



# Reconstrução de Imagens com Masked Autoencoders (MAE)

Aprendizado Não Supervisionado de Máquina (PPGEEC2325)

Autores: Ernane Ferreira e José Augusto

---

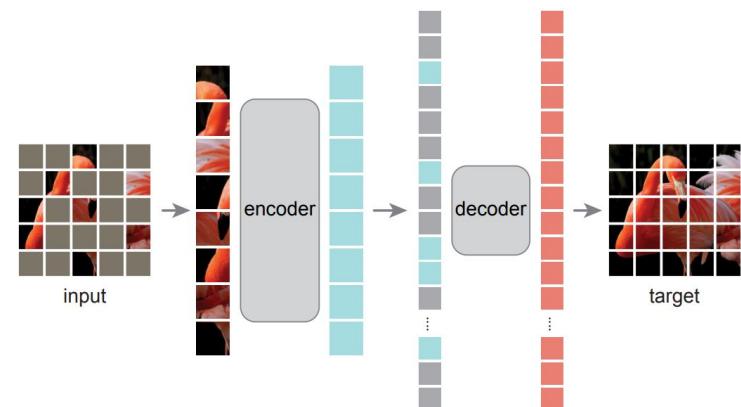
# Sumário

- O que são Masked Autoencoders (MAE)?
- Exemplo de reconstrução 2D
- Por que reconstruir dados 3D?
- Exemplo de reconstrução 3D
- Métricas utilizadas para comparação
- Comparação entre modelos tradicionais e MAE
- Conclusão
- Referências

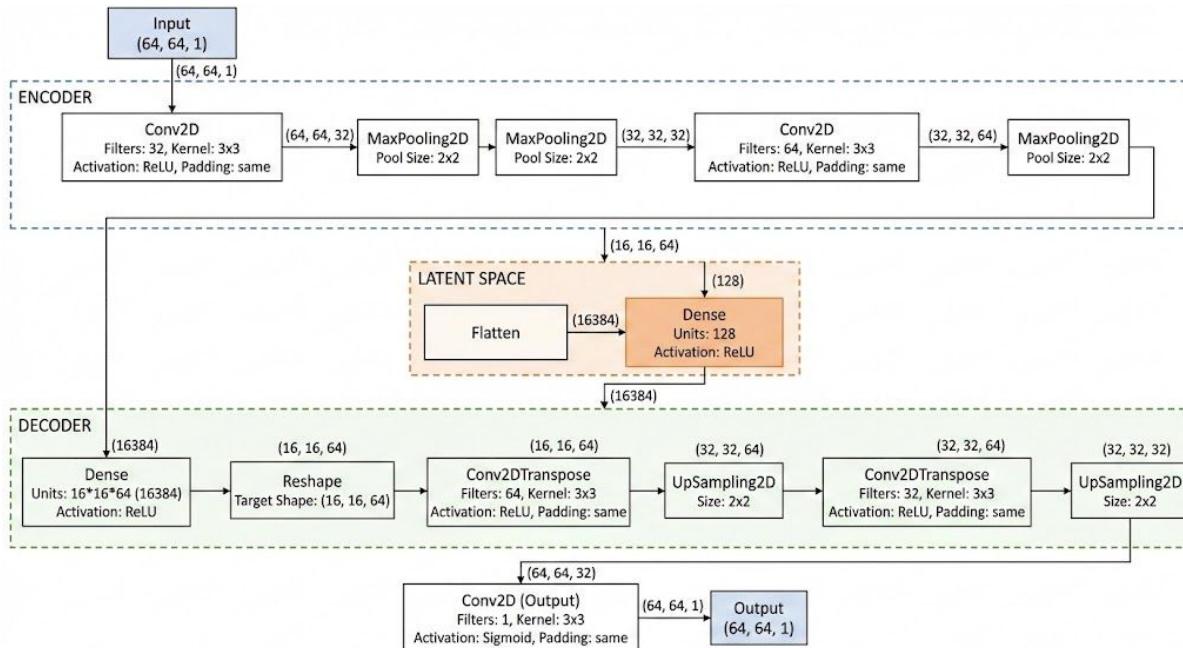
---

# O que são Masked Autoencoders (MAE)?

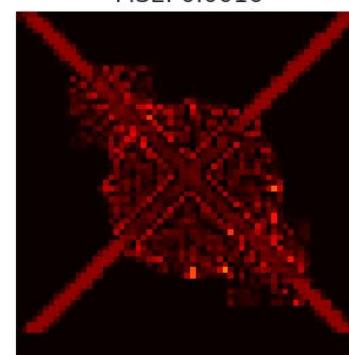
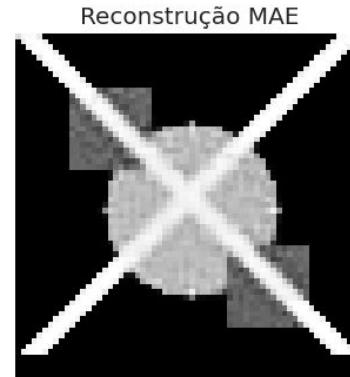
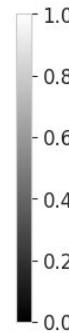
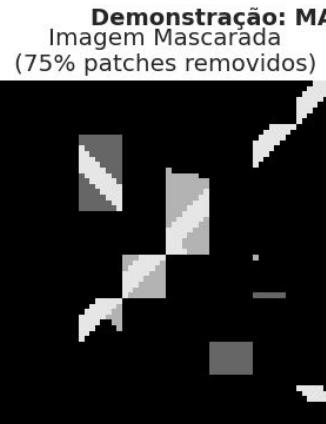
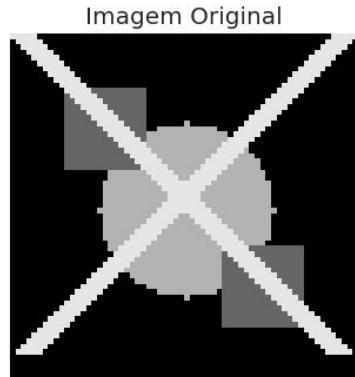
- Aprendem representações removendo partes da entrada.
- Parte da imagem/volume é mascarada (escondida).
- O encoder vê apenas os patches visíveis.
- O decoder reconstrói os patches mascarados.
- É auto-supervisionado.
- Simples, eficiente\* e escalável.
- Excelente para imagens, vídeos e dados 3D (como sísmica).



# Reconstrução de Imagem 2D - Arquitetura Proposta



# Reconstrução de Imagem 2D - Dataset e Resultados



## RESULTADOS DA DEMONSTRAÇÃO 2D

Patches visíveis: 25.0%  
Patches mascarados: 75.0%  
MSE da reconstrução: 0.001604  
PSNR: 27.95 dB

---

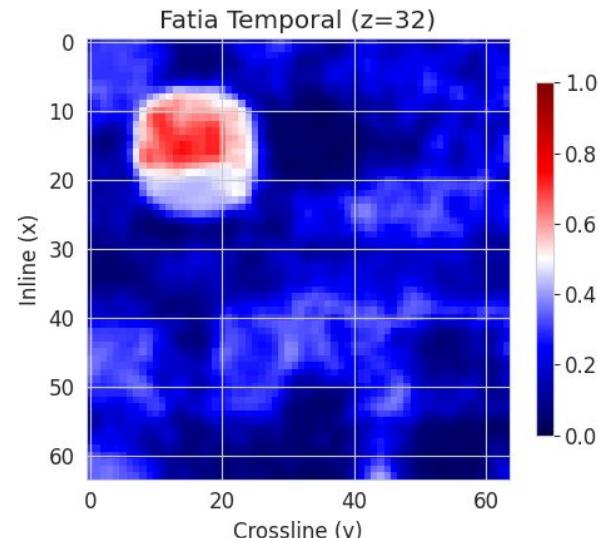
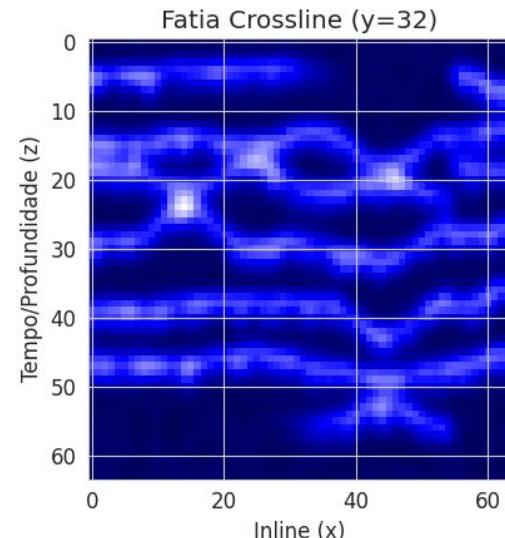
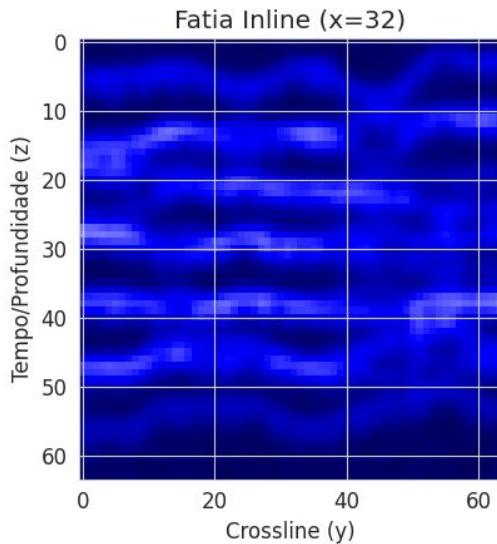
## Por que reconstruir dados 3D?

- Testar técnica de autoencoder (MAE) que parece promissora para a reconstrução de dados 3D.
- Observar a eficácia na área da sísmica.
- Comparar os resultados com abordagens tradicionais: nearest neighbor e interpolação linear

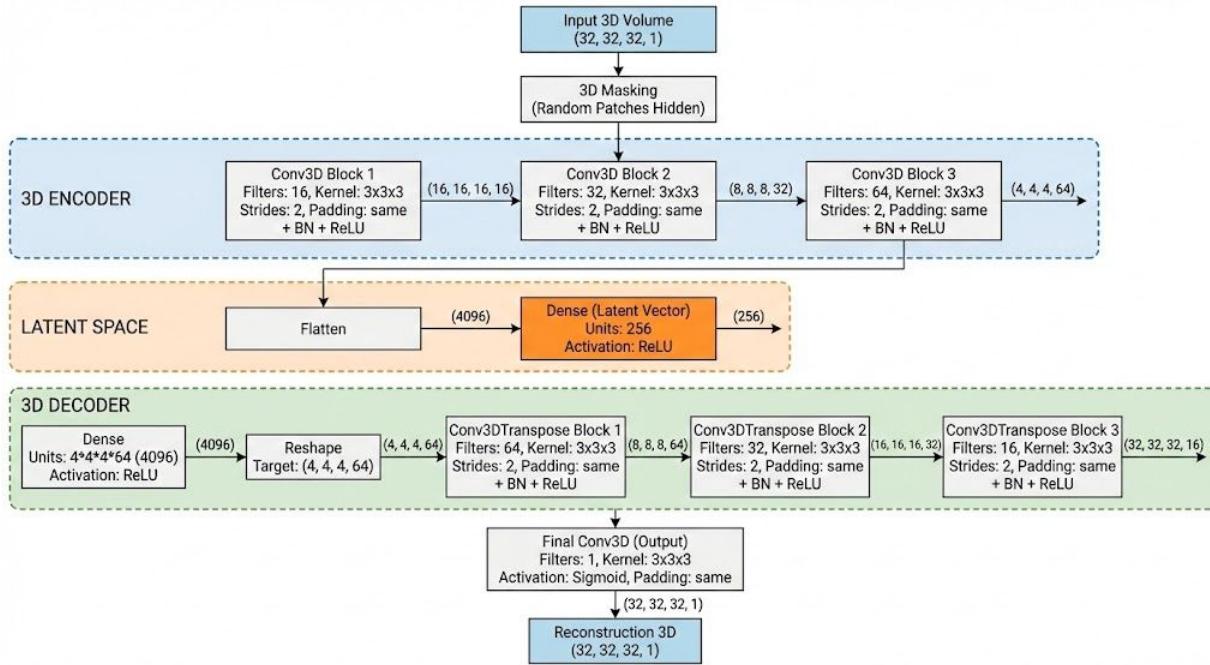


# Reconstrução de Imagem 3D - Exemplo de Dado

Visualização do Volume Sísmico Sintético 3D



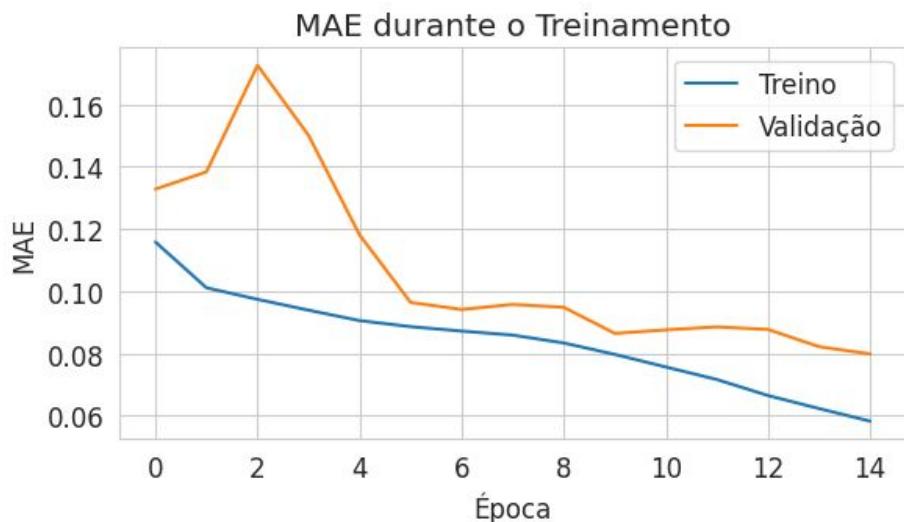
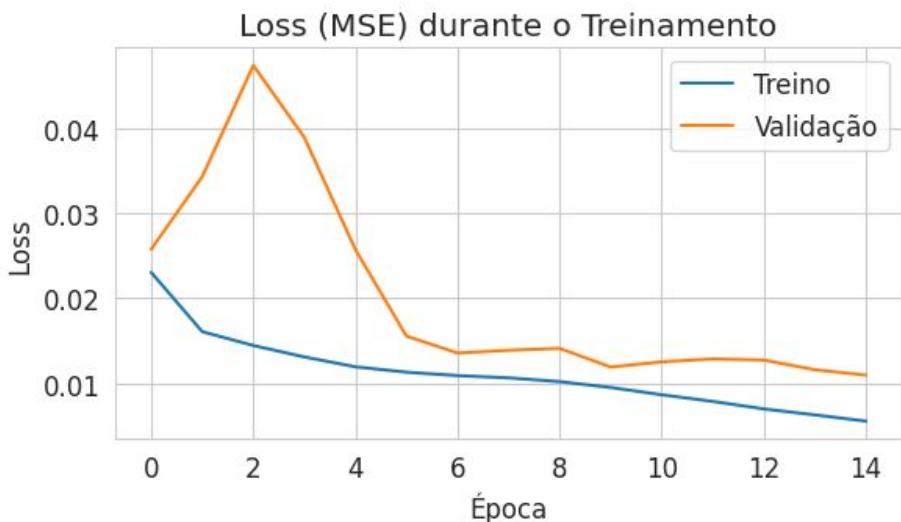
# Reconstrução de Imagem 3D - Arquitetura Proposta





# Reconstrução de Imagem 3D - Treinamento

Progresso do Treinamento do MAE 3D

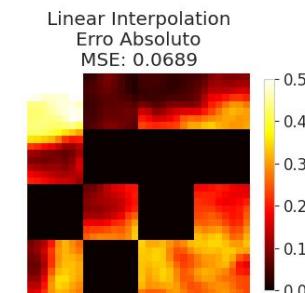
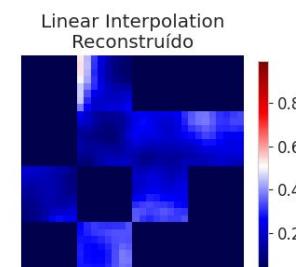
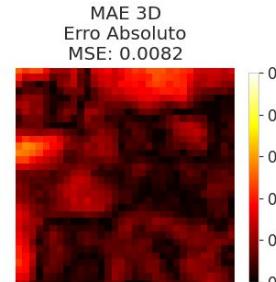
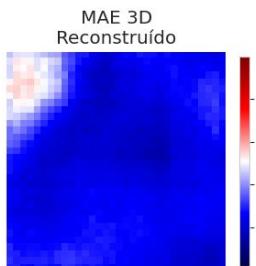
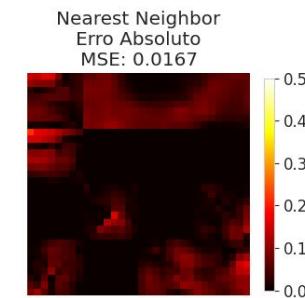
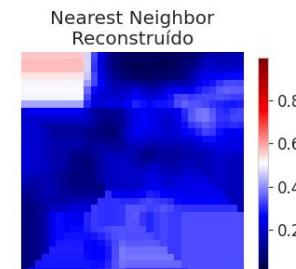
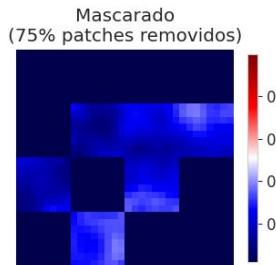
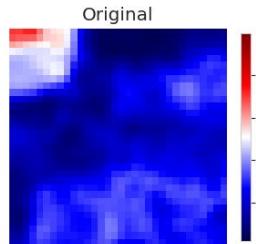




# Métricas Utilizadas para Comparação

Métrica	Quando melhor?	O que mede?	Interpretação	Definição
Mean Square Error (MSE)	menor	erro quadrático	Quão distante está em média; penaliza erros grandes.	$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2$
Mean Absolute Error (MAE)	menor	erro absoluto	Erro absoluto médio; similar mas menos agressivo que o MSE.	$\text{MAE} = \frac{1}{N} \sum_{i=1}^N  x_i - \hat{x}_i $
Signal-to-Noise Ratio (SNR (dB))	maior	relação sinal/ruído	Quanto do sinal "verdadeiro" ainda está presente em relação ao erro.	$\text{SNR} = 10 \log_{10} \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right)$
Peak Signal-to-Noise Ratio (PSNR (dB))	maior	erro relativo ao valor máximo	Versão do SNR que normaliza pela intensidade máxima possível.	$\text{PSNR} = 10 \log_{10} \left( \frac{(\text{MAX})^2}{\text{MSE}} \right)$

# Reconstrução de Imagem 3D - Resultados



# Reconstrução de Imagem 3D - Resultados

```
=====
```

COMPARAÇÃO: MAE 3D vs MÉTODOS TRADICIONAIS

```
=====
```

Preparando cenário de teste...

Volume de teste: (32, 32, 32)

Porcentagem mascarada: 75.0%

1. Reconstrução com MAE 3D...
  2. Interpolação por vizinho mais próximo...
  3. Interpolação linear 3D...
- ```
=====
```

RESULTADOS COMPARATIVOS

```
=====
```

| Método               | MSE      | MAE      | SNR (dB) | PSNR (dB) |
|----------------------|----------|----------|----------|-----------|
| MAE 3D               | 0.008227 | 0.073657 | 9.44     | 20.82     |
| Nearest Neighbor     | 0.016698 | 0.091433 | 6.37     | 17.75     |
| Linear Interpolation | 0.068949 | 0.217880 | 0.21     | 11.59     |

```
=====
```

---

# Conclusão

- O Masked Autoencoder (MAE) apresentou um ótimo desempenho.
- Mostrou resultados melhores que abordagens tradicionais, como o nearest neighbors e a interpolação linear.
- Apresenta um potencial muito alto para aplicação na geofísica, preenchendo dados faltantes.
- O MAE é promissor para aplicação em dados de aquisição sísmica reais:
  - Preencher gaps de dados
  - Data augmentation
  - Compressão de dados sísmicos
- Limitações e desafios:
  - Alto custo computacional para treinamento
  - Necessidade de grandes volumes de dados para treinamento eficaz
  - Sensibilidade a hiperparâmetros da arquitetura
- Pretendemos aplicar o MAE em dados sísmicos reais e testar arquiteturas mais eficientes.

---

# Referências

- He, K., Chen, X., Xie, S., Li, Y., Dollar, P., & Girshick, R. (2021). "Masked Autoencoders Are Scalable Vision Learners". arXiv preprint arXiv:2111.06377.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". International Conference on Medical Image Computing and Computer-Assisted Intervention.
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). "Image quality assessment: from error visibility to structural similarity". IEEE Transactions on Image Processing, 13(4), 600-612.



**Obrigado!**