Rede neural convolucional profunda com aprendizado por transferência para estimativa da wavelet sísmica

Landeta B.,A.V.*, Novais, A.*, Schleicher, J.* (*: UNICAMP & INCT-GP)

Copyright 2024, SBGf - Sociedade Brasileira de Geofísica.

Este texto foi preparado para a apresentação no IX Simpósio Brasileiro de Geofísica, Salvador, 08 a 10 de outubro de 2024. Seu conteúdo foi revisado pelo Comitê Técnico do X SimBGf, mas não necessariamente representa a opinião da SBGf ou de seus associados. É proibida a reprodução total ou parcial deste material para propósitos comerciais sem a prévia autorização da SBGf.

Resumo

O conhecimento da *wavelet* na sísmica de reflexão é essencial na exploração geofísica, na obtenção quantitativa de informações precisas da subsuperfície. Este trabalho adota o *deep learning* para aprimorar a estimativa da *wavelet* sísmica, enfrentando desafios como a disponibilidade limitada de dados de treinamento. Com ênfase no aprendizado por transferência, adaptamos uma rede neural convolucional usando a arquitetura Inception. Os resultados demonstram a eficácia da abordagem, com a rede alcançando uma boa precisão na predição das *wavelets* no formato do pulso e cinemática, usando no treinamento as *wavelets* Ricker e Gabor. Este estudo destaca o potencial do *deep learning* na geofísica, especialmente para o processamento de dados sísmicos.

Introdução

A sísmica de reflexão é uma ferramenta fundamental na exploração geofísica que permite obter informação detalhada sobre a estrutura da subsuperfície da terra. O campo de pressão após a propagação da onda registrado em superfície depende das propriedades da estrutura de subsuperfície, da *wavelet* e do sistema de aquisição. Uma vez conhecendo as caraterísticas e influências da *wavelet* sobre os dados coletados, pode ser extraída informação mais precisa da subsuperfície, pela qual um dos elementos cruciais no processamento dos dados sísmicos é a extração da *wavelet*.

A importância da recuperação correta da *wavelet* sísmica reside no seu papel central no processo de conexão entre a sísmica de subsuperfície e sua geologia, como aprecia-se no processo de amarração de poços (de Macedo et al., 2017). Além disso, a estimativa da *wavelet* é necessária em processos de inversão sísmica nos quais procura-se um modelo físico da subsuperfície, por exemplo no bem conhecido método de inversão da forma da onda completa (FWI, do inglês Full Waveform Inversion).

Métodos determinísticos para a estimativa da *wavelet* usam como principal instrumento a deconvolução, que procura melhorar a resolução dos dados sísmicos (Sacchi, 1997), tendo como limitação a sensibilidade ao ruído (Kazemi and Sacchi, 2013). Modificações à deconvolução tradicional foram implementadas, por exemplo, em situações onde existem múltiplas *wavelets* sobrepostas nos dados sísmicos devido à presença de camadas geológicas complexas ou descontinuidades (Baziw and Ulrych, 2006), ou para expandir a banda de frequência dos dados sísmicos e obter previsões mais precisas, especialmente para reservatórios delgados (Guo et al., 2024). Eles são frequentemente exigentes em termos computacionais, especialmente quando aplicados a conjuntos grandes e complexos de dados sísmicos. Isso pode exigir recursos computacionais significativos e longo tempo de processamento.

Nos últimos anos, o uso do *deep learning* tem sido o foco para a solução de problemas de inversão sísmica (Adler et al., 2021; Zheng et al., 2019; Wu et al., 2018), isto é motivado pela habilidade de treinar redes de alta capacidade usando grandes conjuntos de dados. As redes neurais convolucionais originalmente projetadas para tarefas de computação gráfica, provaram ser eficazes na extração de recursos complexos e na geração de modelos preditivos a partir de dados brutos. Este avanço tecnológico despertou um interesse na aplicação de técnicas do *deep learning* para enfrentar os desafios na interpretação de dados sísmicos (Araya-Polo et al., 2018; Das et al., 2018). Uma dessas técnicas é a aprendizagem por transferência que recentemente ganhou relevância na comunidade geofísica (Hu et al., 2021; Park and Sacchi, 2020). Esta técnica permite aproveitar o conhecimento adquirido pelos modelos treinados em uma tarefa específica e aplicá-lo em outra tarefa relacionada. Neste trabalho, aplica-se a aprendizagem por transferência no problema de estimativa da *wavelet* sísmica.

Metodologia

Exemplos de treino

A geração de dados suficientes para um correto treinamento da inteligência artificial constitui um desafio para áreas de pesquisas que vêem no *deep learning* uma solução para diversos problemas. Isto se deve à falta de bibliotecas livres para os problemas específicos de aplicação. Nesse contexto, esse trabalho começa com a geração dos dados sísmicos para o aprendizado supervisionado da rede.

A propagação do campo de pressão $P(\mathbf{x},t)$ pela injeção de uma fonte $f(t;\mathbf{x_s})$, na posição $\mathbf{x_s}$, é governada pela equação da onda acústica com densidade constante dada por

$$\frac{1}{c(\mathbf{x})^2}\frac{\partial^2}{\partial t^2}P(\mathbf{x},t;\mathbf{x_s}) - \Delta P(\mathbf{x},t;\mathbf{x_s}) = f(t;\mathbf{x_s}), \qquad (1)$$

onde $c(\mathbf{x})$ denota a velocidade de propagação e $\mathbf{x}$ e t representam as variáveis espacial e temporal, respectivamente, sendo $\mathbf{x_s}$ a coordenada da fonte.

O modelamento da propagação de ondas acústicas foi

realizada seguindo um esquema de diferenças finitas usando a aproximação de segunda ordem em tempo e décima segunda ordem no espaço para as derivadas parciais na equação da onda.

Para que a estimativa da *wavelet* respeite as condições físicas realistas, é importante variar, nos dados de treino, todos os parâmetros físicos que podem variar entre diferentes seções sísmicas. Portanto, considera-se a variação do conteúdo de frequência e do formato do pulso. Além disso, deve-se variar o nível de ruído e o modelo de velocidade.

Rede neural

A fim de adaptar técnicas do *deep learning* (como as redes neurais convolucionais) e estratégias de aprendizado por transferência, o treinamento da rede neural usou uma arquitetura Inception (Szegedy et al., 2016). Esta, originalmente treinada para a classificação de objetos, tem como principal vantagem a sua capacidade para detectar objetos relativamente pequenos. Isto se dá devido à estrutura dos seus blocos convolucionais, que além de permitirem obter resultados de alta qualidade com baixo custo computacional, também mantêm boa qualidade mesmo com um tamanho modesto do conjunto de treinamento.

Treinamento

O conjunto para treinamento da rede convolucional foi formado por pares de seção sísmica e *wavelet*, pelo qual adapta-se a camada de entrada e camada de saída da rede para ser compatível com os dados. Para que a rede consiga prever valores negativos de amplitude, foi implementada como função de ativação na última camada a tangente hiperbólica. Como função de custo usamos o logaritmo do cosseno hiperbólico que é uma função totalmente diferenciável que pode ser encontrada frequentemente em problemas de regressão atuais, obtendo bons resultados.

Para o treinamento da rede foi usada a biblioteca Keras do TensorFlow. Usando a arquitetura Inception foram realizados alguns testes de treinamentos da rede variando a quantidade de exemplos de treino, a arquitetura da rede nas últimas camadas e o uso do aprendizado por transferência. Nos primeiros experimentos, na saída da rede Inception foram implementadas duas camadas totalmente conectadas. Para o treinamento da rede neural, foram geradas 100 seções de tiro comum usando como termo fonte a *wavelet* Ricker com frequência pico de 5, 10, 15, 20 e 25 Hz em meios homogêneos com velocidade constante. Desses, foram considerados 85 exemplos no conjunto de treinamento e 15 exemplos no conjunto de validação. Na Figura 1, mostra-se um par de exemplo de treino, i.e., a seção de tiro comum com a sua correspondente *wavelet* Ricker de frequência pico de 5 Hz.

Resultados

Como primeiro teste, iniciamos os pesos da rede seguindo uma distribuição normal. Assim, sem o uso do aprendizado por transferência não foi possível atingir uma boa acurácia, isto devido ao pequeno número de exemplos de treino. Nota-se que a rede, para cumprir sua tarefa original, foi treinada com cerca de 1.3 milhões de exemplos de treino e no nosso caso foi usado um número

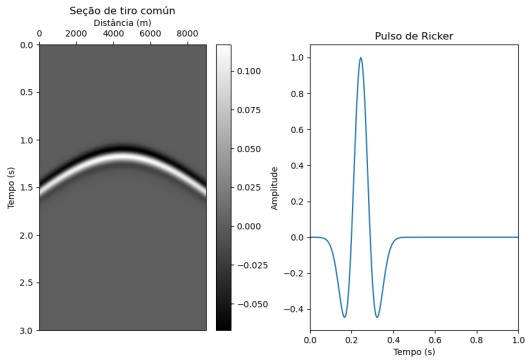

Figure 1: Seção de tiro comum e wavelet Ricker com frequência pico de 5 Hz.

de exemplos de treino bem inferior.

Por este motivo, no segundo experimento a rede foi inicializada com os pesos da rede original e treinando apenas as camadas totalmente conectadas implementadas na saída durante 5 épocas e taxa de aprendizado 0.0001 e logo retreinando a rede toda por 20 épocas aumentando a taxa de aprendizado a 0.001. Desse modo, obtivemos melhores resultados, porém ainda sem atingir uma boa acurácia.

Em um terceiro experimento, foi aplicado a transferência de aprendizado em alguns blocos convolucionais, em outras palavras, os pesos do último bloco convolucional e das duas camadas totalmente conetadas foram treinadas e manteve-se os pesos originais da rede Inception nos outros blocos. Usando 40 épocas e uma taxa de aprendizado de 0.0001 no otimizador Adam obteve-se já uma boa acurácia para predizer *wavelets*, mas só para frequências pico aprendidas pela rede, isto é, usadas nos dados de treinamento. A rede teve problemas para predizer *wavelets* por exemplo a partir de seções de tiro comum modeladas com uma Ricker com frequência pico de 7 Hz.

Para o quarto experimento, incrementou-se o conjunto de treino usando dessa vez pares de seções de tiro comum e *wavelets* com frequências pico escolhidas aleatoriamente entre 5 e 25 Hz, sendo um total de 586 exemplos de treino usando 90% como conjunto de treinamento e 10% como conjunto de validação. Neste caso, partindo com os pesos do último teste de treinamento, mas já com o incremento de dados, a rede foi retreinada em todas as camadas com uma maior taxa de aprendizado igual a 0.001 para procurar outro mínimo local capaz de generalizar os novos exemplos de treino, atingindo assim uma boa acurácia na época 64 usando um tamanho de *batch* de 5.

Em um quinto experimento, na saída da rede Inception foram implementadas um bloco convolucional, fazendo uso da convolução 1D, e uma totalmente conectada como saída. Com o uso dessa arquitetura, os pesos originais da rede Inception foram inicializadas, sendo treinados só os pesos das duas camadas adicionais durante 50 épocas. Obteve-se melhores resultados em comparação com a arquitetura previa, com taxa de aprendizado de 0.001 e tamanho do *batch* de 10.

Por último, para obter uma rede mais robusta na predição em formatos do pulso foram modelados 586 seções de tiro comum usando como termo fonte a *wavelet* Gabor com frequências pico aleatórias no intervalo de 5 a 25 Hz. Na Figura 2 mostra-se uma seção de tiro comum com a sua correspondente *wavelet* Gabor de frequência pico de 25 Hz. Repetiu-se a última arquitetura e estrategia de treino do experimento prévio. Portanto, após o treinamento da mesma rede com ambos os tipos, a rede atingiu uma boa acurácia em 50 épocas.

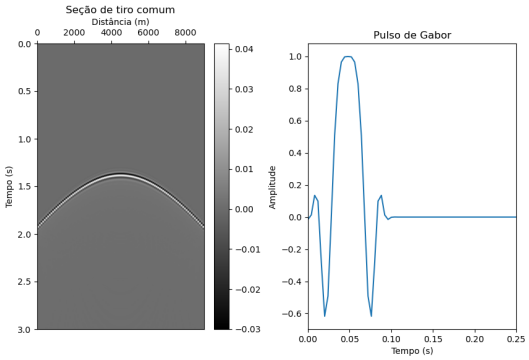

Figure 2: Seção de tiro comum e wavelet Gabor com frequência pico de 5 Hz.

Nas Figuras 3, 4 e 5 mostra-se o resultado da rede após o treinamento final para prever a *wavelet* Ricker de frequência pico 5.11 e 25.0 Hz e a *wavelet* Gabor de frequência pico 16.87 Hz, respectivamente, todas do conjunto de teste. Nota-se uma boa acurácia da rede para prever o formato da *wavelet* e a cinemática da mesma.

Para avaliar se existe perda de conteúdo de frequência na predição, computa-se a transformada de Fourier das *wavelets* sintéticas e preditas para comparação como exibe-se nas Figuras 6 e 7 para as frequências pico de 5.11 e 25.0 Hz da *wavelet* Ricker, respectivamente, e na Figura 8 para a frequência pico de 16.87 Hz da *wavelet* Gabor. Observa-se que existe, em geral, um bom desempenho da rede para predizer conteúdo de frequência das *wavelets*.

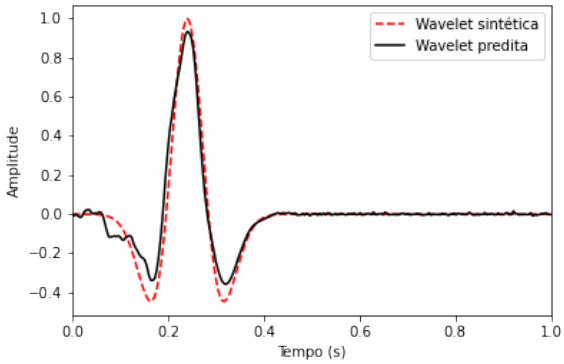

Figure 3: Comparação da wavelet Ricker sintética de 5.11 Hz e a predita pela rede.

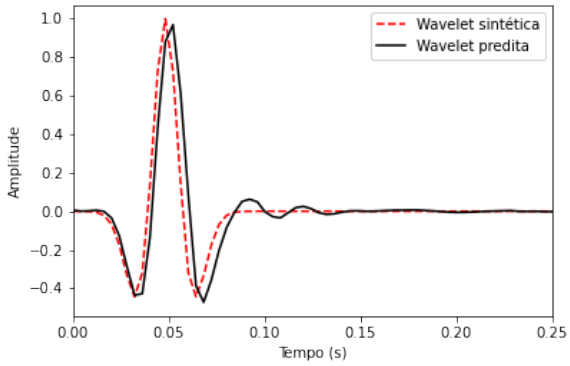

Figure 4: Comparação da wavelet Ricker sintética de 25.0 Hz e a predita pela rede.

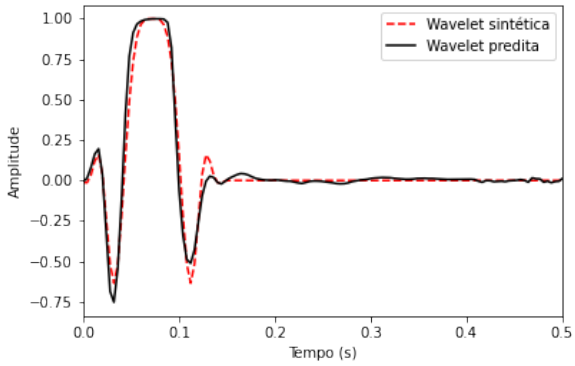

Figure 5: Comparação da wavelet Gabor sintética de 16.87 Hz e a predita pela rede.

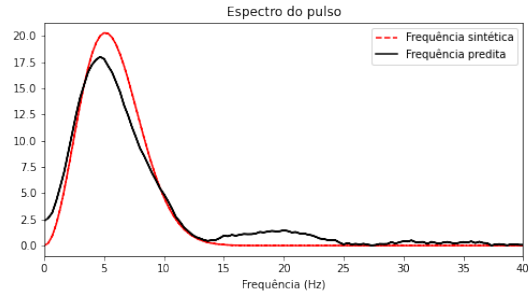

Figure 6: Comparação do conteúdo de frequência da wavelet Ricker sintética de 5.11 Hz e a predita pela rede.

Nas Figuras 9 e 10 apresenta-se a função de custo e a acurácia no treinamento da rede, respectivamente, em ambas figuras exibe-se as curvas para os exemplos de treino e validação variando nas épocas. Para quantificar a acurácia foi implementada a acurácia binaria da biblioteca Keras. Observa-se um bom comportamento em ambas curvas evitando um *overfitting*.

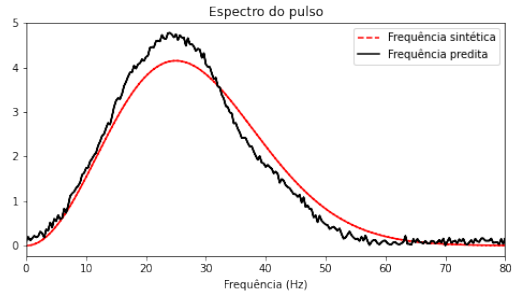

Figure 7: Comparação do conteúdo de frequência da wavelet Ricker sintética de 25.0 Hz e a predita pela rede.

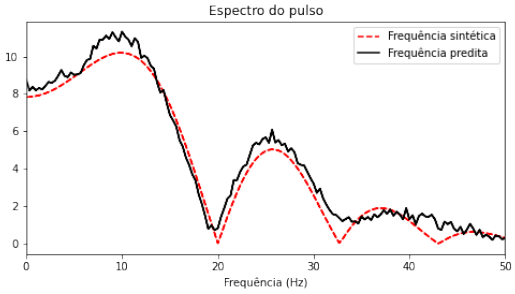

Figure 8: Comparação do conteúdo de frequência da wavelet Gabor sintética de 16.87 Hz e a predita pela rede.

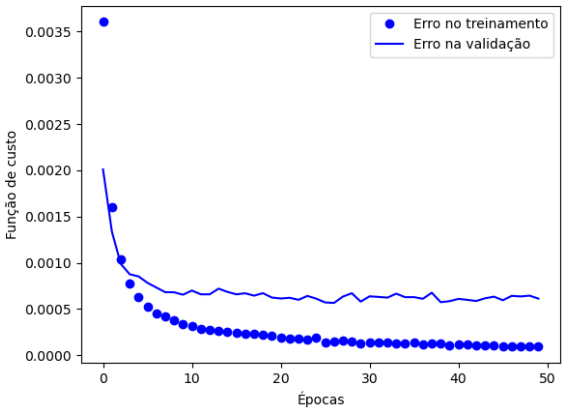

Figure 9: Função de custo.

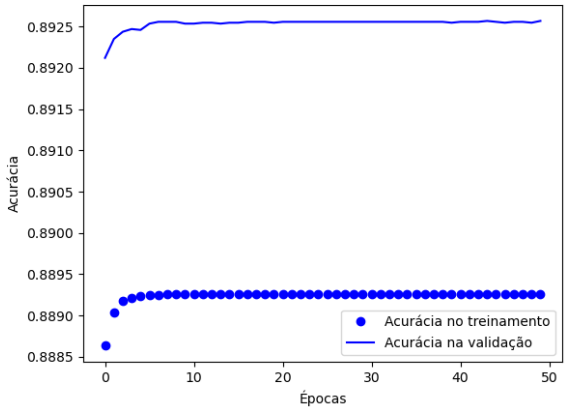

Figure 10: Acurácia non treinamento da rede.

Discussão e Conclusões

Este estudo destaca o potencial do *deep learning* como uma ferramenta eficaz para a inversão sísmica. No entanto, ficou evidente que a eficácia da rede neural em prever *wavelet* sísmicas depende significativamente da diversidade e quantidade de exemplos de treinamento disponíveis. A limitação na quantidade de dados de treinamento resultou em um desempenho aquém do desejado ao se utilizar pesos inicializados aleatoriamente. Nesse contexto, o aprendizado por transferência emergiu como uma abordagem alternativa promissora para melhorar o desempenho da rede neural na disponibilidade de poucos dados de treinamento.

Ao adaptar um modelo pré-treinado, mesmo sendo esse de uma tarefa não relacionada, proveniente da computação gráfica, e ajustar seus pesos com os poucos exemplos de treinamento disponíveis, observamos uma melhoria substancial no desempenho da rede neural. A capacidade de aproveitar o conhecimento prévio de modelos treinados em conjuntos de dados mais amplos foi fundamental para impulsionar o desempenho da nossa rede neural na tarefa da estimativa da *wavelet* sintética. A combinação do incremento de dados e o uso do aprendizado por transferência resultou em um modelo mais robusto e capaz de prever com precisão todas as frequências necessárias para a recuperação da *wavelet* sísmica.

Esses resultados destacam a importância de abordagens adaptativas e estratégias inovadoras na aplicação do *deep learning* em problemas complexos de geofísica, contribuindo assim para avanços significativos na interpretação de dados sísmicos e na caracterização de subsuperfícies geológicas.

Agradecimentos

Os autores agradecem o apoio do nosso trabalho a agências brasileiras como CNPq (apoio ao INCT-GP) e CAPES (bolsa de doutorado).

Referências

Adler, A., Araya-Polo, M., and Poggio, T. (2021). Deep Learning for Seismic Inverse Problems: Toward the Acceleration of Geophysical Analysis Workflows. *IEEE Signal Processing Magazine*, 38(2):89–119.

Araya-Polo, M., Jennings, J., Adler, A., and Dahlke, T. (2018). Deep-learning tomography. *The Leading Edge*, 37(1):58–66.

Baziw, E. and Ulrych, T. J. (2006). Principle phase decomposition: A new concept in blind seismic deconvolution. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8):2271–2281.

Das, V., Pollack, A., Wollner, U., and Mukerji, T. (2018). Convolutional Neural Network for Seismic Impedance Inversion. In *SEG Technical Program Expanded Abstracts*, volume 84, pages 2071–2075.

de Macedo, I. A., da Silva, C. B., de Figueiredo, J. J., and Omoboya, B. (2017). Comparison between deterministic and statistical wavelet estimation methods through predictive deconvolution: Seismic to well tie example from the North Sea. *Journal of Applied Geophysics*, 136:298–314.

Guo, X., Gao, J., Yong, X., Li, S., Gui, J., and Wang, H. (2024). A novel high-resolution imaging method based on sparsity in the time domain and spectral fitting in the frequency domain. *Frontiers in Earth Science*, 12(March):1–13.

Hu, W., Jin, Y., Wu, X., and Chen, J. (2021). Progressive transfer learning for low-frequency data prediction in full-waveform inversion. *GEOPHYSICS*, 86(4):R369–R382.

Kazemi, N. and Sacchi, M. D. (2013). Sparse multichannel blind deconvolution. *Geophysics*, 79(5):V143–V152.

Park, M. J. and Sacchi, M. D. (2020). Automatic velocity analysis using convolutional neural network and transfer learning. *Geophysics*, 85(1):V33–V43.

Sacchi, M. D. (1997). Reweighting strategies in seismic deconvolution. *Geophysical Journal International*, 129(3):651–656.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826. IEEE.

Wu, Y., Lin, Y., and Zhou, Z. (2018). InversionNet: Accurate and Efficient Seismic Waveform Inversion with Convolutional Neural Networks. *SEG Technical Program Expanded Abstracts*, pages 2096–2100.

Zheng, Y., Zhang, Q., Yusifov, A., and Shi, Y. (2019). Applications of supervised deep learning for seismic interpretation and inversion. *Leading Edge*, 38(7):526–533.