Automated seismic interpretation using machine learning and field interpretations

Force Hackathon and Advances of Machine Learning on Subsurface Data

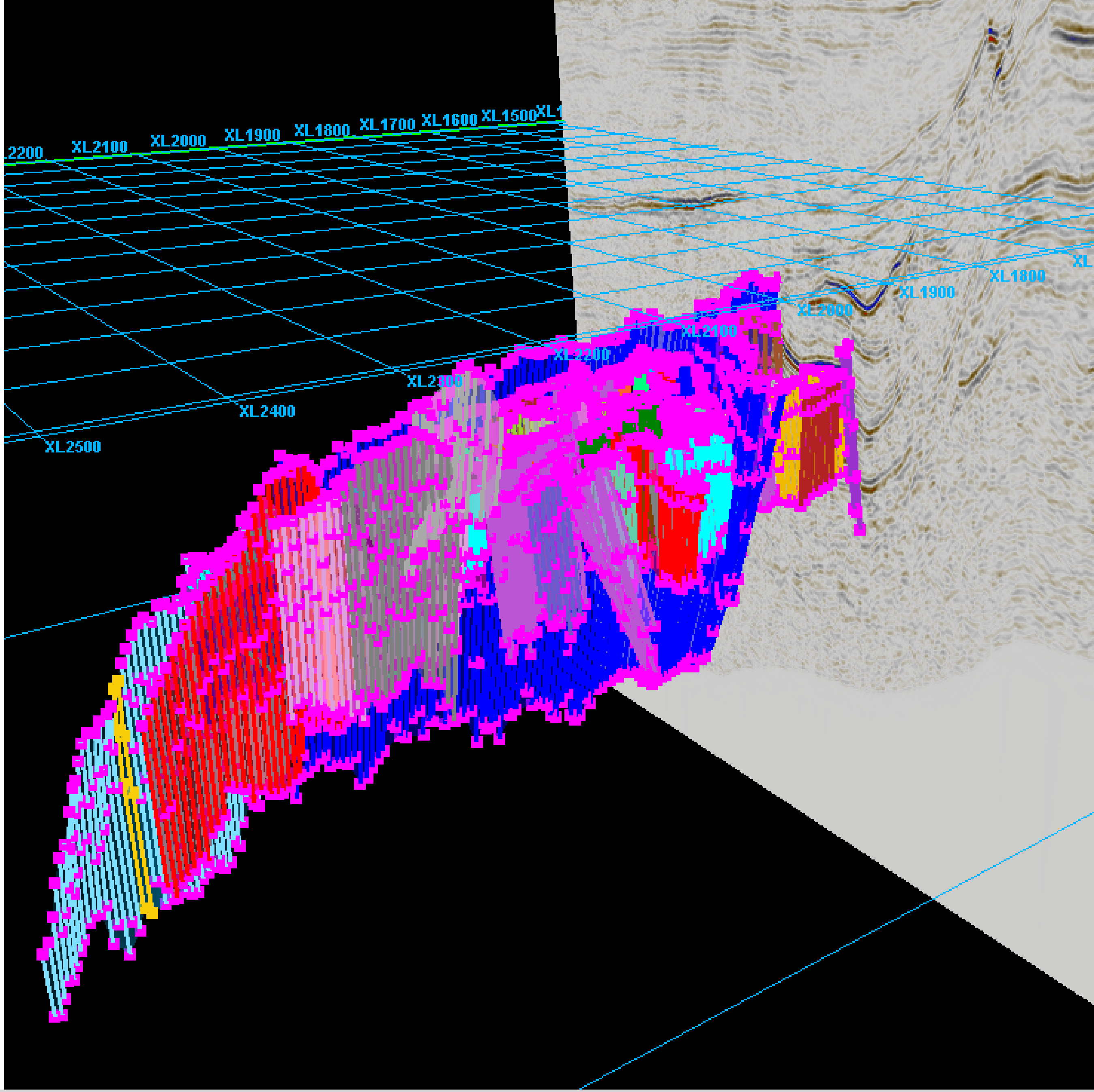
Stavanger, September 20 2018

Espen Birger Raknes and Peder Aursand

Background

Goals

- Proof-of-concept study for predicting **faults** and **horizons**
- Real data needed to be involved
- Train using **human interpretations** from **multiple fields**
  - Exploration and production areas
- **Binary** sample prediction and **pixel-level** prediction (segmentation)
- **2D** and **3D** training data

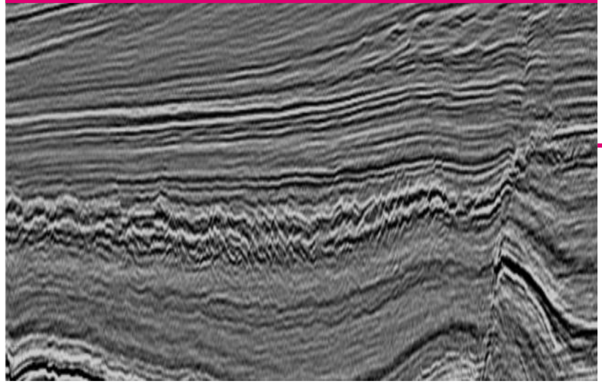


Approach

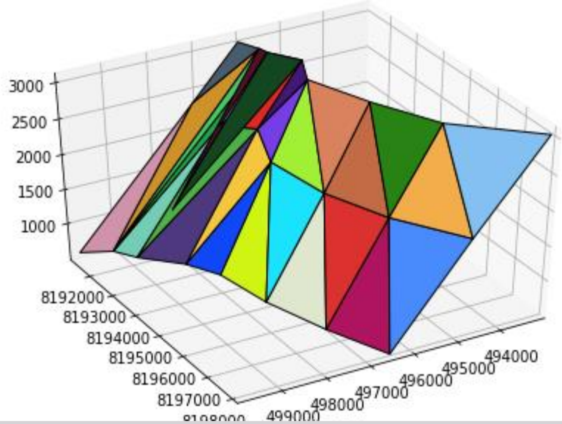
APPROACH

Training data work flow overview

Seismic data sample



Interpretations

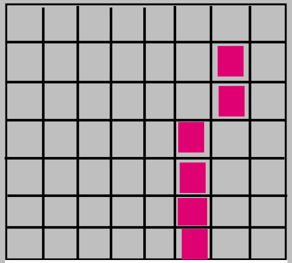


Projection

Target

0	1

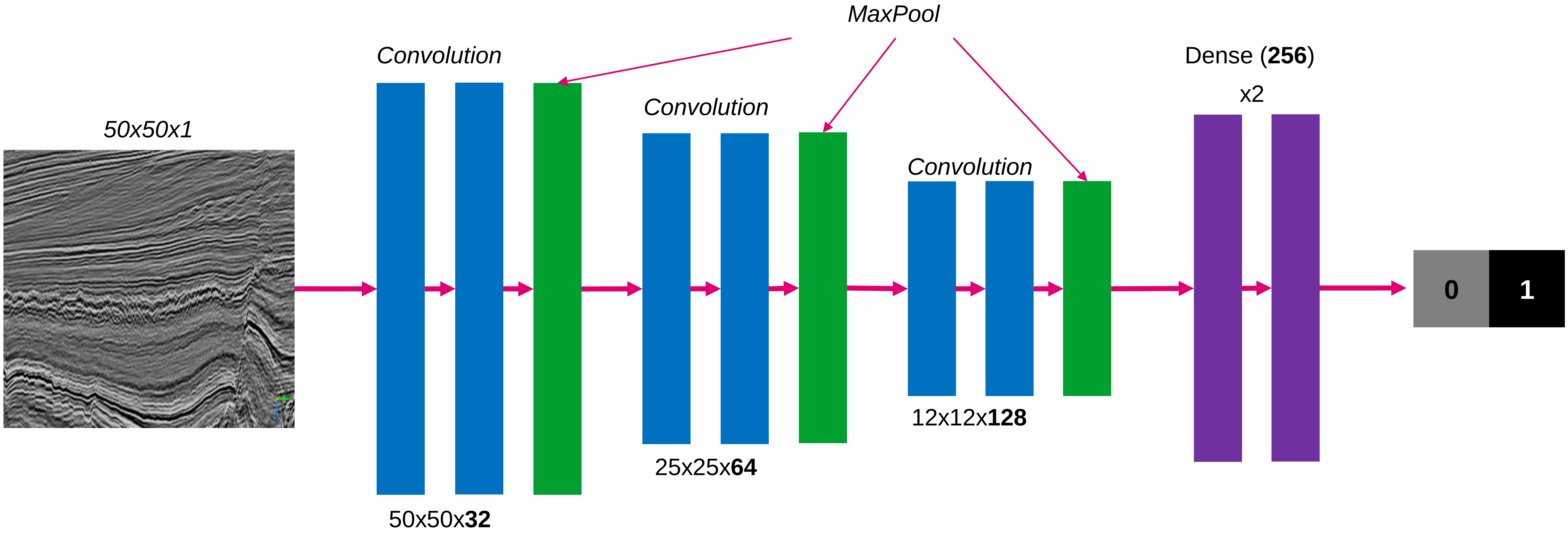
Binary



Pixel

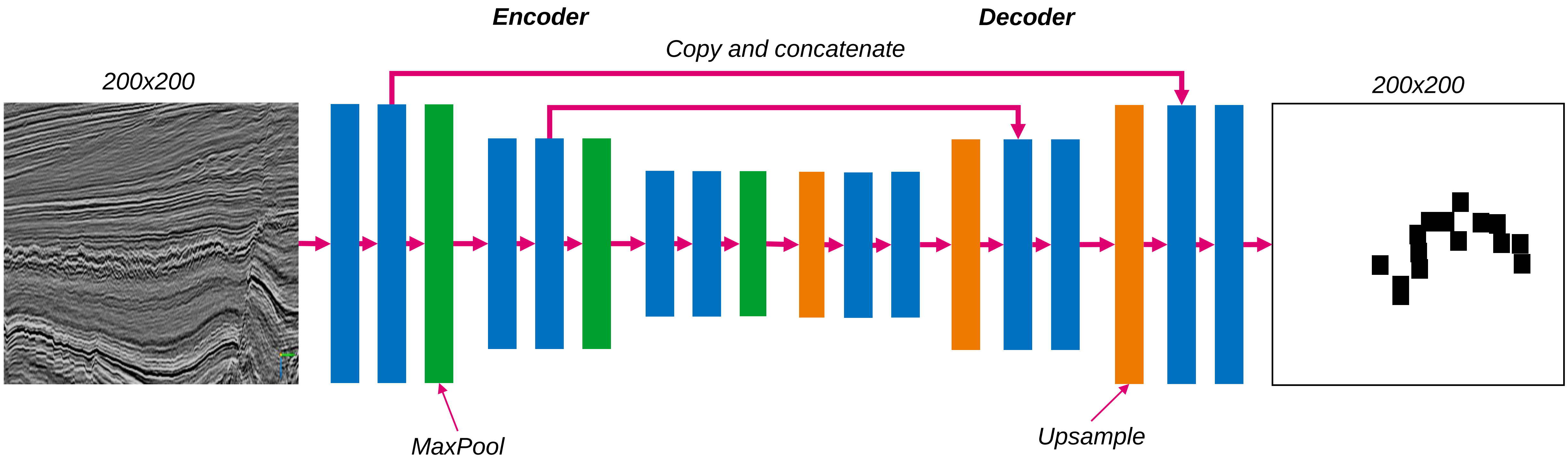
APPROACH

Neural network: Binary sample prediction



Note: Small version of the VGG architecture (*Simonyan and Zisserman, 2014*)

Neural network: Pixel prediction (image segmentation)

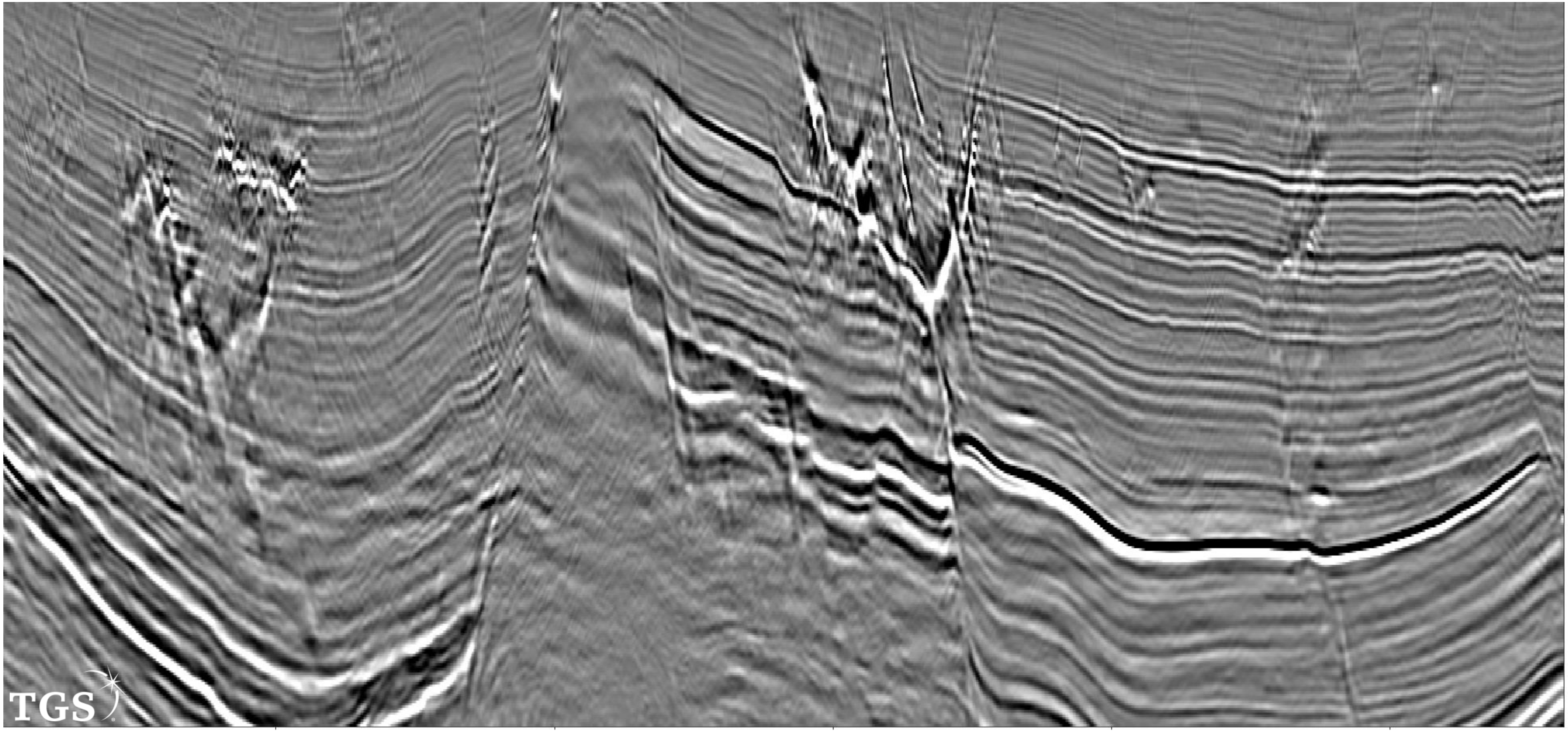


Note: Small version of the SegNet architecture (Badrinarayanan et al., 2015)

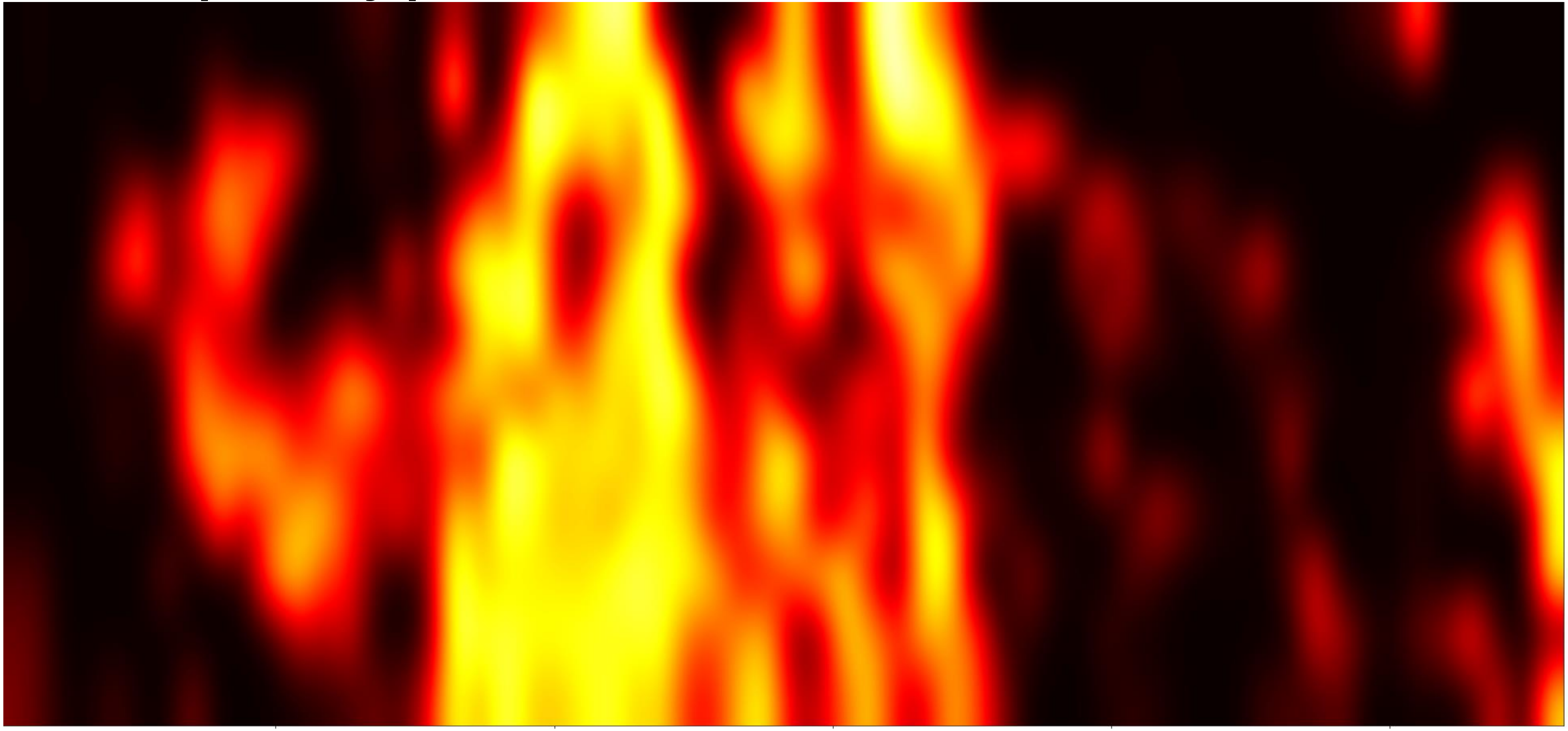
Results

RESULTS: FAULT PREDICTION

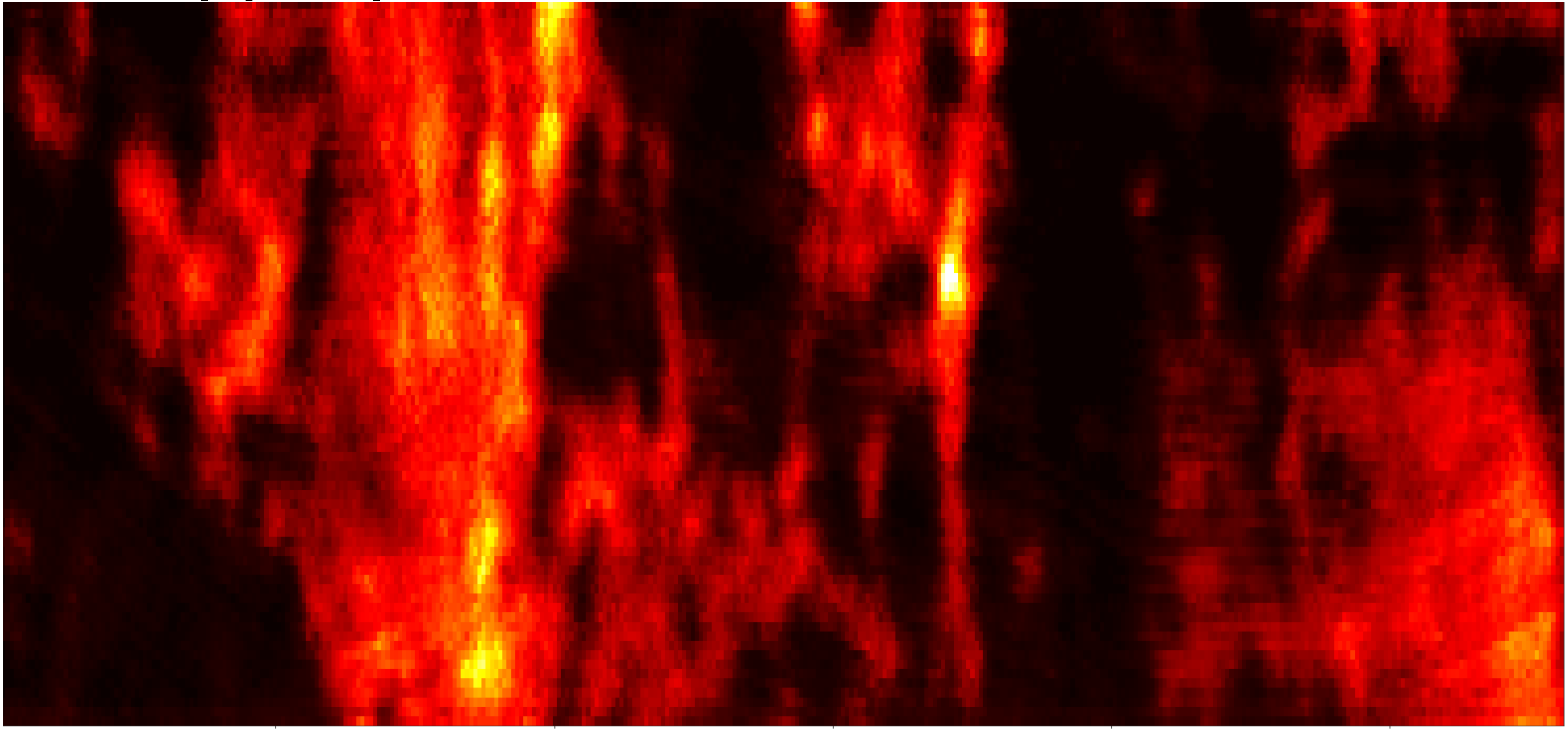
Seismic slice



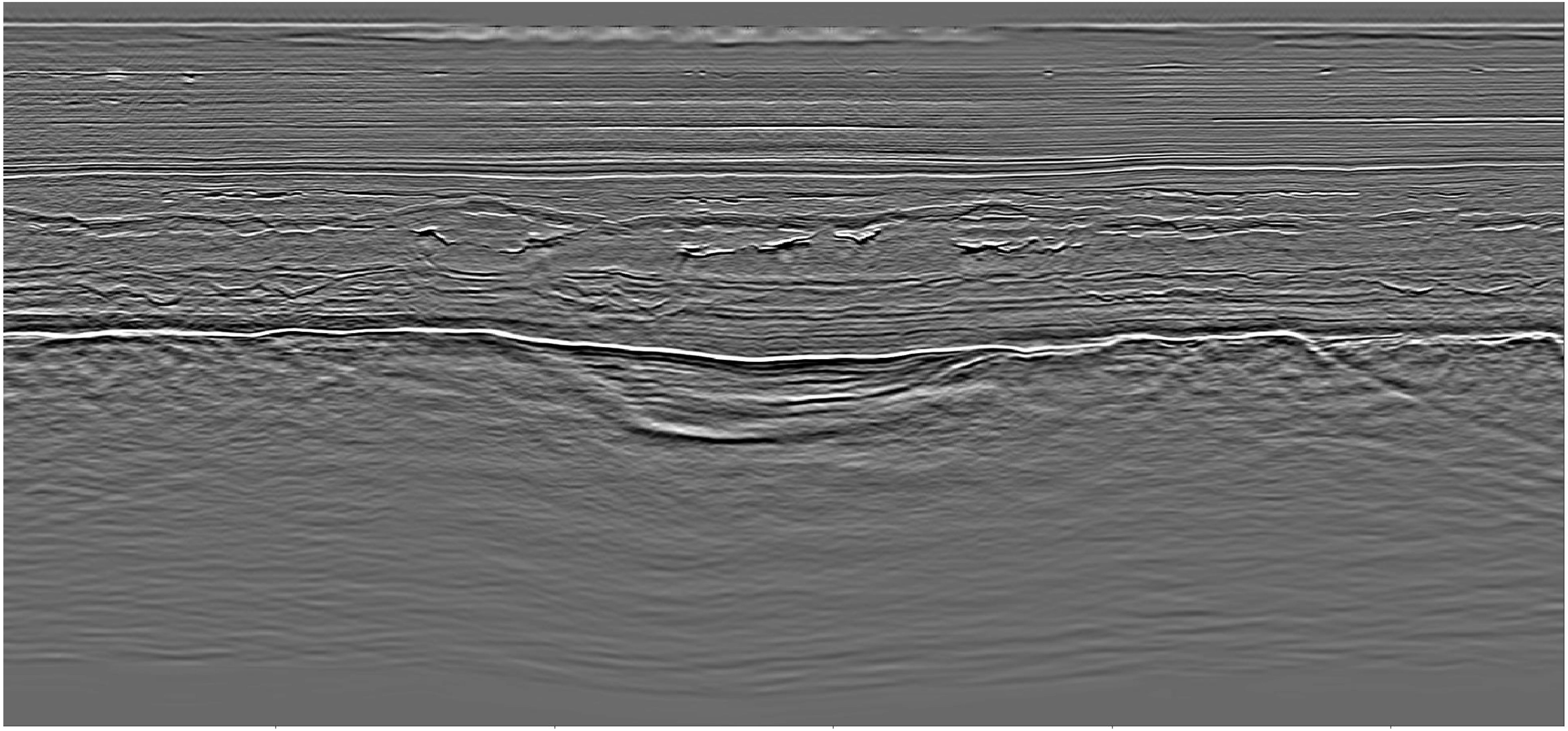
Heat map binary prediction



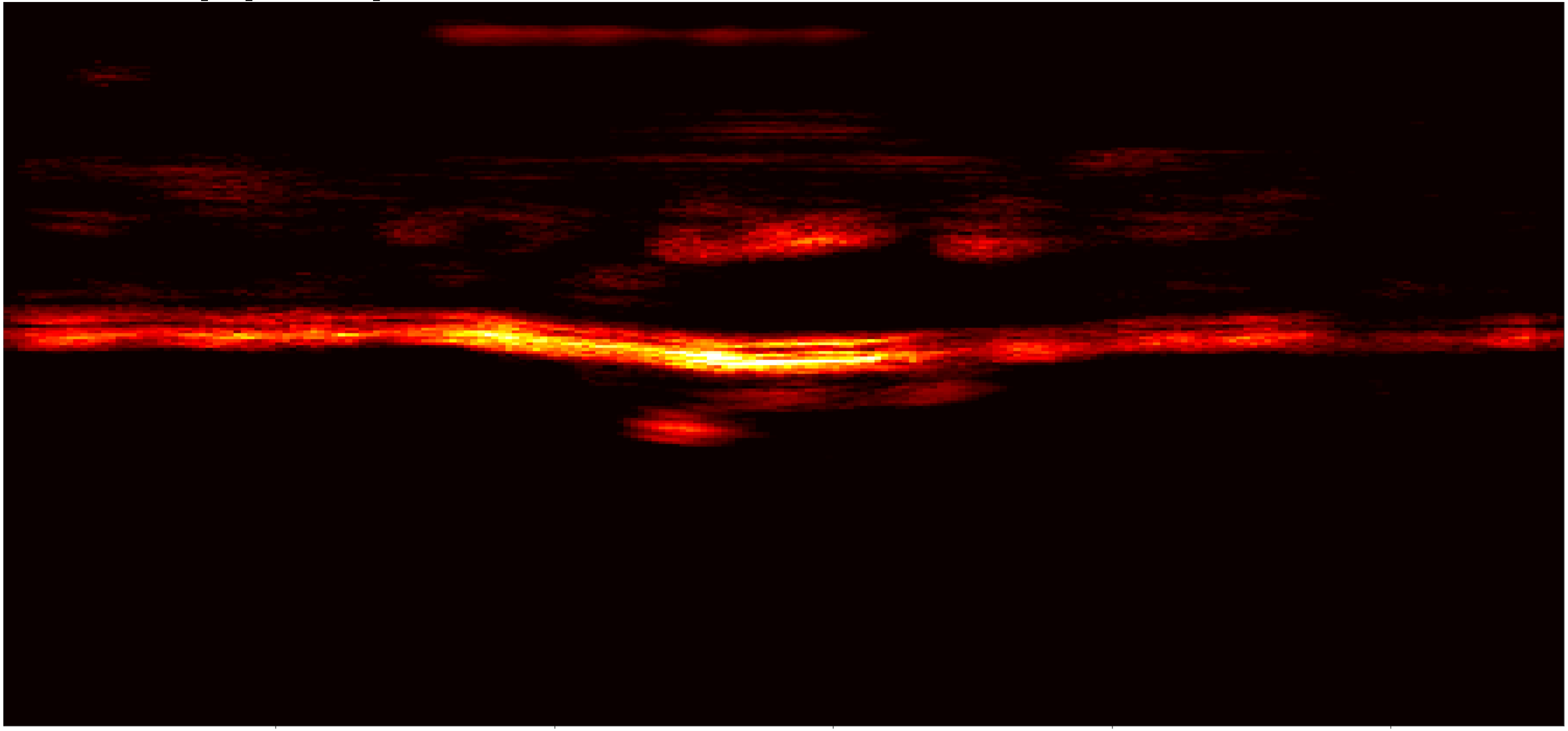
Heat map pixel prediction



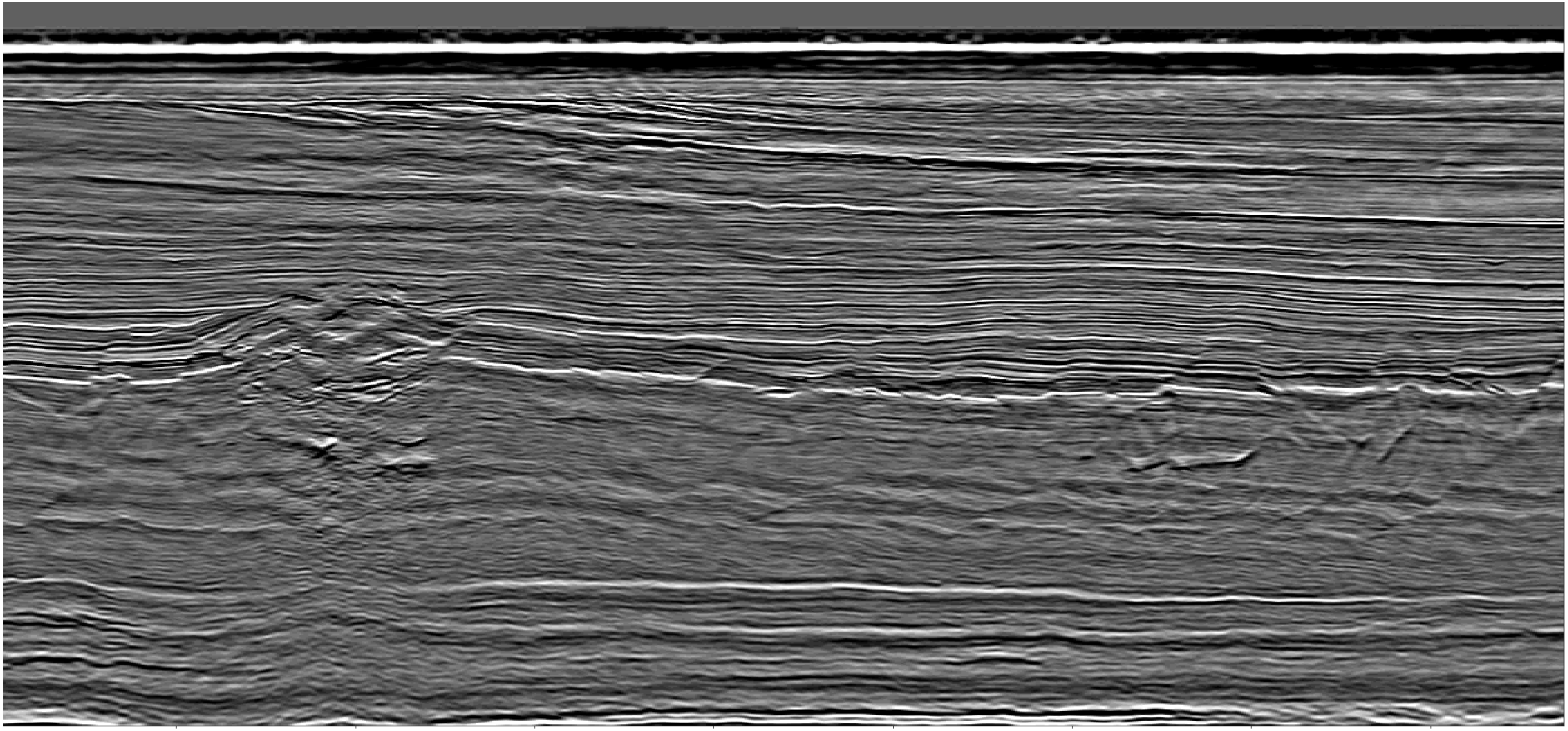
Seismic slice



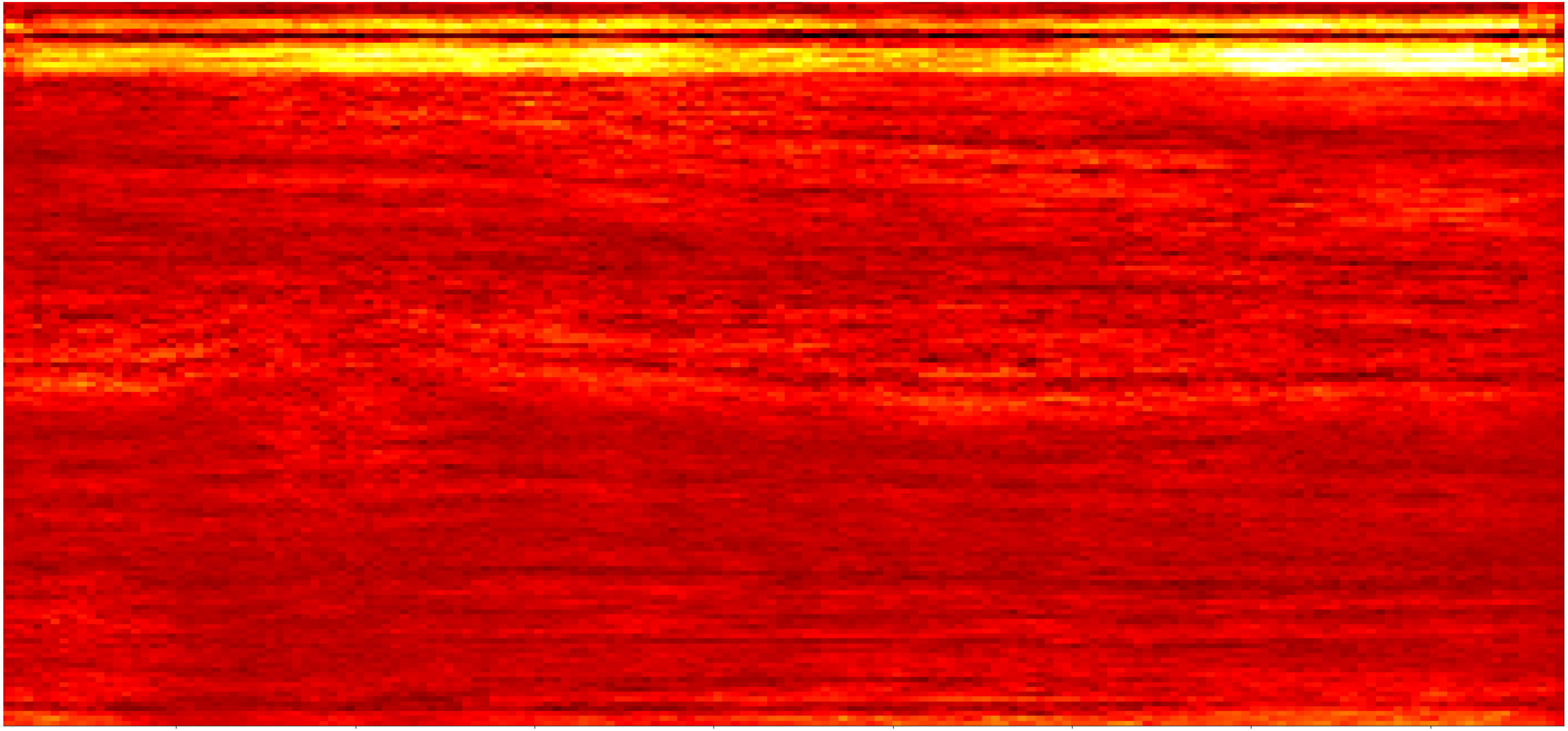
Heat map pixel prediction



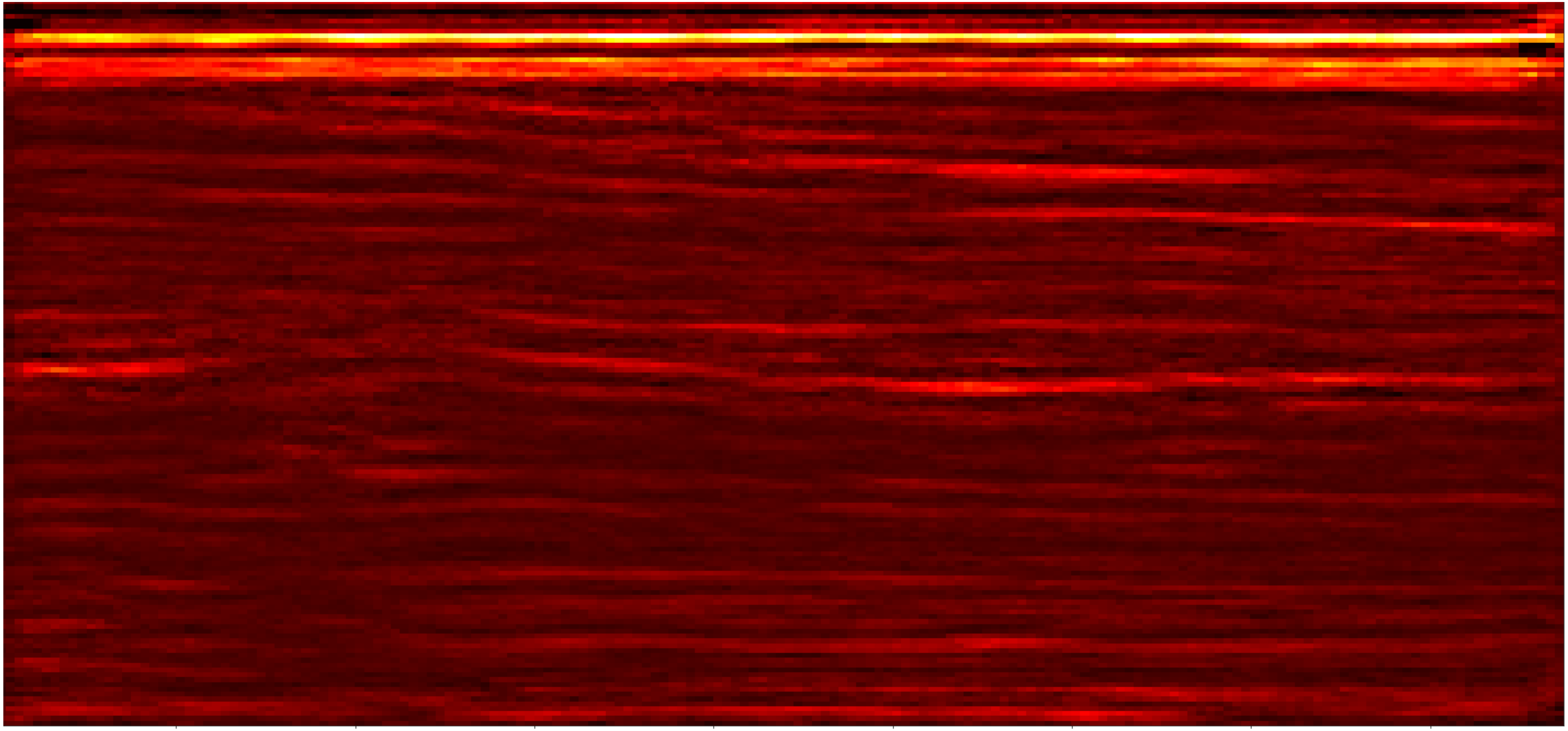
Seismic



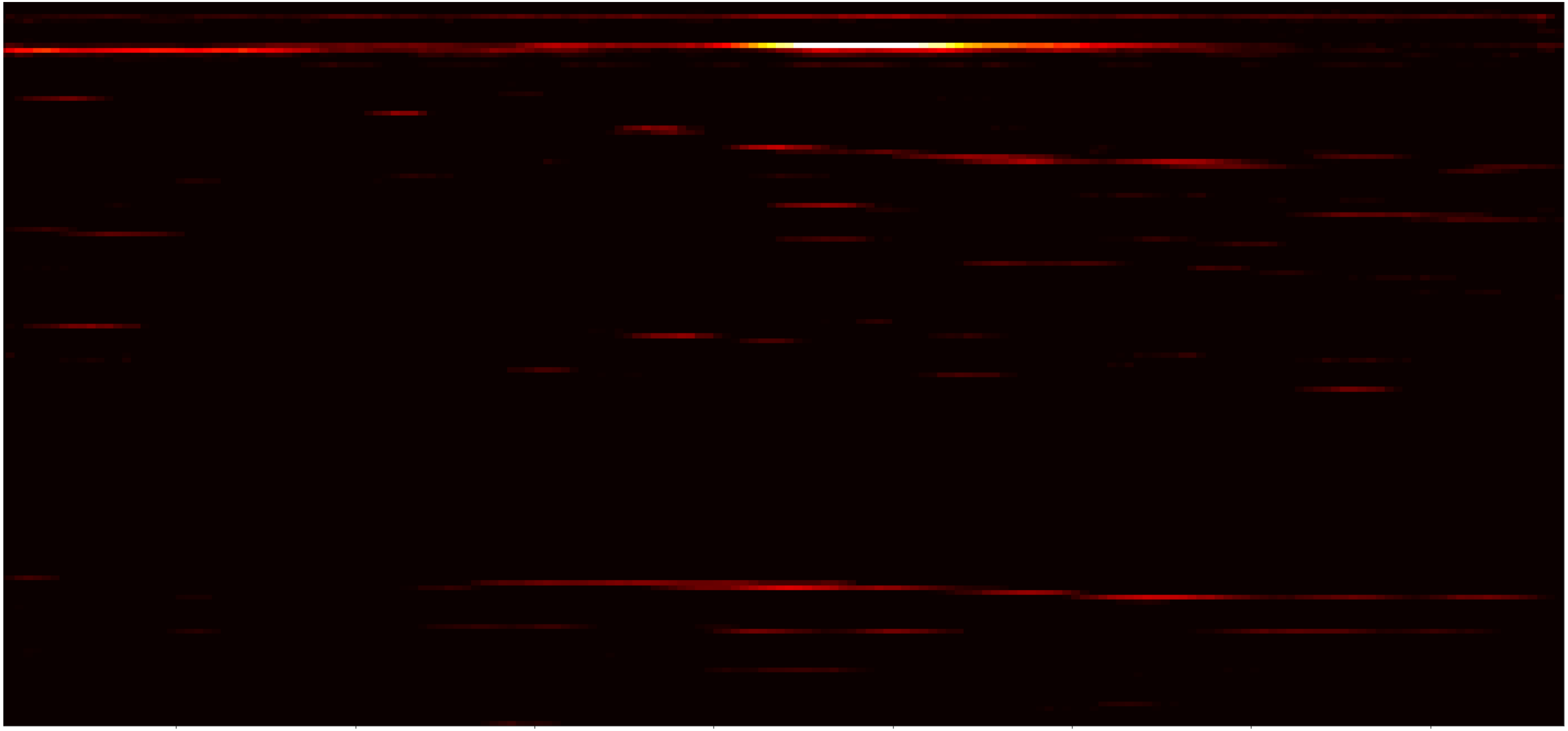
Heat map: Trained on one dataset



Heat map: Trained on another dataset



Heat map: Trained on two datasets



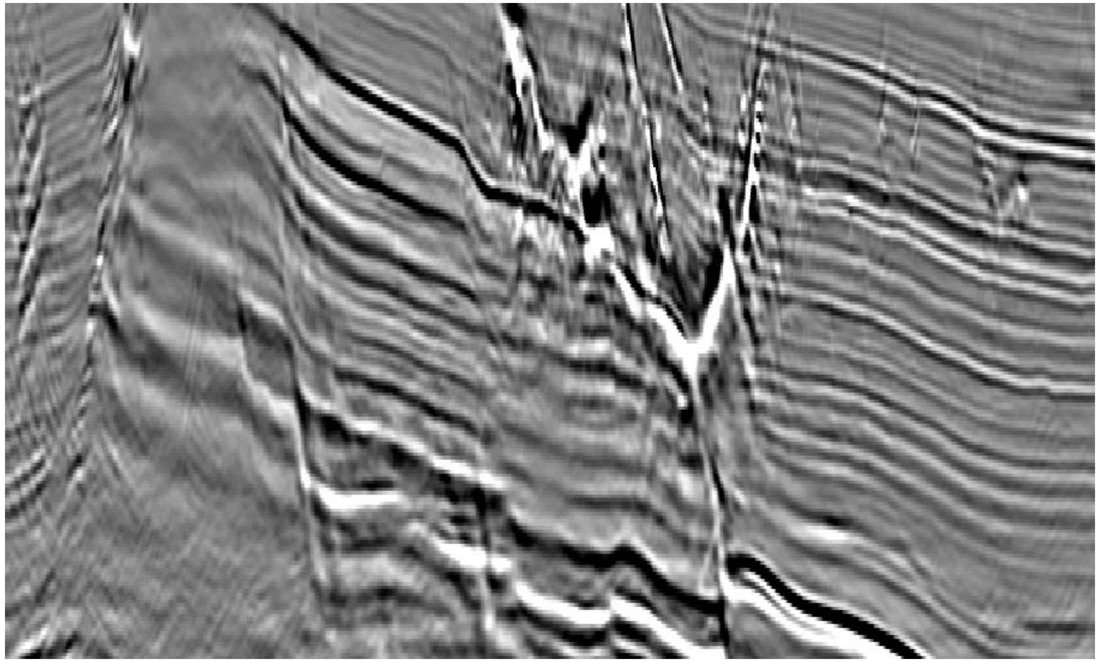
Discussion and conclusion

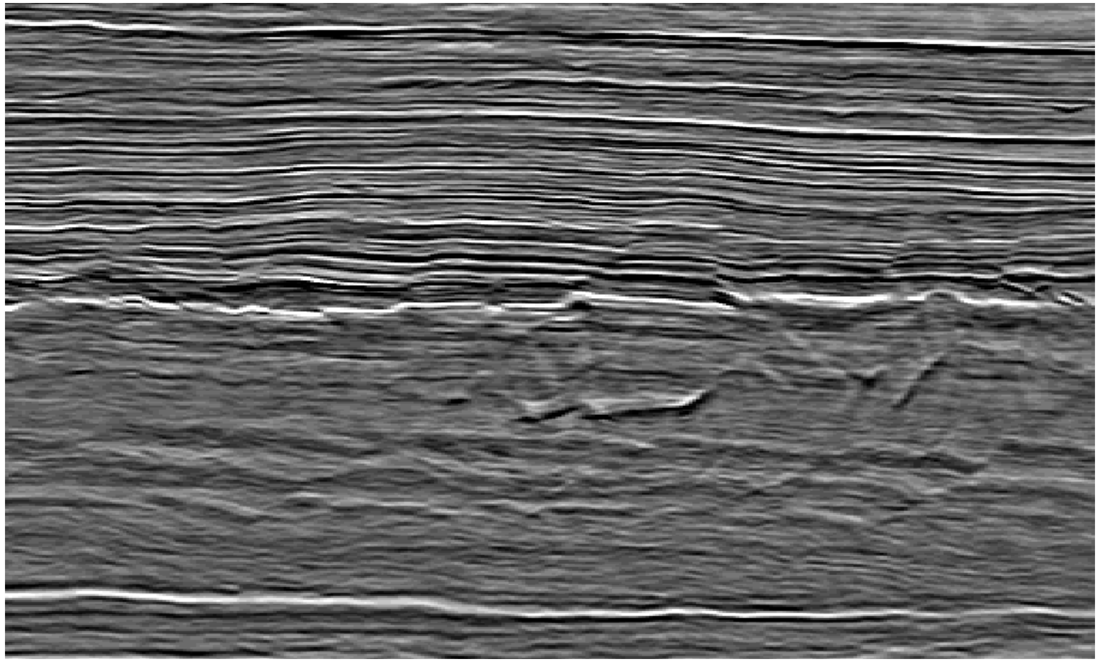
Areas are different

- Differences in
  - Geology
  - Data quality
- Number of interpretations varies dependent on where the area is in the development phase

Suggestions

- More data to train on
- Representative data of the geology
- Model adaption to the area to be used on
- Apply techniques for preventing over-fitting





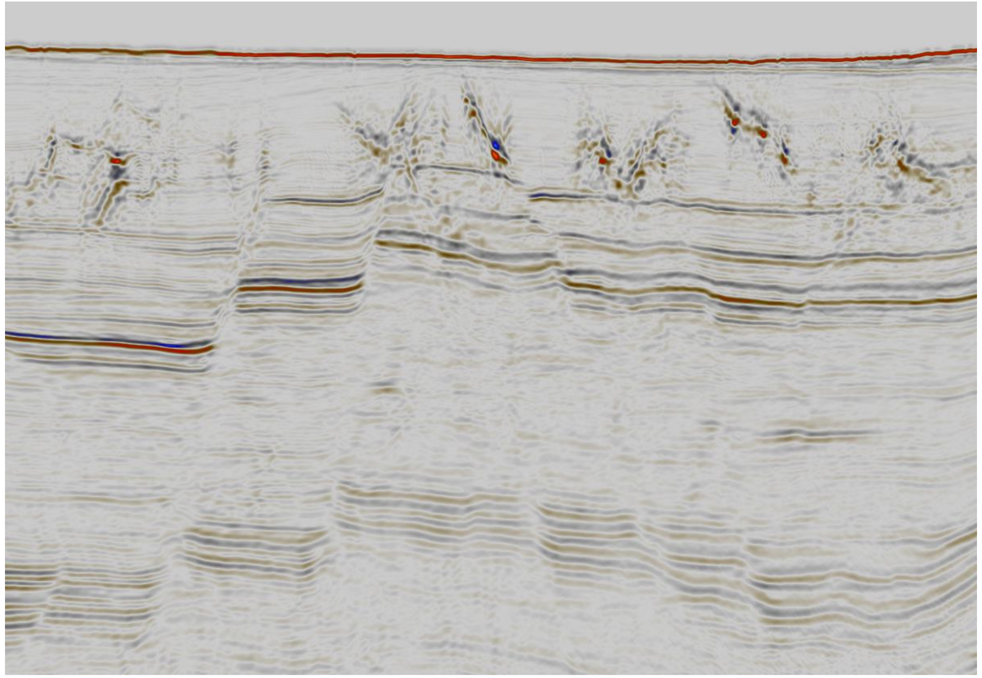
DISCUSSION

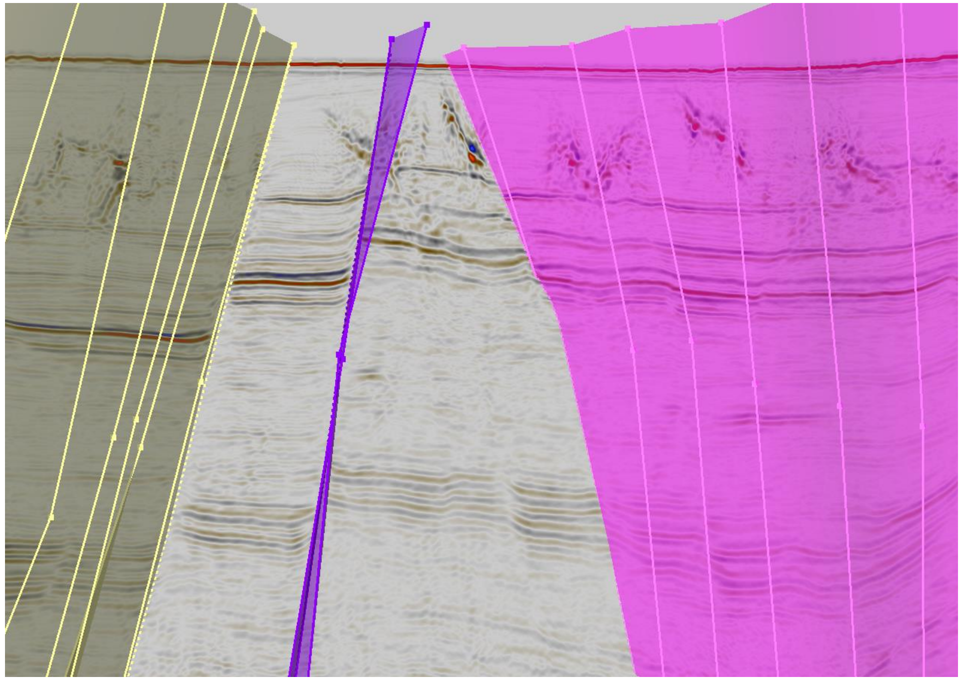
Using human interpretations

- Using human interpretations are not straight forward
- **Many** false negatives
- **Very few** false positives
- Non-symmetric label noise

Suggestions

- Use synthetic data
- Mix of synthetic and human-made data
- Manually QC-ed human-made data
- Use data from areas with enough interpretations

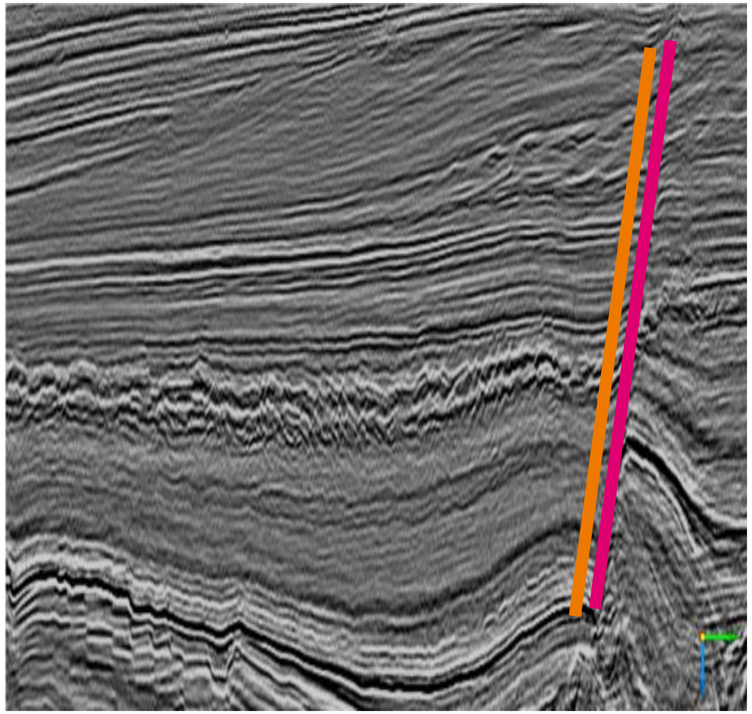


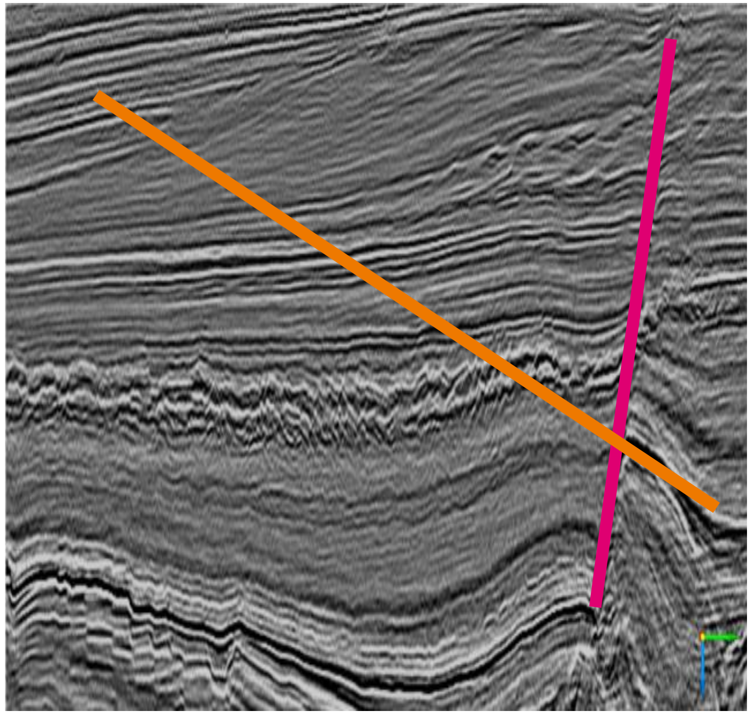


Straight lines are difficult

- Out-of-the box neural networks with standard loss functions struggle with finding straight lines
  - Learning best on objects with some «extent»
- Standard metrics not ideal

- New and better loss functions
- Take topology into account
- Post processing steps on the heat maps





Prediction

True value

Conclusion

- **Summary**
  - Image recognition by training models on human interpretations projected onto seismic samples
  - Features predicted
    - Faults
    - Horizons

- **Challenges**
  - Non-symmetric label noise (false negatives) due to incomplete interpretations (manual data quality check is time-consuming)
  - Prediction accuracy suffers when applying a model to a different field
  - Standard image recognition is not well suited for detecting lines and planes
  - **A reference data set for benchmarking** these kinds of models is badly needed

- **Future work**
  - Transfer learning using pre-trained weights
  - 3D data augmentation (transformations on real data or artificial data)
  - Better metrics for FCNN
  - Train on large combined datasets from multiple fields
  - Pre-stack seismic data

Acknowledgements

We thank CGG, PGS and TGS for permission to present the seismic data.

AkerBP

www.akerbp.com