalmente.

Figura 24 – Imagem de um número escrito à mão com dimensões de 28×28 pixels reorganizado na forma de um vetor de tamanho $28 \times 28 = 784$.



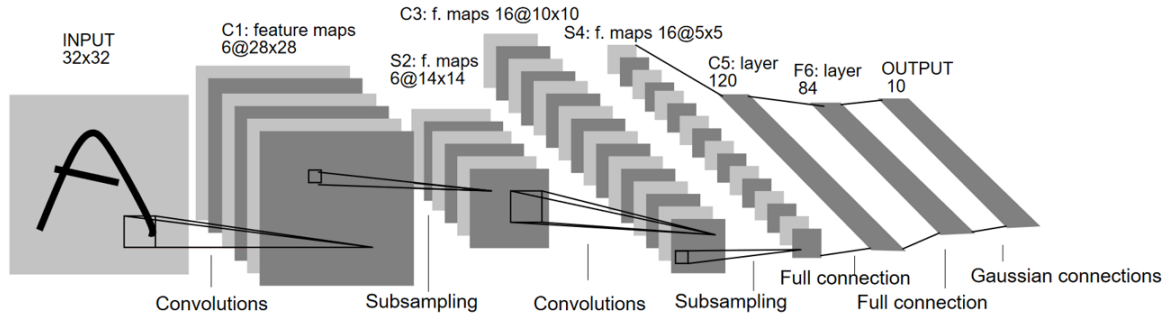
Já uma imagem de 256×256 pixels com três canais vai gerar um vetor de entrada com tamanho $256 \times 256 \times 3 = 196.608$. Para a mesma camada oculta acima existiriam 12.582.977 parâmetros e muito provavelmente seria necessário ampliar o número de neurônios das camadas ocultas com o objetivo de aumentar a capacidade da rede, tornando-a computacionalmente ineficiente.

Diante destas adversidades, LeCun et al. (1998) propuseram a utilização de redes neurais convolucionais. Este tipo de rede consegue capturar as dependências espaciais do

dado e é tolerante à translação, distorção e variações de escala.

As redes neurais convolucionais são compostas por bancos sequenciais de filtros intercalados com outros mecanismos de redução de dimensão da imagem e uma rede classificadora, semelhante ao *perceptron*, ao final. A Figura 25 apresenta a Lenet, rede neural convolucional proposta por LeCun et al. (1998).

Figura 25 – Lenet (LECUN et al., 1998), trabalho pioneiro na utilização de Redes Neurais Convolucionais para classificação de números escritos a mão.



Fonte: (LECUN et al., 1998).

A convolução é uma operação matemática de ponderação entre duas funções, sejam elas discretas ou contínuas. Para o caso de funções unidimensionais, esta operação é realizada da seguinte maneira: uma das funções é espelhada em relação ao eixo y e transladada em direção ao final da outra função. A cada passo da translação é realizada a multiplicação entre as duas funções e o resultado é integrado, no caso contínuo, ou somado, no caso discreto. A convolução entre duas funções é dada pelas equações 2.27 e 2.28 em seu formato contínuo e discreto, respectivamente.

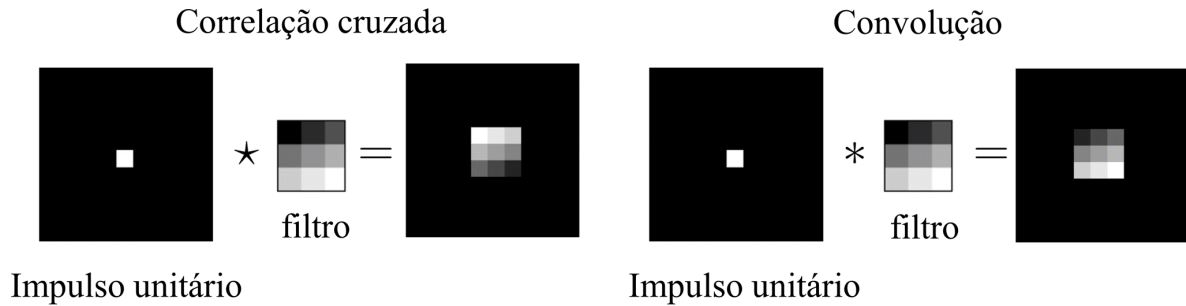
$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau \quad (2.27) \quad (f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n - m] \quad (2.28)$$

Este mesmo processo pode ser estendido para mais dimensões, as mais usuais ao lidar com imagens são as convoluções bi-dimensionais e com dados volumétricos convoluções tri-dimensionais.

No contexto do processamento de imagens, a maioria dos filtros convolucionais são simétricos, portanto espelhar o sinal não produz nenhum resultado. Nessas situações, para sinais reais, a convolução é equivalente à operação de correlação cruzada, tanto no caso contínuo quanto no discreto (Equações 2.29 e 2.30). Mesmo para filtros não simétricos, o resultado obtido seria equivalente ao da convolução, contudo refletido, portanto, operacionalmente, este procedimento não se justifica (Figura 26).

$$(f \star g)(\tau) = \int_{-\infty}^{\infty} f(t)g(t + \tau)dt \quad (2.29) \quad (f \star g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[m + n] \quad (2.30)$$

Figura 26 – Operação de correlação cruzada de um impulso unitário com um filtro à esquerda e convolução com o mesmo filtro à direita. A correlação cruzada retorna o mesmo resultado da convolução, contudo refletido em torno dos eixos x e y .

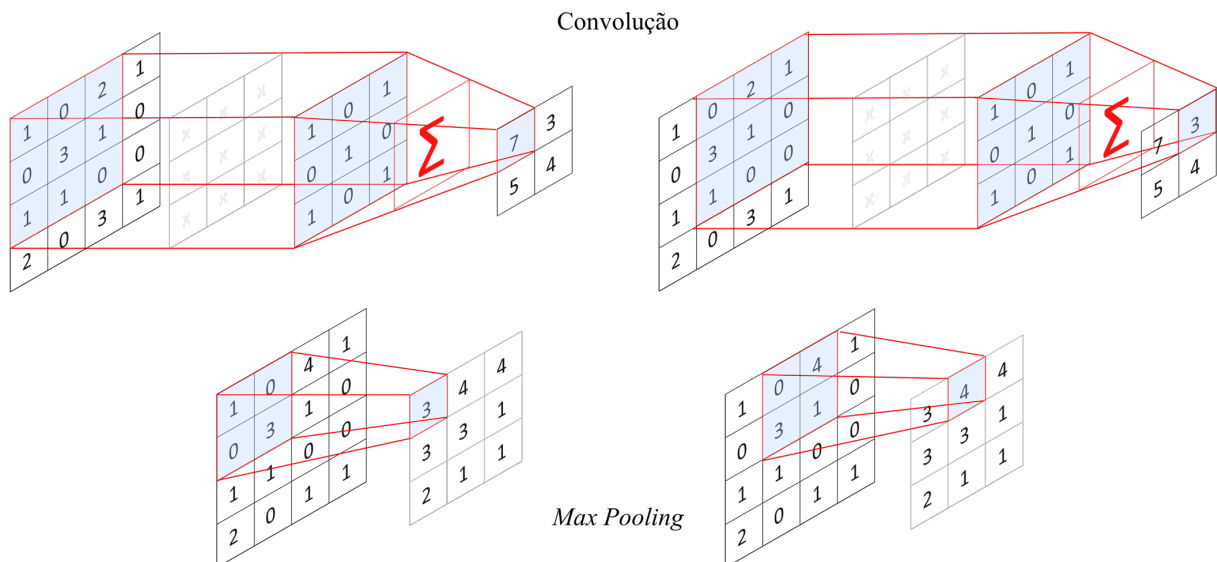


Fonte: Modificado de (ESSA, 2014).

A prática de não reflexão dos filtros foi adotada no âmbito das redes neurais e, apesar do nome, as redes neurais convolucionais realizam a operação matemática de correlação cruzada.

Neste trabalho será adotado o termo convolução para operação de correlação cruzada, sem a reflexão do filtro, como é feito no contexto das redes neurais pela comunidade científica (Figura 27).

Figura 27 – Processo de convolução 2D e *Max Pooling* no contexto das redes neurais convolucionais.



Para reduzir os efeitos de borda durante a convolução, pode-se adicionar pixels ao redor da imagem (*padding*). Opções comumente utilizadas são a reflexão das bordas ou o preenchimento com valores nulos desses pixels adicionais.

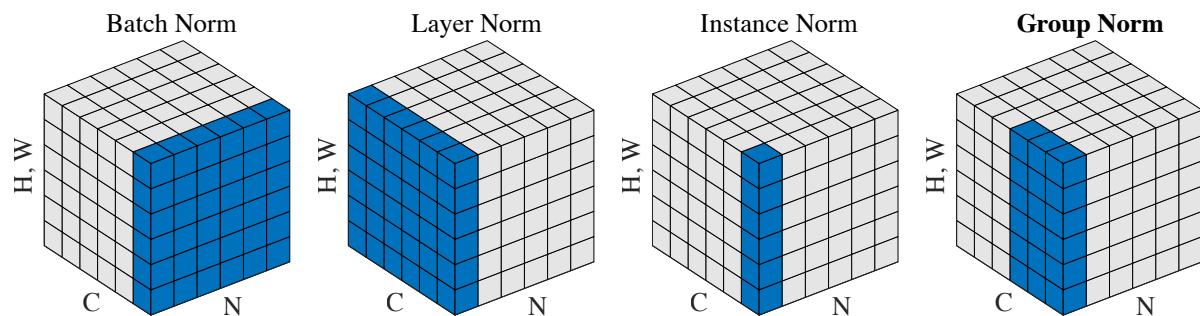
Outra operação necessária de ser definida para a compreensão do funcionamento de uma rede neural convolucional é o *pooling*. Esta operação tem como objetivo reduzir

a dimensão da imagem e, ao mesmo tempo, preservar as informações mais relevantes. A operação de *pooling* consiste em uma janela móvel que desliza sobre a entrada e retorna o maior valor, no caso do *max pooling*, (Figura 27) ou a média aritmética dos valores da janela, no caso do *average pooling*.

Um importante hiper-parâmetro das redes neurais convolucionais é o passo (*stride*) que a janela móvel executa durante as operações de convolução ou *pooling*. Quanto maior o *stride* menor será a dimensão espacial do resultado.

Durante o treinamento de uma rede neural convolucional, o resultado das ativações dos filtros ao longo da rede podem apresentar valores absolutos muito elevados, anulando a contribuição das outras ativações. Da mesma forma que é conveniente normalizar os valores da entrada da rede para que as variáveis possam ter a mesma escala, é interessante normalizar as ativações dos filtros intermediários de uma rede neural convolucional durante o treinamento. As normalizações mais utilizadas são a *batch normalization*, *layer normalization*, *instance normalization* e *group normalization* (Figura 28).

Figura 28 – Tipos de normalizações no contexto das redes neurais convolucionais. N é o eixo do *batches*, C dos canais, e W,H os eixos espaciais.



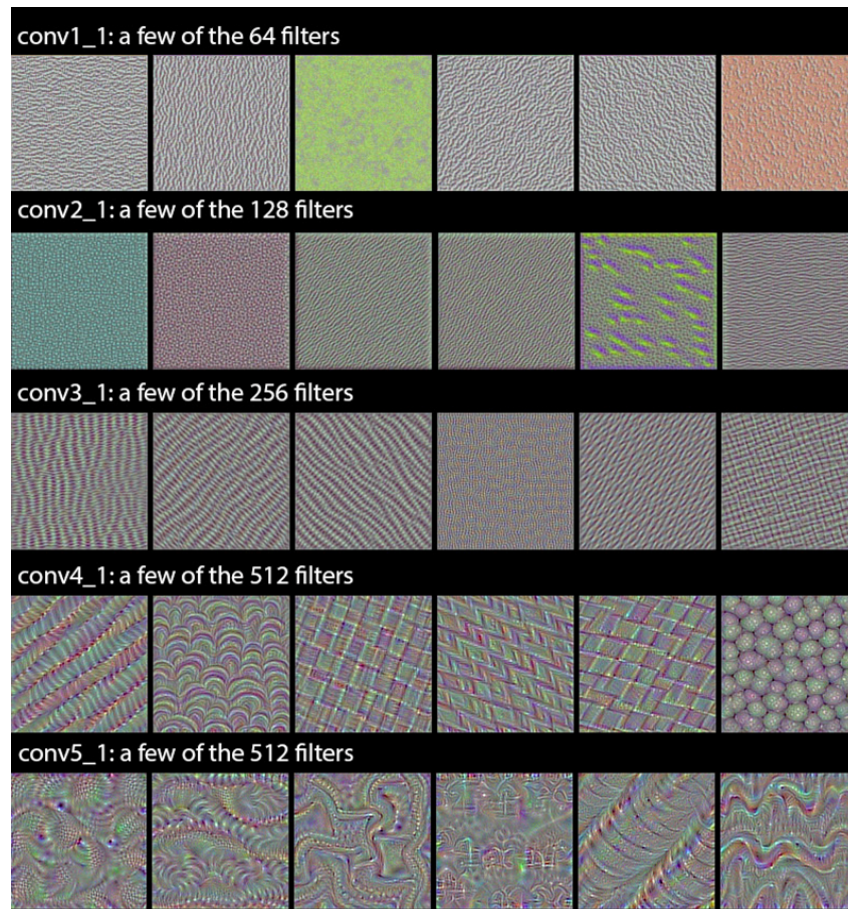
Fonte: (WU; HE, 2019).

O cálculo do gradiente do erro, utilizado durante o treinamento, é realizado por meio da retro propagação do erro. Ao final do processo, por meio de algoritmos de otimização por decida de gradiente, a rede aprende quais filtros convolucionais minimizam o erro entre o resultado previsto e o resultado real.

A Figura 29 apresenta a visualização de alguns filtros da rede VGG16 (SIMONYAN; ZISSERMAN, 2015), nota-se que a complexidade dos filtros vai aumentando no sentido do final da rede. Um padrão semelhante de hierarquização de complexidade foi encontrado nas células do córtex visual do cérebro, sugerindo que um entendimento melhor do processo de visão biológica pode levar a avanços na área de Redes Neurais Convolucionais e vice-versa (LINDSAY, 2021) .

Dentro das arquiteturas de redes neurais convolucionais, as redes auto-codificadoras (*autoencoders*) se destacam pela capacidade de projetar a informação de entrada de forma não linear em um espaço de dimensão reduzida, o chamado espaço latente (WANG;

Figura 29 – Vizualização de alguns filtros da Rede Neural Convolutacional VGG16.



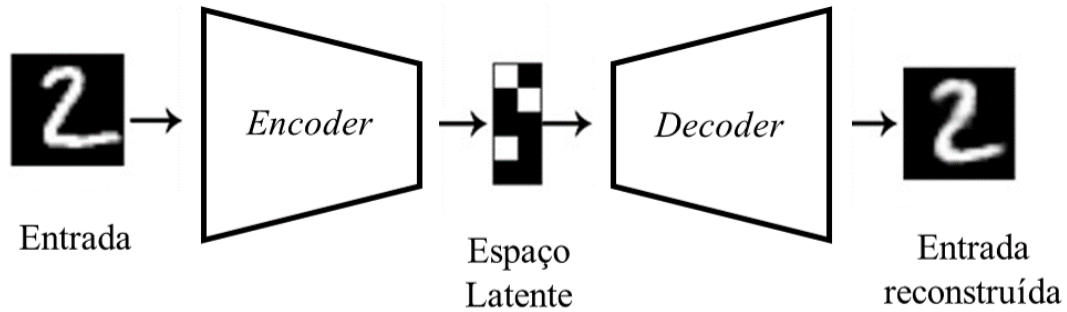
Fonte: Modificado de: (CHOLLET, 2016b).

YAO; ZHAO, 2016). Além do espaço latente, as redes auto-codificadoras possuem em sua constituição um *encoder* (parte da rede na qual o dado de entrada sofre sucessivas reduções de dimensionalidade, passando por funções de ativação não lineares) e um *decoder* (parte da rede na qual o dado oriundo do espaço latente é reconstruído por meio de sucessivos aumentos de dimensionalidade, passando por funções de ativação não lineares). A saída desejada é o próprio dado de entrada, uma vez que o objetivo das redes auto-codificadoras é reconstruir o dado de entrada após a projeção deste no espaço latente (Figura 30). O treinamento se dá por meio de algoritmos de otimização por decida de gradiente utilizando a retro-propagação do erro para o cálculo dos gradientes (GOODFELLOW; BENGIO; COURVILLE, 2016).

As redes auto-codificadoras podem ser implementadas utilizando redes neurais tradicionais ou redes neurais convolucionais. Além da redução não linear de dimensionalidade, estas podem ser utilizadas para tarefas como detecção de anomalias (SAKURADA; YAIRI, 2014), remoção de ruídos (SONG et al., 2020) e completção de imagens (WANG et al., 2020).

Com o objetivo de realizar a tarefa de segmentação de imagens biomédicas Ron-

Figura 30 – Estrutura básica de uma rede auto-codificadora: *encoder*, espaço latente e *decoder*.



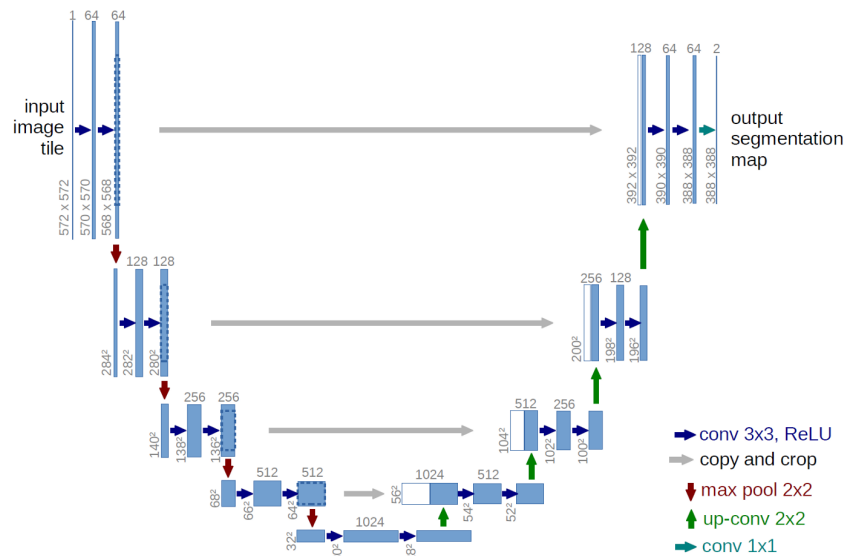
Fonte: Modificado de (CHOLLET, 2016a).

neberger, Fischer e Brox (2015) propuseram a rede U-Net, acrescentando atalhos nas conexões (*skip connections*) entre o *encoder* e o *decoder* de uma rede auto-codificadora convolucional, a fim de permitir um melhor fluxo de gradiente pela rede (Figura 31).

As U-Nets também se mostraram úteis no campo da geofísica. Muller et al. (2022) utilizaram uma U-net para progressivamente estimar modelos de velocidades tomográficas mais detalhados e Wu et al. (2019) utilizaram uma U-net para segmentação de falhas geológicas em volumes sísmicos pós-empilhamento.

A estrutura da U-Net é frequentemente utilizada como rede geradora na arquitetura de redes geradoras adversárias (ISOLA et al., 2017).

Figura 31 – Estrutura de uma U-Net com atalhos nas conexões entre *encoder* e *decoder*.



Fonte: (RONNEBERGER; FISCHER; BROX, 2015).

2.2.4 Redes Neurais Baseadas em Energia

Até o presente parágrafo, foram apresentadas arquiteturas de redes neurais do tipo *feedforward* (propagação para frente), em que dados de entrada são processados sequencialmente, em um fluxo de informação por camadas sucessivas de neurônios conectados por meio dos pesos.

Contudo existem arquiteturas de redes neurais com retroalimentação no fluxo de informação entre os os neurônios. Essas redes são conhecidas como Redes Neurais Recorrentes, ou RNN do inglês *Recurrent Neural Networks*.

As RNN podem ser divididas em dois grupos: as arquiteturas com aprendizado supervisionado, onde existe um conjunto de treinamento composto por entradas e saídas desejadas e o objetivo é encontrar uma regra que mapeie $x \rightarrow y$, e as arquiteturas de aprendizado não supervisionado, onde o conjunto de treinamento é composto apenas por dados de entrada e o objetivo é encontrar relações e extrair padrões associados aos dados. Neste trabalho será dado foco às RNN com aprendizado não supervisionado.

No do grupo das RNN com aprendizado não supervisionado, a Rede de Hopfield, proposta por [Hopfield \(1982\)](#), se destaca por ser a primeira a ter capacidade de apresentar memória associativa mediante a minimização de uma função de energia dado um conjunto de entrada com componentes ortogonais relevantes: $\mathbf{X} = (\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^M)$, $\mathbf{x}^m \in \{-1, 1\}$.

A arquitetura da Rede de Hopfield foi inspirada no modelo matemático para o comportamento dos domínios magnéticos em materiais cristalinos, vinculados à orientação dos *spins* dos elétrons, que podem assumir os estados binários $-\frac{1}{2}$ e $\frac{1}{2}$. Cada um dos *spins* é influenciado pelo campo magnético externo h_{ext} e também pelo campo magnético interno h_{int} , produzido pelos *spins* vizinhos. O campo magnético total, h , pode ser descrito pelo modelo de Ising, dado pela Equação 2.31.

$$h_i = \sum_j w_{ij} S_j + h_{ext} \quad (2.31)$$

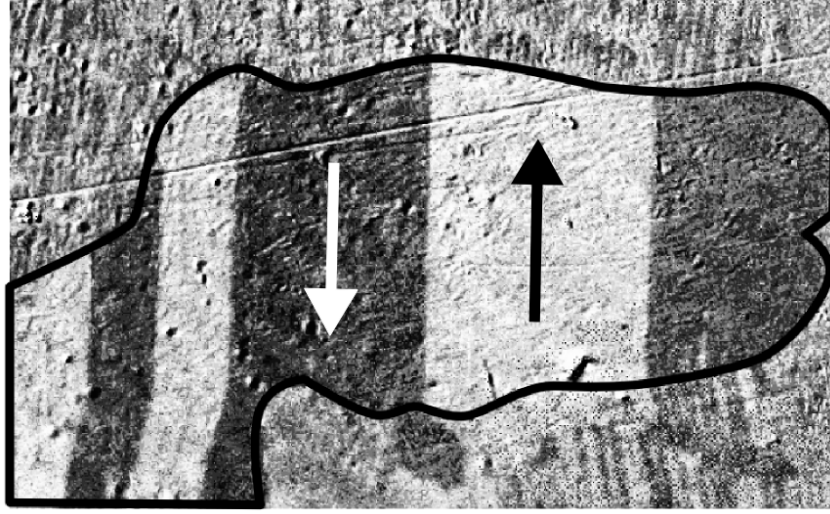
A variável $S_i \in \{-1, 1\}$ representa o estado do *spin* no ponto i em um material cristalino e w_{ij} é uma grandeza que mede a influência de S_j no campo magnético de S_i , chamada de interação de troca entre os *spins* ([HERTZ; KROGH; PALMER, 1991](#)).

A Figura 32 apresenta uma fotomicrografia de efeito Kerr em uma seção polida de aço no qual existe um alinhamento dos *spins* formando domínios magnéticos.

A energia potencial associada as interações entre os *spins* S_i e S_j é dada pela Equação 2.32.

$$E = -\frac{1}{2} \sum_{j \neq i} S_i w_{ij} S_j - h_{ext} \sum_i S_i \quad (2.32)$$

Figura 32 – Microscopia de efeito Kerr mostrando a formação de domínios magnéticos no grão central de uma seção polida de aço.



Fonte: (ZUREKS, 2022).

Quanto mais baixa a energia associada a um determinado estado dos *spins* S_J (padrão dos domínios magnéticos) maior a estabilidade deste estado (HERTZ; KROGH; PALMER, 1991).

Sob baixas temperaturas, os *spins* tendem a se alinhar com o campo magnético externo, $S_i = \text{sign}(h_{\text{ext}})$, isso acontece de forma assíncrona e aleatória. Por outro lado, quando o campo magnético externo é zero, a organização dos domínios magnéticos é dada unicamente pela interação entre os *spins* e a energia potencial depende apenas dos estados de $\mathbf{S}$ (HERTZ; KROGH; PALMER, 1991).

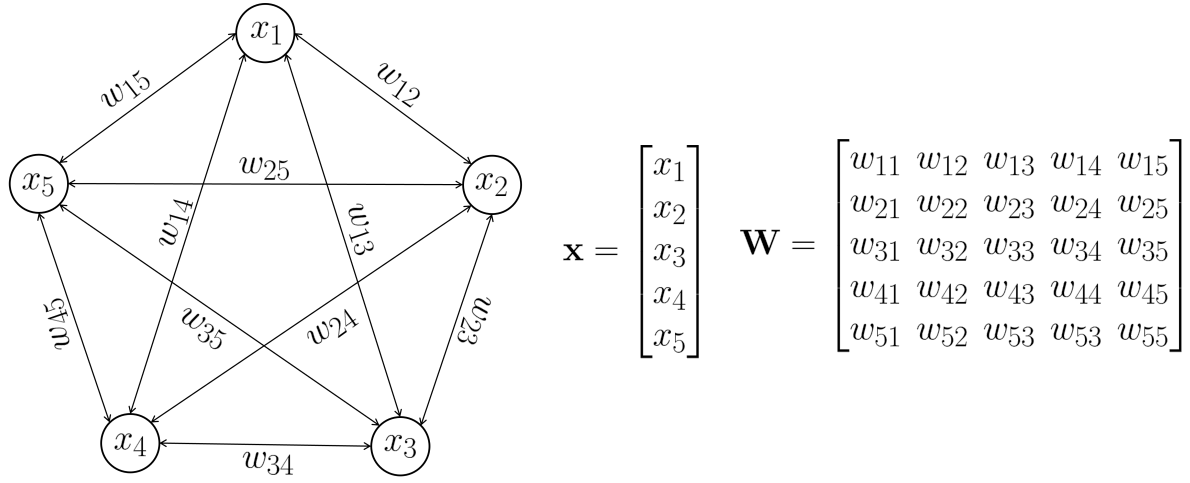
Em analogia ao modelo de Ising, Hopfield (1982) propôs as redes de Hopfield, onde sinais de entrada binários $\mathbf{x} \in \{-1, 1\}$ são os equivalentes dos estados dos *spins* e os pesos são equivalentes à grandeza que mede a interação de troca entre os *spins*. A matriz de pesos $\mathbf{W}$ é simétrica, $w_{ij} = w_{ji}$, com a diagonal principal igual a zero, $w_{i=j} = 0$ (Figura 33).

A Energia da Rede Hopfield também será análoga ao Modelo Ising para o caso onde o campo magnético externo é igual a zero (Equação 2.33).

$$E = -\frac{1}{2} \sum_{j \neq i} x_i w_{ij} x_j = -\frac{1}{2} \mathbf{x}^T \mathbf{W} \mathbf{x} \quad (2.33)$$

Na Rede de Hopfield, a inferência é realizada da seguinte forma: dado um vetor de entrada $\mathbf{x}$, cada neurônio muda o valor de entrada x_i de acordo com a Equação 2.34. Esse procedimento é repetido com a escolha aleatória dos neurônios, até que a Energia da Rede Hopfield, dada pela Equação 2.33, seja minimizada e um estado de equilíbrio seja atingido.

Figura 33 – Rede de Hopfield com cinco entradas, seu vetor de entrada $\mathbf{x}$ e sua matriz de pesos $\mathbf{W}$.



$$x_i = \text{sign} \left(\sum_j x_j w_{ij} \right) = \begin{cases} -1, & \text{se } \sum_j x_j w_{ij} < 0 \\ 1, & \text{se } \sum_j x_j w_{ij} \geq 0 \end{cases} \quad (2.34)$$

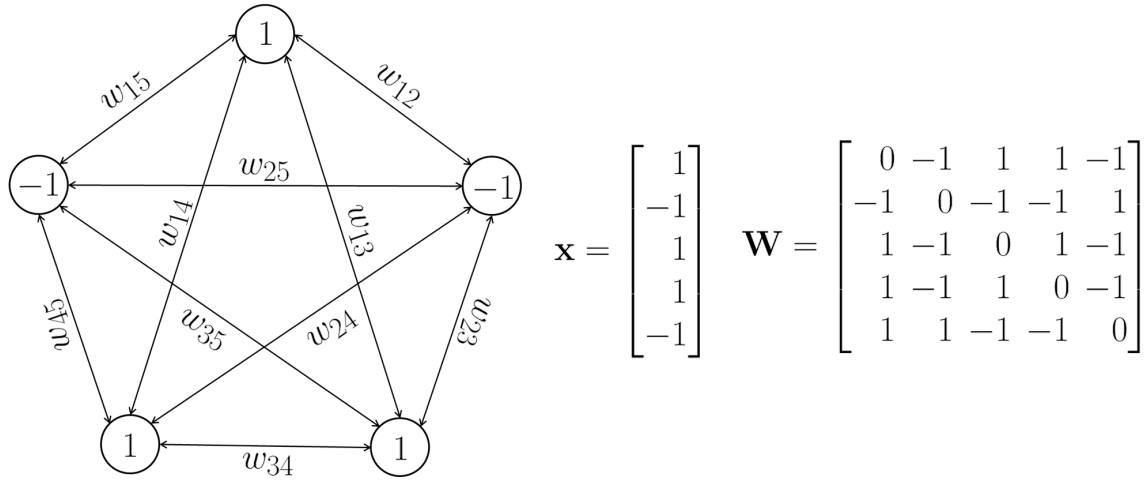
O processo de treinamento da Rede de Hopfield pode ser realizado por meio da aprendizagem hebbiana, introduzida por [Hebb \(1949\)](#) na tentativa de descrever o fenômeno de plasticidade sináptica em redes neurais biológicas. [Hebb \(1949\)](#) propôs o seguinte: neurônios que disparam juntos são mais propensos a se conectarem e neurônios que disparam fora de sincronia são menos propensos a se conectarem.

Em um exemplo simples, considerando-se apenas uma entrada $\mathbf{x} = [1, -1, 1, 1, -1]^T$, o processo de treinamento da Rede de Hopfield mostrado na Figura 33 pode ser realizado por meio da aprendizagem hebbiana da seguinte maneira: para dois neurônios vizinhos no mesmo estado (em sincronia) $w_{ij} \leftarrow 1$; para dois neurônios vizinhos em estados diferentes (fora de sincronia) $w_{ij} \leftarrow -1$, o que é equivalente a $w_{ij} = x_i x_j$ ([DENNIS, 1997](#)). A Figura 34 apresenta os valores de $\mathbf{W}$ tal que a entrada $\mathbf{x}$ seja memorizada pela rede.

Estendendo o raciocínio da aprendizagem hebbiana para um conjunto de treinamento $\mathbf{X} = (\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^M)$, com $\mathbf{x}^m \in \{-1, 1\}$, os valores de w_{ij} são dados deterministicamente, pela Equação 2.35 ou em sua forma vetorial pela Equação 2.36.

$$w_{ij} = \frac{1}{M} \sum_{m=1}^M x_i^m x_j^m \quad \text{para } i \neq j$$

$$w_{ij} = \frac{1}{M} \sum_{m=1}^M x_i^m x_j^m - 1 = 0 \quad \text{para } i = j \quad (2.35)$$

Figura 34 – Rede de Hopfield treinada para um único padrão de entrada $\mathbf{x}$.

$$\mathbf{W} = \frac{1}{M} \sum_{m=1}^M \mathbf{x}^m (\mathbf{x}^m)^T - \mathbf{I} \quad (2.36)$$

Apos o treinamento, a Rede de Hopfield é capaz de realizar tarefas como detecção de erros, remoção de ruídos ou completar sequências com termos faltantes (DENNIS, 1997). A figura 35 apresenta alguns exemplos na aplicação de remoção de ruídos e completção de imagens utilizando as Redes de Hopfield.

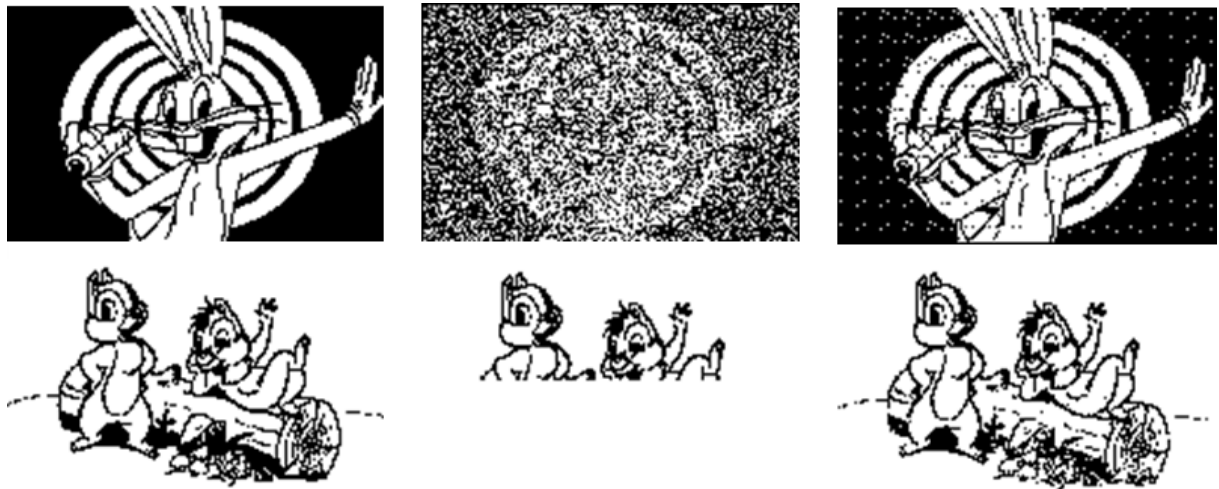
Apesar de suas limitações, como o fato das entradas $\mathbf{X}$ necessitarem de componentes ortogonais relevantes e do número de padrões armazenados ser limitados a $0.15M$, onde M é o numero de neurônios (HOPFIELD, 1982), a Rede de Hopfield foi um grande avanço na área das redes neurais e formam a base conceitual para outras redes baseadas em energia, como as Maquinas de Boltzmann e Máquinas de Boltzmann Restritas.

As redes baseadas em energia são modelos geradores, ou seja, são capazes de aprender uma distribuição de probabilidades mediante um conjunto de dados de treinamento. Após o treinamento é possível recuperar amostras que honram a distribuição aprendida (SALAKHUTDINOV, 2015).

A otimização da Rede de Hopfield considera apenas interações entre pares de neurônios. Estender a arquitetura da rede para interações mais complexas é possível, contudo, o treinamento e a inferência se tornam difíceis de se realizar. Uma solução para este problema foi adicionar à rede novos neurônios que não são usados para representar os dados de entrada, cuja a função é acrescentar complexidade ao sistema. Esses neurônios são conhecidos como neurônios ocultos e atuam como intermediários entre os neurônios visíveis. Como resultado tem-se a Máquina de Boltzmann (HUEMBELI et al., 2022).

As Maquinas de Boltzmann são uma generalização para a Rede de Hopfield, quando a interação entre os neurônios ocultos e os visíveis é igual a zero. Outra forma de diminuir

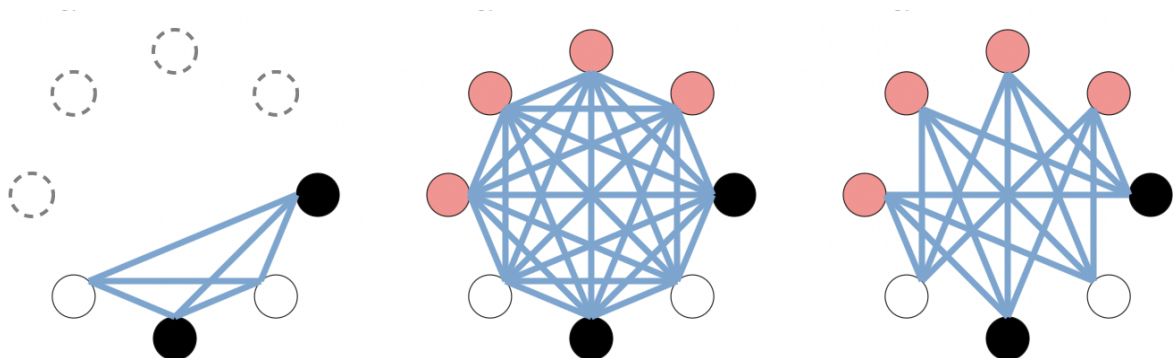
Figura 35 – Rede de Hopfield utilizada para reconstrução de uma imagem degradada, acima, e incompleta, abaixo.



Fonte: (DENNIS, 1997).

a complexidade de maneira a simplificar o treinamento das Maquinas de Boltzmann é eliminar as conexões entre os neurônios da camada de entrada e as conexões entre os neurônios da camada oculta, permitindo apenas conexões entre os neurônios ocultos e os de entrada (HUEMBELI et al., 2022). Essa Arquitetura de rede é conhecida como Máquina de Boltzmann Restrita (Figura 36).

Figura 36 – Rede de Hopfield à esquerda (o numero de neurônios é igual ao tamanho da dimensão do dado de entrada), Máquina de Boltzmann ao centro (existem conexões entre todos os neurônios, os de entrada e os ocultos) e Máquina de Boltzmann Restrita à direita (apenas conexões entre os neurônios ocultos e os de entrada).



Fonte: (HUEMBELI et al., 2022).

2.2.5 Redes Adversárias Geradoras

O cálculo do gradiente do erro para redes baseadas em energia mais complexas é um problema intratável, e a aproximação por métodos estocásticos nem sempre é bem sucedida (GOODFELLOW et al., 2014).

As redes neurais convolucionais profundas, juntamente com o aumento da capacidade de processamento das Unidades de processamento gráfico, do inglês *graphics processing units* (GPUs) e da abundância de dados gerados nas últimas décadas com o advento e popularização da internet, causaram uma verdadeira revolução na área de classificação de imagens, gerando acurácias elevadas, comparáveis à acurácia performada por seres humanos (KRIZHEVSKY; SUTSKEVER; HINTON, 2012; SIMONYAN; ZISSERMAN, 2015; HE et al., 2016). Contudo, esta performance elevada ficou restrita ao campo das redes discriminativas, as redes geradoras não acompanharam este progresso de imediato.

No trabalho intitulado *Generative Adversarial Nets* (GAN) ou em português: Redes Adversárias Geradoras, Goodfellow et al. (2014) argumentam que, se uma rede discriminativa consegue obter uma acurácia tão elevada, de alguma forma os parâmetros dessa rede contém as estruturas e características inerentes às categorias classificadas por ela. Esses autores propuseram uma nova metodologia de treinamento, onde uma rede discriminativa, $D(x)$, fornece o gradiente utilizado para o treinamento da rede geradora, $G(z)$, em uma estratégia de competição entre as redes. A variável z é um vetor de ruído aleatório e x é a entrada do discriminador.

A duas redes são treinadas simultaneamente, o propósito de G é capturar a distribuição do dado de treinamento de forma a maximizar o erro de D , que tem como função estimar a probabilidade de uma amostra x ser oriunda dos dados de treinamento ou ter sido produzida por G .

Esse tipo de estratégia permite o uso do algoritmo de retro-propagação do erro para o cálculo dos gradientes e atualização dos parâmetros por meio de algoritmos de descida de gradiente. A função de perda $\mathcal{L}_{\text{GAN}}(G, D)$ é um problema do tipo *minmax* (Equação 2.37).

$$\arg \min_G \max_D \mathcal{L}_{\text{GAN}}(G, D) = \mathbb{E}_x[\log D(x)] + \mathbb{E}_z[1 - \log D(G(z))] \quad (2.37)$$

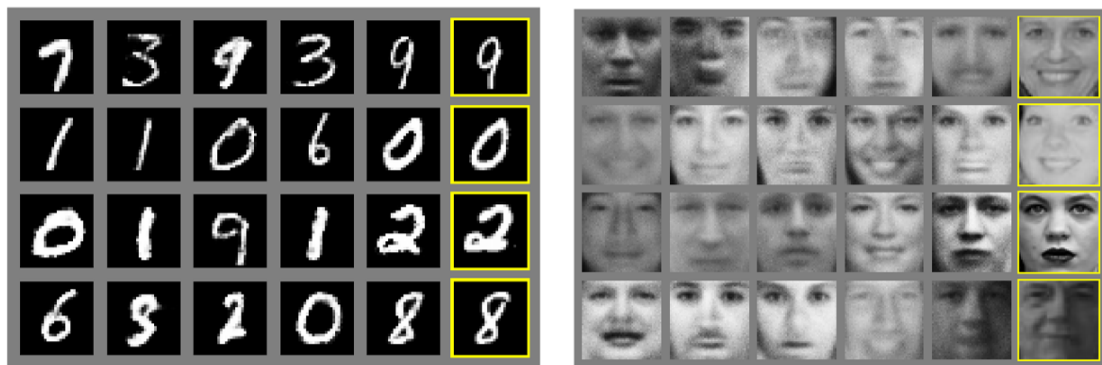
Onde $D(x)$ é treinada para maximizar a probabilidade da classificação correta entre amostras do dado de treinamento e amostras oriundas da rede geradora e $G(z)$ é treinada para minimizar $\log(1 - D(G(z)))$ como mostra a Equação 2.37 (GOODFELLOW et al., 2014).

A função $\mathcal{L}_{\text{GAN}}(G, D)$ é aprendida durante o treinamento adversarial e, dessa forma, pode ser adaptada a uma diversidade de dados sem a necessidade de intervenção de especialistas para elaborar funções objetivos diferentes (ISOLA et al., 2017)

Goodfellow et al. (2014) utilizaram, na maior parte dos experimentos, geradores e discriminadores compostos por *perceptrons* multi-camadas. A Figura 37 exibe exemplos de resultados para os *datasets* MINIST e TFD. Contudo é possível implementar os geradores

e discriminadores utilizando redes neurais convolucionais, culminando em resultados mais realistas e na possibilidade de realizar operações no espaço latente ([RADFORD; METZ; CHINTALA, 2016](#)).

Figura 37 – Visualização de amostras produzidas por redes adversárias geradoras adversárias. À direita de cada imagem, com borda em amarelo, imagens reais, utilizadas para treinamento.



Fonte: ([GOODFELLOW et al., 2014](#)).

Um avanço importante na arquitetura das GANs foi o condicionamento do discriminador e do gerador, proposto por [Mirza e Osindero \(2014\)](#), permitindo um maior controle sobre as amostras geradas. A Figura 38, retirada do trabalho acima, exemplifica este processo utilizando o *dataset* MNIST condicionado por categoria.

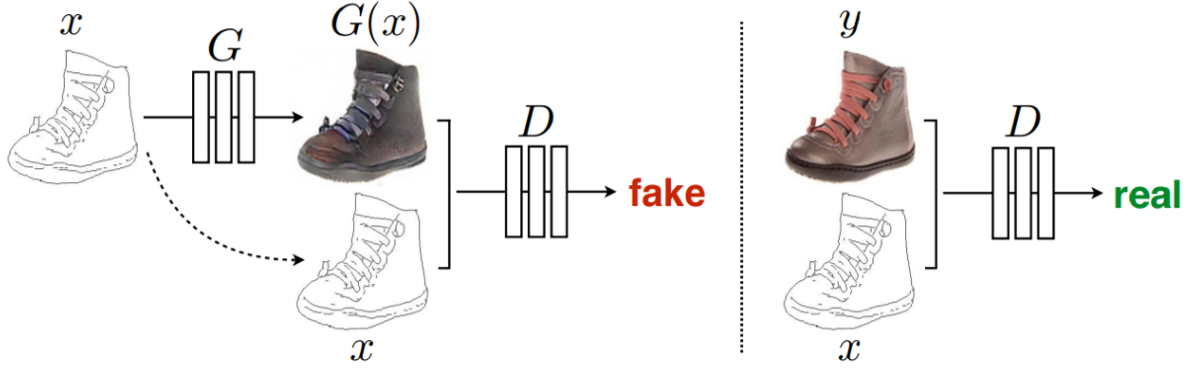
Figura 38 – Visualização de amostras produzidas por redes adversárias geradoras condicionais. A amostra de cada linha foi condicionada por sua respectiva categoria, no caso, os números de 0 a 9 ([MIRZA; OSINDERO, 2014](#)).



Fonte: ([MIRZA; OSINDERO, 2014](#)).

O condicionamento das GANs não precisa ficar restrito à rótulos. [Isola et al. \(2017\)](#) propuseram uma rede chamada *pix2pix* na qual o condicionamento do gerador e do discriminador é feito por imagens (Figura 39), com o intuito de realizar o processo de mapear uma imagem de entrada x em uma imagem de saída y .

Figura 39 – Processo de treinamento de uma GAN condicional treinada para mapear bordas para fotos. O discriminador aprende a classificar entre imagens sintetizadas pelo gerador, $G(x)$, e imagens reais y . O objetivo do gerador é produzir imagens que não possam ser distinguidas de imagens reais. Tanto o gerador quanto o discriminador são condicionados pela imagem de entrada x .



Fonte: (ISOLA et al., 2017).

A função de erro de uma GAN condicional é dada pela Equação 2.38. No caso da rede *pix2pix*, além da função de erro $\mathcal{L}_{\text{cGAN}}(G, D)$, os autores acharam benéfico adicionar a função de perda $L1$ (Equação 2.39) com o objetivo de diminuir a distância entre as imagens sintetizadas pelo gerador e as imagens reais. Dessa forma, o objetivo final da rede *pix2pix* é dado pela Equação 2.40.

$$\mathcal{L}_{\text{cGAN}}(G, D) = \mathbb{E}_{x,y} [\log D(x, y)] + \mathbb{E}_{x,z} [\log(1 - D(x, G(x, z)))] \quad (2.38)$$

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z} [\|y - G(x, y)\|_1] \quad (2.39)$$

$$G^* = \arg \min_G \max_D \mathcal{L}_{\text{cGAN}}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad (2.40)$$

A Figura 40 apresenta pares de imagens condicionantes e imagens geradas pela rede *pix2pix*.

Determinar qual a melhor métrica para avaliar a qualidade das imagens geradas pelas GANs é uma tarefa ainda em aberto (BORJI, 2019). Métricas tradicionais como o erro quadrático médio dos pixels não conseguem capturar a dependência espacial e a estatística conjunta do dado gerado utilizando a função de perda adversarial $\mathcal{L}_{\text{GAN}}(G, D)$ (ISOLA et al., 2017).

Métricas utilizadas para análise de imagens como a medida do índice de similaridade estrutural, do inglês Structural Similarity Index Measure (SSIM) ou baseadas em filtros de redes neurais convolucionais pré-treinadas como o *Fréchet Inception Distance* (FID) têm apresentado resultados interessantes (BORJI, 2019).

Figura 40 – Exemplos de imagens produzidas pela rede *pix2pix* tendo como entrada uma imagem condicionante, imagem de satélite para mapas, bordas para fotos, dia para noite e foto em tons de cinza para foto colorida.



Fonte: Modificado de (ISOLA et al., 2017).

2.3 Trabalhos Relacionados

Trabalhos que utilizam métodos de aprendizado de máquina, em especial métodos de aprendizados profundos, baseados em redes neurais convolucionais, para resolução de problemas sísmicos inversos tem se tornado cada vez mais abundantes. Adler, Araya-Polo e Poggio (2021) apresentam uma visão global destes trabalhos que utilizam soluções baseadas em dados para construção de modelos de velocidade sísmica, impedância, refletividade e expansão de largura de banda do dado sísmico.

Saraiva et al. (2021) utilizaram a GAN condicional *pix2pix* (ISOLA et al., 2017) e dados sísmicos reais, pós empilhamento, juntamente com modelo de velocidade advindo do processo de tomografia, para gerar modelos de velocidades 2D tão acurados quanto os modelos gerados pelo método de FWI tradicional.

Puzyrev et al. (2019) mostraram a viabilidade técnica de se realizar inversões sísmicas 1D utilizando redes neurais, tendo como entrada dados sintéticos.

Wu, Lin e Zhou (2018) propuseram uma rede auto-codificadora convolucional capaz de produzir modelos de velocidade utilizando para o treinamento dados sintéticos 2D e os respectivos *gathers*. Li et al. (2020) aprimoraram esta arquitetura utilizando uma função de perda híbrida, composta pela norma L_2 e o SSIM (Structural Similarity Index Measure).

Araya-Polo, Farris e Florez (2019) empregaram uma GAN para gerar modelos de velocidades de forma a aumentar o conjunto de dados utilizado para treinamento de uma rede convolucional com o intuito de mimetizar o operador de tomografia sísmica.

Zheng et al. (2019) utilizaram dados sintéticos 1D, similar aos dados de poços, e seus *gathers* pré empilhamento correspondentes para treinar redes neurais supervisionadas

capazes de produzir dados de Velocidade Acústica, Velocidade Elástica e densidade. [Das et al. \(2019\)](#) empregaram método similar para gerar impedâncias acústicas, porém utilizando uma rede composta por convoluções 1D e tendo como entrada o sismograma sintético.

O trabalho de [Mosser et al. \(2018\)](#) propôs a utilização da rede StyleGAN ([ZHU et al., 2017](#)), desenvolvida inicialmente para realizar transferência de estilos, para resolver problemas sísmicos diretos e inversos utilizando dados sintéticos pós-empilhamento sem a necessidade do pareamento dos dados de entrada/saída desejada utilizados durante o treinamento.

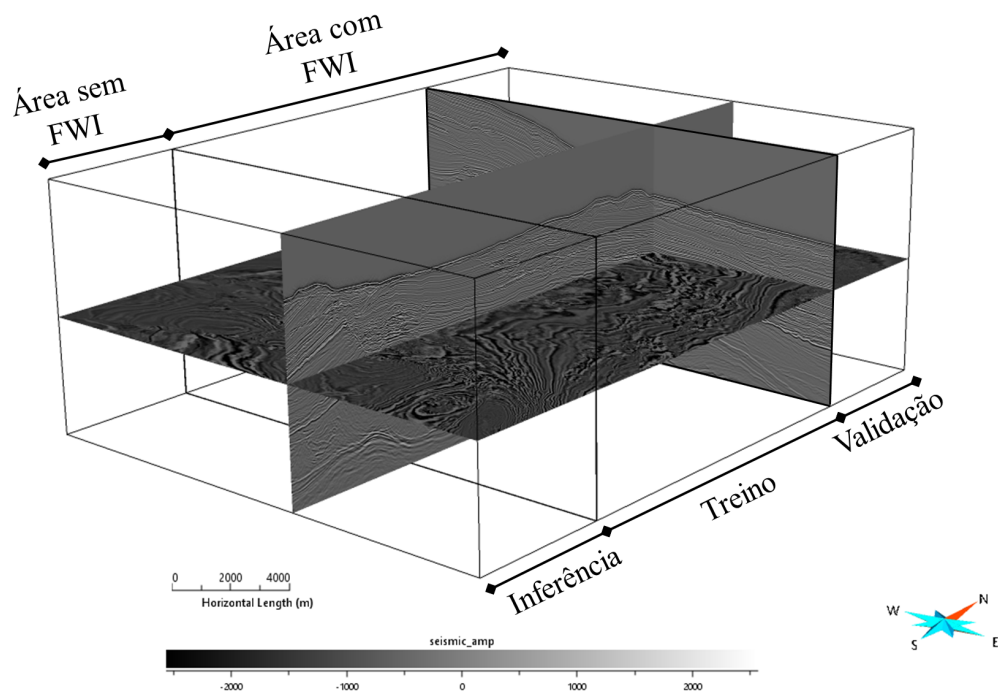
3 Materiais e Métodos

3.1 Dados disponíveis

3.1.1 Dados Brutos

Para realização deste trabalho foram disponibilizados pela Petrobras dados referentes a uma aquisição sísmica *streamer* 3D, que cobriu uma área de 885 km^2 na Bacia de Campos, Brasil. O processamento sísmico inicial foi realizado utilizando um modelo de velocidades intervalares estimado por meio de tomografia e uma Migração Kirchhoff pré-empilhamento em profundidade. Posteriormente, em uma área limitada a 680 km^2 , foi realizado um reproprocessamento sísmico utilizando a técnica de FWI para a estimativa do campo de velocidades intervalares (Figura 41).

Figura 41 – Visualização 3D do volume sísmico de amplitudes disponibilizado para estudo. Todo o dado possui uma modelagem de velocidades por tomografia, contudo a modelagem FWI fica restrita a porção norte.



Foram utilizados quatro volumes para o treinamento e validação, restritos à área com modelo de velocidade FWI:

- Volume de Amplitude Sísmica.
- Volume de Velocidade Média - Tomografia.

- Gride de Tempo Duplo de Trânsito (TWT).
- Volume de Velocidade Intervalar - FWI.

Estes dados, originalmente no formato seg-y ([HAGELUND; LEVIN, 2017](#)), foram carregados no ambiente python, utilizando a biblioteca *seggyio* ([EQUINOR, 2022](#)) de forma a retornar o dado no formato de matriz *numpy*. A Tabela 3 apresenta a estatística descritiva dos dados disponíveis.

Tabela 3 – Estatística descritiva dos dados disponíveis.

	Mínimo	Máximo	Média	Desvio padrão
Amplitude Sísmica	-34999,7	33855,2	-0,15748	995,142
Velocidade Média - Tomografia (m/s)	1492,35	2745,45	1715,11	263,74
Gride TWT (s)	1500	4500	3000	867,18
Velocidade Intervalar - FWI (m/s)	1458,57	5413,21	2568,35	1059,43

Cada volume sísmico possui 1451 *inlines*, 3001 *crosslines* e 751 amostras verticais em tempo duplo de trânsito. O espaçamento entre as *inlines* é de 12,5 m, entre as *crosslines* de 12,5 m e a amostragem vertical é de 0.004 s. O pré-processamento do dado consistiu no seguinte fluxo:

- Limitar a variabilidade dos dados em duas vezes e meia o desvio padrão com o intuito de eliminar *outliers*.
- Dividir os dados em porção de treino (70%) e validação (30%).
- Normalizar os dados de treino e validação entre -1 e 1, utilizando como referência os dados de treino.

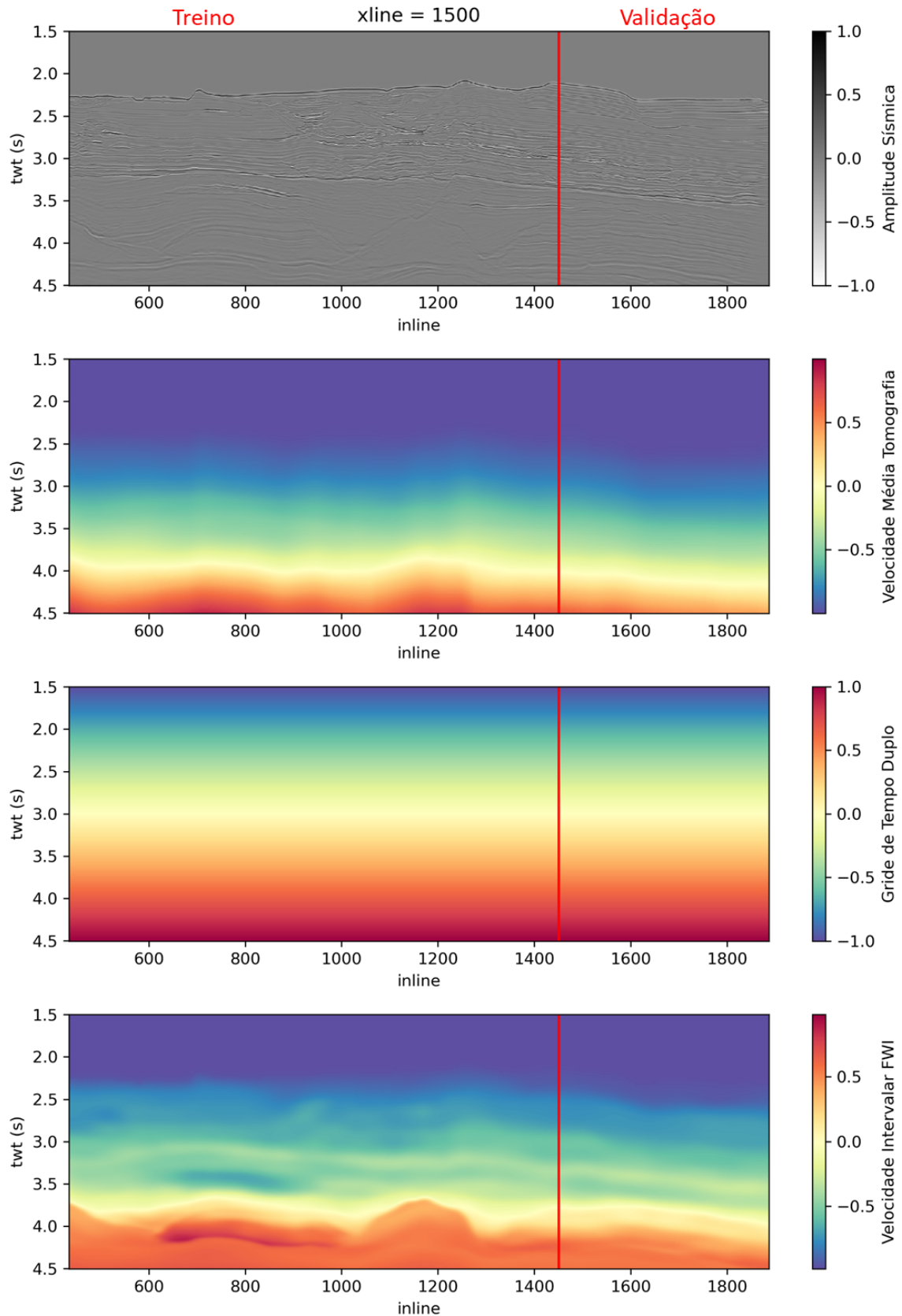
O dado sísmico possui uma similaridade e interdependência espacial entre as amostras, fato inerente aos processos deposicionais das rochas sedimentares investigadas pelo método sísmico. Desta forma, a divisão entre dado de treino e validação foi realizada de maneira similar ao procedimento adotado para divisão de séries temporais ([PEDREGOSA et al., 2011](#)), separando o dado em porções contíguas, a fim de evitar vazamento de informação entre os dados de treino e validação.

A Figura 42 apresenta a *xline* central dos dados utilizados para treinamento e validação após normalização.

3.1.2 Base de dados

Durante o treinamento a inferência da rede se faz necessário extrair sub-volumes dos dados de entrada devido ao grande volume de informação (a soma dos dados de treino e validação é da ordem de 50GB) e a limitação de memória das placas gráficas. Para tanto,

Figura 42 – *Xline* central dos dados disponíveis após normalização: Amplitude Sísmica, Velocidade Média (Tomografia), Gríde de Tempo Duplo (TWT), Velocidade Intervalar (FWI).



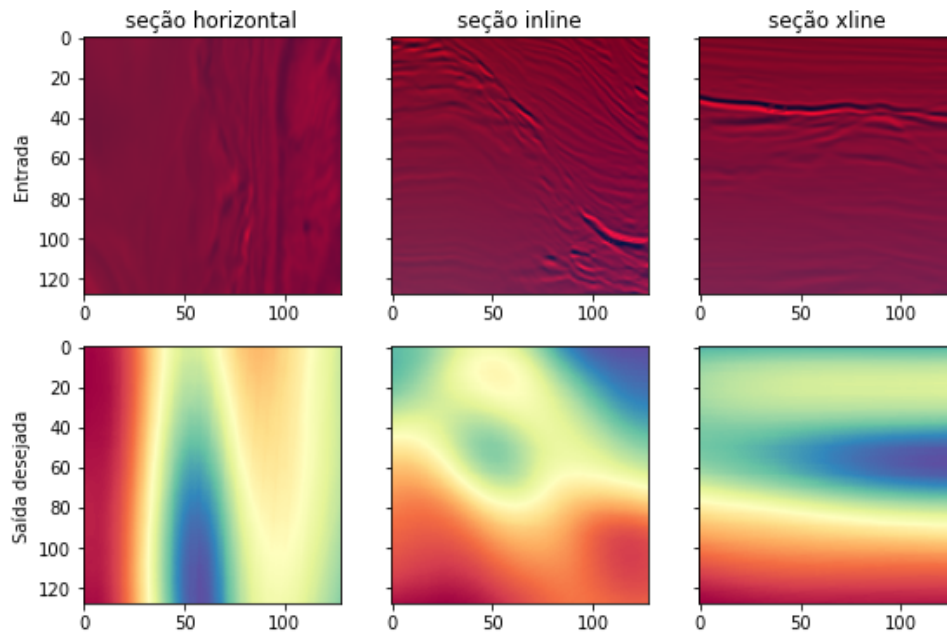
foi implementado, utilizando como referência a classe `torch.utils.data.Dataset` do `pytorch`, dois tipos de `dataset`, um para o dado de treinamento e outro para o dado de validação. Nos dois casos os volumes de entrada para a rede (Amplitude Sísmica, Velocidade Média Tomografia e Gride de Tempo Duplo de Trânsito) foram empilhados resultando em um dado com três canais. A saída desejada é o volume com a Velocidade Intervalar FWI.

Para o `dataset` de treino foram extraídos, de forma aleatória, 2048 sub-volumes com dimensão $128 \times 128 \times 128$. Esta quantidade de sub-volumes garante que existirá sobreposição de dados. Foi realizado também o processo de *data-augmentation* por meio de operações aleatórias de reflexão do dado (*random flip*) e ampliação de até 20% com o auxílio da biblioteca `torchIO` (PÉREZ-GARCÍA; SPARKS; OURSELIN, 2021).

A avaliação do dado de validação requer a reconstrução do Volume de Velocidade Intervalares à partir dos sub-volumes produzidos pela Rede Geradora, para tanto, foram extraídos 3036 sub-volumes com dimensão $128 \times 128 \times 128$, de forma sequencial, com *stride* de 64. O *stride* de 64 foi utilizado de maneira a garantir uma sobreposição de 50% entre os sub-volumes, o que foi conveniente durante a etapa de reconstrução (Sub-capítulo 4.3).

A Figura 43 apresenta seções de uma amostra dos sub-volumes de entrada no formato de uma imagem RGB (*red, green and blue*, com a Amplitude Sísmica no Canal R, a Velocidade Média Tomografia no canal G e o Gride de Tempo Duplo no canal B). A saída desejada, Velocidade Intervalar FWI, possui apenas um canal e foi plotada utilizando escala de cor.

Figura 43 – Exemplo de sub-volumes de treinamento, acima a entrada da rede, com três canais e abaixo a saída desejada, com apenas um canal.



O mesmo procedimento para a construção do `dataset` de validação foi adotado para

o *dataset* de inferência, contudo este ultimo possui apenas os dados de entrada, de onde foram extraídos 3036 sub-volumes com dimensão $128 \times 128 \times 128$, de forma sequencial, com *stride* de 64.

3.2 Bibliotecas Numéricas

Para o pré-processamento dos dados foram utilizadas as bibliotecas *seggio*, *numpy* e *sklearn*. A implementação das redes Geradora e Discriminadora, do *loop* de treinamento e do *dataset* foi realizada utilizando a biblioteca *pytorch*. O processo de *data augmentation* foi realizado por meio da biblioteca *torchIO*. O volume de saída, gerado pela rede, foi salvo no formato *.seggy* utilizando a biblioteca *seggio*.

3.3 Equipamentos

O pré-processamento do dado, o treinamento e inferência da rede foram realizados utilizando um nó de um servidor com a seguinte configuração:

- CPU Intel(R) Xeon(R) Gold 6240 CPU @ 2.60GHz com 18 núcleos e 394 GB de memória RAM
- Quatro placas gráficas NVIDIA Tesla V100 SXM2 com 32 GB de memória.

4 Metodologia Experimental

Utilizando a biblioteca *pytorch* (PASZKE et al., 2019) foi implementada uma GAN condicional capaz de utilizar dados tri-dimensionais (3D) durante o treinamento. Essa arquitetura foi inspirada no trabalho de Isola et al. (2017), que lidava com imagens 2D, e no trabalho de Cirillo, Abramian e Eklund (2020) que utilizou uma cGAN 3D para realizar segmentação semântica de imagens médicas. A rede proposta utiliza convoluções 3D e convoluções transpostas 3D na Rede Geradora e convoluções 3D na Rede Discriminadora. Os resultados foram avaliados qualitativamente de acordo com análise visual e quantitativamente utilizando duas métricas, o Erro Percentual Absoluto Médio, do inglês *Mean Absolute Percentage Error* (MAPE) e a Medida do Índice de Similaridade Estrutural, do inglês *Structural Similarity Index Measure* (SSIM).

4.1 Arquitetura da Rede

4.1.1 Rede Geradora

A Rede Geradora é uma U-NET 3D composta por quatro blocos com operações de convoluções 3D, *instance normalization* e função de ativação do tipo *Leaky ReLU* em que o dado sofre redução da dimensionalidade, partindo de uma entrada com dimensões de $128 \times 128 \times 128 \times 3$, em direção ao gargalo da rede, com dimensão de $8 \times 8 \times 8 \times 512$ (Figura 44).

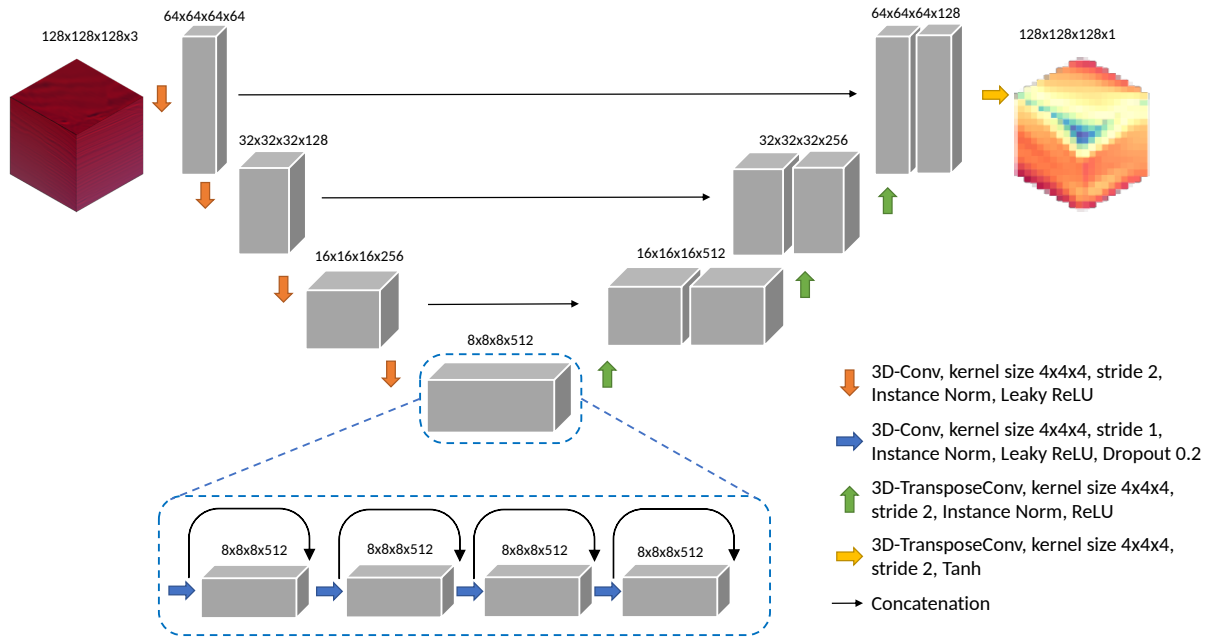
No gargalo ocorrem quatro blocos de operações, com preservação das dimensões, compostos de convoluções 3D, *instance normalization*, função de ativação do tipo *Leaky ReLU* e *Dropout* com propabilidade de 20% (Figura 44).

Em seguida o dado oriundo do gargalo passa por uma sequência de três blocos de convolução transposta 3D, *instance normalization* e função de ativação do tipo *Leaky ReLU*, sofrendo um aumento de dimensionalidade de $8 \times 8 \times 8 \times 512$ para $64 \times 64 \times 64 \times 128$, quando finalmente passa por um bloco de convolução transposta 3D e função de ativação do tipo tangente hiperbólica, resultando em uma saída com dimensão $128 \times 128 \times 128 \times 1$ (Figura 44).

4.1.2 Rede Discriminadora

A Rede Discriminadora é uma rede convolucional 3D, do tipo PatchGAN (ISOLA et al., 2017), composta por cinco blocos de convolução 3D, *instance normalization*, sendo que nos quatro primeiros blocos a função de ativação é do tipo *Leaky ReLU* e no último

Figura 44 – Arquitetura da Rede Geradora.



Fonte: Modificado de Cirillo, Abramian e Eklund (2020).

do tipo Sigmóide .

A entrada do discriminador consiste na concatenação do dado condicionante com o sub-volume de velocidade intervalar FWI correspondente, seja este produzido pela Rede Geradora ou advindo do conjunto de dados reais. A saída da Rede Discriminadora é uma matriz 3D de dimensão $4 \times 4 \times 4 \times 1$. Cada valor dessa matriz possui um vínculo com uma parte da entrada (linhas azuis sobre o sub-volume de entrada na Figura 45) e retorna a probabilidade desta parte pertencer ao dado real ou produzido pela Rede Geradora (Figura 45).

4.1.3 GAN condicional

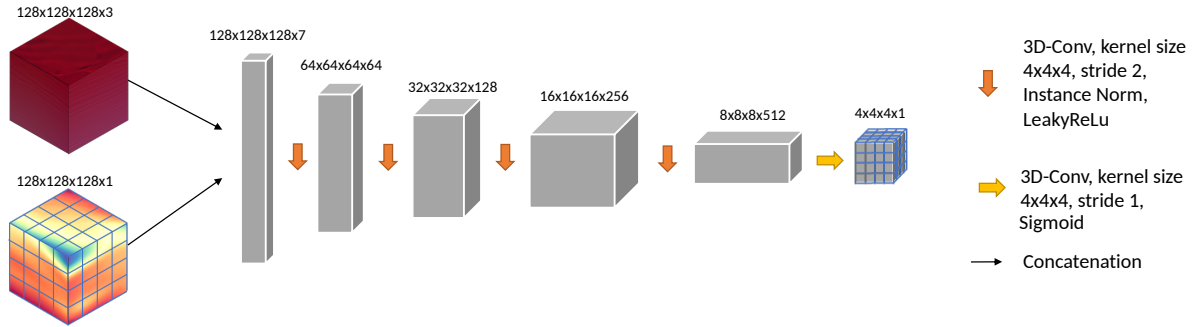
A Figura 46 resume a arquitetura final da GAN condicional proposta, composta pelas redes Geradora e Discriminadora.

4.2 Treinamento

A inicialização dos pesos das redes Geradoras e Discriminadora foi realizada com amostras de uma distribuição normal de $\mu = 0$ e $\sigma = 0.2$.

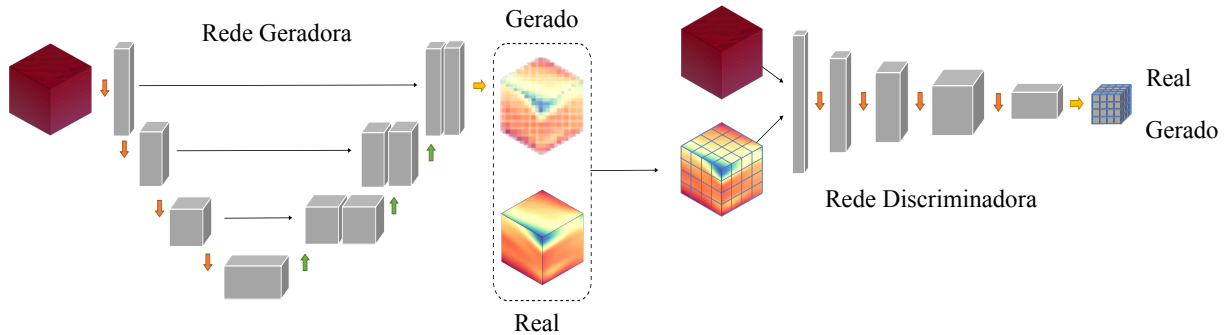
A função de perda é composta por dois termos, um advindo da PatchGAN (Equação 2.38) e outro composto pela distância L1 ponderado pelo hiperparâmetro $\lambda = 100$ (Equação 2.39). Foi utilizado um *batch size* = 1.

Figura 45 – Arquitetura de Rede Discriminadora PatchGAN, originalmente proposta por [Isola et al. \(2017\)](#) e posteriormente implementada utilizando convoluções 3D por [Cirillo, Abramian e Eklund \(2020\)](#). O dado de entrada é subdividido em volumes de dimensão $32 \times 32 \times 32 \times 1$ e classificado em dado real, oriundo do conjunto de treinamento, ou falso, produzido pela Rede Geradora.



Fonte: Modificado de [Cirillo, Abramian e Eklund \(2020\)](#).

Figura 46 – Arquitetura de GAN condicional proposta: a Rede Geradora produz um sub-volume de velocidade intervalar tendo como entrada o dado condicionante, a Rede Discriminadora classifica se o sub-volume é oriundo do conjunto de dados reais ou dos dados gerados.



Fonte: Modificado de [Cirillo, Abramian e Eklund \(2020\)](#).

Foram realizadas 200 épocas com um *batch size* = 1 e uma taxa de aprendizado constante de $\eta = 0.0002$ nas 100 primeiras épocas seguido por um decaimento linear de η até atingir um valor nulo ao final da última época. O algoritmo de otimização utilizado foi o *adaptive moment estimation* (Adam) com parâmetros de momento $\beta_1 = 0.5$ e $\beta_2 = 0.99$.

As redes Discriminadora (D) e Geradora (G) são treinadas alternadamente, começando por D. A quantidade de épocas de D em relação a G pode variar a fim de evitar o *mode collapse*, quando, para diferentes entradas, G retorna amostras com baixa variabilidade ou evitar *vanishing gradients*, quando o erro de G fica excessivamente pequeno, não permitindo o cálculo do gradiente para atualização dos parâmetros ([ISOLA et al., 2017](#)). Contudo, neste trabalho, o mesmo número de épocas para G e D se mostrou eficaz.

Com a finalidade monitorar a evolução do treinamento, os parâmetros de G e D foram salvos ao final de cada época. O Algoritmo 2 apresenta a metodologia utilizada para

o treinamento da rede.

Algoritmo 2: *Loop* de treinamento

Entrada: $\mathbf{X}, \mathbf{y}$ (Dado condicionante e saída desejada)

Saída: $\hat{\mathbf{y}}$

```

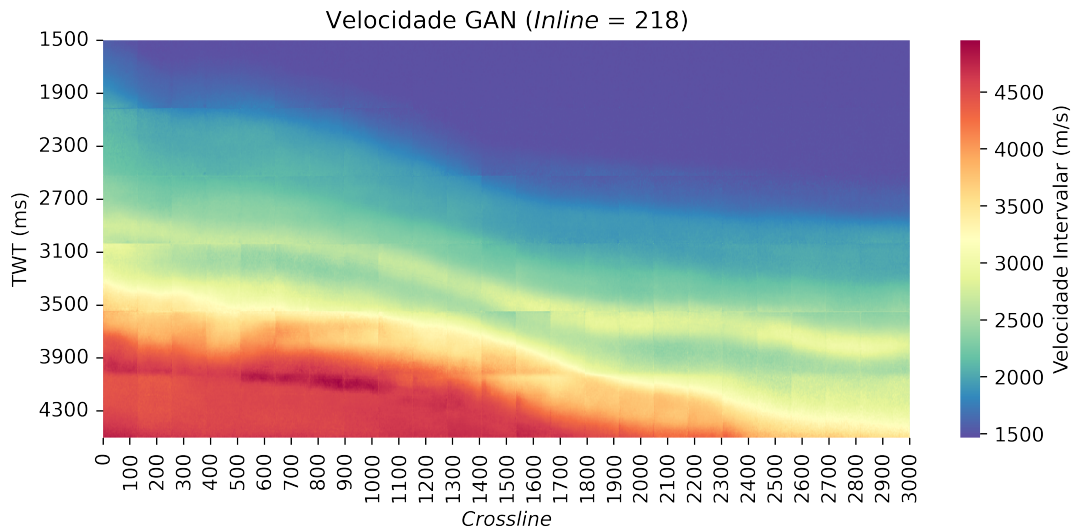
1 início
2    $\mathbf{w}_G \leftarrow \mathcal{N}(0, 0.2)$ 
3    $\mathbf{w}_D \leftarrow \mathcal{N}(0, 0.2)$ 
4    $\eta = 0.0002$ 
5    $\lambda = 100$ 
6   para época  $\leftarrow 1$  até 200 faça
7      $\eta = \eta \cdot (1 - \max(0, \text{época} - 100)/100)$ 
8     para todo  $(\mathbf{x}, y)$  em  $(\mathbf{X}, \mathbf{y})$  faça
9
10      ***Treinamento Discriminador***
11       $\hat{y} = G(\mathbf{x})$ 
12       $Erro\_D_{Real} = \text{EntropiaCruzadaBinária}(D(\mathbf{x}, y), 1)$ 
13       $Erro\_D_{Gerado} = \text{EntropiaCruzadaBinária}(D(\mathbf{x}, \hat{y}), 0)$ 
14       $Erro\_D = (Erro\_D_{Real} + Erro\_D_{Gerado})/2$ 
15       $\mathbf{w}_D \leftarrow \text{atualiza}_W(\mathbf{w}_D, Erro\_D)$ 
16
17      ***Treinamento Gerador***
18       $Erro\_G_{GAN} = \text{EntropiaCruzadaBinária}(D(\mathbf{x}, \hat{y}), 1)$ 
19       $Erro\_L1 = |\hat{y} - y|$ 
20       $Erro\_G = Erro\_G_{GAN} + \lambda * Erro\_L1$ 
21       $\mathbf{w}_G \leftarrow \text{atualiza}_W(\mathbf{w}_G, Erro\_G)$ 
22      salva  $\mathbf{w}_D$ 
23      salva  $\mathbf{w}_G$ 
24    fim
25  fim
26 fim
```

4.3 Inferência e reconstrução do volume

A reconstrução do volume sísmico de velocidades FWI requer a reorganização dos sub-volumes de saída da Rede Geradora de forma que tenham a mesma dimensão do dado original. Este processo acarretou na presença de artefatos na junção dos sub-volumes (Figura 47).

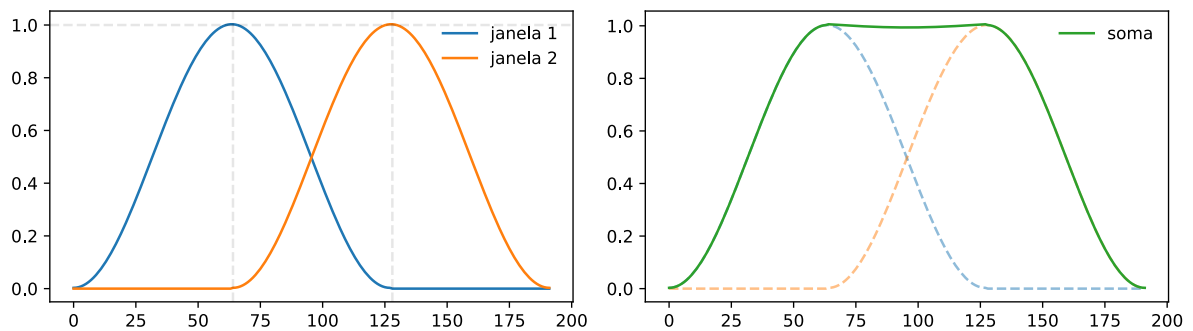
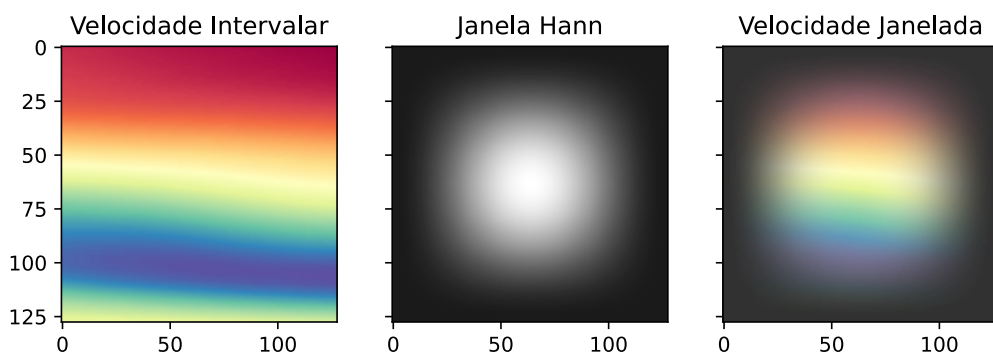
A solução adotada para remover os artefatos nas junções foi utilizar uma sobreposição de 50% nos sub-volumes de entrada e aplicar um janelamento Hann (VIRTANEN et al., 2020) tri-dimensional, em cada amostra, antes da reorganização (Figuras 48 e 49). Após a reconstrução do volume de validação, aplicou-se uma máscara com intuito de compensar

Figura 47 – Marcas provocadas pela junção dos sub-volumes sem sobreposição .



as anomalias de amplitudes referentes às sobreposições das janelas e às bordas¹.

Figura 48 – Soma de duas janelas Hann com sobreposição de 50%.

Figura 49 – Janela Hann 3D aplicada a um sub-volume produzido pela rede Geradora (corte central na direção *inline*)

Depois da reconstrução do volume de validação, os valores passaram pelo processo

¹ Solução proposta pelo Geofísico Erick Costa e Silva Talarico durante conversa em 04/04/2022.

inverso da normalização inicial, de forma a retornar valores condizentes com a distribuição de velocidades intervalares FWI.

4.4 Métricas

Os resultados foram avaliados de acordo com três métricas:

- Análise visual realizada por um geofísico familiarizado com o dado sísmico.
- Erro Percentual Absoluto Médio, do inglês *Mean Absolute Percentage Error* (MAPE)
- Medida do Índice de Similaridade Estrutural (WANG et al., 2004), do inglês *Structural Similarity Index Measure* (SSIM)

A análise visual foi realizada no volume de velocidades reconstruído a partir dos sub-volumes produzidos pela Rede Geradora. Foi verificado se o resultado final conseguiu reproduzir as principais feições geológicas e distribuição dos valores de velocidade presentes no dado de validação.

O erro percentual absoluto médio, dado pela Equação 4.1, foi adotado por ser uma métrica utilizada para controle de qualidade durante o processo convencional de construção do modelo de velocidades e por ser de fácil interpretação.

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|y - \hat{y}|}{\max(\epsilon, |y|)} \quad (4.1)$$

O SSIM é um índice composto (Equação 4.2) que mede a similaridade entre dois sinais tendo como base a comparação entre suas médias (Equação 4.3), variâncias (Equação 4.4) e coeficiente de correlação (Equação 4.5). Foi utilizada a implementação 3D do SSIM Park (2019), que realiza este cálculo para informações volumétricas.

$$\text{SSIM}(y, \hat{y}) = f(l(y, \hat{y}), c(y, \hat{y}), s(y, \hat{y})), \quad (4.2)$$

$$l(y, \hat{y}) = \frac{2\mu_y\mu_{\hat{y}} + C_1}{\mu_y^2 + \mu_{\hat{y}}^2 + C_1}, \quad (4.3) \quad c(y, \hat{y}) = \frac{2\sigma_y\sigma_{\hat{y}} + C_2}{\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2}, \quad (4.4) \quad s(y, \hat{y}) = \frac{\sigma_{y\hat{y}} + C_3}{\sigma_y\sigma_{\hat{y}} + C_3}, \quad (4.5)$$

Onde C_1 , C_2 , e C_3 são constantes para evitar divisão por zero. Fazendo $C_3 = \frac{C_2}{2}$, é possível obter a equação do SSIM (4.6).

$$\text{SSIM}(y, \hat{y}) = \frac{(2\mu_y\mu_{\hat{y}} + C_1) + (2\sigma_{y\hat{y}} + C_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + C_1)(\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2)}. \quad (4.6)$$

Tanto o MAPE quanto o SSIM foram aplicados antes da reconstrução do dado e são representativos dos valores médios de todos os sub-volumes.

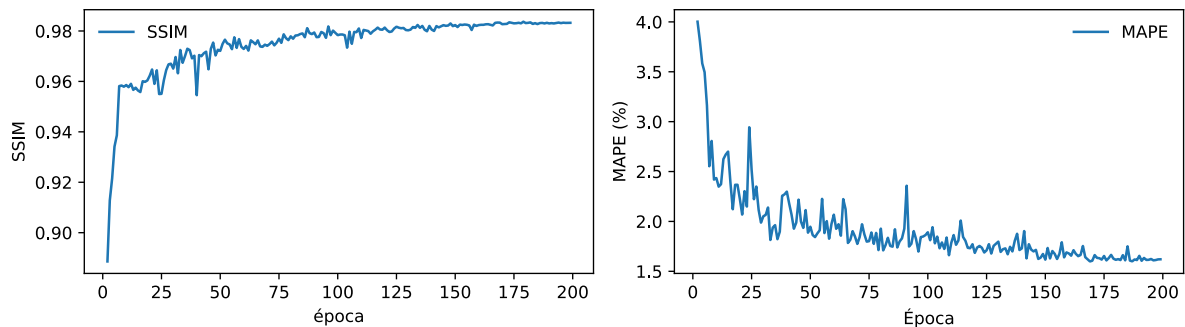
5 Resultados e discussões

5.1 Resultados

O tempo para o treinamento da rede foi de aproximadamente 13 horas. O tempo gasto para gerar os resultados referentes aos volumes de validação e inferência foi de dois minutos e 40 segundos para cada volume.

As métricas quantitativas utilizadas (MAPE e SSIM) apresentaram resultados coerentes entre si e com a análise visual. Os gráficos da Figura 50 ilustram a evolução das métricas quantitativas ao longo das épocas, notou-se uma rápida convergência, com uma melhora significativa já nas primeiras épocas. A última época, de número 200, foi a que apresentou melhores resultados para a análise visual e métricas quantitativas, SSIM= 0.98 e MAPE= 1.6%.

Figura 50 – SSIM e MAPE do dado de validação por época de treinamento.



As distribuições de velocidades intervalares do dado FWI e do dado gerado pela GAN apresentaram uma grande similaridade (Figura 51).

Os resultados produzidos pela Rede Geradora são visualmente condizentes com o dados utilizado para validação. As figuras 52 e 53 apresentam uma sessão *crossline* e *inline* comparando o dado de velocidades intervalares FWI com o dado gerado pela GAN utilizando o modelo gerador referente à última época. Foi realizada também uma análise comparando as amostras verticais desse mesmo dado (Figura 54).

As regiões mais profundas, principalmente na região do sal (velocidades intervalares em torno de 4500 m/s), apresentaram um resultado pior que as regiões mais rasas. Este fato pode ser notado nas figuras 52, 53 e 54, bem como na comparação entre as distribuições de velocidades intervalares (Figura 51).

Ao analisar o resultado do dado de inferência (Figura 41), equivalente à região onde foi realizado a estimativa de velocidades intervalares utilizando apenas a tomografia

Figura 51 – Comparação entre as distribuições da velocidade intervalar FWI e a velocidade intervalar produzida pela GAN.

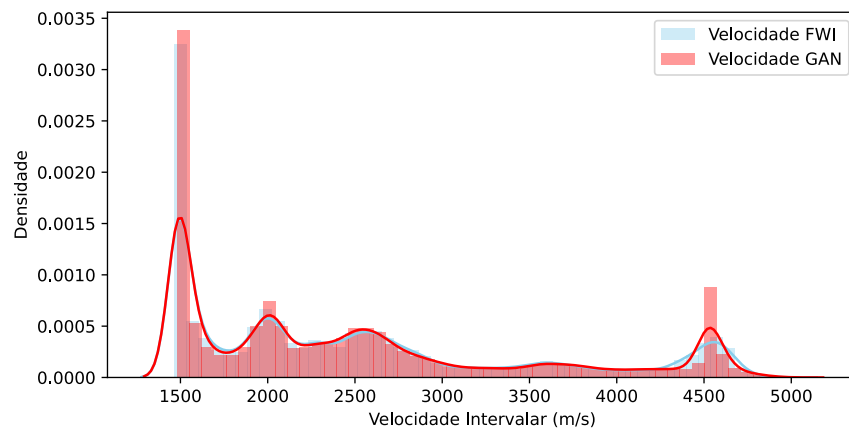
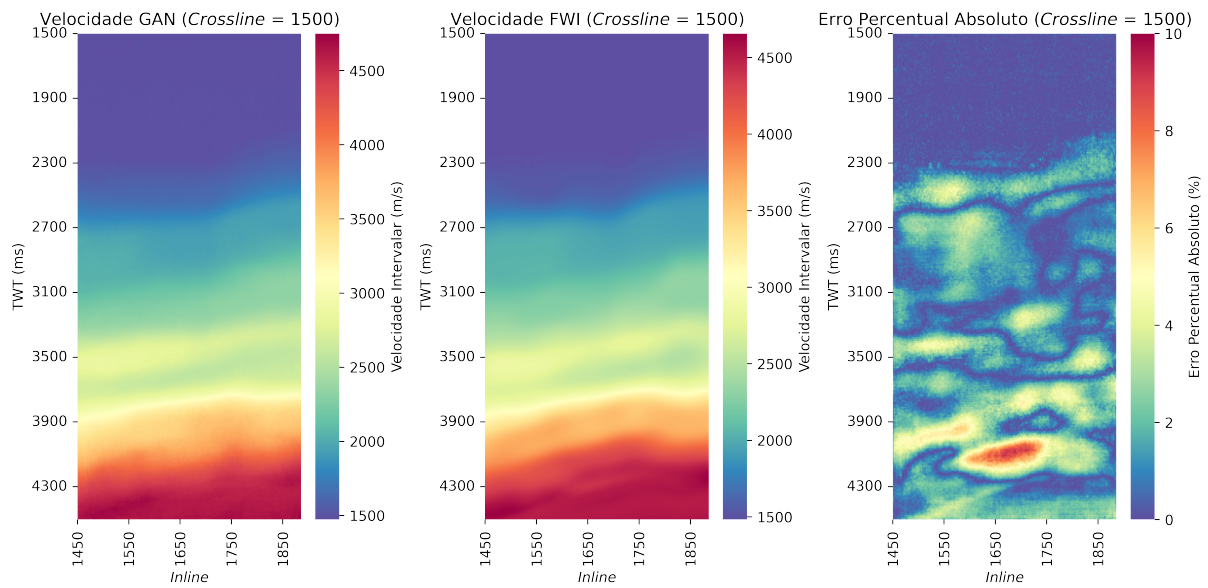


Figura 52 – *Crossline* 1500: Velocidade estimada pelo Modelo Gerador, Velocidade FWI real e Erro Percentual Absoluto entre os volumes reais e gerados pela cGAN.



sísmica, notou-se uma grande coerência do dado gerado com as principais estruturas presentes no dado de amplitude sísmica (figuras 55 e 56).

Figura 53 – *Inline 670*: Velocidade estimada pelo Modelo Gerador, Velocidade FWI real e Erro Percentual Absoluto entre os dois volumes.

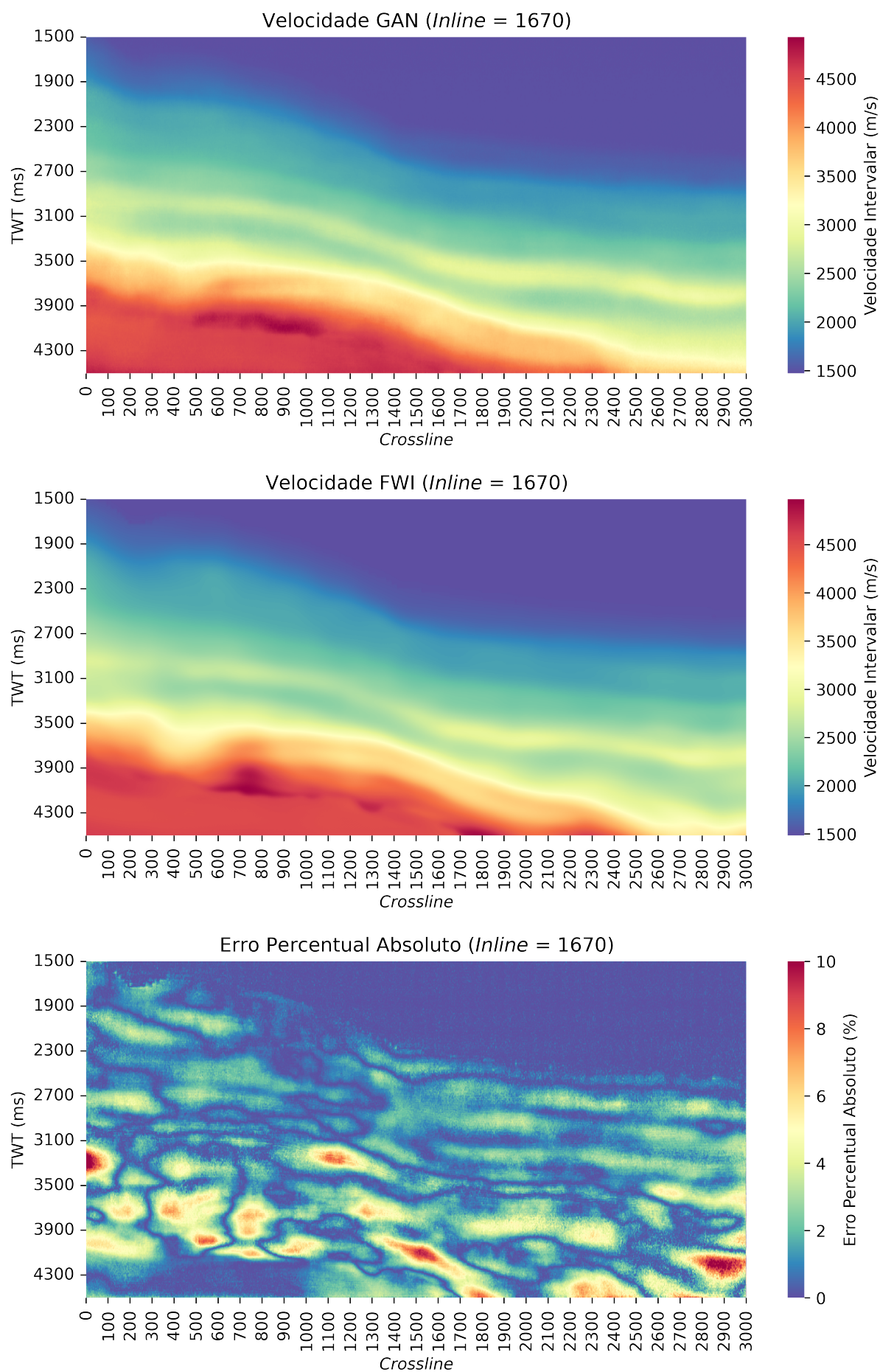


Figura 54 – Amostras verticais retiradas na região do volume de validação.

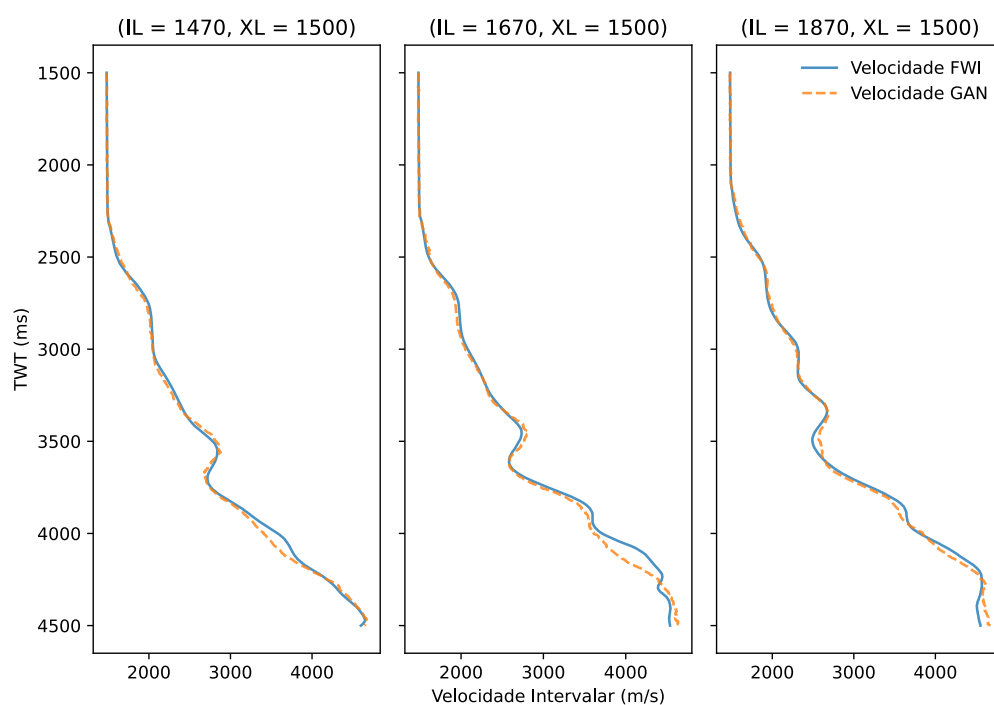
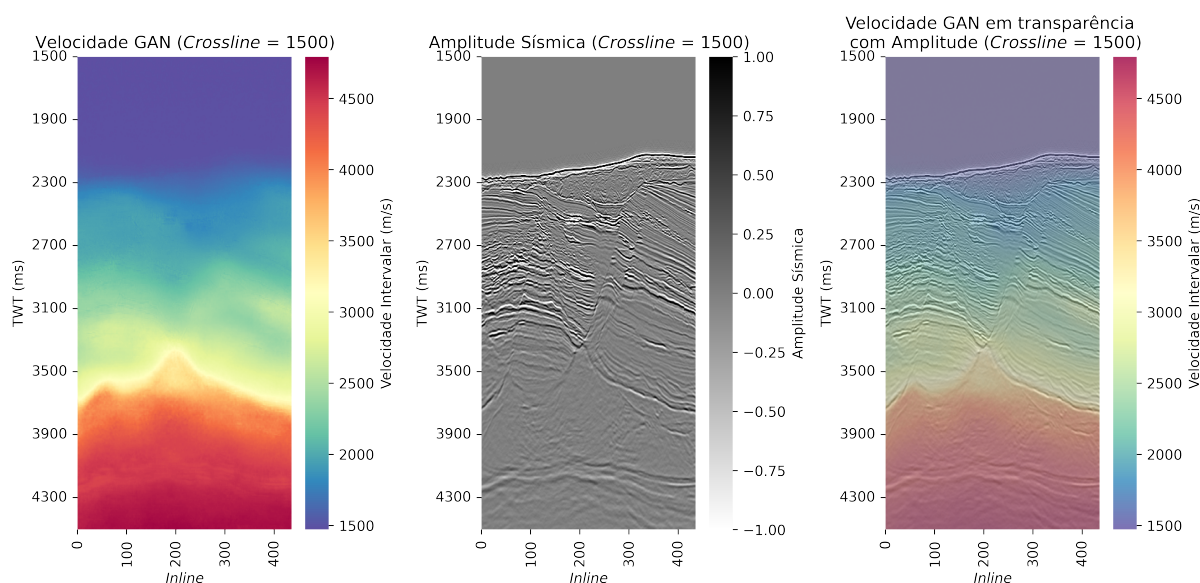
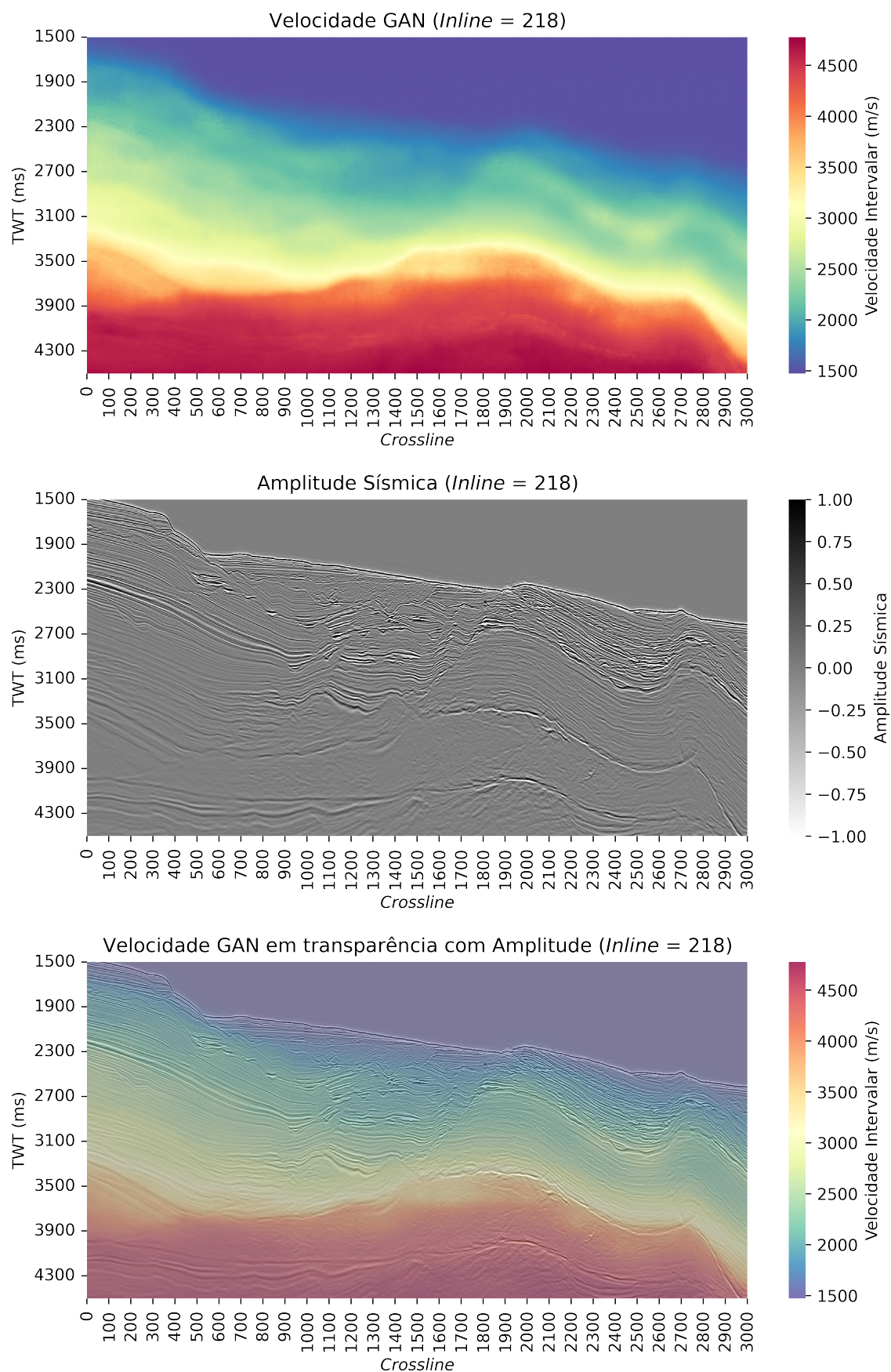
Figura 55 – *Crossline* 1500: Velocidade estimada pelo Modelo Gerador, Amplitude Sísmica e Velocidade estimada pelo Modelo Gerador em transparência com a Amplitude Sísmica. Nota-se uma coerência das principais estruturas geológicas presentes nos dois volumes.

Figura 56 – *Inline* 218: Velocidade estimada pelo Modelo Gerador, Amplitude Sísmica e Velocidade estimada pelo Modelo Gerador em transparência com a Amplitude Sísmica. Nota-se uma coerência das principais estruturas geológicas presentes nos dois volumes.



5.2 Discussões

Trabalhos anteriores mostraram a viabilidade de se gerar modelos de velocidades intervalares por meio de redes neurais tendo como entrada dados sísmicos sintéticos 1D e 2D empilhados (MOSSER et al., 2018; PUZYREV et al., 2019; LI et al., 2020). Os resultados, apesar de promissores, foram moderados, apresentando uma qualidade inferior aos modelos de velocidade gerados pelos métodos tradicionais de inversão sísmica.

Saraiva et al. (2021) utilizaram dados reais para o treinamento de uma cGAN 2D e mostraram que somente o dado sísmico empilhado não possui informação suficiente que permita a rede neural encontrar uma regra para a geração de modelos de velocidade intervalares equivalentes aos produzidos utilizando a técnica FWI. Foi necessário adicionar informações complementares, no caso, a velocidade média advinda da tomografia sísmica e o gride de tempo duplo de trânsito, para que a cGAN fosse capaz de gerar modelos de velocidades intervalares com qualidade equivalente aos gerados utilizando a técnica FWI. Saraiva et al. (2021) argumentam que redes neurais são melhores interpoladores do que extrapoladores (BARNARD; WESSELS, 1992; XU et al., 2020) e as informações adicionais, de baixa frequência, permitem a interpolação com o dado sísmico empilhado, com um conteúdo de frequências maior, de forma similar ao que é realizado, de forma implícita, utilizando o método FWI tradicional (Figura 12).

Este trabalho utilizou uma cGAN 3D e a mesma estratégia de Saraiva et al. (2021), utilizando dados adicionais de velocidade média tomográfica e *grid* de tempo duplo de trânsito, obtendo um resultado de qualidade similar, contudo volumétrico.

A tomografia sísmica é um problema mal-posto (JONES, 2010), existem múltiplas soluções de velocidades intervalares que honram a mesma velocidade média. O uso de restrições impostas pelo modelo geológico conceitual e dados de poços visam favorecer a convergência para um modelo de velocidades intervalares geologicamente plausível, o que nem sempre é possível. Desta forma o uso da velocidade média foi adotado, por ser uma informação mais confiável. Além disso, a velocidade média em um ponto do volume trás consigo informação sobre a velocidade das camadas sobrejacentes, por consequência, os sub-volumes de treinamento extraídos de forma aleatória carregam informação indireta sobre seu posicionamento, auxiliando na convergência da rede.

Assim como a velocidade média, o gride de tempo duplo de trânsito também foi adicionado com o intuito de fornecer informação espacial, pois arquiteturas de redes como *transformes* (VASWANI et al., 2017) e *visual transformers* (DOSOVITSKIY et al., 2021) utilizam estratégia similar com sucesso, ao adicionar informação com referência espacial ao dado de entrada por meio dos *positional encodings* e *position embeddings*, respectivamente.

A referência espacial fornecida pelo tempo duplo de trânsito, junto com a informação de velocidade média pode ser utilizada pela rede para obter informação de profundidade.

Como existe uma relação direta entre aumento de velocidade e espessura de soterramento, devido ao efeito de compactação, essa informação pode ser útil.

Os canais adicionais atuam como regularizadores, limitando o espaço de soluções possíveis e permitem que a cGAN capture a distribuição espacial da velocidade intervalar FWI utilizada para treinamento.

A estimativa de um modelo de velocidades intervalares utilizando a técnica FWI fornece melhores resultados em regiões mais rasas, onde os cabos da aquisição sísmica possuem comprimento suficiente para capturar as *diving waves* (ZHOU et al., 2015), fora dessas regiões os resultados são descartados e substituídos pelo dado da velocidade intervalar oriundo da tomografia sísmica. Este fenômeno foi capturado durante o treinamento, e os valores de velocidade intervalares produzidos pela Rede Geradora apresentaram uma melhor qualidade nas regiões mais rasas (figuras 53, 52, 51 e 54).

Cada dado sísmico possui uma assinatura característica, essa assinatura vai depender da forma como o dado foi adquirido e do fluxo de processamento adotado. Uma forma de expressar essa assinatura é pela estimativa da *wavelet* média necessária para deconvolver o traço sísmico de forma a retornar um modelo de refletividades sísmica (Figura 6). O uso da Rede Geradora em um dado sísmico com assinatura sísmica muito diferente da assinatura sísmica do dado utilizado para treinamento pode resultar em resultados insatisfatórios.

Após treinamento, a Rede Geradora é capaz de gerar volumes de velocidades intervalares em questão de minutos. Em contraste, o processo de construção de um modelo de velocidades intervalares utilizando a FWI pode levar meses e até mesmo anos devido a grande quantidade de simulações numéricas necessárias para a atualização iterativa do modelo de velocidades intervalares inicial (TARANTOLA, 1984). Contudo, para realizar uma comparação justa, seria necessário utilizar algoritmos de FWI implementados para a execução em GPUs, o que ainda está em fase de pesquisa (WANG et al., 2011; MAO; WU; WANG, 2012; FANG et al., 2020) e não disponíveis em escala industrial.

6 Conclusões e trabalhos futuros

A proposta deste trabalho, de utilizar uma GAN condicional 3D do tipo *pix2pix*, para gerar modelos de velocidades intervalares de alta resolução se mostrou eficaz. Modelos de velocidades intervalares, equivalentes aos produzidos pelo método FWI, foram produzidos pela Rede Geradora utilizando os dados de validação e de inferência em questão de minutos. A qualidade do dado foi atestada por métricas qualitativas (análise visual) e quantitativas (SSIM e MAPE).

A utilização dos três canais (Amplitude Sísmica, Velocidade Média Tomográfica e Gride de Tempo Duplo de Trânsito) como entrada foi vital para que a Rede Geradora fosse capaz de produzir os resultados pretendidos.

A implementação da cGAN 3D e a estratégia de reconstrução do volume de velocidades utilizando a sobreposição e janelamento Hann dos sub-volumes de saída permitiu gerar resultados com alta qualidade e sem a presença de marcas ou artefatos.

Trabalhos futuros podem investigar o efeito de se inserir mais dados nos canais de entrada, a exemplo: interpretação sísmica de horizontes e falhas, informação de longitude e latitude para cada ponto do gride e a utilização de volumes de empilhamento parcial *near*, *mid* e *far*.

É necessário avaliar o efeito de treinar a rede proposta utilizando um *dataset* composto por dados de múltiplas aquisições e processamentos sísmico.

Outras arquiteturas de redes geradoras têm ganhado notoriedade na tarefa de síntese de imagens e merecem ser comparadas com a solução proposta neste trabalho, como por exemplo de as redes auto-codificadoras variacionais e os modelos de difusão.

Referências

- AAPG. *Hydrocarbon migration*. 2022. Disponível em: <https://wiki.aapg.org/Hydrocarbon_migration>. Citado na página 22.
- ABBASI, B.; GOLDENHOLZ, D. M. Machine learning applications in epilepsy. *Epilepsia*, v. 60, p. 2037 – 2047, 2019. Citado na página 19.
- ADLER, A.; ARAYA-POLO, M.; POGGIO, T. A. Deep learning for seismic inverse problems: Toward the acceleration of geophysical analysis workflows. *IEEE Signal Processing Magazine*, v. 38, p. 89–119, 2021. Citado na página 60.
- ANP. *BRAZILIAN EXPLORATORY OPPORTUNITIES STRATEGIES TO IMPROVE ATTRACTIVENESS*. 2020. Disponível em: <<https://www.gov.br/anp/pt-br/centrais-de-conteudo/apresentacoes-e-palestras/2020/arquivos/02-07-2020-brazilian-exploratory-opportunities-strategies.pdf>>. Citado na página 34.
- ARAYA-POLO, M.; FARRIS, S.; FLOREZ, M. A. Deep learning-driven velocity model building workflow. *The Leading Edge*, 2019. Citado na página 60.
- ARIAS, P. A. et al. Technical summary. *Climate Change*, p. 33–144, 2021. Citado na página 19.
- BARNARD, E.; WESSELS, L. Extrapolation and interpolation in neural network classifiers. *IEEE Control Systems*, v. 12, p. 50–53, 1992. Citado na página 79.
- BARON, D. Machine learning in astronomy: A practical overview. *arXiv preprint arXiv:1904.07248*, 2019. Citado na página 19.
- BIONDI, B. L. *3D seismic imaging*. [S.l.]: Society of Exploration Geophysicists, 2006. Citado 3 vezes nas páginas 26, 27 e 30.
- BISHOP, C. M.; NASRABADI, N. M. Pattern recognition and machine learning. *J. Electronic Imaging*, v. 16, p. 049901, 2007. Citado 2 vezes nas páginas 40 e 41.
- BISWAS, R. S. et al. Estimating normal moveout velocity using the recurrent neural network. *Interpretation*, 2019. Citado na página 31.
- BORJI, A. Pros and cons of gan evaluation measures. *Comput. Vis. Image Underst.*, v. 179, p. 41–65, 2019. Citado na página 59.
- BRUHN, C. H. et al. Campos and santos basins: 40 years of reservoir characterization and management of shallow-to ultra-deep water, post-and pre-salt reservoirs-historical overview and future challenges. In: ONEPETRO. *OTC Brasil*. [S.l.], 2017. Citado na página 22.
- CALDERÓN-MACÍAS, C.; SEN, M. K.; STOFFA, P. L. Automatic nmo correction and velocity estimation by a feedforward neural network. *Geophysics*, v. 63, p. 1696–1707, 1998. Citado na página 31.
- CENGEL, Y.; CIMBALA, J. *EBOOK: Fluid Mechanics Fundamentals and Applications*

- (*SI units*). [S.l.]: McGraw Hill, 2013. Citado na página 27.
- CHAURIS, H.; NOBLE, M. Testing the behavior of differential sem blance for velocity estimation. In: *SEG Technical Program Expanded Abstracts 1998*. [S.l.]: Society of Exploration Geophysicists, 1998. p. 1305–1308. Citado na página 31.
- CHOLLET, F. *Building Autoencoders in Keras*. 2016. Disponível em: <<https://blog.keras.io/building-autoencoders-in-keras.html>>. Citado na página 51.
- CHOLLET, F. *How convolutional neural networks see the world*. 2016. Disponível em: <<https://blog.keras.io/how-convolutional-neural-networks-see-the-world.html>>. Citado na página 50.
- CIESLEWICZ, D.; LAWTON, D. C. Receiver notching in a linear $v(z)$ near-surface medium. *CREWES Research Report*, v. 10, p. 3–1, 1998. Citado na página 27.
- CIRILLO, M. D.; ABRAMIAN, D.; EKLUND, A. Vox2vox: 3d-gan for brain tumour segmentation. *ArXiv*, abs/2003.13653, 2020. Citado 4 vezes nas páginas 15, 67, 68 e 69.
- CLAERBOUT, J. F. *Imaging the earth's interior*. [S.l.]: Blackwell scientific publications Oxford, 1985. v. 1. Citado na página 33.
- COVERT, T.; GREENSTONE, M.; KNITTEL, C. R. Will we ever stop using fossil fuels? *Journal of Economic Perspectives*, v. 30, n. 1, p. 117–38, 2016. Citado na página 19.
- DAS, V. et al. Convolutional neural network for seismic impedance inversion. *GEOPHYSICS*, 2019. Citado na página 61.
- DENNIS, S. *The Hopfield Network: Descent on an Energy Surface*. 1997. Disponível em: <<https://staff.itee.uq.edu.au/janetw/cmc/chapters/Hopfield/>>. Citado 3 vezes nas páginas 54, 55 e 56.
- DEVLIN, J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018. Citado na página 35.
- DOSOVITSKIY, A. et al. An image is worth 16x16 words: Transformers for image recognition at scale. *ArXiv*, abs/2010.11929, 2021. Citado na página 79.
- DROEGE, P. *Urban energy transition: from fossil fuels to renewable power*. [S.l.]: Elsevier, 2011. Citado na página 19.
- DROUIN, E.; PILOQUET, P.; PÉREON, Y. The first illustrations of neurons by camillo golgi. *The Lancet Neurology*, Elsevier, v. 14, n. 6, p. 567, 2015. Citado 2 vezes nas páginas 36 e 37.
- EARLE, S. A. *Physical Geology – 2nd Edition*. British Columbia: BCcampus Open Education, 2015. Citado 2 vezes nas páginas 22 e 23.
- EGMONT-PETERSEN, M.; RIDDER, D. de; HANDELS, H. Image processing with neural networks—a review. *Pattern recognition*, Elsevier, v. 35, n. 10, p. 2279–2301, 2002. Citado na página 35.
- EQUINOR. *segio*. [S.l.]: GitHub, 2022. <<https://github.com/equinor/segio>>. Citado na página 63.

- EREN, L. Bearing fault detection by one-dimensional convolutional neural networks. *Mathematical Problems in Engineering*, Hindawi, v. 2017, 2017. Citado na página 35.
- ESSA, I. *CS 6475: Computational Photography*. 2014. Disponível em: <<https://github.com/vishalpolley/Computational-Photography/blob/master/Slides/IrfanEssa-CP-02-5-Filtering.pdf>>. Citado na página 48.
- FANG, J. et al. Elastic full-waveform inversion based on gpu accelerated temporal fourth-order finite-difference approximation. *Comput. Geosci.*, v. 135, p. 104381, 2020. Citado na página 80.
- FERREIRA, R. S. et al. Automatic velocity analysis using a hybrid regression approach with convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, v. 59, p. 4464–4470, 2021. Citado na página 31.
- GHODDUSI, H.; CREAMER, G. G.; RAFIZADEH, N. Machine learning in energy economics and finance: A review. *Energy Economics*, Elsevier, v. 81, p. 709–727, 2019. Citado na página 19.
- GOLDBERG, Y. Neural network methods for natural language processing. *Synthesis lectures on human language technologies*, Morgan & Claypool Publishers, v. 10, n. 1, p. 1–309, 2017. Citado na página 35.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <<http://www.deeplearningbook.org>>. Citado 2 vezes nas páginas 44 e 50.
- GOODFELLOW, I. et al. Generative adversarial nets. In: GHAHRAMANI, Z. et al. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2014. v. 27. Disponível em: <<https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf>>. Citado 4 vezes nas páginas 36, 56, 57 e 58.
- GRAY, E. G. Axo-somatic and axo-dendritic synapses of the cerebral cortex: An electron microscope study. *Journal of Anatomy*, v. 93, p. 420–433, 1959. Citado na página 36.
- HAGELUND, R.; LEVIN, S. A. Seg-y_r2. 0: Seg-y revision 2.0 data exchange format. *Society of Exploration Geophysicists, Houston*, 2017. Citado na página 63.
- HASTIE, T. J.; TIBSHIRANI, R.; FRIEDMAN, J. H. The elements of statistical learning. In: . [S.l.: s.n.], 2001. Citado na página 45.
- HAVSKOV, J.; ALGUACIL, G. Instrumentation in earthquake seismology. *Instrumentation in Earthquake Seismology*, 2005. Citado na página 23.
- HE, K. et al. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 770–778, 2016. Citado na página 57.
- HEBB, D. O. *The organization of behavior: A neuropsychological theory*. [S.l.]: Joh Wiley Sons, Inc., 1949. Citado na página 54.
- HERTZ, J. A.; KROGH, A.; PALMER, R. G. Introduction to the theory of neural computation. In: *The advanced book program*. [S.l.: s.n.], 1991. Citado 2 vezes nas páginas 52 e 53.

- HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, v. 79 8, p. 2554–8, 1982. Citado 4 vezes nas páginas 36, 52, 53 e 55.
- HORTSCH, M. A short history of the synapse – golgi versus ramón y cajal. In: . [S.l.: s.n.], 2009. Citado na página 36.
- HUEMBELI, P. et al. The physics of energy-based models. *Quantum Machine Intelligence*, v. 4, p. 1–13, 2022. Citado 3 vezes nas páginas 36, 55 e 56.
- IRENA. World energy transitions outlook: 1.5°C pathway. In: INTERNATIONAL RENEWABLE ENERGY AGENCY,. *World Energy Transitions Outlook 2022: 1.5°C Pathway*. [S.l.], 2022. Citado na página 19.
- ISOLA, P. et al. Image-to-image translation with conditional adversarial networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 5967–5976, 2017. Citado 9 vezes nas páginas 15, 20, 51, 57, 58, 59, 60, 67 e 69.
- JONES, I. F. Tutorial: Velocity estimation via ray-based tomography. *First Break*, v. 28, 2010. Citado na página 79.
- KADRA, A. et al. Well-tuned simple nets excel on tabular datasets. In: RANZATO, M. et al. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2021. v. 34, p. 23928–23941. Disponível em: <<https://proceedings.neurips.cc/paper/2021/file/c902b497eb972281fb5b4e206db38ee6-Paper.pdf>>. Citado na página 45.
- KRAMES, C. *Chemical synapse schema - domínio público*. 2022. Disponível em: <https://commons.wikimedia.org/wiki/File:Chemical_synapse_schema_cropped.jpg>. Citado na página 38.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, v. 60, p. 84 – 90, 2012. Citado 2 vezes nas páginas 35 e 57.
- LAILLY, P.; BEDNAR, J. The seismic inverse problem as a sequence of before stack migrations. 1983. Citado na página 32.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proc. IEEE*, v. 86, p. 2278–2324, 1998. Citado 4 vezes nas páginas 13, 35, 46 e 47.
- LI, S. et al. Deep-learning inversion of seismic data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 58, p. 2135–2149, 2020. Citado 2 vezes nas páginas 60 e 79.
- LINDSAY, G. W. Convolutional neural networks as a model of the visual system: Past, present, and future. *Journal of Cognitive Neuroscience*, MIT Press - Journals, v. 33, n. 10, p. 2017–2031, sep 2021. Disponível em: <https://doi.org/10.1162%2Fjocn_a_01544>. Citado na página 49.
- LINER, C. L. Concepts of normal and dip movement. *Geophysics*, v. 64, p. 1637–1647, 1999. Citado na página 34.
- LIU, G. et al. Stacking seismic data using local correlation. *Geophysics*, v. 74, 2009. Citado na página 32.

- LUČIĆ, V. Cryo-electron tomography and synaptic transmission in the brain. *Research Outreach*, 2020. Citado na página 37.
- MAGOON, L. B.; DOW, W. G. The petroleum system-from source to trap. *AAPG Bulletin (American Association of Petroleum Geologists);(United States)*, v. 75, n. CONF-910403-, 1994. Citado na página 22.
- MAO, J.; WU, R. shan; WANG, B. Multiscale full waveform inversion using gpu. *Seg Technical Program Expanded Abstracts*, 2012. Citado na página 80.
- MASLAND, R. H. Neuronal cell types. *Current Biology*, v. 14, p. R497–R500, 2004. Citado na página 36.
- MAUL, A. et al. Benefits of inserting salt stratification to detail velocity model prior to least-squares reverse-time migration. *Journal of Applied Geophysics*, Elsevier, v. 195, p. 104469, 2021. Citado na página 34.
- MAUL, A. R. Caracterização sísmica das estratificações da seção evaporítica salina e suas aplicações nos projetos de exploração, desenvolvimento e produção de hidrocarbonetos. Niterói, 2020. Citado 2 vezes nas páginas 19 e 34.
- 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 5 vezes nas páginas 13, 35, 37, 38 e 39.
- MINSKY, M.; PAPERT, S. Perceptrons - an introduction to computational geometry. In: . [S.l.: s.n.], 1969. Citado 2 vezes nas páginas 40 e 41.
- MIRZA, M.; OSINDERO, S. Conditional generative adversarial nets. *ArXiv*, abs/1411.1784, 2014. Citado 2 vezes nas páginas 14 e 58.
- MOSSER, L. et al. Rapid seismic domain transfer: Seismic velocity inversion and modeling using deep generative neural networks. *ArXiv*, abs/1805.08826, 2018. Citado 2 vezes nas páginas 61 e 79.
- MULLER, A. P. O. et al. Deep-tomography: iterative velocity model building with deep learning. In: . [S.l.: s.n.], 2022. Citado na página 51.
- NASSIF, A. B. et al. Speech recognition using deep neural networks: A systematic review. *IEEE access*, IEEE, v. 7, p. 19143–19165, 2019. Citado na página 35.
- NATIONAL MUSEUM OF THE U.S. NAVY. *MARK I perceptron*. 1960. Disponível em: <[https://commons.wikimedia.org/wiki/File:330-PSA-80-60_\(USN_710739\)_\(20897323365\).jpg](https://commons.wikimedia.org/wiki/File:330-PSA-80-60_(USN_710739)_(20897323365).jpg)>. Citado na página 41.
- NIELSEN, M. A. *Neural networks and deep learning*<<http://neuralnetworksanddeeplearning.com/>>. [S.l.]: Determination press San Francisco, CA, USA, 2019. Citado 2 vezes nas páginas 42 e 43.
- NOBEL PRIZE OUTREACH. *The Nobel Prize in Physiology or Medicine 1906*. 2022. Disponível em: <<https://www.nobelprize.org/prizes/medicine/1906/summary/>>. Citado na página 36.
- PARK, J. *pytorch-ssim for 3D images*. 2019. Disponível em: <<https://github.com/>>

[jinh0park/pytorch-ssim-3D](#)>. Citado na página 72.

PASZKE, A. et al. Pytorch: An imperative style, high-performance deep learning library. In: WALLACH, H. et al. (Ed.). *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019. p. 8024–8035. Disponível em: <<http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>>. Citado na página 67.

PEDREGOSA, F. et al. *sklearn.model_selection.TimeSeriesSplit*. 2011. Disponível em: <https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.TimeSeriesSplit.html>. Citado na página 63.

PENNA, R. et al. Challenges in seismic imaging and reservoir characterization of presalt oilfields in offshore brazil. In: ONEPETRO. *Offshore Technology Conference*. [S.l.], 2013. Citado na página 34.

PEREIRA-DIAS, B. et al. Exploring inversion strategies in image domain least squares migration. In: EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS. *First EAGE/SBGf Workshop on Least-Squares Migration*. [S.l.], 2018. v. 2018, n. 1, p. 1–5. Citado na página 34.

PÉREZ-GARCÍA, F.; SPARKS, R.; OURSELIN, S. Torchio: a python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Computer Methods and Programs in Biomedicine*, p. 106236, 2021. ISSN 0169-2607. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169260721003102>>. Citado na página 65.

PETROBRAS. *Caderno de Mudança do Clima*. 2022. Disponível em: <<https://petrobras.com.br/pt/sociedade-e-meio-ambiente/meio-ambiente/mudancas-do-clima/>>. Citado na página 19.

PUZYREV, V. et al. Seismic inversion with deep neural networks: a feasibility analysis. *81st EAGE Conference and Exhibition 2019*, 2019. Citado 2 vezes nas páginas 60 e 79.

RADFORD, A.; METZ, L.; CHINTALA, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *CoRR*, abs/1511.06434, 2016. Citado na página 58.

RAIKAR, S.; ADAMSON, S. *Renewable energy finance: theory and practice*. [S.l.]: Academic Press, 2019. Citado na página 19.

RAISSI, M. Deep hidden physics models: Deep learning of nonlinear partial differential equations. *The Journal of Machine Learning Research*, JMLR. org, v. 19, n. 1, p. 932–955, 2018. Citado na página 20.

RAISSI, M.; PERDIKARIS, P.; KARNIADAKIS, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, Elsevier, v. 378, p. 686–707, 2019. Citado 2 vezes nas páginas 20 e 35.

RAMÓN Y CAJAL, S. *Drawing of Purkinje cells (A) and granule cells (B) from pigeon cerebellum - domínio público*. 1899. Disponível em: <<https://commons.wikimedia.org/wiki/File:PurkinjeCell.jpg>>. Citado na página 37.

- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. *ArXiv*, abs/1505.04597, 2015. Citado na página 51.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, American Psychological Association, v. 65, n. 6, p. 386, 1958. Citado 5 vezes nas páginas 13, 35, 39, 40 e 41.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, v. 323, p. 533–536, 1986. Citado 3 vezes nas páginas 41, 42 e 44.
- RUSSELL, B. H. *Introduction to seismic inversion methods*. [S.l.]: SEG Books, 1988. Citado na página 28.
- SAKURADA, M.; YAIRI, T. Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: *MLSDA '14*. [S.l.: s.n.], 2014. Citado na página 50.
- SALAKHUTDINOV, R. Learning deep generative models. *Annual Review of Statistics and Its Application*, Annual Reviews, v. 2, p. 361–385, 2015. Citado na página 55.
- SANTOS, J. G. dos; CATAPANG, A. N.; REYTA, E. D. Understanding the fundamentals of earthquake signal sensing networks. In: . [S.l.: s.n.], 2019. Citado na página 24.
- SARAIVA, M. et al. Data-driven full-waveform inversion surrogate using conditional generative adversarial networks. *2021 International Joint Conference on Neural Networks (IJCNN)*, 2021. Citado 3 vezes nas páginas 27, 60 e 79.
- SCHULTZ, P. S. The seismic velocity model as an interpretation asset. In: . [S.l.: s.n.], 1998. Citado na página 19.
- SCHUSTER, G. T. *Seismic inversion*. [S.l.]: Society of Exploration Geophysicists, 2017. Citado na página 31.
- SHARMA, R. et al. A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Computers & Operations Research*, Elsevier, v. 119, p. 104926, 2020. Citado na página 19.
- SHERIFF, R. E. *Encyclopedic dictionary of applied geophysics*. [S.l.]: Society of exploration geophysicists, 2002. Citado 3 vezes nas páginas 23, 26 e 31.
- SIDIROPOULOU, K.; PISSADAKI, E.; POIRAZI, P. Inside the brain of a neuron. *EMBO reports*, v. 7, p. 886–92, 10 2006. Citado na página 37.
- SIMM, R.; BACON, M.; BACON, M. *Seismic amplitude: An interpreter's handbook*. [S.l.]: Cambridge University Press, 2014. Citado na página 34.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2015. Citado 2 vezes nas páginas 49 e 57.
- SONG, H. et al. Seismic random noise suppression using deep convolutional autoencoder neural network. *Journal of Applied Geophysics*, Elsevier, v. 178, p. 104071, 2020. Citado na página 50.
- SUBSURFWIKI. *Gather*. 2022. Disponível em: <<https://subsurfwiki.org/wiki/Gather>>. Citado na página 30.

- TARANTOLA, A. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, v. 49, p. 1259–1266, 1984. Citado 3 vezes nas páginas 20, 32 e 80.
- THOMAS, J. E. *Velocidades Sísmica*. [S.l.: s.n.], 2003. Citado na página 23.
- THOMAS, J. E. et al. *Fundamentos de Engenharia de Petróleo*. [S.l.: s.n.], 2001. Citado 2 vezes nas páginas 22 e 23.
- UNFCCC. The paris agreement. In: *Report of the Conference of the Parties to the United Nations Framework Convention on Climate Change (21st Session, 2015: Paris)*. Retrived December. [S.l.: s.n.], 2015. v. 4, p. 2017. Citado na página 19.
- UPSEIS. *Seismology*. 2022. Disponível em: <<https://www.mtu.edu/geo/community/seismology/learn/seismology-study/>>. Citado na página 24.
- VASWANI, A. et al. Attention is all you need. *ArXiv*, abs/1706.03762, 2017. Citado na página 79.
- VIGH, D. et al. Efficiently using full-waveform inversion in the velocity model building flow. In: EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS. *78th EAGE Conference and Exhibition 2016*. [S.l.], 2016. v. 2016, n. 1, p. 1–5. Citado na página 32.
- VIRIEUX, J. et al. An introduction to full waveform inversion. In: *Encyclopedia of exploration geophysics*. [S.l.]: Society of Exploration Geophysicists, 2017. p. R1–1. Citado na página 33.
- VIRTANEN, P. et al. *scipy.signal.windows.hann*. 2020. Disponível em: <<https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.windows.hann.html>>. Citado na página 70.
- WANG, B. li et al. Cuda-based acceleration of full waveform inversion on gpu. *Seg Technical Program Expanded Abstracts*, 2011. Citado na página 80.
- WANG, Y. et al. Seismic trace interpolation for irregularly spatial sampled data using convolutional autoencoder. *Geophysics*, Society of Exploration Geophysicists, v. 85, n. 2, p. V119–V130, 2020. Citado na página 50.
- WANG, Y.; YAO, H.; ZHAO, S. Auto-encoder based dimensionality reduction. *Neurocomputing*, Elsevier, v. 184, p. 232–242, 2016. Citado na página 50.
- WANG, Z. et al. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, v. 13, p. 600–612, 2004. Citado na página 72.
- WNUK, A. et al. Brain facts: A primer on the brain and nervous system. 2018. Citado 2 vezes nas páginas 36 e 37.
- WU, X. et al. Faultseg3d: Using synthetic data sets to train an end-to-end convolutional neural network for 3d seismic fault segmentation. *GEOPHYSICS*, 2019. Citado na página 51.
- WU, Y.; HE, K. Group normalization. *International Journal of Computer Vision*, v. 128, p. 742–755, 2019. Citado na página 49.

- WU, Y.; LIN, Y.; ZHOU, Z. Inversionet: Accurate and efficient seismic-waveform inversion with convolutional neural networks. *SEG Technical Program Expanded Abstracts 2018*, 2018. Citado na página 60.
- XU, K. et al. How neural networks extrapolate: From feedforward to graph neural networks. *ArXiv*, abs/2009.11848, 2020. Citado na página 79.
- YILMAZ, Ö. *Seismic data analysis: Processing, inversion, and interpretation of seismic data*. [S.l.]: Society of exploration geophysicists, 2001. Citado 5 vezes nas páginas 26, 30, 31, 33 e 34.
- ZHANG, R.; LIU, Y.; SUN, H. Physics-guided convolutional neural network (phycnn) for data-driven seismic response modeling. *Engineering Structures*, Elsevier, v. 215, p. 110704, 2020. Citado 2 vezes nas páginas 20 e 35.
- ZHENG, Y. et al. Applications of supervised deep learning for seismic interpretation and inversion. *The Leading Edge*, 2019. Citado na página 60.
- ZHOU, H.-W. et al. Reverse time migration: A prospect of seismic imaging methodology. *Earth-Science Reviews*, v. 179, p. 207–227, 2018. Citado na página 34.
- ZHOU, W. et al. Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation. *Geophysical Journal International*, v. 202, p. 1535–1554, 2015. Citado na página 80.
- ZHU, J.-Y. et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. *2017 IEEE International Conference on Computer Vision (ICCV)*, p. 2242–2251, 2017. Citado na página 61.
- ZHU, Y. et al. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *ArXiv*, abs/1901.06314, 2019. Citado 2 vezes nas páginas 20 e 35.
- ZUREKS, C. V. *Magnetic domain with arrows by Zureks*. 2022. Disponível em: <https://commons.wikimedia.org/wiki/File:Magnetic_domain_with_arrows_by_Zureks.png>. Citado na página 53.

Apêndices

APÊNDICE A – Publicação 1

Trabalho apresentado na *International Joint Conference on Neural Networks 2021* (IJCNN) intitulado *Data-driven Full-waveform Inversion Surrogate using Conditional Generative Adversarial Networks* (doi: 10.1109/IJCNN52387.2021.9534128) e parte integrante desta dissertação.

Data-driven Full-waveform Inversion Surrogate using Conditional Generative Adversarial Networks

Marcus Saraiva^{a,c,†}, Avelino Forechi^{b,†}, Jorcy de Oliveira Neto^{c,†}, Antônio DelRey^{c,†}, Thomas Rauber^{a,*}

^a*Departamento de Informática – Centro Tecnológico
Universidade Federal do Espírito Santo
Vitória, Espírito Santo, Brazil
thomas.rauber@ufes.br

^b*Coordenadoria de Engenharia Mecânica
Instituto Federal do Espírito Santo
Aracruz, Espírito Santo, Brazil
†avelino.forechi@ifes.edu.br*

^c*Exploração e Produção
Petróleo Brasileiro S.A.
Vitória, Espírito Santo, Brazil
‡{mvsaraiva, jorcyneto, delrey}@petrobras.com.br*

Abstract—In the Oil and Gas industry, estimating a subsurface velocity field is an essential step in seismic processing, reservoir characterization, and hydrocarbon volume calculation. Full-waveform inversion (FWI) velocity modeling is an iterative advanced technique that provides an accurate and detailed velocity field model, although at a very high computational cost due to the physics-based numerical simulations required at each FWI iteration. In this study, we propose a method of generating velocity field models, as detailed as those obtained through FWI, using a conditional generative adversarial network (cGAN) with multiple inputs. The primary motivation of this approach is to circumvent the extremely high cost of full-waveform inversion velocity modeling. Real-world data were used to train and test the proposed network architecture, and three evaluation metrics (percent error, structural similarity index measure, and visual analysis) were adopted as quality criteria. Based on these metrics, the results evaluated upon the test set suggest that the GAN was able to accurately match real FWI generated outputs, enabling it to extract from input data the main geological structures and lateral velocity variations. Experimental results indicate that the proposed method, when deployed, has the potential to increase the speed of geophysical reservoir characterization processes, saving on time and computational resources.

Index Terms—Conditional generative adversarial network (cGAN), full-waveform inversion (FWI), seismic velocity modeling.

I. INTRODUCTION

Oil industry applications utilize seismic reflection as a primary indirect method for investigating the properties of subsurface strata. This method uses a controlled source of acoustic waves and receiver arrays (geophones or hydrophones) to record the transit time and amplitude of acoustic waves reflected by contrasting rock type interfaces (Fig. 1).

During seismic acquisition, the same subsurface region is sampled multiple times from different offsets, resulting in data redundancy. The vertical projection of this region on the surface is called the common midpoint (CMP). Sorting the

We thank Petrobras S.A. for providing computational resources and the data used in this publication.

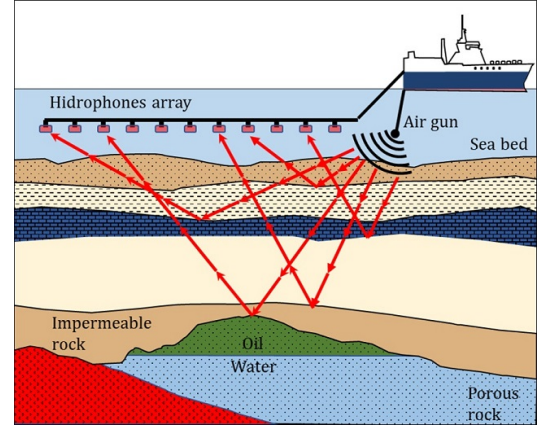


Fig. 1. Schematic representation of a marine seismic acquisition.

data by CMP, we can then obtain the so-called CMP gathers (Fig. 2).

To obtain an image of the subsurface area, it is then necessary to process the recorded data (transit time and amplitude of the reflected acoustic waves). The estimation of an accurate velocity model is an essential step in seismic processing. This model is used in several steps of the seismic processing, including seismic migration, normal moveout (NMO) correction, and time-to-depth conversion.

Seismic migration is a procedure that aims to "migrate" seismic events to their correct position and collapse reflections and diffractions, creating a seismic image of the subsurface [1]. NMO correction adjusts all CMP seismic traces to a theoretical zero offset (source and receiver at the same position) [2]. This procedure enables the use of data redundancy to increase signal-to-noise ratio through CMP-corrected gather stacking (Fig. 2). The velocity model is also important for time-to-depth conversion, a necessary step for reservoir characterization and oil volume estimation [3].

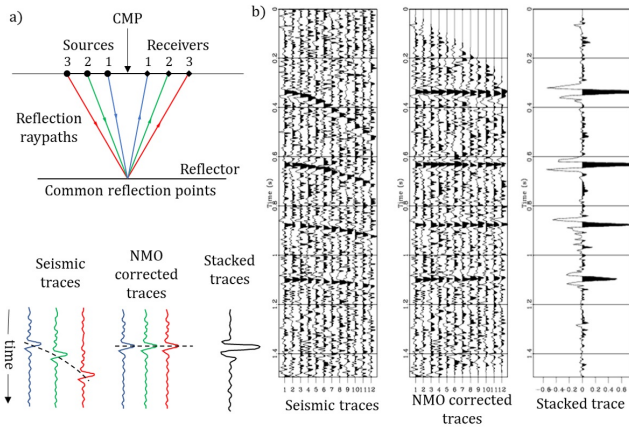


Fig. 2. a) Schematic representation of normal moveout correction; b) Process of stacking a synthetic CMP gather after normal moveout correction to improve the signal-to-noise ratio (Adapted from [4]).

Complex geological settings such as salt domes and folded or faulted regions require detailed velocity models in order to form a satisfactory seismic image. Currently, velocity modeling using full-waveform inversion (FWI) is an advanced technique providing an accurate and detailed velocity model [5].

In this paper, we present a method using an image-to-image translator (pix2pix) [6] based on a conditional generative adversarial network (cGAN) to translate an input (post-stack seismic image, average tomographic velocity, and two-way time grid) into an interval velocity model as detailed as FWI.

A. Seismic data and velocity model

The primary outcome of 3D seismic processing is a seismic image volume. Seismic volume slices in the acquisition direction are conventionally referred to as inlines, slices perpendicular to the acquisition direction as crosslines, slices perpendicular to the vertical axis as time slices, and a vertical sample vector as a seismic trace (Fig. 3).

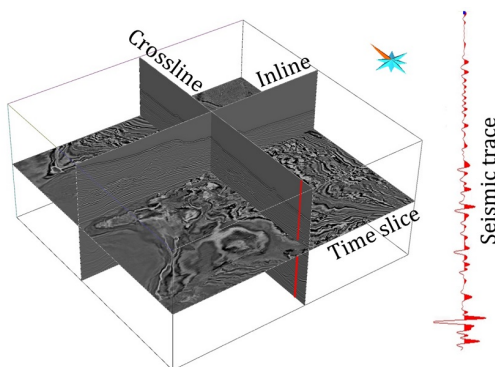


Fig. 3. 3D representation of seismic image and main slicing directions: inline (parallel to seismic acquisition direction), crossline (perpendicular to seismic acquisition direction) and time slice (perpendicular to vertical axis). The red line shows a seismic trace (vertical sample vector).

Seismic volume is among the most important inputs to characterizing hydrocarbon reservoirs. After the data's correlation with well logs, it is possible to map the main seismic horizons and delimit the top and bottom of hydrocarbon reservoirs.

A typical workflow for seismic velocity estimation consists of three steps, as follows.

- 1) **Normal moveout (NMO) velocity estimation** is performed by manually picking velocities that best horizontalize hyperbola reflections on CMP gathers (Fig. 4). This process returns a very smooth average NMO velocity model that can be approximated by a root mean square velocity for small offsets.
- 2) **Ray-based or grid tomography** is an iterative technique using seismic reflection's travel time measurements and associated amplitudes to calculate seismic velocity [7]. Seismic tomography delivers a low-frequency seismic acoustic velocity model, typically up to 2–3 Hz, and commonly takes the NMO velocity as the first input.
- 3) **Full-waveform inversion (FWI)** is a seismic inversion technique, initially proposed by [8] and [9], using pre-stack seismic data (raw data from seismic acquisition) to output a detailed interval velocity model.

FWI takes an initial low-resolution velocity model as input to generate synthetic pre-stack seismic data by wave equation numerical modeling. This synthetic model is then compared with real data from seismic acquisition. FWI iteratively updates this velocity model in order to minimize the difference between real and synthetic data (Fig. 5).

As FWI is performed with pre-stack seismic data (up to three orders of magnitude greater than the traditional seismic volume), it becomes computationally expensive and time-consuming. The use of this technique has only become industrially feasible in recent years, with an increase in processing power, although the theory was established in the 1980s [8], [9].

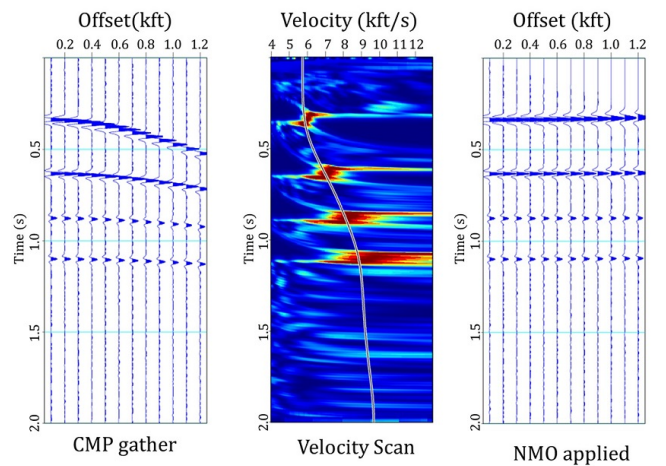


Fig. 4. Normal moveout velocity picking. The line passing by the red spots on the central image was manually selected to define the velocity that best horizontalize the reflections at the CMP gather (adapted from [10]).

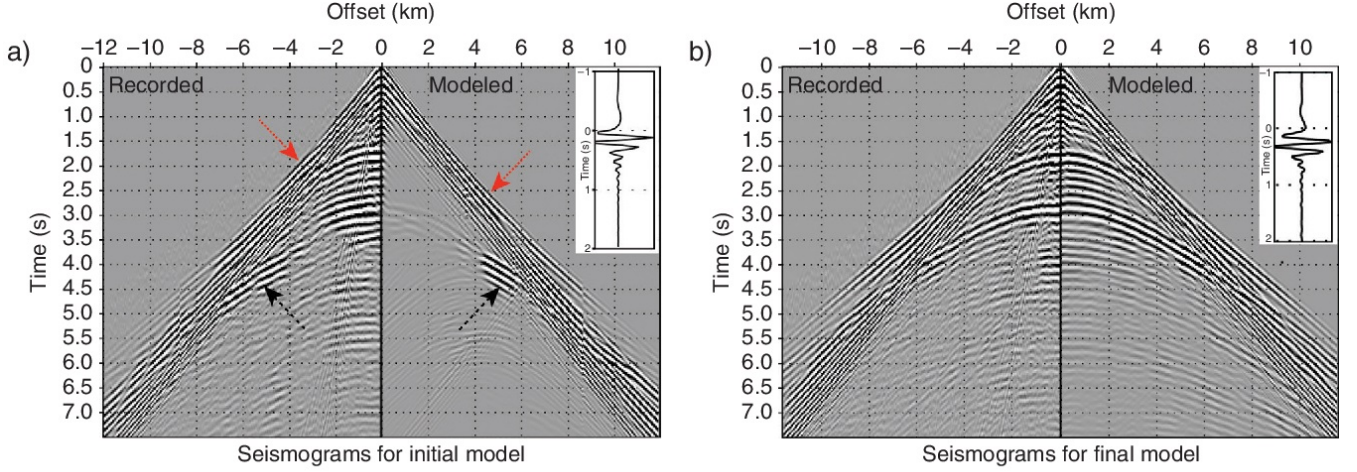


Fig. 5. FWI modeling: a) Real seismogram compared with synthetic seismogram generated from the initial velocity model and b) Real seismogram compared with synthetic seismogram generated from the final velocity model (From [11]).

Building an FWI velocity model may require from three months up to a year of construction, depending on the seismic acquisition size and geological subsurface complexity. Theoretically, FWI can estimate a high-resolution velocity model using as initial input a very smooth velocity model. In practice, to support the convergence and reduce the number of iterations, it is conventional to take a tomographic velocity model as input.

The FWI velocity model fills the gap between low-frequency tomographic velocity model and the seismic reflection image volume (Fig. 6), increasing the useful data bandwidth.

B. Generative Adversarial Networks (GAN)

Convolutional neural networks (CNN) are getting more and more popular in geophysical applications, typically in classification and segmentation tasks [12], [13], [14], [15], and [16]. Moreover, recent studies have demonstrated the feasibility of geophysical modeling using CNNs [17], [18], and [19].

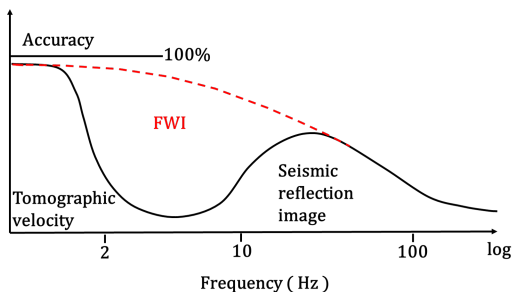


Fig. 6. Simplified representation of the information gap between tomographic velocity model and seismic reflection image volume. FWI velocity model fills this gap (Adapted from [20]).

Among CNN architectures, generative adversarial networks (GAN) go beyond traditional classification and segmentation tasks; GANs stand out for their ability to learn how to mimic data provided to them by capturing their statistical and spatial distributions [21].

GANs are generative models that learn to generalize a rule of how to transform a random noise vector z to output image y ($G : z \rightarrow y$) [21]. The pix2pix GAN architecture, proposed by [6] and adopted in this paper, is a conditional GAN (cGAN), which has as its input the sum of the random noise vector z with an image x , to find a rule of how to translate this input to image y ($G : \{x, z\} \rightarrow y$).

To accomplish this task, pix2pix uses two networks competing against each other. While one network generates synthesized images (a generator), the other network judges whether these images are valid (a discriminator). Generator G is trained to produce output images that are indistinguishable from the actual images by a discriminator D , which in turn is trained to improve its performance on detecting the generator's fake images.

In cGAN, the generator output y is connected directly to the discriminator input, together with the image x . Through the backpropagation process, discriminator classification in real or fake images provides a signal for the generator to update its weights and bias, thus improving performance. The objective function of a cGAN can be expressed by (1).

$$\mathcal{L}_{\text{cGAN}}(G, D) = \mathbb{E}_{x,y} [\log D(x, y)] + \mathbb{E}_{x,z} [\log(1 - D(x, G(x, z)))] , \quad (1)$$

where the generator G tries to minimize this objective function against an adversarial discriminator D that tries to maximize it (2).

$$G^* = \arg \min_G \max_D \mathcal{L}_{\text{cGAN}}(G, D). \quad (2)$$

Previous work [22] shows that L1 regularization is beneficial to this process, helping the generator G to create sharper images (3).

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z} \left[\|y - G(x, z)\|_1 \right]. \quad (3)$$

Thus, the final objective function for the pix2pix conditional GAN used in this work is (4).

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G). \quad (4)$$

Many image processing problems involve transforming one image into another. These problems are often addressed using algorithms specific to each situation. Conditional GANs are a general solution with good performance on a variety of problems [6]. Fig. 7 shows some generic examples using pix2pix cGAN.

II. MATERIALS AND METHODS

A. Data

From a marine seismic acquisition of approximately 680 km^2 in the Campos Basin, Brazil, we used four seismic volumes:

- Post stack seismic image
- Average tomographic velocity
- Two-way time (TWT) grid
- FWI interval velocity

We loaded the data, originally in seg-y format, as a 3D matrix in a Python environment using the Segyio library [23]. The shape of each matrix was $1451 \times 3001 \times 751$, where the x and y samples were spaced 12.5 m apart and z samples spaced every 0.004 s .

Preprocessing involved limiting data variability in two standard deviations to eliminate outliers, partially cropping regions without reflections (sea water thickness), and later data normalization using a global maximum and minimum as parameters (5).

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}. \quad (5)$$



Fig. 7. Examples using pix2pix GAN: aerial photos to maps, edges to photos, day to night, and gray scale to color (Adapted from [6]).

We divided the data into training and test sets at a ratio of 70% to 30%, respectively (Fig. 8). For the training set, we randomly sampled 2000 patches (512×512). For the test set, we randomly sampled 800 patches (512×512), both in the x direction (direction of acquisition).

The post-stack seismic image, average tomographic velocity, and TWT grid were combined to form a three-channel input image. The training pair was composed of this image and the FWI interval velocity (Fig. 9).

B. Network

We used the pix2pix conditional GAN implemented in Pytorch with the UNet [24] as a generator and PatchGAN [6] as the discriminator counterpart.

UNet is an encoder-decoder architecture with skip connections, and PatchGAN is a special type of classifier network subdividing the image into 70×70 patches and classifying each of them as real or fake (Fig. 10). The GAN loss is the average loss of all patches. This strategy focuses on the high frequency of the image data, whereas the low-frequency content is captured by L1 loss.

C. Experiments

All experiments were performed using an HP Z8 workstation with a 24 GB memory Quadro P6000 GPU. We began the experiment using only seismic images and associated FWI velocity as a training pair, but the generator produced artificial high frequency structure images with artifacts (Fig. 11). We conclude that seismic images alone did not provide enough information for the network to converge.

To obtain better results, we added two additional channels, the average tomographic velocity and the TWT grid, resulting in a three-channel input image (Fig. 9).

Although the tomographic interval velocity was inaccurate, it provided a reliable average velocity. Hence, we used average velocity rather than interval velocity. TWT information supported the network in better estimating spatial positions after we extracted the patches randomly, once there is a direct relation between velocity and overburden thickness.

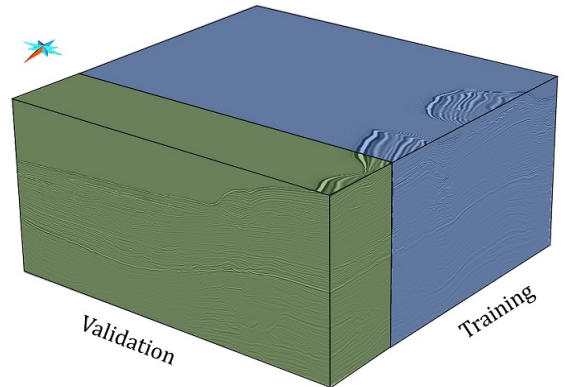


Fig. 8. Seismic volume showing the spatial distributions of training and test sets.

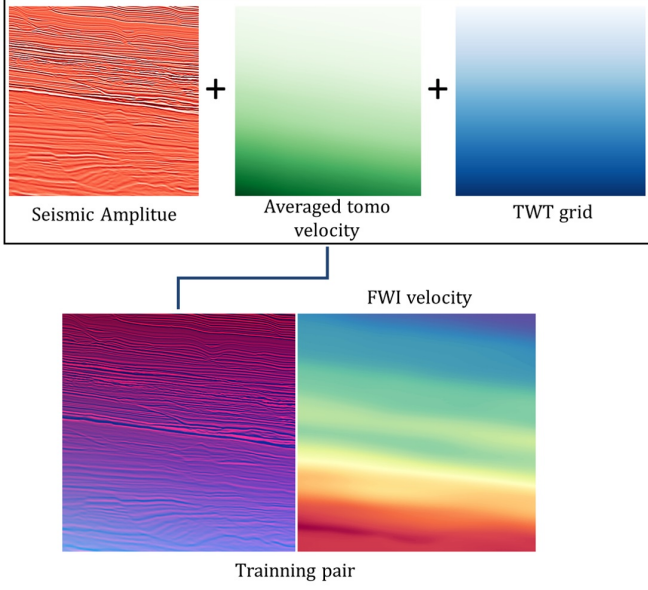


Fig. 9. Three channels combined input and FWI velocity patches producing a training pair.

Our assumption was that the tomographic average velocity and the TWT grid would function as boundary conditions, helping the network converge to the right interval velocity. The training time for each experiment was approximately four hours.

The generator and discriminator were trained using the same learning rate ($lr = 2 \times 10^{-5}$) which was established empirically, based on the visual quality of the training set results. We used the stochastic gradient descent method and the Adam solver with momentum parameters $\beta_1 = 0.5$ and $\beta_2 = 0.999$. The $L1$ multiplier λ was set to 100.

The results were evaluated based on three criteria: visual analysis, percent error (PE), and structural similarity index measure (SSIM) [25].

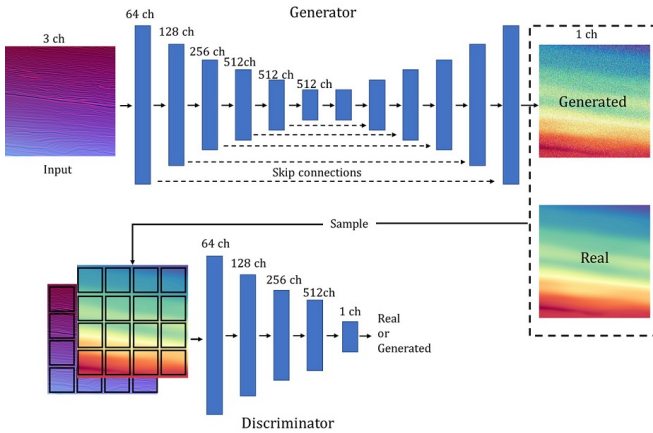


Fig. 10. Schematic representation of the pix2pix network architecture.

Percent error is a common metric for quality control in seismic processing; it is defined as (6).

$$PE = \frac{1}{N} \sum_{i=1}^N \frac{|y(i) - y'(i)|}{y(i)} * 100. \quad (6)$$

SSIM is a composite index (7) measuring the similarity between two signals based on a comparison of their mean (8), variance (9), and correlation coefficient (10).

$$SSIM(y, y') = f(l(y, y'), c(y, y'), s(y, y')), \quad (7)$$

$$l(y, y') = \frac{2\mu_y\mu_{y'} + C_1}{\mu_y^2 + \mu_{y'}^2 + C_1}, \quad (8)$$

$$c(y, y') = \frac{2\sigma_y\sigma_{y'} + C_2}{\sigma_y^2 + \sigma_{y'}^2 + C_2}, \quad (9)$$

$$s(y, y') = \frac{\sigma_{yy'} + C_3}{\sigma_y\sigma_{y'} + C_3}, \quad (10)$$

where C_1 , C_2 , and C_3 are constants to avoid division by zero. Setting $C_3 = \frac{C_2^2}{2}$, we obtain the SSIM equation (11).

$$SSIM(y, y') = \frac{(2\mu_y\mu_{y'} + C_1) + (2\sigma_{yy'} + C_2)}{(\mu_y^2 + \mu_{y'}^2 + C_1)(\sigma_y^2 + \sigma_{y'}^2 + C_2)}. \quad (11)$$

III. RESULTS

Our pix2pix GAN approach produced an accurate velocity model using three channels (post-stack seismic image, average tomographic velocity, and two-way time) as input (Fig. 12). The chosen metrics show excellent results on the test set. The average percent error was 1.5%, average SSIM was 0.994, and the generated velocity was visually identical to the FWI velocity, without artifacts, verifying that the geological structures were incorporated into the GAN velocity model. The two distributions are very close, confirming the metric results (Fig. 13).

The criterion to stop training was set by visual analysis of the training set results and later confirmed using the percent error and the SSIM results based on the test set. Both metrics stabilized between 60 and 80 epochs (Fig. 14 and Fig. 15).

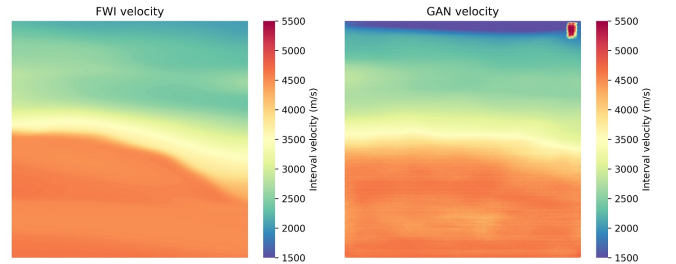


Fig. 11. FWI velocity compared with one channel GAN velocity.

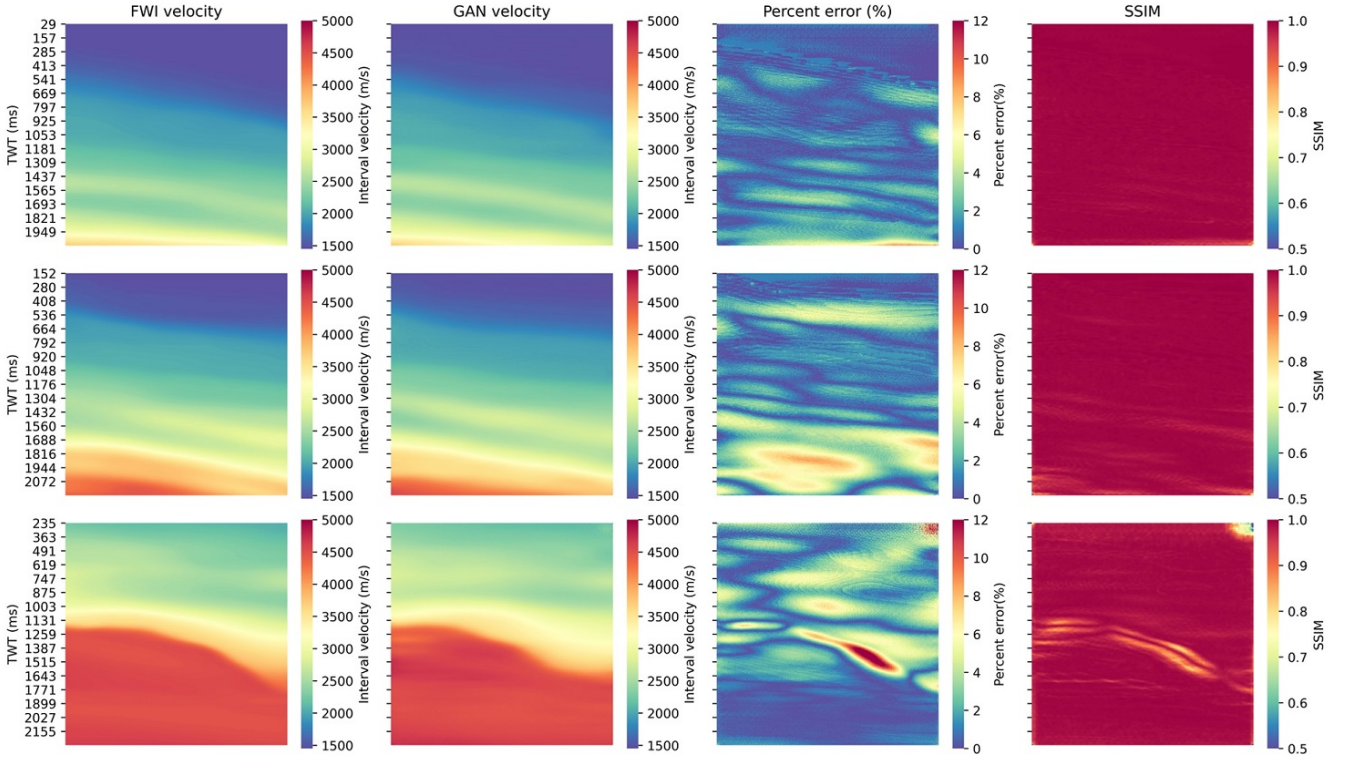


Fig. 12. FWI velocity, GAN velocity, percent error, and SSIM map over three representative patches (shallow, intermediate, and deep regions from top to bottom, respectively).

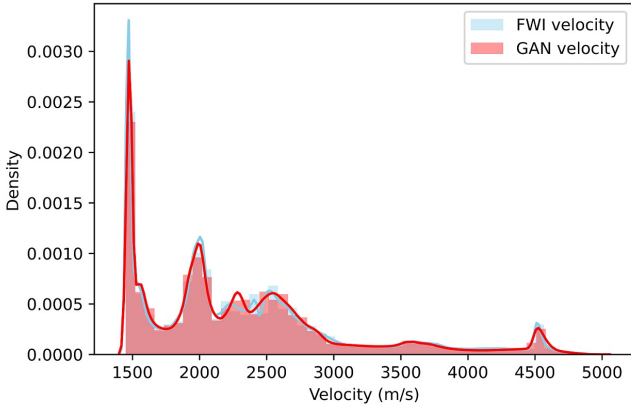


Fig. 13. FWI and GAN velocity distributions of test set.

Unexpectedly, the one-channel model experiment showed satisfactory SSIM and PE results when compared with the three-channel model, although the visual analysis showed very poor results.

Shallower regions presented better results than deeper regions on all metrics (Fig. 12).

IV. DISCUSSION

Previous works [26] demonstrated the feasibility of using CNNs for one-dimensional (1D) synthetic seismic inversion,

and [27] used CNNs to perform seismic inversion with synthetic pre-stack seismic data. Conditional unpaired CycleGANs [28] were employed to perform forward and inverse seismic modeling with post-stack seismic data and velocity models, while also using synthetic data with a single input channel (post-stack seismic image). They achieved only qualitatively moderate results. GANs were also used to generate prior models to constrain solutions in the traditional FWI algorithm [29].

In this study, we apply a conditional pix2pix GAN to generate highly accurate velocity models using real post-stack seismic data with multiple inputs.

Neural networks are understood to perform better at interpolation than extrapolation tasks [30], [31]; therefore, it is challenging for a cGAN to extrapolate to lower frequencies using only the post-stack seismic image as input. However, it becomes easier to interpolate the missing frequencies (Fig. 6) if we combine lower frequency data as input (in this case, average tomographic velocity and TWT grid). This explains why adding more information helped to improve the results.

Providing a TWT grid and the average tomographic velocity enabled a depth estimation for each sample, which is useful information for estimating the interval velocity as it increases with depth.

Ordinary physics-based FWI provides better results in shallower regions, where the acquisition cables have sufficient length to capture diving waves [32] and worse results in

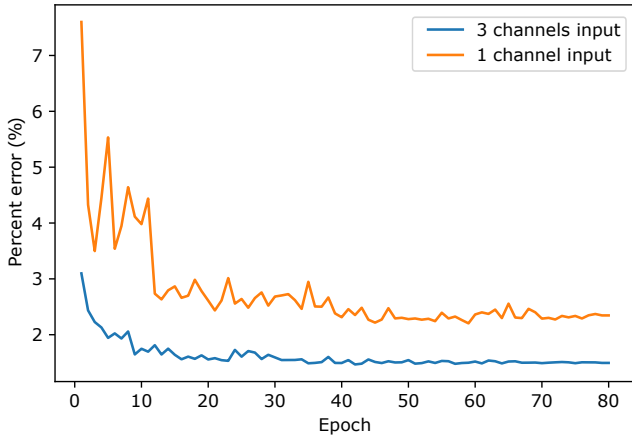


Fig. 14. Percent error versus epochs using test set.

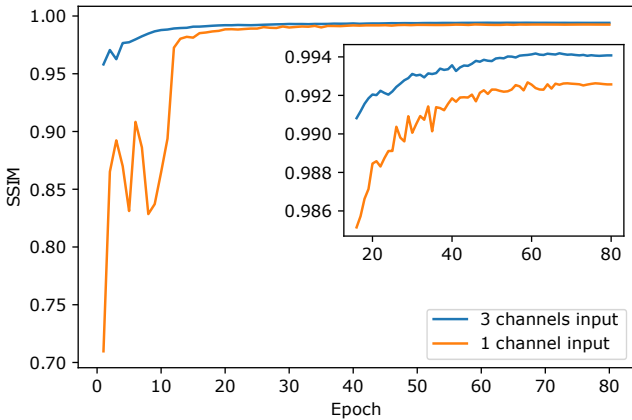


Fig. 15. SSIM versus epochs using test set.

deeper regions. As expected, the pix2pix GAN captured this inaccuracy and also produced worse results in deeper regions (Fig. 12).

The $L1$ regularization loss acts as a global metric, minimizing the average loss of all pixels, whereas the GAN loss, provided by the PatchGAN discriminator, focuses on high-frequency structures. We suspect that the reason the experiment using one channel presented good PE and SSIM values was that the generator produced a velocity field that honors the low frequency associated to average velocity (good PE) and the high frequency variations (good SSIM), however producing poorly accurate visual results. This result reveals the difficulty of finding reliable metrics for evaluating GANs [33].

The major threat to the method is the usage of poor-quality training data. Despite the robustness of FWI velocity modeling, examples where FWI could not converge to an acceptable velocity model because of seismic acquisition limitations and low signal-to-noise ratio are common [34]. If this data were used to train the model, the pix2pix cGAN will also produce a poor-quality velocity model.

GANs are normally difficult to train, they are prone to reduce the output variability or a complete mode collapse [35]. Different architectures such as Variational Autoencoders and Autoencoder with perceptual loss need to be tested in order to avoid such problems.

V. CONCLUSIONS

We have presented a method, based on a pix2pix conditional GAN with multiple input, that can be used to surrogate FWI velocity modeling. The three inputs (post-stack seismic image, average tomographic velocity, and TWT grid) were vital, as the experiment using only one channel (post-stack seismic image) failed.

Further development of new network architectures focused on seismic data particularities and considering three or more spatial dimensions may generate an improvement in the quality of our results.

The test data presented excellent results based on PE, SSIM, and visual analysis. The main geological structures and lateral velocity changes were captured by the resulting velocity model.

Once we train the GAN, it was really fast the generation of a FWI interval velocity model surrogate, almost real time. Therefore, our proposed method can replace the FWI modeling or providing an initial velocity model that will need fewer iterations to converge during traditional FWI, similar strategy adopted by [36]

A fair comparison between the FWI and our cGAN surrogate model should consider the implementation of the FWI algorithm in parallel, using GPU hardware. Although some papers have proposed GPU implementations of the FWI algorithm [37], [38] and [39], they only used 2D synthetic data, mainly due to GPU memory constraints. At present, FWI algorithm processing on GPU hardware has yet to become industrially feasible.

The proposed method, when deployed, has the potential to speed up geophysical reservoir characterization processes, saving time and reducing computational resource requirements.

REFERENCES

- [1] J. Gazdag and P. Sguazzero, "Migration of seismic data," *Proceedings of the IEEE*, vol. 72, no. 10, pp. 1302–1315, 1984.
- [2] O. Yilmaz, *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. Society of Exploration Geophysicists, 01 2001. [Online]. Available: <https://doi.org/10.1190/1.9781560801580>
- [3] P. Schultz, *The seismic velocity model as an interpretation asset*. Society of Exploration Geophysicists, 1998.
- [4] G. Liu, S. Fomel, L. Jin, and X. Chen, "Stacking seismic data using local correlation," *Geophysics*, vol. 74, 2009.
- [5] S. Kapoor, D. Vigh, H. Li, and D. Derharoutian, "Full waveform inversion for detailed velocity model building," in *74th EAGE Conference and Exhibition incorporating EUROPEC 2012*. European Association of Geoscientists & Engineers, 2012, pp. cp–293.
- [6] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," 2018.
- [7] I. Jones, "Tutorial: Velocity estimation via ray-based tomography," *First Break*, vol. 28, 2010.
- [8] P. Lailly, "The seismic inverse problem as a sequence of before stack inversions," in *Conference on Inverse Scattering—Theory and Application*, vol. 11. Siam, 1983, p. 206.

- [9] A. Tarantola, "Inversion of seismic reflection data in the acoustic approximation," *Geophysics*, vol. 49, pp. 1259–1266, 1984.
- [10] Madagascar Development Team, *Madagascar Software, Version 1.4*, <http://www.ahay.org/>, 2012.
- [11] J. Virieux, A. Asnaashari, R. Brossier, L. Métivier, A. Ribodetti, and W. Zhou, "An introduction to full waveform inversion," in *Encyclopedia of exploration geophysics*. Society of Exploration Geophysicists, 2017, pp. R1–1.
- [12] Z. Geng and Y. Wang, "Automated design of a convolutional neural network with multi-scale filters for cost-efficient seismic data classification," *Nature Communications*, vol. 11, 2020.
- [13] W. Xiong, X. Ji, Y. Ma, Y. Wang, N. M. AlBinHassan, M. N. Ali, and Y. Luo, "Seismic fault detection with convolutional neural network," *Geophysics*, vol. 83, 2018.
- [14] Y. Shi, X. Wu, and S. Fomel, "Saltseg: Automatic 3d salt segmentation using a deep convolutional neural network," *Interpretation*, vol. 7, 2019.
- [15] X. Wu, L. Liang, Y. Shi, and S. Fomel, "Faultseg3d: Using synthetic data sets to train an end-to-end convolutional neural network for 3d seismic fault segmentation," *Geophysics*, vol. 84, 2019.
- [16] A. B. Guazzelli, M. Roisenberg, and B. Rodrigues, "Efficient 3d semantic segmentation of seismic images using orthogonal planes 2d convolutional neural networks," *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, 2020.
- [17] Y. Ren, X. Xu, S. Yang, L. Nie, and Y. Chen, "A physics-based neural-network way to perform seismic full waveform inversion," *IEEE Access*, vol. 8, pp. 112 266–112 277, 2020.
- [18] B. Moseley, A. Markham, and T. Nissen-Meyer, "Solving the wave equation with physics-informed deep learning," *arXiv: Computational Physics*, 2020.
- [19] R. Zhang, Y. Liu, and H. Sun, "Physics-guided convolutional neural network (phycnn) for data-driven seismic response modeling," *ArXiv*, vol. abs/1909.08118, 2019.
- [20] J. F. Claerbout, *Imaging the earth's interior*. Blackwell scientific publications Oxford, 1985, vol. 1.
- [21] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio, "Generative adversarial nets," in *NIPS*, 2014.
- [22] D. Pathak, P. Krähenbühl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2536–2544, 2016.
- [23] Equinor. segyio documentation. [Online]. Available: <https://seggyio.readthedocs.io/en/latest/index.html>
- [24] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *ArXiv*, vol. abs/1505.04597, 2015.
- [25] Z. Wang, A. C. Bovik, H. R. Sheikh, S. Member, E. P. Simoncelli, and S. Member, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, pp. 600–612, 2004.
- [26] V. Puzyrev, A. Egorov, A. Pirogova, C. Elders, and C. Otto, "Seismic inversion with deep neural networks: A feasibility analysis," in *81st EAGE Conference and Exhibition 2019*, vol. 2019, no. 1. European Association of Geoscientists & Engineers, 2019, pp. 1–5.
- [27] Y. Wu, Y. Lin, and Z. Zhou, "Inversionnet: Accurate and efficient seismic waveform inversion with convolutional neural networks," *Seg Technical Program Expanded Abstracts*, 2018.
- [28] L. Mosser, W. Kimman, J. Dramsch, S. Purves, A. D. L. Fuente, and G. Ganssle, "Rapid seismic domain transfer: Seismic velocity inversion and modeling using deep generative neural networks," *ArXiv*, vol. abs/1805.08826, 2018.
- [29] L. Mosser, O. Dubrule, and M. J. Blunt, "Stochastic Seismic Waveform Inversion Using Generative Adversarial Networks as a Geological Prior," *Mathematical Geosciences*, vol. 52, no. 1, pp. 53–79, jan 2020. [Online]. Available: <http://arxiv.org/abs/1806.03720http://link.springer.com/10.1007/s11004-019-09832-6>
- [30] E. Barnard and L. Wessels, "Extrapolation and interpolation in neural network classifiers," *IEEE Control Systems*, vol. 12, pp. 50–53, 1992.
- [31] K. Xu, J. Li, M. Zhang, S. Du, K. Kawarabayashi, and S. Jegelka, "How neural networks extrapolate: From feedforward to graph neural networks," *ArXiv*, vol. abs/2009.11848, 2020.
- [32] W. Zhou, R. Brossier, S. Operto, and J. Virieux, "Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation," *Geophysical Journal International*, vol. 202, pp. 1535–1554, 2015.
- [33] A. Borji, "Pros and cons of gan evaluation measures," *Comput. Vis. Image Underst.*, vol. 179, pp. 41–65, 2019.
- [34] G. Yao, D. Wu, and S.-X. Wang, "A review on reflection-waveform inversion," *Petroleum Science*, vol. 17, no. 2, pp. 334–351, 2020.
- [35] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," *arXiv preprint arXiv:1606.03498*, 2016.
- [36] A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Židek, A. W. Nelson, A. Bridgland *et al.*, "Improved protein structure prediction using potentials from deep learning," *Nature*, vol. 577, no. 7792, pp. 706–710, 2020.
- [37] B. li Wang, J. Gao, H. Zhang, and W. Zhao, "Cuda-based acceleration of full waveform inversion on gpu," *Seg Technical Program Expanded Abstracts*, 2011.
- [38] J. Mao, R. shan Wu, and B. Wang, "Multiscale full waveform inversion using gpu," *Seg Technical Program Expanded Abstracts*, 2012.
- [39] J. Fang, H. Chen, H. Zhou, Y. Rao, P. Sun, and J. Zhang, "Elastic full-waveform inversion based on gpu accelerated temporal fourth-order finite-difference approximation," *Comput. Geosci.*, vol. 135, p. 104381, 2020.

APÊNDICE B – Publicação 2

Trabalho apresentado na *Seventeenth International Congress of the Brazilian Geophysical Society 2021* intitulado *Use of Generative Adversarial Networks to Mimic Seismic Inversions* e parte integrante desta dissertação.



Use of Generative Adversarial Networks to Mimic Seismic Inversions

Marcus Saraiva (Petrobras, UFES); Alexandre Maul (Petrobras); Avelino Forechi (UFES); Jorcy de Oliveira Neto (Petrobras); Antônio Del Rey (Petrobras); Thomas Rauber (UFES)

Copyright 2021, SBGf - Sociedade Brasileira de Geofísica.

This paper was prepared for presentation during the 17th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 16-19 August 2021.

Contents of this paper were reviewed by the Technical Committee of the 17th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

Abstract

Hydrocarbon industry works with seismic reflection as a chief indirect method for investigating subsurface data. Seismic volume provides inaccurate geometries of geological structures, inferring rock properties based on seismic amplitude. One considers it as an ill-posed inverse problem. In order to obtain reliable information, we need to perform seismic inversion processes, as they are very important for correct reservoir characterizations, although expensive and time-consuming. The aim of the acoustic seismic inversion process is to obtain an impedance volume through a recursive process by minimizing the error between the original seismic data and a synthetic seismic data. Progressively, updates occur in this model until the error reaches a desired threshold. Incorporation of well log information and interpreted seismic horizons as boundary conditions enable the full process to expand original seismic volume frequency bandwidth. Iterdec (Iterative Deconvolution of Seismic Data) is a process that targets to improve reflectors definition, increasing resolution. It is an important first-order approximation, applicable to a seismic trace in which the pulse phase is null and the value of reflection coefficient is proportional to the amplitude of the same central peak. Each of seismic trace local maximum and local minimum represents an approximation of a valid reflection coefficient. Full Waveform Inversion (FWI) is another seismic inversion technique. However, it uses pre-stacked seismic data to obtain a high-resolution interval velocities model. As FWI takes place with pre-stacked seismic data (three orders of magnitude greater than the traditional seismic volume), it becomes computationally very expensive and time-consuming. Its use has only become industrially feasible recently, with the increase in available processing. Convolutional neural networks (CNN) are getting popular in many applications, typically in classification and segmentation tasks. Moreover, recent studies have demonstrated the feasibility of geophysical modeling using CNNs. Among CNN architectures, generative adversarial networks (GAN) go beyond the traditional classification and segmentation tasks; GANs stand out for their ability to learn how to mimic data provided to them by capturing their statistical and spatial distributions. GANs are generative models that learn to generalize rules of how to transform a random noise vector z to output image y ($G:z \rightarrow y$). The pix2pix GAN architecture is a conditional GAN (cGAN), which has as its input the sum of the random noise vector z with an image x , to find a rule to translate this input to image y ($G:\{x,z\} \rightarrow y$). To accomplish this task, pix2pix uses two networks competing against each other. While one network generates synthesized images (generator), the other judges whether these images are valid (discriminator). The training of Generator G produces output images that are indistinguishable from the actual images by a discriminator D , which in turn is trained to improve its performance on detecting the generator's fake images. In this work, we present a method using an image-to-image translator (pix2pix) based on a conditional generative adversarial network (cGAN) to translate an input data (post-stack seismic image combined with some seismic processing subproducts) into three seismic inversion volumes: Acoustic Seismic Inversion, Iterdec and FWI, for a faster reservoir characterization.