ML4Seismic Partners Meeting 2023

What Uncertainties do we need in Deep Learning for Seismic Interpretation?

Ryan Benkert, Mohit Prabhushankar, and Ghassan AlRegib

Nov. 7th – Nov. 9th

OLIVES @GeorgiaTech

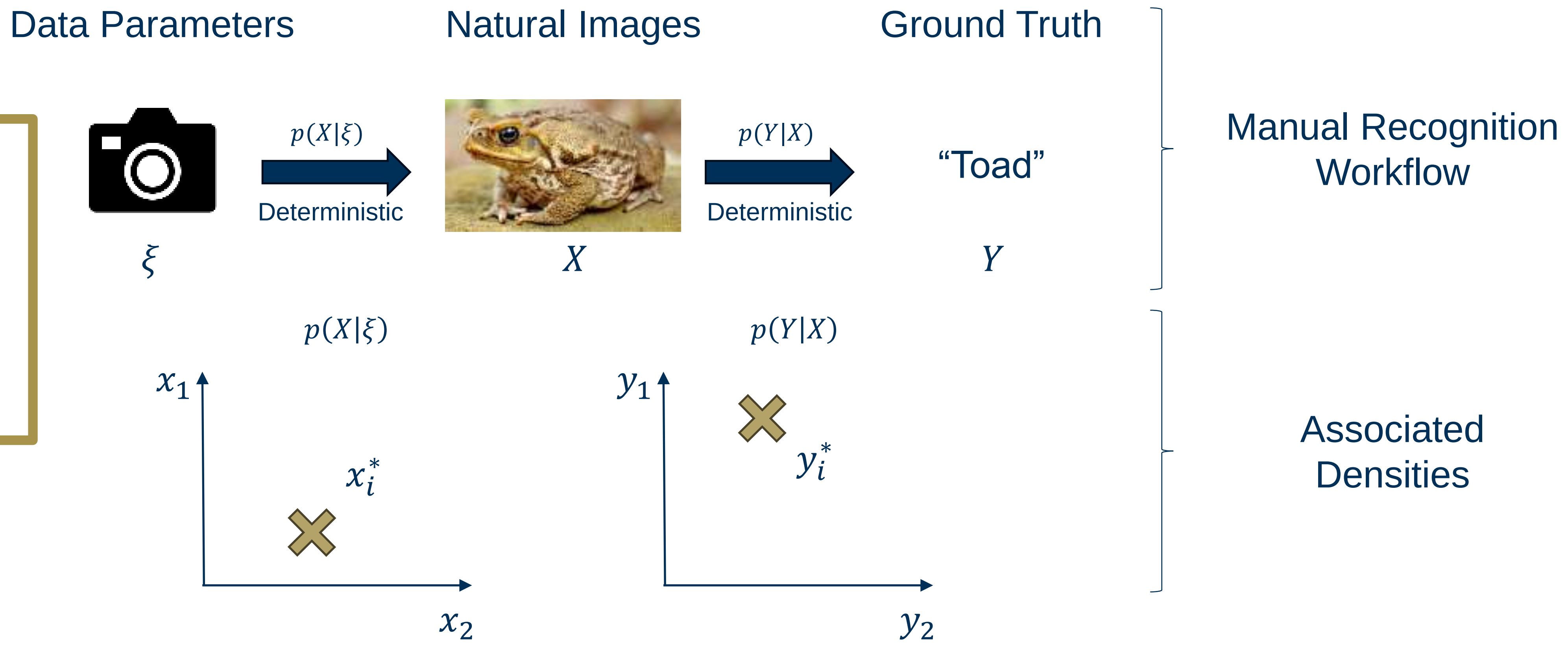
Georgia Tech

Assumptions for Natural Image Data

Labels and Image Curation are Deterministic, and Ground Truth Information is Available

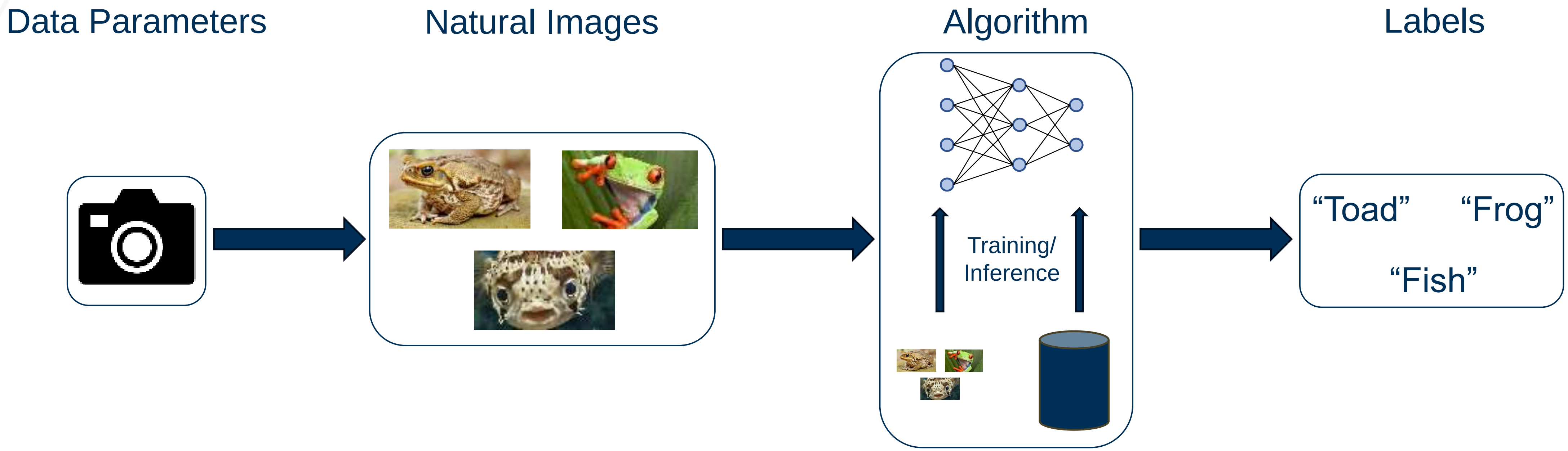
Probabilistic Characteristics of Natural Images

Data Assumptions:
- Deterministic **image curation**
- Deterministic **labels**
- **Labels are ground truth**



Assumptions of Current Uncertainty Frameworks are Reasonable for Natural Images

Uncertainty Frameworks for Natural Images



Data | Images | Ground Truth

$$\xi \xrightarrow[\text{Deterministic}]{\delta(x - x_i^*)} X \xrightarrow[\text{Uncertainty}]{p(Y|X,W)} Y$$

Assumptions for Seismic Image Data

Labels and Image Curation are Probabilistic, and no Ground Truth Information is Available

Probabilistic Characteristics of Seismic Images

Data Assumptions:

- Probabilistic **image curation**
- Probabilistic **labels**
- **Labels are interpretations**

Data Parameters

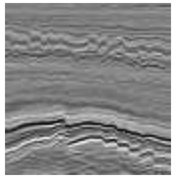
e.g. Sensors

ξ

$p(X|\xi)$

Probabilistic

Seismic Images



X

$p(\alpha|X)$

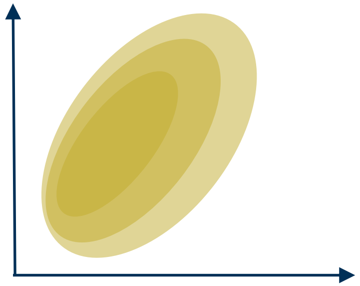
Probabilistic

Interpretations

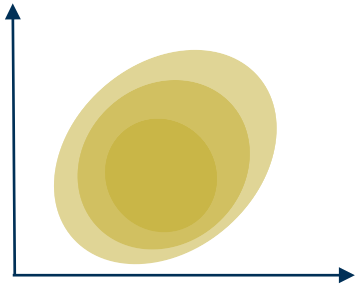
$\alpha \sim p(\alpha|X)$

Manual Interpretation Workflow

$p(X|\xi)$



$p(\alpha|X)$



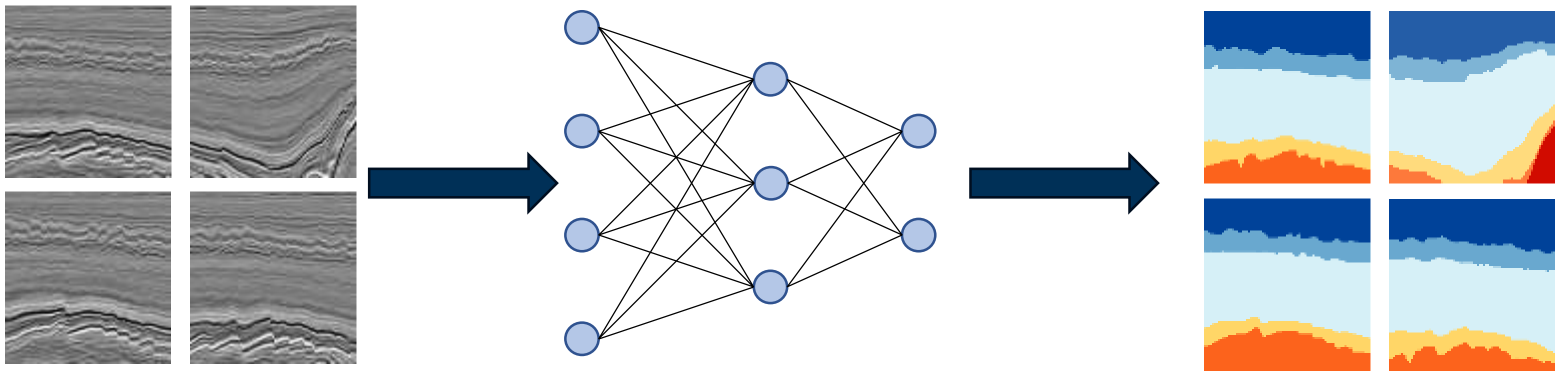
Associated Densities

Current Uncertainty Frameworks in Machine Learning

Current Frameworks Estimate Uncertainty within the ML Pipeline, and not the Interpretation Pipeline

Current Uncertainty Frameworks in Seismic Interpretation

Seismic Data | Algorithm | Labels



Data | Ground Truth

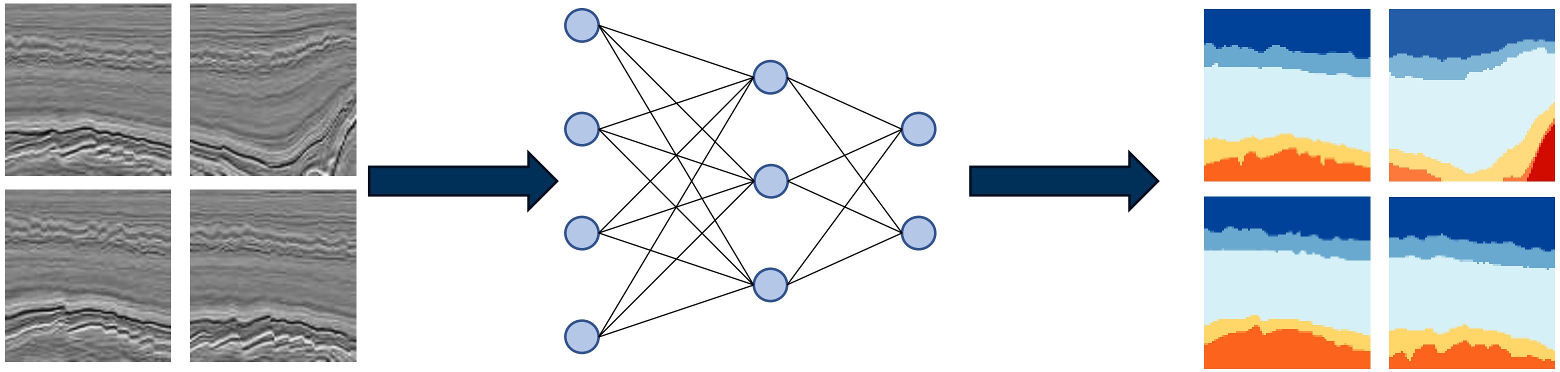
$$X \xrightarrow[\text{Uncertainty}]{p(Y|X,W)} Y$$

Current Uncertainty Frameworks in Machine Learning

Current Frameworks Estimate Uncertainty within the ML Pipeline, and not the Interpretation Pipeline

Current Uncertainty Frameworks

Deterministic! Seismic Data   Algorithm   Labels Deterministic!



Data   **Ground Truth**

$$X \xrightarrow{p(Y|X,W)} Y$$

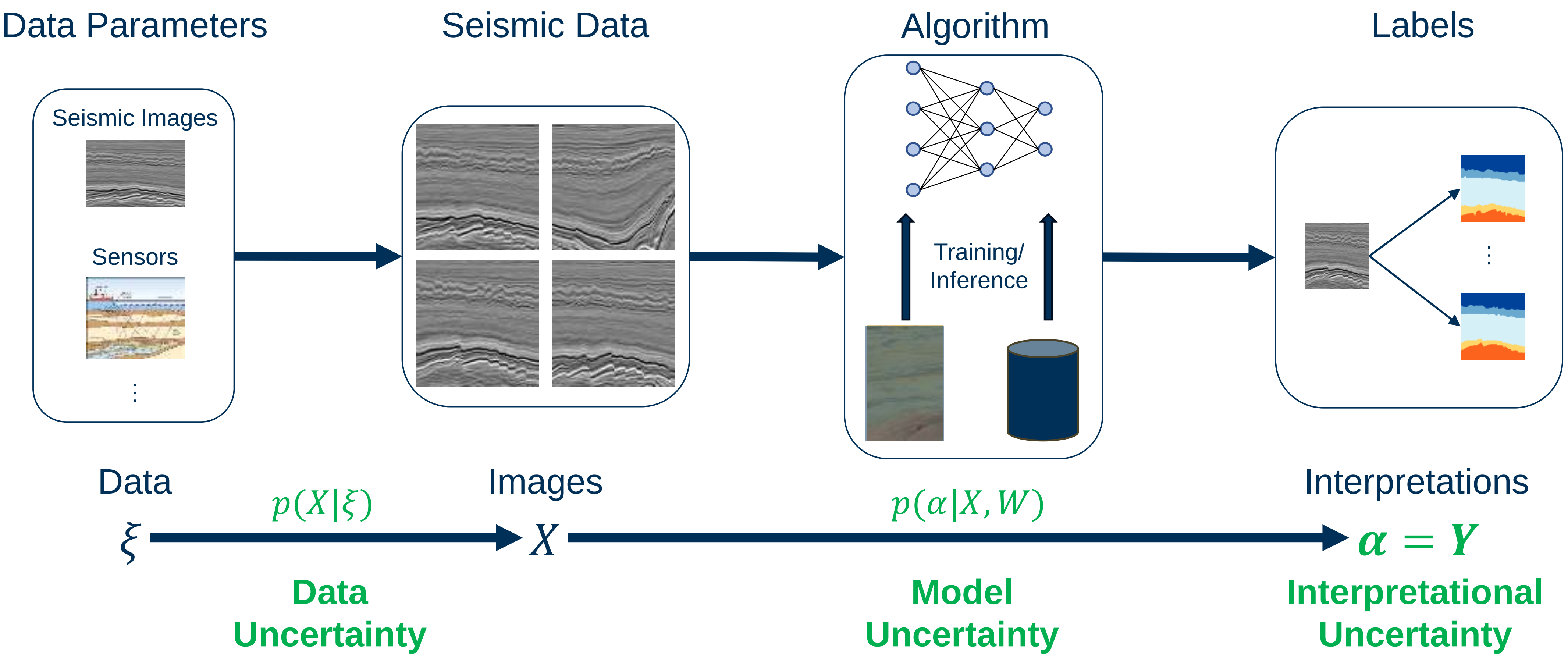
Uncertainty

Current frameworks **assume deterministic data and ground truth labels** and does not reflect the practical interpretation pipeline

Proposed Uncertainty Framework

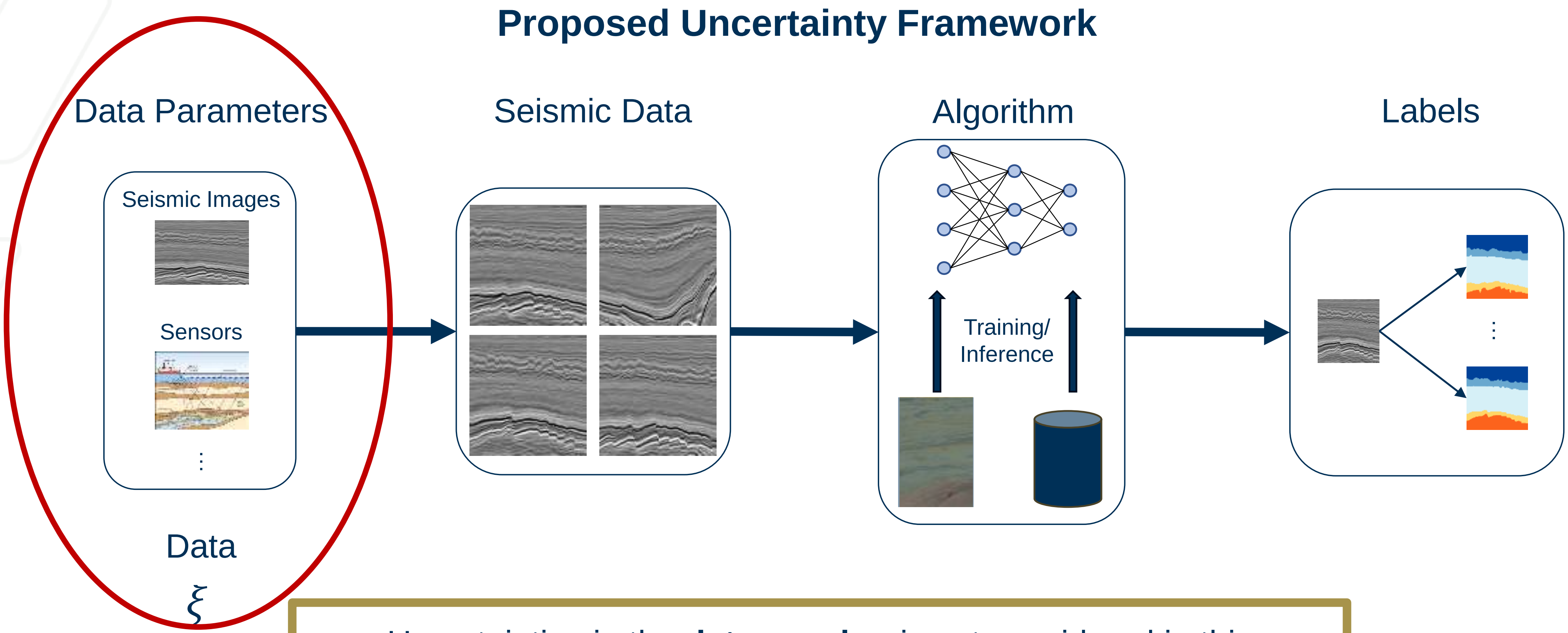
Our Framework Reflects the Processing Steps of the Entire Interpretation Pipeline

Proposed Uncertainty Framework



Proposed Uncertainty Framework

Our Framework Reflects the Processing Steps of the Entire Interpretation Pipeline



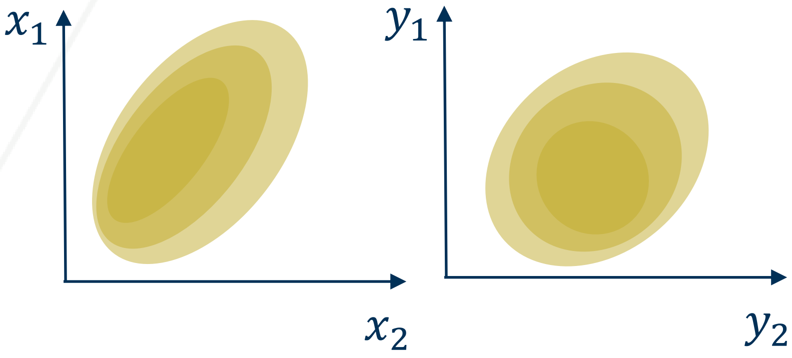
Uncertainties in the **data sensing** is not considered in this framework. We focus on uncertainties **after data collection**.

Presentation Structure and Outline

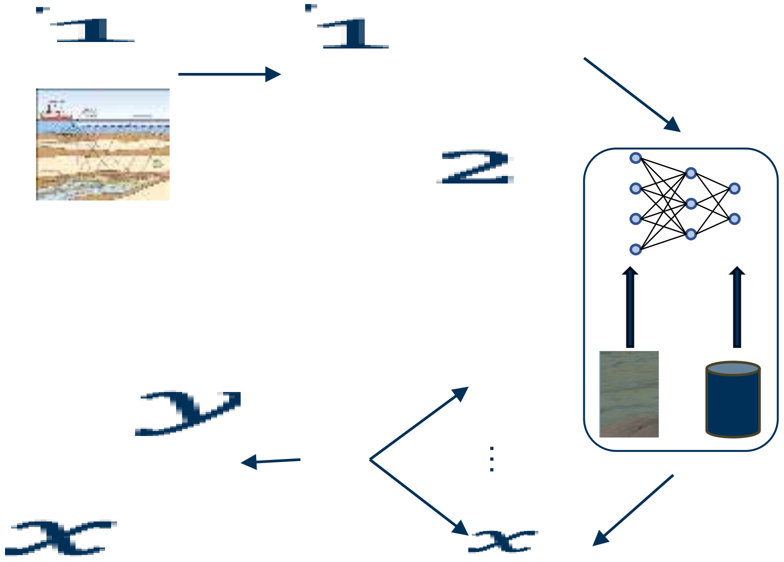
Discussion Topics are Active Learning, ATLAS, and Results

Contributions: Propose Uncertainty Framework for Seismic Interpretation

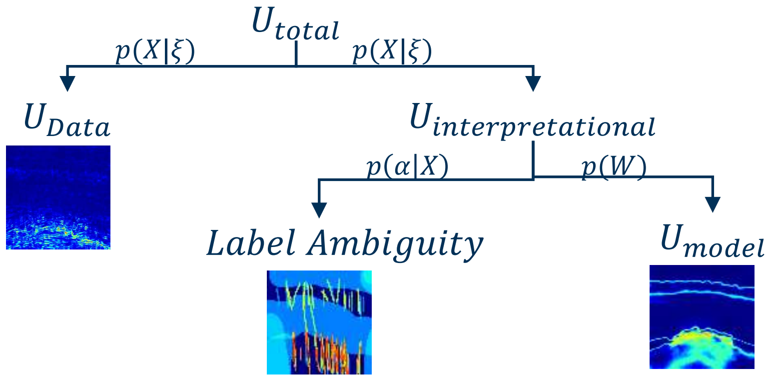
Background: Probabilistic Data



Framework: Uncertainty Sources



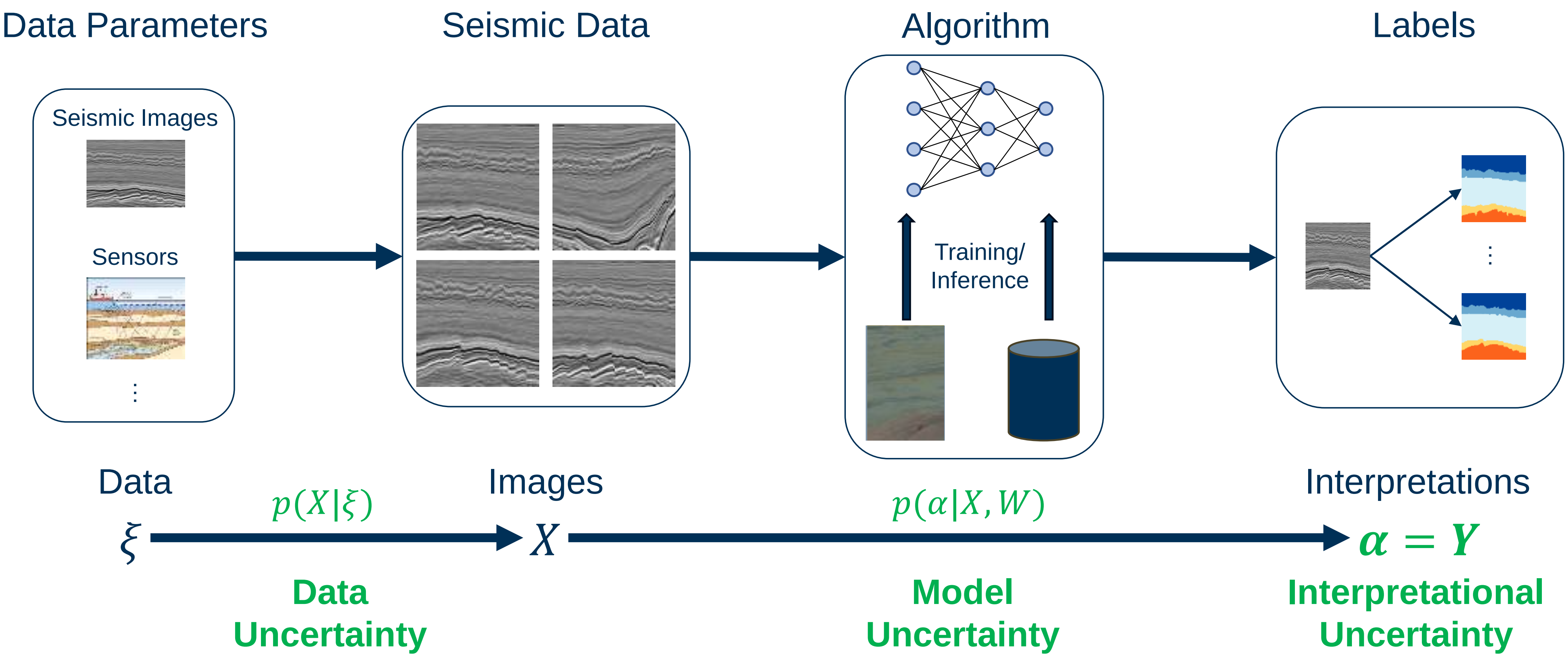
Uncertainty Relationships



Proposed Uncertainty Framework

Our Framework Reflects the Processing Steps of the Entire Interpretation Pipeline

Proposed Uncertainty Framework



Framework: Observational Uncertainty

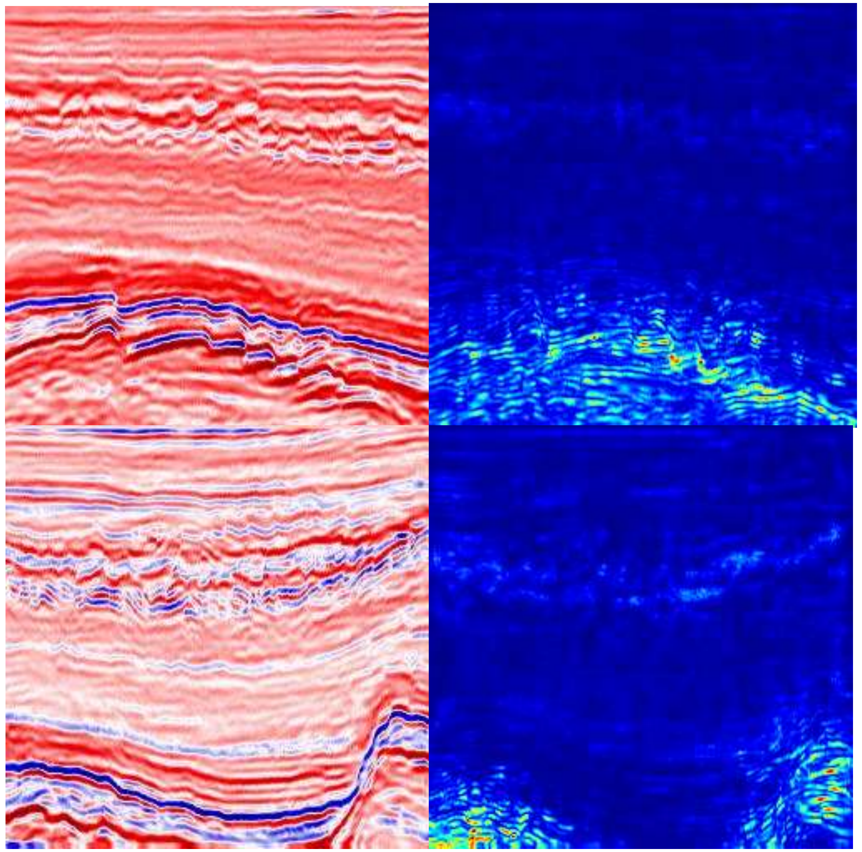
Observational Uncertainty Captures Stochasticity in the Seismic Images

Observational Uncertainty

- Observational uncertainty captures **stochasticity in the images**
- Relevant density:

$$p(X|\xi)$$

- **Measure uncertainty through the spread in $p(X|\xi)$**
- **Propagates** to later stages of the pipeline



Toy Example

Framework: Model Uncertainty

Model Uncertainty Captures Stochasticity in the Neural Network Parameters

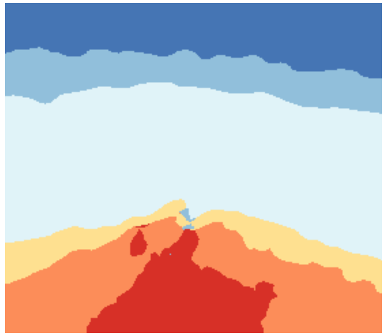
Model Uncertainty

- Algorithmic uncertainty models uncertainty in the **algorithmic parameters**
- Can be **reduced** with more **data**
- Source of uncertainty:

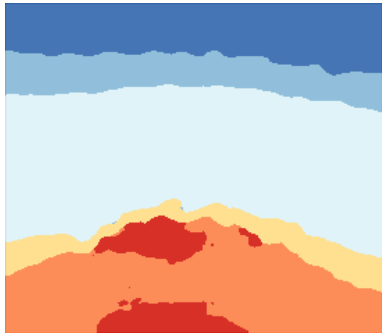
$$p(\mathrm{W})$$

- **Measure uncertainty through the stochasticity in $p(W)$**

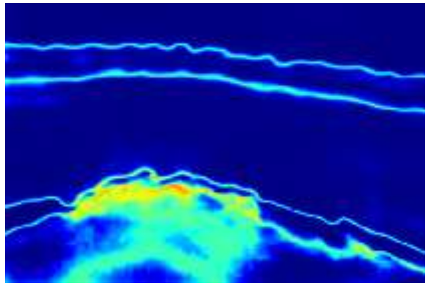
$p(\alpha|X, W)$



⋮



U_{Model}



Framework: Interpretational Uncertainty

Interpretational Uncertainty Captures Stochasticity in the Annotations

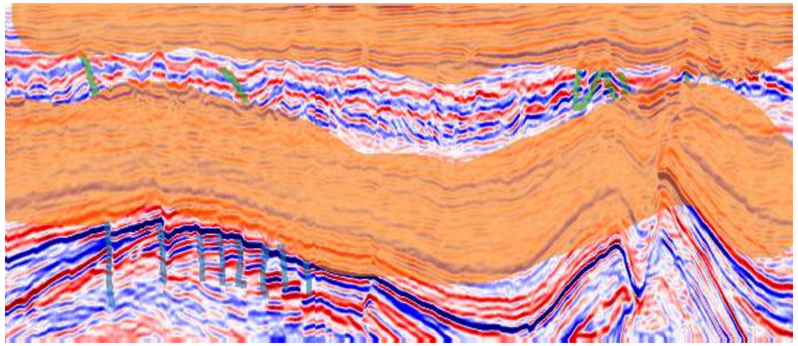
Interpretational Uncertainty

- Interpretational uncertainty captures **stochasticity in the output interpretations**
- Relevant density:

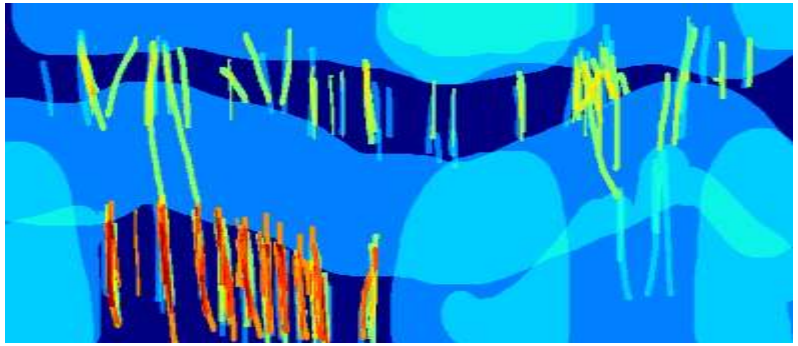
$$p(\alpha|X)$$

- **Measure uncertainty through the spread in $p(\alpha|X)$**
- Is **influenced by model uncertainty** if interpretations are predicted by a neural network

$p(\alpha|X), U_{Model} = 0$

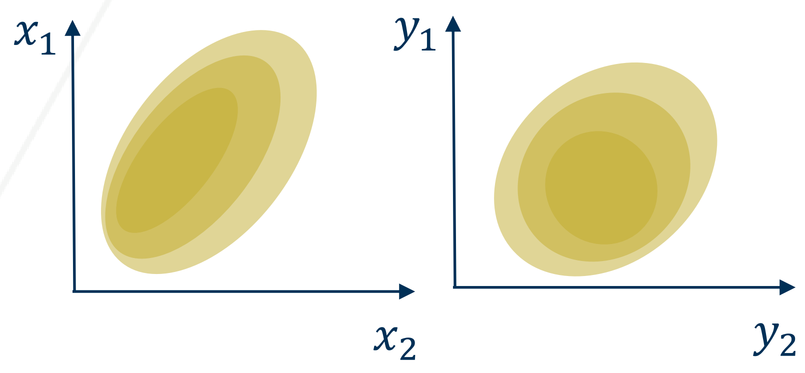


$Label\ Ambiguity$

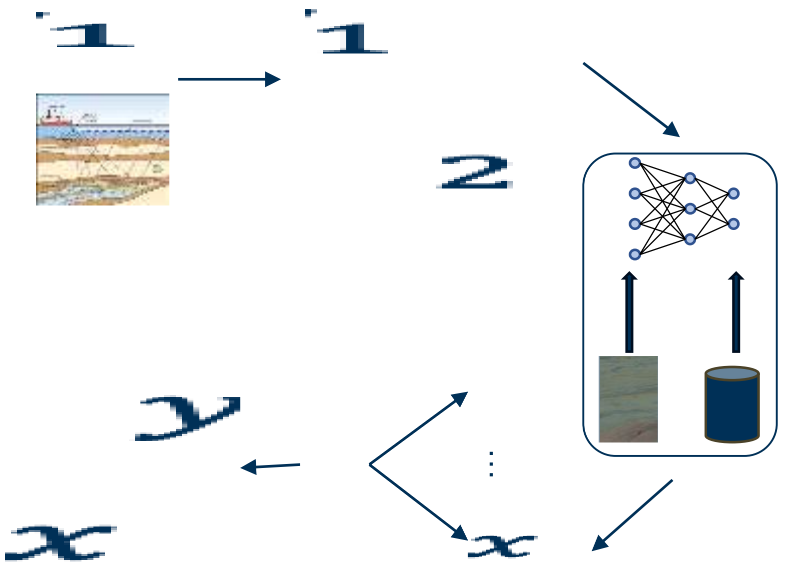


Contributions: Propose Uncertainty Framework for Seismic Interpretation

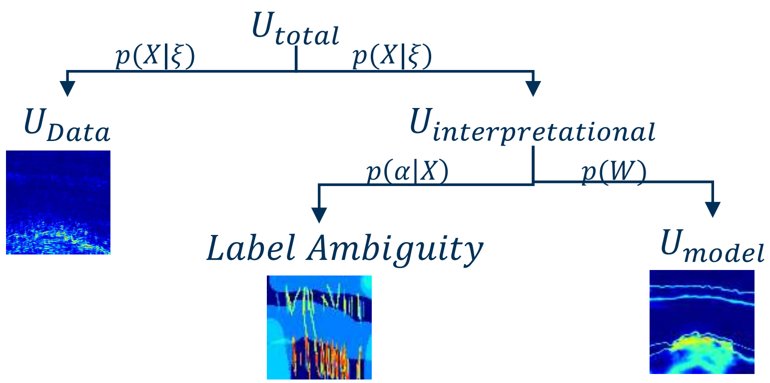
Background: Probabilistic Data



Framework: Uncertainty Sources



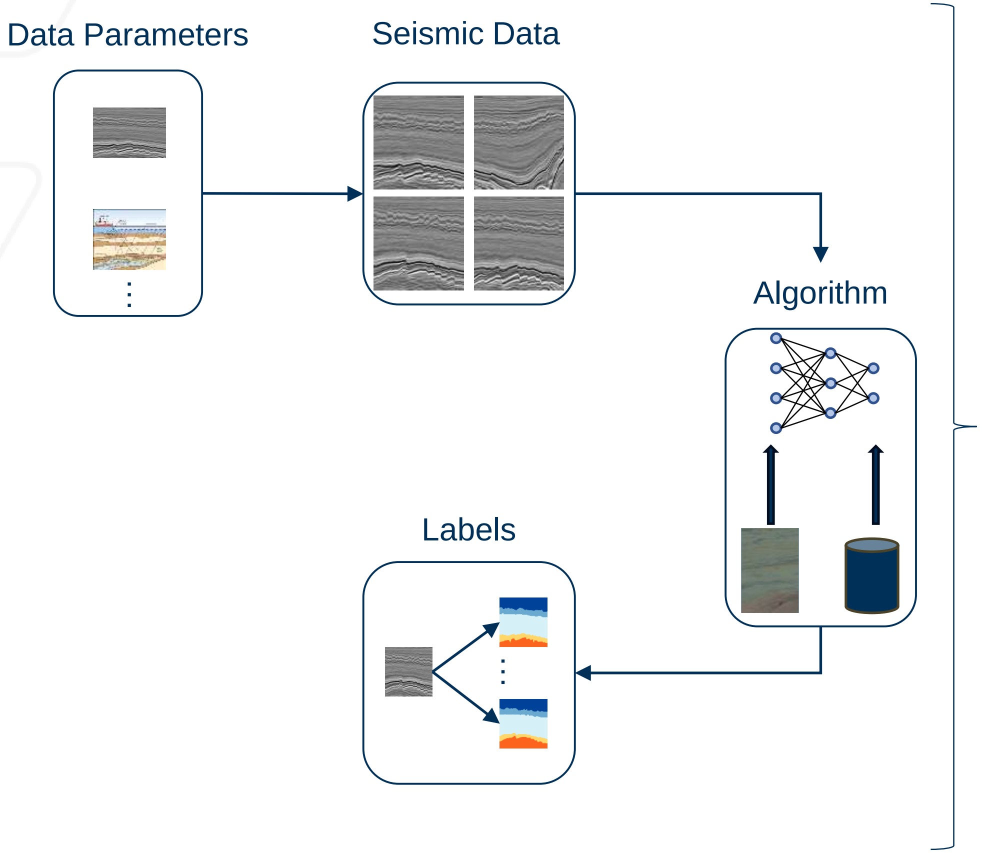
Uncertainty Relationships



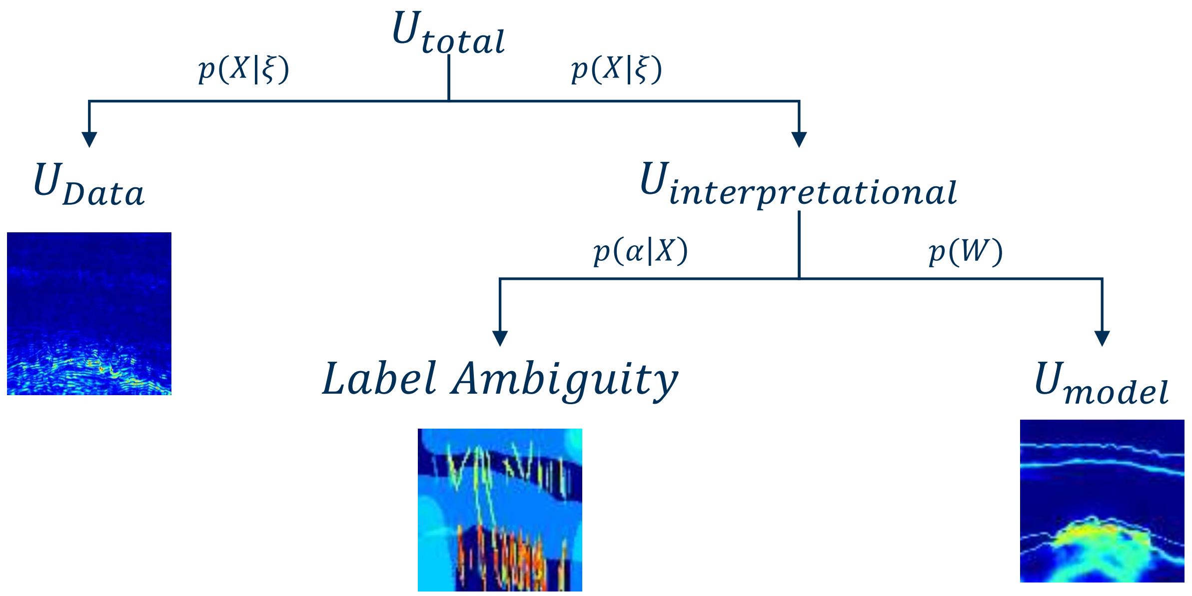
Relationships of Uncertainty
Uncertainty Relationships Propagate Throughout the Pipeline

PLACEHOLDER Reflective Uncertainty

Interpretation Workflow



Uncertainty Relationships



Uncertainty propagates through the interpretation pipeline.

Relationships of Uncertainty

Data Uncertainty can be Calculated from the Total Uncertainty

Data Uncertainty

- Full interpretation pipeline $\boldsymbol{p(\alpha|\xi)}$ **can be decomposed** in sub-posteriors:

$$p(\alpha|\xi) = \int p(\alpha|X)p(X|\xi)\, dX$$

- Data uncertainty is measurable in the **spread over $\boldsymbol{p(X|\xi)}$:**

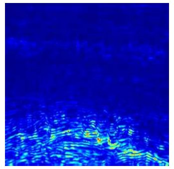
$$U_{Data} = \underbrace{H[p(\alpha|\xi)]}_{U_{total}} - \underbrace{\mathbb{E}_{p(X|\xi)} H[p(\alpha|X)]}_{U_{Interpretational}}$$

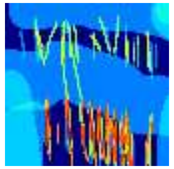
where $\boldsymbol{H}$ **is the entropy** of a distribution

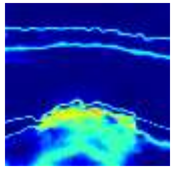
U_{total}

$p(X|\xi)$ $p(X|\xi)$

U_{Data} $U_{interpretational}$



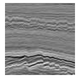




Alternative Explanation

$$U_{Data} = \underbrace{H[p(\alpha|\xi)]}_{} - \underbrace{\mathbb{E}_{p(X|\xi)} H[p(\alpha|X)]}_{}$$

 … 

 … 

Full Pipeline

Remaining Pipeline after data processing

Relationships of Uncertainty

Model Uncertainty Captures Stochasticity in the Neural Network Parameters

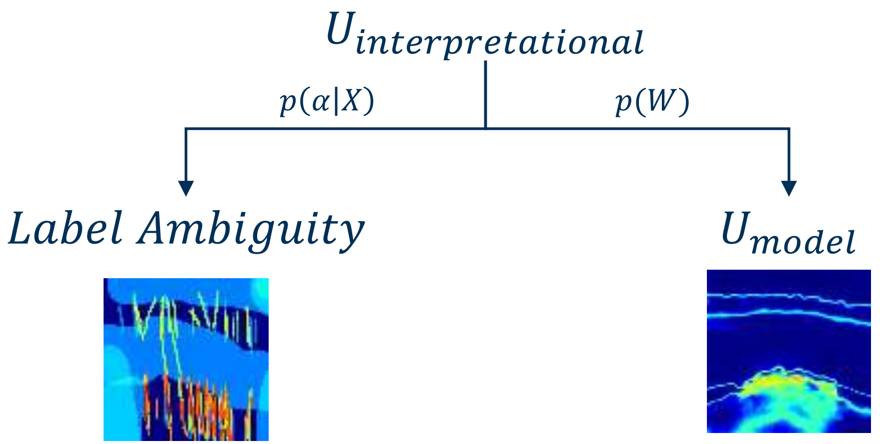
Model Uncertainty

- Remaining interpretation pipeline $\boldsymbol{p(\alpha|X)}$ **can be decomposed** in sub-posteriors:

$$p(\alpha|X) = \int \underbrace{p(\alpha|X,W)}_{\text{Label Ambiguity}}\underbrace{p(W)}_{\text{Model}}\, dX$$

- Data uncertainty is measurable in the **spread over $\boldsymbol{p(W)}$:**

$$U_{Model} = H\left[\mathbb{E}_{p(W)}H[p(\alpha|X,W)]\right] - \mathbb{E}_{p(W)}H[p(\alpha|X,W)]$$



Relationships of Uncertainty

Interpretational Uncertainty is a Function of Model Uncertainty and Label Ambiguity

Interpretational Uncertainty

- Uncertainty in the **interpretation** is given by the **posterior $\boldsymbol{p(\alpha|X)}$:**

$$U_{Interpretational} = \mathbb{E}_{p(X|\xi)} H[p(\alpha|X)]$$

- When interpretations are given by a **model:**

$$U_{Interpretational} = \mathbb{E}_{p(X|\xi)} H[p(\alpha|X)]$$

$$= \mathbb{E}_{p(X|\xi)} H\left[\mathbb{E}_{p(W)} p(\alpha|X,W)\right]$$

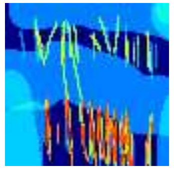
$$= \mathbb{E}_{p(X|\xi)}[U_{Model} + \mathbb{E}_{p(W)} H[p(\alpha|X,W)]]$$

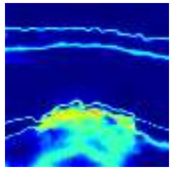
$$\mathbb{E}_{p(X|\xi)}[U_{Model} + \mathbb{E}_{p(W)} H[p(\alpha|X,W)]]$$

$U_{interpretational}$

$p(\alpha|X)$ $p(W)$

Label Ambiguity U_{model}





Explicit Relationship

$$U_{Interpretational} = \mathbb{E}_{p(X|\xi)}[U_{Model} + \underbrace{\mathbb{E}_{p(W)} H[p(\alpha|X,W)]]}_{\text{Label Ambiguity}}$$

Relationships of Uncertainty

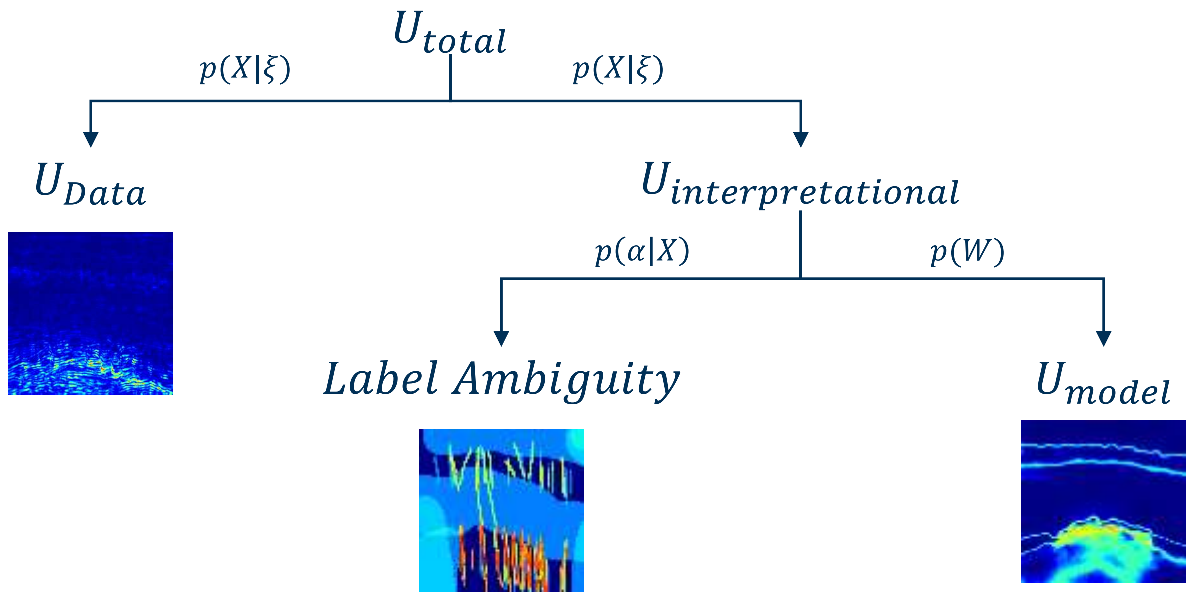
All Uncertainty Sources together Constitute Total Uncertainty

PLACEHOLDER Reflective Uncertainty

Explicit Relationship

$$U_{total} = U_{Data} + \mathbb{E}_{p(X|\xi)}[U_{Model} + Label\ Ambiguity]$$

Uncertainty Relationships



Uncertainty propagates through the interpretation pipeline.

Thanks for Listening

Questions?

ML4SEISMIC

OLIVES
@GeorgiaTech

Publications **Code**

Preliminaries: Probabilistic Input and Output Data

Input and Output Data is Probabilistic in Seismic Interpretation Pipelines

Data Characteristics for Natural Image Pipelines

- Conventional ML frameworks assume **fixed data and labels**

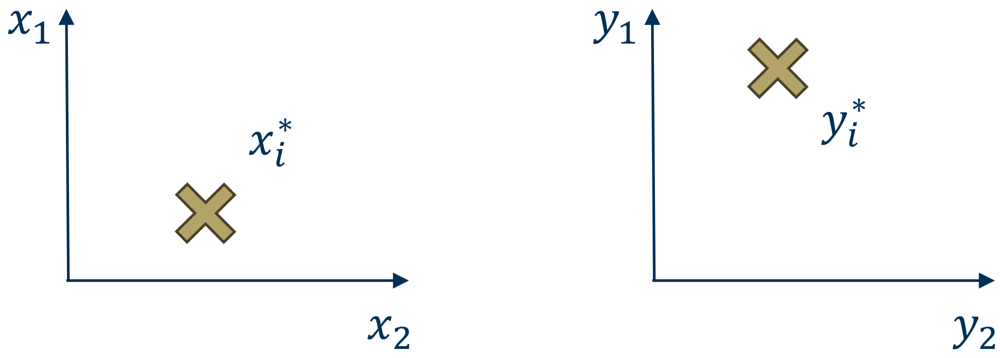
- Underlying **deterministic** densities:

$$y_i \sim \delta(y - y_i^*)$$
$$x_i \sim \delta(x - x_i^*)$$

- **Realistic assumption for natural images** where both image generation and label assignment are deterministic

- **Assumption problematic for seismic as label and data are probabilistic**

Example for Natural Images



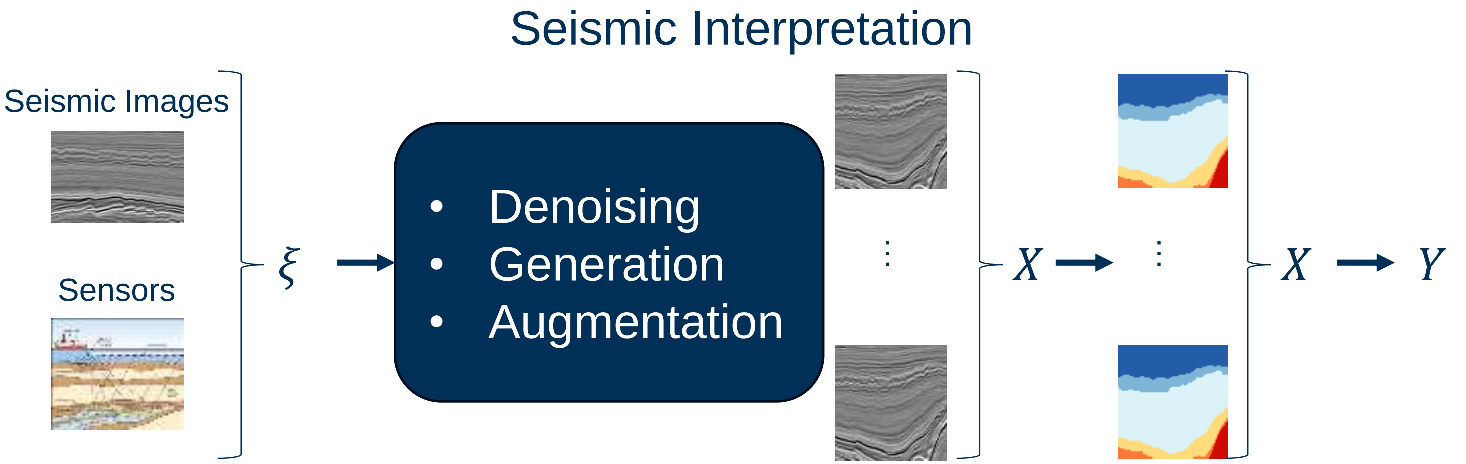
Preliminaries: Probabilistic Input and Output Data

Input and Output Data is Probabilistic in Seismic Interpretation Pipelines

Data Characteristics for Seismic Interpretation Pipelines

- **Data and annotations are stochastic for seismic interpretation**
- Underlying **stochastic** densities:

$$y_i \sim p(y|\alpha)$$
$$x_i \sim p(x|\xi)$$



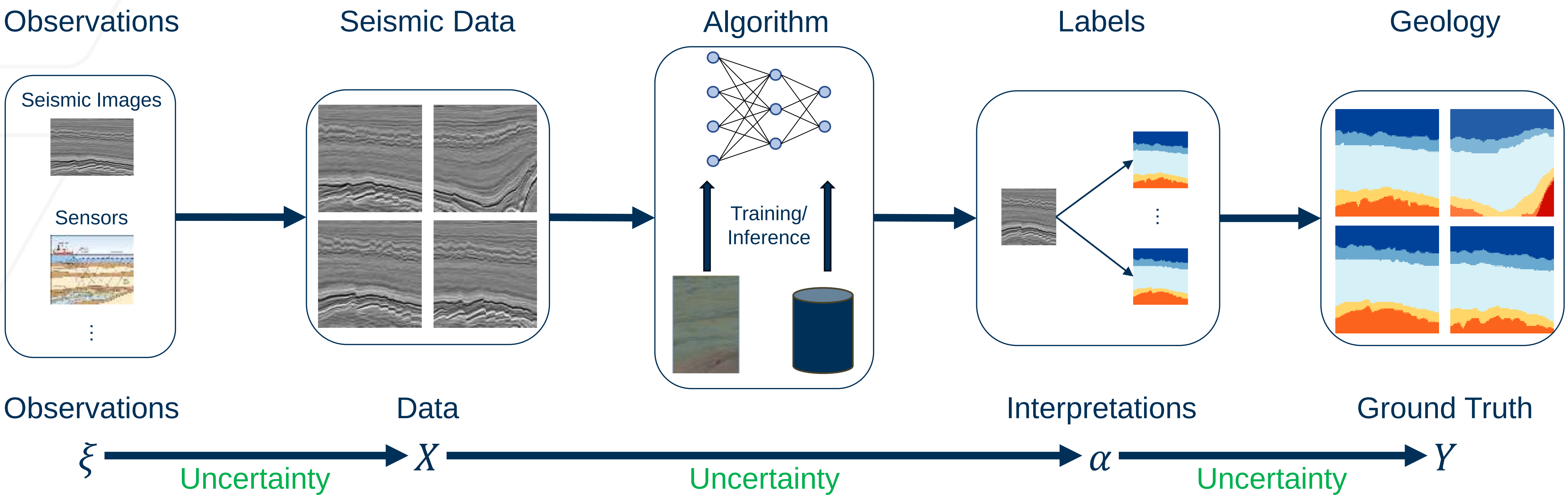
- **Examples** of image stochasticity:
  - Generation
  - Augmentation
  - Processing

- **Label stochasticity stems from image and expert ambiguity**

Proposed Uncertainty Framework

Our Framework Reflects the Processing Steps of the Entire Interpretation Pipeline

Proposed Uncertainty Framework



Framework: PLACEHOLDER – Reflective Uncertainty
Reflective Uncertainty Mirror a Set of Transfromations and Enable Modularity

PLACEHOLDER Reflective Uncertainty

Consider a pipeline

$$h_{pipeline} = h_1 \circ h_2 \ldots \circ h_N,$$

where $h_{pipeline}$ is a series of concatenated functions. We say a uncertainty **framework is reflective** if it, when recursively applied, **reflects the individual pipeline components**:

$$u_{pipeline} = u_1 \circ u_2 \ldots \circ u_3$$

$$h_{pipeline}(x) = h_1(h_2(h_3(x))) \qquad u_{pipeline}(x) = u_1(u_2(u_3(x)))$$

Desirable Property for seismic due to **modularity and separability**

Framework: PLACEHOLDER

Find Relationship between Uncertainty Scores and Uncertainty in the Total Pipeline

Our Uncertainty Framework is Reflective

- Observational Uncertainty

$$p(y|\xi) = \int p(y|x)p(x|\xi)dx$$

- Interpretational Uncertainty:

$$p(y|x) = \int p(y|\alpha)p(\alpha|x)d\alpha$$

- Algorithmic Uncertainty:

$$p(\alpha|x) = \int p(\alpha|x,w)p(w)dw$$

$$\boldsymbol{p(y|\xi) = \int \underbrace{p(y|\alpha)}_{U_{int}} p(\alpha|x,w) \underbrace{p(w)}_{U_{alg}} \underbrace{p(x|\xi)}_{U_{obs}} dw d\alpha dx}$$

$$\boldsymbol{U(y|\xi) = \cdots = f(U_{int}, U_{obs}, U_{alg})}$$

OLIVES @GeorgiaTech

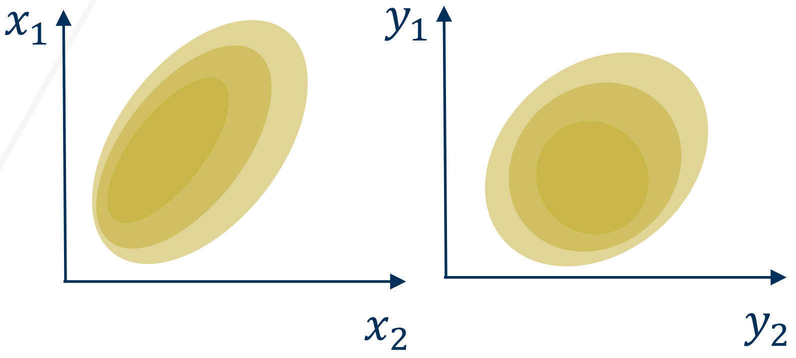
Georgia Tech

Presentation Structure and Outline

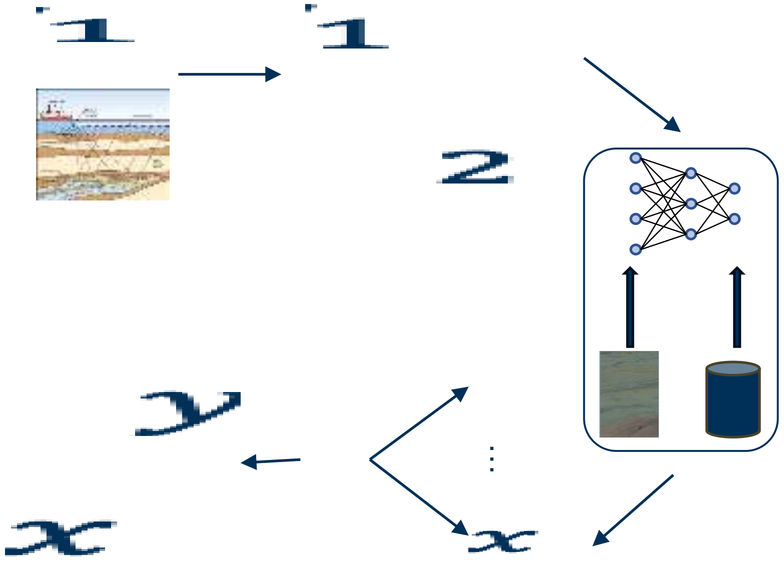
Discussion Topics are Active Learning, ATLAS, and Results

Contributions: Propose Uncertainty Framework for Seismic Interpretation

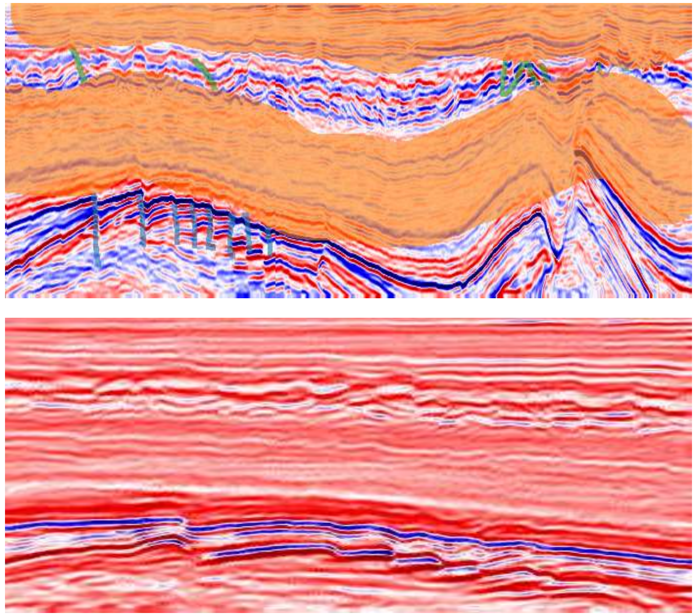
Background: Probabilistic Data



Framework: Uncertainty Sources



Results: Visuals

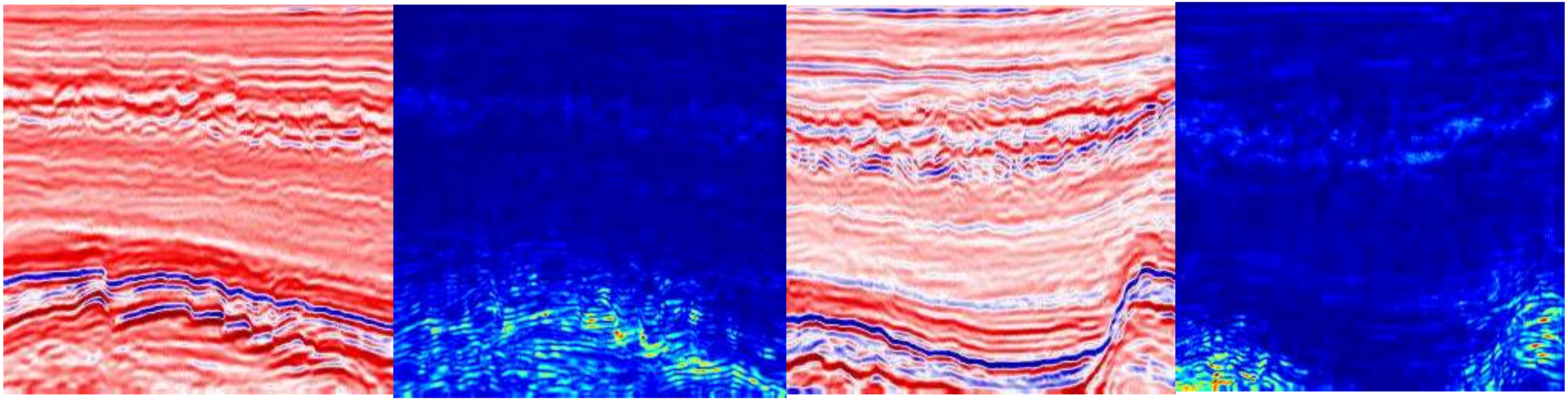


Results: Observational Uncertainty

Find Relationship between Uncertainty Scores and Uncertainty in the Total Pipeline

Observational Uncertainty

$p(X|\xi_1)$ $\quad U(X)$ $\quad p(X|\xi_2)$ $\quad U(X)$



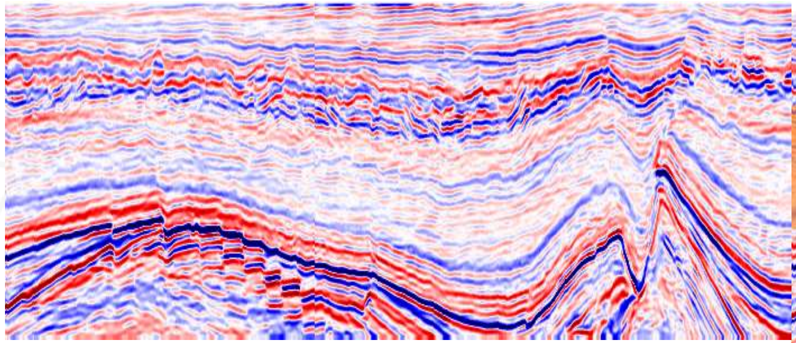
$$\mathrm{U_{obs}}(X) \approx \mathbb{E}_{p(X|\xi)}\, X^2 - \left(\mathbb{E}_{p(X|\xi)}\, X\right)^2$$

Results: Interpretational Uncertainty

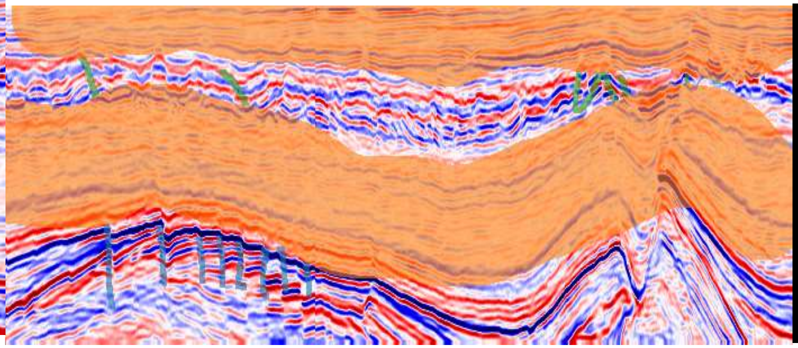
Find Relationship between Uncertainty Scores and Uncertainty in the Total Pipeline

Interpretational Uncertainty

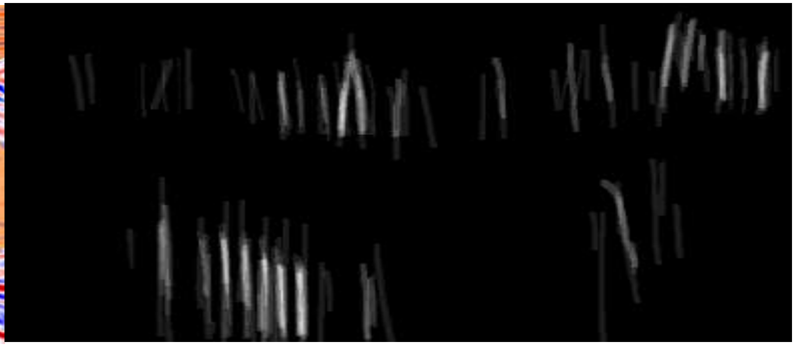
X_1



$p(\alpha|X_1)$



$U(\alpha)$



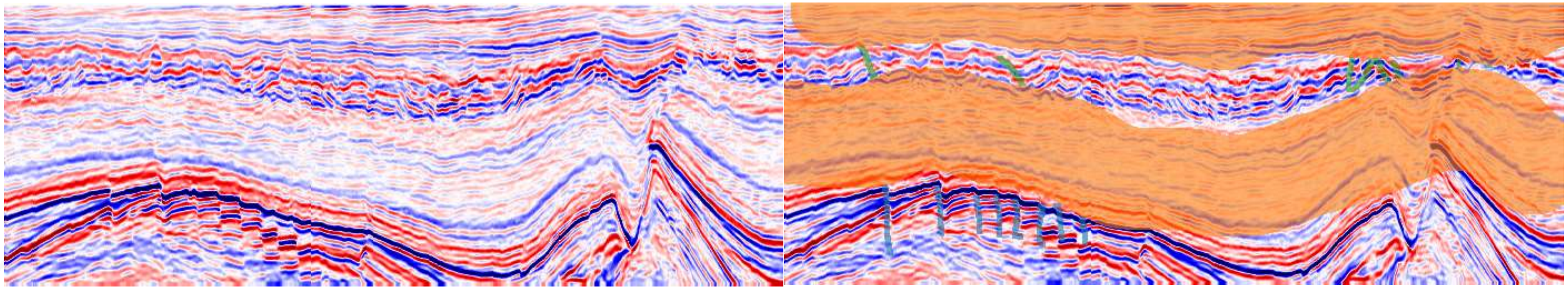
$$\mathrm{U_{int}}(\alpha) \approx \mathbb{E}_{p(\alpha|X)}\ \alpha^2 - \left(\mathbb{E}_{p(\alpha|X)}\ \alpha\right)^2$$

Results: Algorithmic Uncertainty

Find Relationship between Uncertainty Scores and Uncertainty in the Total Pipeline

Algorithmic Uncertainty PLACEHOLDER

X_1 $p(\alpha|X_1)$ $U(X)$



$$\mathrm{U}_{\mathrm{int}}(\alpha) \approx \mathbb{E}_{p(\alpha|X)}\, \alpha^2 - \left(\mathbb{E}_{p(\alpha|X)}\, \alpha\right)^2$$