GEOPHYSICAL FOUNDATION MODEL: UNMASKING THE TRACES

EXECUTIVE SUMMARY

ThinkOnward's Geophysical Foundation Model is a pre-trained Vision Transformer trained on 450 synthetically generated Synthoseis 3D seismic volumes. We use a new elastic architecture and trace masking process to fine-tune the Geophysical Foundation Model for the downstream task of seismic interpolation. We use 50 3D seismic volumes from the Patch the Planet Challenge, hosted by ThinkOnward as our benchmark hold-out dataset. Using a Structural Similarity Index Metric (SSIM) to compare results we document that the Geophysical Foundation Model is 2-3 times better than Shang et al. (2023), and similar to Lasscock et al. (2024).

INTRODUCTION

Sheng et al.'s (2023) innovative Seismic Foundation Model (SFM) paper lays the groundwork for the first peer reviewed geophysics based foundational model. The model builds upon the Vision Transformer Masked Autoencoder proposed by Meta in He et al. (2021). Sheng et al. (2023) curated a large seismic dataset containing more than two million images from various basins around the world. Sheng et al. (2023) used random masking with 14x14 pixel size patches to pre-train their Vision Transformer, and used the encoder weights to fine-tune for specific downstream tasks, like interpolation.

PROBLEM AND APPROACH

There are two key issues with the SFM. First, seismic input images are required to be 224x224 pixels in size, due to the architecture of the model. These are not realistic image sizes for 2D seismic data. The second issue is the masking method for seismic images during pre-training. The method depends upon the size of the images and is limited to square patches.

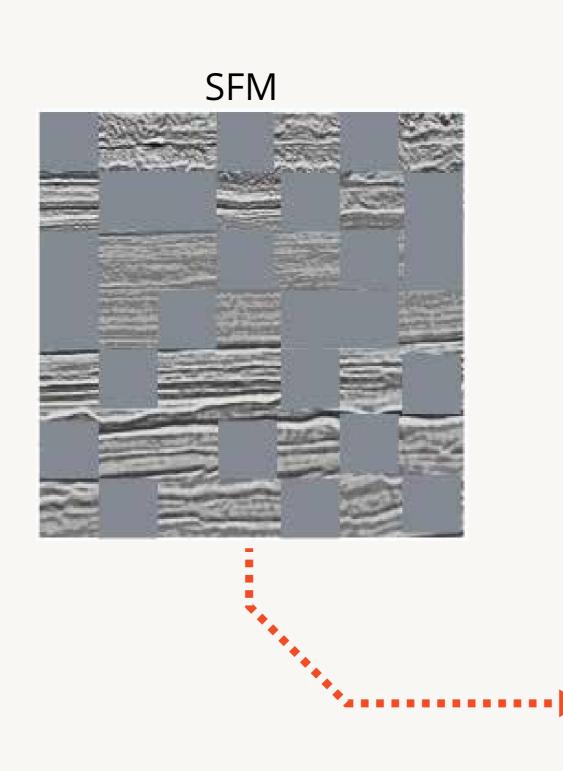
Conceptual encoder/decoder architecture for the trace based masking

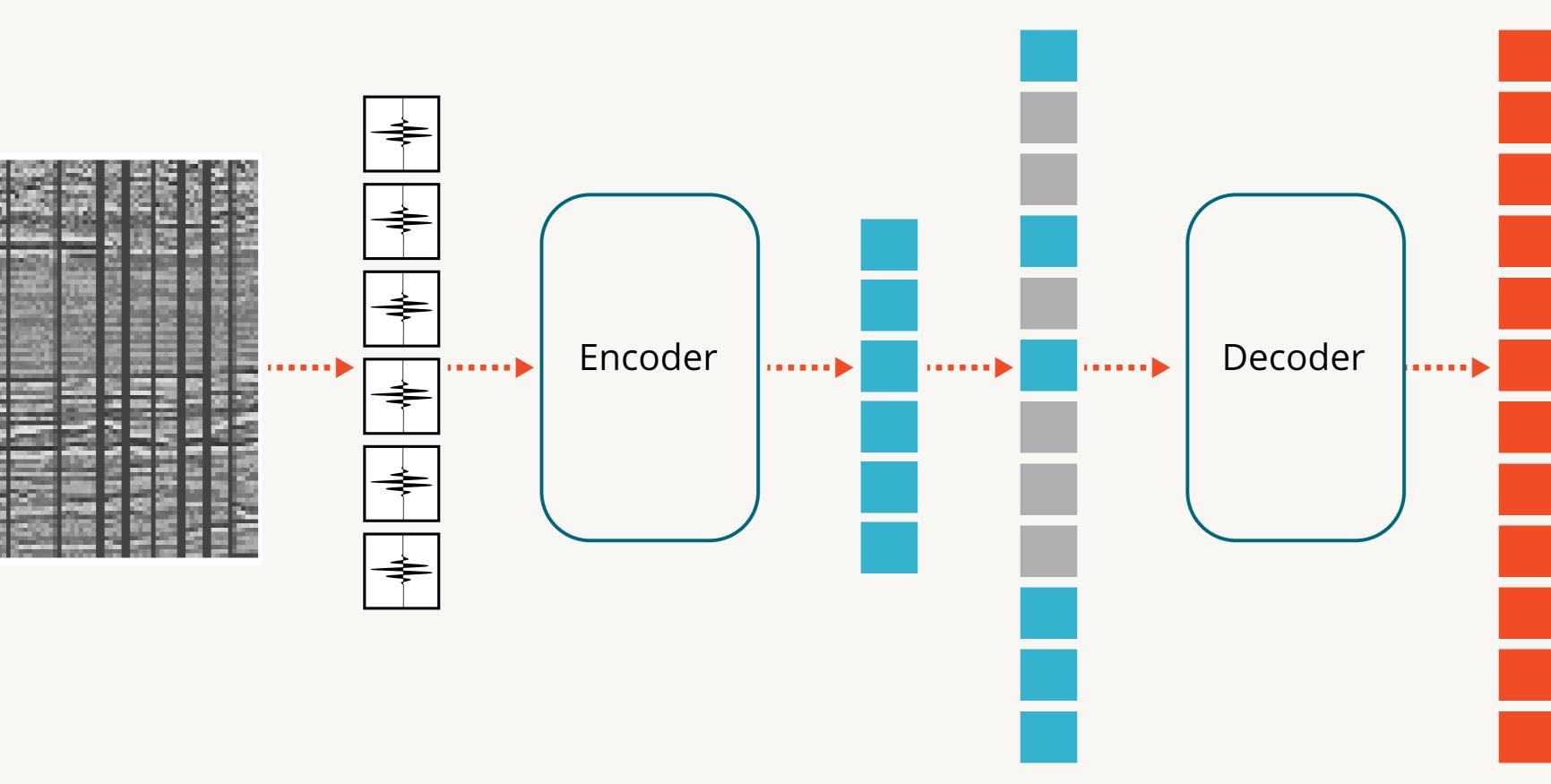
10	8	8	5	12
	B	ř.	ä.	- 52
		H	5	23
12	E	H	Ξ.	2
85	E	ŝ	8	13
	B	8		
2	i.	k	8	8
12	H	8	8	18
	i.	E	2	12
56	ie,	2		5
	8	E	5	
2	5			10
	Į,			
		r,	ŝ.	8
2	R	в	i.	3

ThinkOnward's Geophysical Foundation Model addresses the image size and masking issues by rethinking the entire process. We considered the question: "What is the fundamental aspect of a seismic dataset?". Our answer to this question is the seismic trace, a 1D representation of the data recorded for 1 channel. By adjusting the data masking process to traces and making adjustments to the internal architecture, the model not only accepts seismic images of any size for pre-training and fine-tuning, but also honors the fundamental aspect of seismic data. With the trace-based masking and data loading pipeline we train the backbone on 450 synthetic seismic volumes released as part of the Patch the Planet Challenge. After pre-training the backbone, we then designed and built a regression head for the Vision Transformer so that we can then fine tune for a variety of downstream tasks. For this poster we are interested in interpolating missing traces, with an example of one additional downstream task







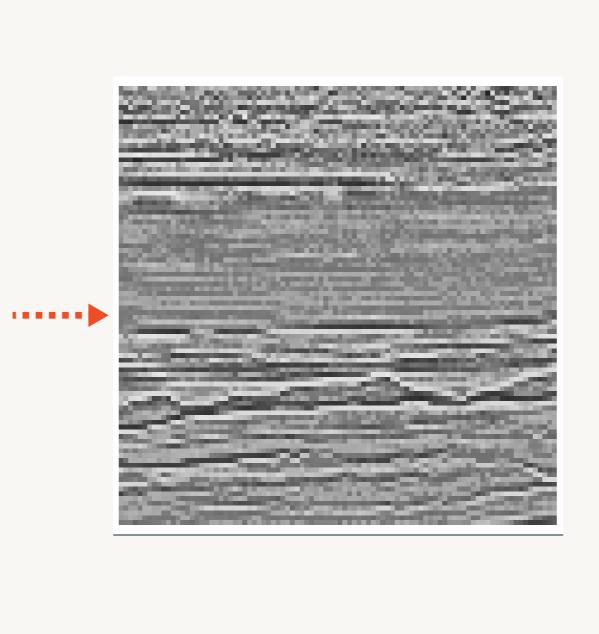


Credit to: Mike McIntire, Ognjen Tanovic, Jakub Mazura, Nate Suurmeyer, Jesse R. Pisel Shoutouts to the following for making this work possible: Shell plc, Menal Gupta, Nataliia Manakova, Vitalii Panchyk, Rodrigo Assar

Masking Techniques

ThinkOnward

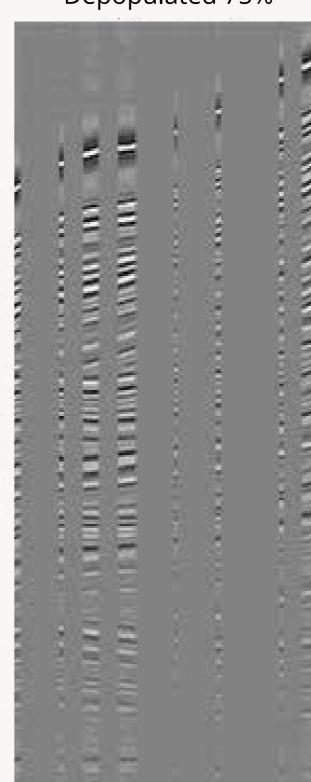
		100
		22
191123133	CONTRACTOR & TOMA TO ANY DR. CONTRACTOR	18 II.
	Party and a start of the start	÷
		82
and the second second		1
and it is not the state	LA MAR INC	
State I and the state of	The second of the second	
and a set one has	the second states and the second s	
density of the subscription	wants publicated arrange of the later had	÷.
the second se	Contraction of the local section of the local secti	
the second second second second	and the second se	
and the second second		
and the second second second	the R. Longian Andrews The Tax Man And All and	
And the Party of t	THE REAL PROPERTY AND	
The second s	The second is the second of th	
	wanted with the second s	
and the support	and a second of the second of	
TLE STREET STREET	A second comments what we have been and the second se	
and the second second	and the second s	
and the second se	COMPANY AND	
The second second	Contraction of the local data with the local line and	
And A DE CONTRACTOR	and some start and the start and the	
Contraction of the owner of the owner of the	the second design of the second s	-
	Comparison and a second s	
and the local division of the	Many succession without the last the last the last the	
THE R P. LEWIS CO., NAME OF	states in the second states in the second states of the second states and	at 1
and a first with the local sector		
	the same and the same and the same and the same and	
	the second	
and the second second	and a second state of the	8
and the local division of the local division	and the second states when the second states in the	
and the second se	THE ADDRESS TO BE ADDRESS TO ADDR	
State of the local division of the local div	And Annual and and the second second	
Deres in the later water	the second secon	
and a state of	stand of the second of the second sec	
NAME OF TAXABLE PARTY.		
of the local division of the local division of the	and have been a series of the second s	



RESULTS / LESSONS LEARNED

For the interpolation downstream task we find that the Geophysical Foundation Model performs reasonably well even with highly depopulated seismic volumes.

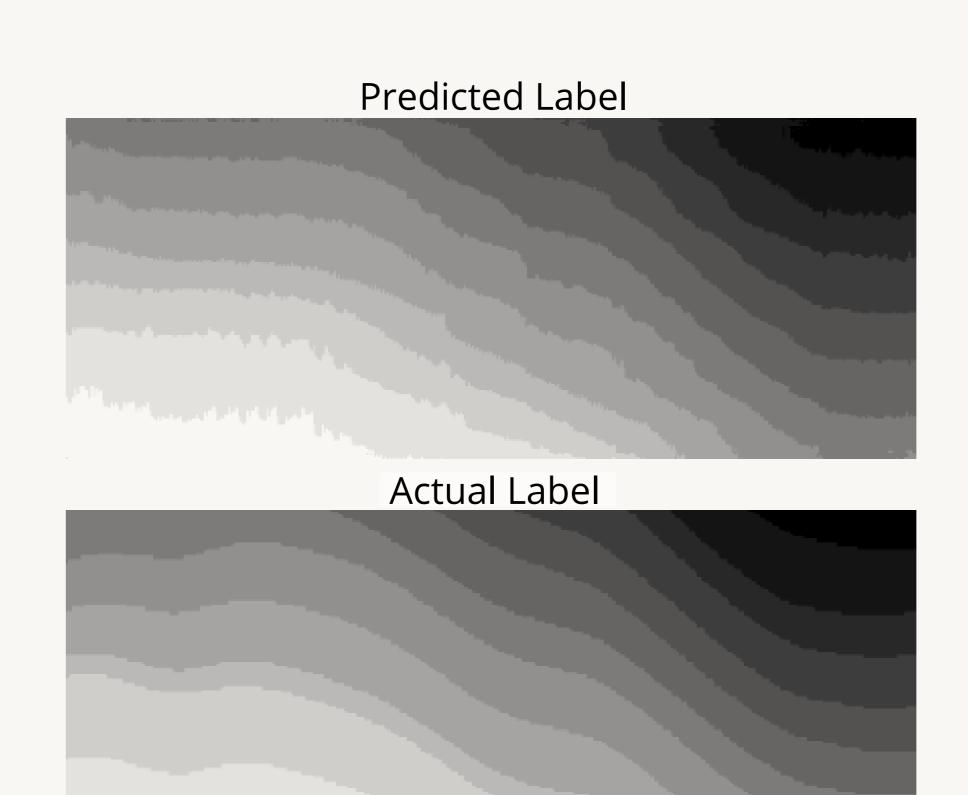




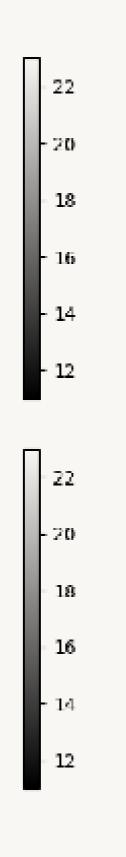


Masking traces increases overall performance over masking patches

For the segmentation downstream task we find that the Geophysical Foundation Model performs reasonably well at segmenting seismic horizons with a classification head.



ML Reconstructed Seismic



RECOMMENDATIONS

- Trace based masking should be the standard for subsurface vision models
- SSIM should be used for ground truth regression head task evaluations
- •A variety of downstream tasks need to be evaluated

Continued development for the regression and classification heads are key for different downstream tasks.

Regression head downstream tasks include:

- De-noising
- Interpolation
- Velocity Modeling

Classification head downstream tasks include:

- Horizon segmentation
- Geobody segmentation
- Fault segmentation

Additional evaluation on how traces are masked (random, uniform, or clustered) will also help inform future models in this space.

