

DISTANCE TRANSFORM LOSS: Boundary-aware Segmentation of Seismic Data

Rafael Henrique Vareto^a, Ricardo Szczerbacki^b, Luiz A. Lima^b, Pedro O. S. Vaz-de-Melo^a and William Robson Schwartz^a

^a *Uncertainty in Artificial Intelligence Lab - Federal University of Minas Gerais*

Avenida Presidente Antônio Carlos, 6627, Belo Horizonte, Brazil 31270-901 - {rafaelvareto,olmo,william}@dcc.ufmg.br

^b *Exploration and Production Division (E&P) - Petróleo Brasileiro S.A.*

Avenida República do Chile 65, Rio de Janeiro, Brazil 20031-170 - {ricardo.s,lual}@petrobras.com.br

ARTICLE INFO

Keywords:

Cost Function
Distance Transformation
Seismic Data
Boundary Awareness
Semantic Segmentation
Deep Learning.

Abstract

The segmentation of seismic data is a challenging exercise given the complexity and high variability of subsurface sources. This arduous task is effective in the identification of geological features, including facies classification, fault detection, and horizon interpretation. As a result, this work introduces a new cost function entitled Distance Transform Loss (DTL) that punishes deep networks when class boundaries are misclassified in exchange for more accurate contour delineations, an important aspect in the geological field. DTL consists of four key steps: contour detection, distance transform mapping, pixel-wise multiplication, and the summation of all grid elements. We conduct a comprehensive evaluation of deep convolutional architectures using publicly available seismic datasets, demonstrating that the proposed approach consistently enhances semantic segmentation performance. The results highlight DTL as a robust and architecture-agnostic loss function, capable of addressing class imbalance and boundary delineation challenges that commonly arise in seismic interpretation tasks.

1. Introduction


Given the remarkable evolution and widespread adoption of convolutional neural networks (CNNs), geophysical researchers are increasingly turning to deep-learning models for interpreting seismic volumes and enhancing geological analysis (Monteiro et al., 2024). Leveraging CNNs in geophysics allows for unprecedented accuracy and efficiency in detecting and characterizing earth features. These methods are particularly useful in analyzing seismic images, which are used in the exploration of gas, minerals and oil; as well as an essential mechanism toward groundwater or aquifer studies and natural hazard assessment (Sun et al., 2022; Tao et al., 2022; Haggerty et al., 2023), to name a few. Deep learning techniques, for instance, have proven highly effective in seismic segmentation tasks due to their ability to learn complex patterns. They can automate the interpretation of subsurface geology, reducing the time and effort required in manual analysis and facilitating the identification and understanding of diverse underground formations.

Semantic segmentation of geological data involves classifying each voxel within a seismic volume into discrete categories representing distinct subsurface elements, such as fault structures, horizon boundaries, lithologies, reservoirs, or salt bodies (An et al., 2021; Xu et al., 2021; Lee et al., 2024). This task not only assigns labels to voxels but also aims to accurately interpret the complex geological structures captured in seismic surveys. By precisely delineating object boundaries, automated segmentation offers a detailed and comprehensive understanding of subsurface features. This granularity is

crucial, as it can reveal small-scale features that may be overlooked in traditional analysis methods. Ultimately, segmentation enhances the precision of geological analysis, leading to significant cost reductions in exploration and empowering scientists with more informed decision-making and insights for resource exploration and extraction alongside an improved geological understanding.

Convolutional neural networks, along with the broader suite of deep learning methods, have transformed image analysis and are now commonly employed in automated seismic facies segmentation. These networks are particularly well-suited for geology tasks given their ability to automatically learn hierarchical features from the input data (Li et al., 2020; Zhang et al., 2021). Neural networks are highly dependent on *cost functions*, also called *loss functions*¹, which serve as the guiding hand for the training process, determining how well the network is performing and where adjustments are needed (Vareto et al., 2024). More precisely, a loss function quantifies the difference between the network's predictions and the ground truth, providing a measure of error that the network aims to minimize during training (Jadon, 2020). The goal is to find the minimum penalty score of the cost function, which corresponds to the best set of weights and biases that make accurate predictions. Without this mechanism, a neural network would lack feedback to assess its performance, refine parameters and improve the predictions during training (Taghanaki et al., 2021).

The inherent design of neural networks presents additional challenges for accurately preserving boundaries in seismic segmentation. Downsampling layers, often employed to reduce computational complexity and expand the recep-

 rafaelvareto@dcc.ufmg.br (Rafael Henrique Vareto)

ORCID(s): 0000-0002-0431-5945 (Rafael Henrique Vareto);

0000-0002-9749-0151 (Pedro O. S. Vaz-de-Melo); 0000-0003-1449-8834 (William Robson Schwartz)

¹In this manuscript, we use the terms *cost* and *loss functions* interchangeably.

tive field, can inadvertently lead to a loss of fine-grained details, including critical boundary information. Limited receptive fields can further exacerbate this issue by restricting the model's ability to capture global context, leading to blurred or inaccurate boundary predictions. Additionally, the tendency of neural networks to smooth spatial edges during feature extraction can diminish the sharpness of class boundaries, making it increasingly challenging to distinguish between closely aligned or intricately connected regions (Wang et al., 2017). As detailed in the Section 2, the foregone challenges have led to the development of specialized loss functions and architectural modifications that aim to preserve sharp boundaries while maintaining high overall segmentation accuracy.

To address the persistent issue of boundary misclassification in seismic segmentation, we propose a novel cost function termed Distance Transform Loss (DTL). This method builds upon the concept of distance transform mapping in digital image processing (Fabbri et al., 2008; Felzenszwalb and Huttenlocher, 2012; Gonzalez and Woods, 2018), a technique that quantifies the spatial proximity of each pixel to its nearest boundary. DTL introduces penalties that are proportional to the distance of a misclassified pixel from the closest ground-truth boundary, effectively encouraging the network to focus on achieving sharper and more accurate boundary delineations. These penalties are calculated using the L_1 Manhattan distance, a measure that captures the minimal number of grid steps required to reach the nearest true boundary pixel. Unlike traditional loss functions that treat all pixels uniformly, DTL prioritizes regions of higher boundary sensitivity, ensuring that errors in these critical areas are more heavily penalized.

To further distinguish DTL from existing boundary-aware loss functions, we clarify that our approach performs pixel-wise multiplication between the predicted contour map and the ground-truth distance map, producing a discrete and differentiable supervision matrix. This formulation is designed to support traditional pixel-wise losses, like cross-entropy loss, and can be easily integrated with them in a multi-term objective function. Unlike generic boundary-aware losses, DTL is tailored originally for seismic data, where accurately capturing subtle and complex geological boundaries — such as faults, salt bodies, and stratigraphic interfaces — is of critical importance.

For a clear illustration, Figure 1 provides an overview of how DTL detects contours, calculates the distance transformation, and finally returns the corresponding penalty score based on the calculated distances. By construction, DTL ensures that the estimated loss is proportional to the amount of boundary misalignment, holding larger penalties when predictions deviate further from the true boundaries. Observe that the designed criterion is not only visually intuitive but also conceptually straightforward, making it easy to implement through existing deep learning frameworks. The introduction of DTL represents a significant step forward in enhancing the robustness of neural networks in tasks that require detailed and accurate segmentation, particularly in complex domains such as seismic data analysis.

There are a few seismic facies datasets available for semantic segmentation, such as Netherlands' F3 BLOCK, Nova Scotia's PENOBSCOT, and New Zealand's PARIHAKA (Alaudah et al., 2019; Baroni et al., 2019; Inc, 2020). To the best of our knowledge, there is no record of studies encompassing experiments using these publicly available volumes simultaneously. In fact, this seems to be the first study to provide both quantitative and qualitative analyses encompassing experiments conducted on the aforementioned datasets. We incorporate them into our semantic segmentation analysis, acknowledging their valuable resources for tasks like horizon interpretation, facies classification, and fault detection. Employing multiple datasets endorses the reliability, generalizability, and broader applicability of the models being evaluated.

It is important to note that the primary goal of this study is to isolate and analyze the effect of the proposed DTL on boundary accuracy within controlled, well-annotated environments. While our experiments demonstrate that DTL generalizes well across datasets of varying geological settings, a detailed analysis of noise robustness and transfer learning remains beyond the scope of this work and is considered a promising direction for future research. In summary, the major contributions of this study are:

- DTL, a novel cost function designed to guide deep neural networks in predicting sharper and more precise boundaries for seismic structures. DTL is particularly effective in enhancing the accuracy of boundary delineation, addressing a common challenge in image segmentation tasks by minimizing boundary misclassifications and improving the overall fidelity of segmentation outputs.
- A batch of experiments analyzing different cost functions and deep learning architectures using specific metrics, such as Boundary F1 Score (BF1S) and Mean Intersection over Union (MIOU). These metrics were chosen to provide a balanced assessment of boundary accuracy and overall segmentation quality, allowing for a detailed comparison of distinct approaches and demonstrating DTL's superiority in handling complex segmentation tasks.
- An evaluation of different seismic volumes showcasing the robustness and adaptability of the proposed DTL cost function across diverse domains (datasets). The results highlight DTL's ability to generalize well across varying seismic datasets, underscoring its potential as a reliable tool for segmentation in heterogeneous data environments.

2. Related Works

Semantic segmentation has become a crucial task in geoscientific applications, especially for interpreting complex subsurface structures from seismic data. At the heart of most segmentation models lies the loss function, which directly influences the quality of the predicted output (Vareto et al.,

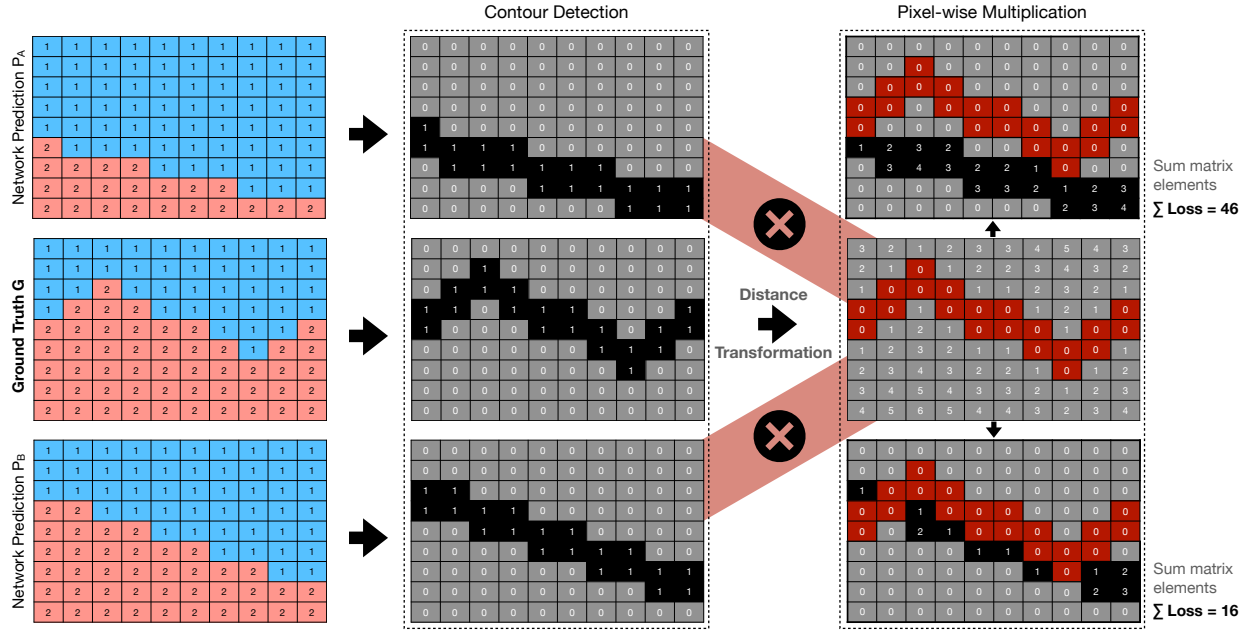


Figure 1: DISTANCE TRANSFORM LOSS: Pipeline depicts how the ground-truth image G and two different network predictions P_A and P_B are processed to obtain their respective error scores (loss). Note that the approach starts detecting the contour among class labels *one* and *two* (blue and coral-painted pixels, respectively). Then, it proceeds with the distance transform estimation for the ground truth only and continues with the pixel-wise multiplication between the prediction contours and the ground-truth distance map. It ends up with the summation of all matrix elements to provide a conclusive loss score.

2023). However, in the context of geophysical interpretation, where accurately delineating the boundaries between geological structures is paramount, traditional cost functions often fall short. This has prompted a growing amount of research focused on developing edge-aware loss formulations and metrics that better align with the needs of semantic facies segmentation. In this section, we review key contributions in this area, highlighting their strengths and limitations, with particular attention to their applicability to seismic data and their capacity to preserve boundary precision.

The pixel-wise Cross-Entropy Loss (CEL) has emerged as one of the most widely used loss functions in seismic segmentation tasks (Jadon, 2020; Wang et al., 2021). CEL is favored for its straightforward approach as it encourages models to predict the correct class for each pixel by measuring the dissimilarity between estimated class probabilities and ground-truth labels. Despite its popularity, CEL overlooks a critical aspect of semantic segmentation: prioritizing the accurate delineation of edges and boundaries between distinct geological features or regions. Boundaries between geological features often represent the most critical information in seismic data, as they delineate reflection transitions and stratigraphic changes between distinct subsurface structures. CEL does not impose additional penalties for errors occurring near the boundaries of different classes but treats all pixels uniformly regardless of their spatial context. As a result, neglecting the importance of boundary precision can significantly compromise the accuracy of seismic segmentation and present profound implications for the interpretation

of subsurface geology.

Familiar with the limitations of CEL in handling boundary precision, several researchers have explored edge-aware strategies to enhance segmentation accuracy. Csurka et al. (2013), for instance, adapted the Berkeley contour matching score (Martin et al., 2004) to propose the Boundary F1 Score (BF1S), specifically designed to evaluate the accuracy of reconstructed contours in segmentation problems. Bokhovkin and Burnaev (2019) introduced the Boundary F1 Loss (BFL), which integrates both BF1S and Intersection over Union (IOU) to penalize errors along object borders in binary classification of remote sensing data. The BF1S metric provides a more refined assessment of edge quality compared to traditional pixel-wise accuracy, whereas IOU has proven effective in contexts such as satellite imagery, where overall region overlap is critical. Although BFL improves boundary delineation, it is prone to higher computational cost and increased sensitivity to hyperparameters, especially with high-resolution imagery. Its adaptation to seismic data is further limited by the ambiguity of boundaries in highly textured or overlapping regions. In addition, its vulnerability to noise and acquisition artifacts may cause overfitting or imprecise contour predictions, undermining its reliability.

Audebert et al. (2019) reframed segmentation as a regression problem, leveraging distance maps to capture spatial relationships and encourage smooth, continuous outputs. This method predicts signed distance fields rather than binary masks, allowing the model to learn spatially coherent object representations. While effective for objects with regular ge-

ometry, it is less capable of representing the irregular and heterogeneous patterns common in subsurface data. Moreover, signed distance maps are highly sensitive to ground-truth precision, and noisy annotations can sharply degrade performance. Caliva et al. (2019) proposed a cost function that weights pixel-level errors according to their proximity to object boundaries. This approach prioritizes edge accuracy, attempting to produce sharper object delineations. However, in seismic data, where formations exhibit complex layering and poorly defined boundaries, signed distance maps may induce numerical instabilities near noisy interfaces, often requiring smoothing or post-processing for stability.

In medical imaging, Xue et al. (2020) introduced a shape-aware segmentation method that also predicts signed distance maps to improve organ boundary localization. Although effective in domains with clear structural priors and regular geometries, this method assumes a degree of shape consistency that does not extend to seismic interpretation, where structures are frequently irregular, discontinuous, and noisy. Furthermore, signed distance maps are computationally heavier than unsigned ones, particularly in 3D or high-resolution volumes. Similarly, Gerl et al. (2020) targeted biological tissue segmentation using a loss based on signed distance maps. Their formulation was designed for single continuous surfaces with nearly horizontal layering and limited class diversity. The method incorporates prior knowledge of skin thickness to enforce anatomical plausibility, which improves segmentation stability, but such assumptions are not transferable to seismic data characterized by faults, unconformities, and abrupt lithological changes.

More recently, Wang et al. (2022) introduced the Active Boundary Loss (ABL), which enhances edge localization by computing a direction vector for each pixel along the predicted boundary and guiding it towards the nearest point on the ground truth. This iterative correction mechanism allows the model to refine its predictions dynamically, improving contour alignment. ABL has shown promising results in urban image segmentation tasks, especially where object boundaries are sharp and spatially consistent. However, in seismic applications, boundaries such as faults, unconformities, and stratigraphic contacts are often weak, noisy, and discontinuous, unlike those in natural images. This complexity makes the estimation of reliable direction vectors difficult, since many boundaries are fragmented, uncertain, or even absent in the annotations. Thus, while ABL is theoretically well-suited for edge refinement, its dependency on boundary integrity and consistent spatial structure can hinder robustness and limit its generalization in seismic facies segmentation.

Although the aforementioned approaches have made notable progress in improving edge delineation in semantic segmentation, they are not without limitations. Many of these methods introduce increased computational overhead and exhibit heightened sensitivity to hyper-parameter tuning, particularly when applied to high-resolution images and large-scale datasets. Furthermore, despite their effectiveness in natural image domains, most of these loss functions (Zhao

et al., 2019; Chen et al., 2020; Borse et al., 2021; Liang et al., 2022; Tian et al., 2022) have not been rigorously evaluated on seismic datasets or released their implementation to the research community. Seismic data present unique challenges, such as low signal-to-noise ratios, discontinuous and subtle boundaries, and significant intra-class variability, which may not be adequately addressed by models calibrated on conventional benchmarks. Moreover, the absence of standardized evaluation protocols and benchmarking practices within the geophysical domain complicates fair comparisons and raises concerns regarding the generalizability and robustness of these methods when transferred to more complex and domain-specific tasks.

The effectiveness of the aforementioned loss functions depends on two interdependent factors: (i) the choice of an appropriate deep learning architecture for semantic segmentation and (ii) the availability of high-quality annotated datasets. In fact, the generalization capability of deep learning models is heavily reliant on the quality and quantity of annotated datasets as they provide the essential ground truth required for training models to accurately distinguish and classify complex subsurface geological structures. Fortunately, publicly labeled datasets – such as F3 BLOCK, PENOBSCOT, and PARIHAKA – serve as essential benchmarks, enabling reproducible evaluation of how cost functions perform under domain-specific challenges such as class imbalance and ambiguous stratigraphic transitions.

The Netherlands' F3 BLOCK dataset (Alaudah et al., 2019) was built using 3D seismic data acquired in 1987, supplemented with information from 26 well logs and guided by a detailed geological study of the region. The dataset represents a fully interpreted geological model of the F3 Block, located offshore in the North Sea — an area well-known for its hydrocarbon potential. To simplify the original stratigraphic complexity for machine learning applications, the benchmark merges seven lithostratigraphic units (Upper, Middle, and Lower North Sea Groups; Chalk; Rijnland; Scruff; Zechstein) into six facies classes by combining the Rijnland and Chalk groups. The seismic dataset covers an area of approximately 386 km² and is organized as a 602 × 902 × 255 volume, comprising 602 inlines (100–701), 902 crosslines (300–1201), and 255 depth samples. Basically, each voxel represents a cube of 25 × 25 × 25 meters in size, offering a high-resolution 3D representation of the subsurface.

The PENOBSCOT dataset from Nova Scotia (Baroni et al., 2019) is derived from the publicly available *Penobscot 3D seismic survey*, acquired on the Scotian Shelf, Canada. Although the original dataset contained five interpreted horizons and well log data, it was reinterpreted for machine learning purposes to include seven horizons and four faults, which together define eight seismic facies intervals (Shallow Cover, Deep Parallel, Mississauga Deltaics, Transitional Parallel, Subtle Deep Marine, Variable Subparallel, Slump Deposit Zone, and High-Energy Truncated). The full seismic volume spans roughly 87 km², consisting of 601 inlines and 482 crosslines, with a total recording time of 6000 ms sampled at 2 ms intervals. The acquisition bin size is 12.5 meters in the

inline direction and 25 meters in the crossline direction. For model development and evaluation, the dataset is split into training and testing subsets: 289 slices are used for training and 192 for testing in the crossline direction, whereas 358 slices are used for training and 238 for testing in the inline direction.

The PARIHAKA dataset from New Zealand (Inc, 2020) is based on a 3D seismic survey conducted offshore Taranaki, covering an area of approximately 1520 km². Dual-source arrays were fired alternately every 25 meters during acquisition, and approximately six seconds of seismic reflection data were recorded for each shot, with a sampling interval of 2 ms. The dataset was released by *New Zealand Petroleum and Minerals (NZPM)* and comprises a seismic volume along with a corresponding labeled volume, where each voxel is assigned to a seismic facies class. The annotation volume categorizes each voxel into one of six facies categories: Basement, Slope Mudstone A, Mass Transport Deposit, Slope Mudstone B, Slope Valley, and Submarine Canyon System. For machine learning applications, the data is formatted into a volume of dimensions 590 × 782 × 1006, representing 590 inlines, 782 crosslines, and 1006 time samples per trace, as standardized in prior benchmarks (Kaur et al., 2023).

F3 BLOCK, PENOBSCOT, and PARIHAKA provide a diverse foundation for benchmarking seismic segmentation methods. In fact, each dataset offers unique geological settings, stratigraphic complexities, acquisition parameters, and a detailed set of annotations, allowing for a comprehensive evaluation of model performance across different geological conditions. Unfortunately, most publicly available seismic datasets lack annotated volumes, which constrains the scalability and generalizability of supervised deep learning models. Labeled seismic data are notably scarce due to the inherent complexity and cost associated with generating accurate annotations, requiring domain expertise from geoscientists who must analyze volumetric data and identify subtle geological structures with high precision. Although the present study focuses on a novel cost function tailored for supervised learning, we recognize the importance of unsupervised and self-supervised strategies that exploit the intrinsic structure of seismic data to pre-train models in the absence of manual labels. Such approaches not only reduce the dependency on expert-generated annotations but also improve model generalization across unseen data.

Recent developments such as SALT3DNET (Yang et al., 2024), the self-supervised learning framework proposed by Monteiro et al. (2022), and the few-shot segmentation approach introduced by Saad et al. (2022), demonstrate the potential of these techniques to advance seismic interpretation with minimal supervision, bridging the gap between data availability and model performance. While self-supervised learning has become a widely adopted strategy for pretraining models on large-scale unlabeled seismic data, downstream specialization still relies on task-specific fine-tuning and optimization to achieve high segmentation accuracy. This is where the proposed Distance Transform Loss provides distinct advantages: by incorporating boundary-aware supervi-

sion, our approach promotes sharper delineation of complex geological features — such as faults, channel edges, and salt-body boundaries — thus enhancing the effectiveness of fine-tuning in seismic segmentation tasks.

3. Boundary-Aware Cost Function

Distance Transform (DT), also referred to as Distance Mapping or Distance Transformation, is a method for computing the distance from each non-target pixel in an image (*background*) to the nearest pixel belonging to a target object (*foreground*). DT produces a distance map usually employing L_1 , L_2 , or L_∞ norm, respectively known as Manhattan, Euclidean, and Chebyshev. The DT computation relies on the prior execution of image translation and contour detection: the former applies a transformation matrix to displace pixels in x and y directions and generate shifted images, whereas the latter subtracts an image from its shifted version so that edge-lying pixels store non-null values and the remaining ones are filled with zeros. DT lays the foundation for the proposed Distance Transform Loss, a criterion that can be attached to versatile cost functions in favor of punishing deep neural networks when region or object boundaries are poorly delineated.

3.1. Distance Transformation

The Distance Transform (DT) is an image processing technique that transforms a binary image $I \in \mathbb{R}^{m \times n}$ with m rows and n columns into a new representation where each background pixel at coordinates (x, y) is assigned a value corresponding to its distance from the nearest foreground (target) pixel located at (\check{x}, \check{y}) (Felzenszwalb and Huttenlocher, 2012). To formalize this process, let $D, \mathcal{P} : \mathbb{I}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ denote functions applied over image I , where D computes the Manhattan distance between a pixel (x, y) and its closest foreground pixel (\check{x}, \check{y}) , and \mathcal{P} is a predicate function that returns zero when $(\check{x}, \check{y}) \in F$, with F representing the set of all foreground pixels. The resulting distance map provides a continuous and differentiable gradient signal that is particularly useful for tasks involving boundary-sensitive segmentation.

$$D(I) = \min_{(\check{x}, \check{y}) \in I} (|x - \check{x}| + |y - \check{y}| + \mathcal{P}(\check{x}, \check{y})), \forall (x, y) \in I \quad (1a)$$

$$\mathcal{P}(\check{x}, \check{y}) = \begin{cases} 0 & \text{if } (\check{x}, \check{y}) \in F \\ \infty & \text{otherwise} \end{cases} \quad (1b)$$

Instinctively, \mathcal{P} secures that the mapping function D only considers the distance between the target and non-target points by aggressively penalizing the ones not enrolled in F as they become very large to satisfy the minimization function. Note that the first term in Equation (1a) does not depend on the y column and, therefore, it can be rewritten as follows in order to fix row to index x so that it iteratively searches for the minimum separation between reference point (x, y) and all the other pixels lying in coordinates $(x, \check{y}) \forall \check{y} \in I^n$:

$$D(I) = \min_{\check{x} \in I} (|x - \check{x}| + \min_{\check{y} \in I} (|y - \check{y}| + \mathcal{P}(\check{x}, \check{y}))), \forall (x, y) \in I \quad (2a)$$

$$D(I) = \min_{\check{x} \in I} (|x - \check{x}| + D|_{\check{x}}(y)), \forall (x, y) \in I \quad (2b)$$

The restriction $D|_{\check{x}}(y)$ applied to distance function D refers to a single-directional column-wise distance transform limited to row indexed by \check{x} . In other words, a two-dimensional Manhattan L_1 distance mapping consists of m one-dimensional computations along each column times n transformations along each row. Algorithm 1 computes the distance mapping under the L_1 metric for an array of size n taking $\mathcal{O}(n)$ time (Felzenszwalb and Huttenlocher, 2012)². For $m \times n$ binary images, a single-directional distance mapping can be attained by repeating Algorithm 1 a total of m times. Then, a multi-directional approximation to the L_1 norm can be attained in linear time with respect to the number of pixels, i.e. $\mathcal{O}(d \times m \times n)$, where d is the number of orientations usually assuming a constant value $d = 2$. The first execution covers the vertical direction whereas the second one handles the horizontal orientation. As a result, the algorithm produces an optimal $m \times n$ distance transformation matrix seeing that boundaries can resemble any sort of shape or structure.

Algorithm 1 One-dimensional Distance Transform (D)

Require: $I \in \{0, 1\}^n \quad \triangleright$ binary image ($I_{[i]} = 1$ if $i \in F$)
 $D \leftarrow \bar{I} * \infty \quad \triangleright$ image complement
for $i = 1$ **to** n **do**
 $D_{[i]} \leftarrow \min(D_{[i]}, D_{[i-1]} + 1) \quad \triangleright$ forward pass
end for
for $i = n - 1$ **to** 0 **do**
 $D_{[i]} \leftarrow \min(D_{[i]}, D_{[i+1]} + 1) \quad \triangleright$ backward pass
end for

3.2. Distance Metric Efficiency

Essentially, the adoption of the Manhattan distance is generally motivated by both computational efficiency and its alignment with the discrete grid structure of seismic data. The L_1 metric preserves directional uniformity on axis-aligned grids, which is consistent with how class boundaries typically appear in seismic volumes. It ensures more conservative and interpretable penalization near class boundaries, which is crucial in tasks such as seismic facies classification where geological boundary delineation is paramount. Other metrics like L_2 , also known as Euclidean distance, tends to underestimate the diagonal units of distance maps, potentially leading to overly optimistic gradient signals during backpropagation.

The L_1 -based distance transform can be optimally implemented via dynamic programming, as demonstrated in Algorithm 1, due to its inherent monotonicity property. More precisely, the L_1 metric considers only horizontal and vertical distances, the minimum distance values propagate consistently across the grid, enabling the reuse of previously

computed values during forward and backward passes. In contrast, the L_2 (Euclidean) distance does not possess the same monotonicity since it involves diagonal relationships and square root operations, requiring more complex algorithms that do not benefit from the very same straightforward dynamic programming optimizations. Furthermore, the selection of L_1 is not only computationally motivated but also implementation-friendly for GPU-based training pipelines where runtime efficiency is critical.

3.3. Distance Transform Loss

The Distance Transform Loss (DTL) is designed to drive a machine-learning model toward predicting sharp and accurate boundaries by penalizing networks for misclassified region delineations. Figure 1 illustrates how the method works when distinct predictions, P_a and P_b (top/bottom), are contrasted with the same reference label G (middle row). DTL serves the purpose of deriving a penalty score that reflects the Manhattan distance between the predicted boundary and its associated ground truth. Equation 3a provides a definition for DTL, which depends on contour detection (3b), binarization (3c), geometric translation (3d), followed by a pixel-wise multiplication (Gonzalez and Woods, 2018). Component $P \in \mathbb{R}^{m \times n \times K}$ represents the predicted segmentation image whereas $G \in \{\mathbb{B}^{m \times n \times K} \mid \mathbb{B} = \{0, 1\}\}$ denotes its corresponding binary ground truth, K indicates the number of training classes, and (x, y) designates the pair of pixel coordinates. Essentially, DTL can be expressed as:

$$\mathcal{L}_{\text{DTL}}(P, G) = \sum_{k=0}^K \sum_{(x,y)}^{m \times n} [C(P, k) \odot D(C(G, k))]_{[x,y]} \quad (3a)$$

$$C(I, k) = [I^{(k)} - \check{I}^{(k)}] \neq 0, \text{ where:} \quad (3b)$$

$$I_{[x,y]}^{(k)} = \begin{cases} 1 & \text{if } \arg \max_z (I_{[x,y,z]}) = k \\ 0 & \text{otherwise} \end{cases}, \forall (x, y) \in I \quad (3c)$$

$$\check{I}_{[\check{x}, \check{y}]}^{(k)} = I_{[x,y]}^{(k)} \mid [\check{x} \ \check{y}] = [x \ y \ 1] \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ t_x & t_y \end{bmatrix}, \forall (x, y) \in I^{(k)} \quad (3d)$$

The proposed DTL function, denoted as $\mathcal{L}_{\text{DTL}}(P, G)$, takes instances P and G as inputs and immediately employs sub-functions $C(P, k)$ and $C(G, k)$ for each image independently, considering class k . In fact, $C(I, k)$ corresponds to the process of obtaining contours through the subtraction of the binary image $I^{(k)}$ from its shifted version $\check{I}^{(k)}$, an agile strategy to highlight edge-details prominently. Component $I^{(k)}$ represents the flattening and binarization of image I with respect to class k . Basically, it iterates over all pairs of pixel coordinates (x, y) in the three-dimensional grid I and checks

²Big \mathcal{O} notation is a well-known way to describe the performance of an algorithm as the input size grows. In terms of distance transformation, it provides an upper bound of how the algorithm's runtime increases as the input image dimension varies.

whether the index along the z -axis holding the highest probability distribution matches class k . When it holds true, the corresponding position in the resulting two-dimensional matrix $I^{(k)} \in \mathbb{R}^{m \times n}$ is set to 1 or 0, otherwise. Component $\check{I}^{(k)}$ refers to the transformation of coordinates where (x, y) are pixels positions in the binarized image $I^{(k)}$ and (\check{x}, \check{y}) are their respective coordinates on the translated image \check{I} . In the proposed approach, the translation matrix holds parameters $t_x = 1$ and $t_y = 1$ responsible for the shifting amount on the x and y -axis, respectively; sliding all pixels one unit to the right and one unit downward.

Upon the estimation of both components, function $C(I, k)$ calculates the underlying contour for class k between binary images $I^{(k)}$ and $\check{I}^{(k)}$ by subtracting one from the other. As illustrated in Figure 1, the contour detection process generates a new image where pixels located along the class boundaries accumulate non-zero values, whereas the background region is filled with pixels having a value of zero. Fundamentally, the Boolean operator \neq ensures that non-negative values persist on the grid as it is transformed into a binary matrix.

Back to Equation (3a), function $\mathcal{L}_{\text{DTL}}(P, G)$ invokes the distance transform function D to compute the map associated with the ground-truth boundary contours. It is accomplished by iteratively assigning each background pixel the L_1 distance to the nearest target point lying on the foreground class boundary. Symbol \odot indicates an element-wise multiplication conducted between the binarized contour $C(P, k)$ and the obtained distance mapping $D(C(G, k))$, which results in a transient image $I' \in \mathbb{R}^{m \times n}$. Observe that the inner-right sigma notation designates the overall sum of all pixel values in the resulting sparse matrix by looping through all pixel coordinates (x, y) taken from image I' . The outer-left summation aggregates the error score of each class $k \in K$ separately. Basically, a low error score announces that the predicted segmentation image achieves a near-perfect match to its respective ground-truth sample.

3.4. Aggregation of Cost Functions

The proposed cost function focuses on guiding the network toward the demarcation of more accurate class frontiers as it considers the Manhattan geometry among pixel coordinates. Unlike recent boundary-aware cost functions (Bokhovkin and Burnaev, 2019; Zhu et al., 2018; Liang et al., 2022; Borse et al., 2021; Wang et al., 2022), DTL has been specifically designed for multinomial semantic segmentation. As a matter of fact, DTL can be used in conjunction with other loss functions that measure the discrepancy between the predicted probability distribution of a neural network model and the actual class labels of the ground-truth data, such as the pixel-wise CEL (Wang et al., 2021), as shown below in the form of \mathcal{L}_{CEL} :

$$\mathcal{L}_{\text{CEL}}(P, G) = \sum_{(x,y)} \sum_{k=0}^{K-1} \left(-\log \left(P_{[x,y,k]} \right) * G_{[x,y,k]} \right) \quad (4a)$$

$$\mathcal{L}(P, G) = \mathcal{L}_{\text{CEL}}(P, G) + (w * \mathcal{L}_{\text{DTL}}(P, G)^e) \quad (4b)$$

The combined loss, denoted as $\mathcal{L}(P, G)$, is the one we have adopted to train deep network models for seismic semantic segmentation in the experimental section. Not only does it encompass \mathcal{L}_{DTL} but also \mathcal{L}_{CEL} , whose error score decreases when the predicted per-pixel probability distribution P approaches the target distribution G , encouraging the model to predict the correct class for each pixel. Function $\mathcal{L}(P, G)$ contains two parameters, w and e , that have been adopted to equalize the order of magnitude between CEL and DTL. The former acts as a conventional weight whereas the latter constitutes an exponentiation mark. Concisely, the addition of \mathcal{L}_{CEL} and \mathcal{L}_{DTL} seeks the proper multi-class classification at the pixel level and simultaneously adds an extra penalty to mismatched boundaries. Therefore, minimizing \mathcal{L}_{CEL} and \mathcal{L}_{DTL} during the training stage leads to a model that better classifies pixels and segments regions more accurately.

3.5. Computational Cost Analysis

The computational complexity of loss functions is a key consideration in the training efficiency of deep learning models. The pixel-wise CEL is one of the most widely used cost functions in semantic segmentation due to its simplicity and efficiency. As indicated in Equation 4a, CEL independently computes the negative log-likelihood between the predicted class probability and the ground-truth label for each pixel. Given a 2D grayscale or multi-channel image of size $m \times n$, this operation involves a softmax activation followed by a logarithm and multiplication, all of which are constant-time operations per pixel. Therefore, the total computational complexity of CEL scales linearly with the number of pixels, resulting in a time complexity of $\mathcal{O}(m \times n)$.

Additionally, DT introduces a spatially-aware component by encoding the distance of each pixel to the nearest object boundary. The proposed approach uses the L_1 Manhattan distance metric, which is particularly well-suited for grid-based image data. For a 2D image, the DT can be computed efficiently via two directional passes — horizontal and vertical — resulting in a total complexity of $\mathcal{O}(d \times m \times n)$, where $d = 2$ is a constant number of scan directions. When applied per category in a K -class segmentation task, the complexity increases linearly with K , becoming $\mathcal{O}(K \times m \times n)$. Despite this additional cost, our efficient dynamic programming implementations described in Algorithm 1 makes DT mapping computationally manageable during training, adding no more than 20% runtime per iteration.

The overall computational cost is a combination of both components as we integrate DT with CEL. Specifically, DTL performs a pixel-wise multiplication between the distance map and the predicted class probabilities, effectively enforcing a spatially-aware penalty for misclassified pixels near object boundaries. As CEL operates at $\mathcal{O}(m \times n)$ and DT mapping runs at $\mathcal{O}(K \times m \times n)$, the complexity of aggregating both becomes $\mathcal{O}((K + 1) \times m \times n)$. Even though DTL introduces a measurable overhead compared to CEL alone, this increase is modest relative to the potential improvements in boundary delineation. Furthermore, modern GPU-accelerated deep learning frameworks can efficiently

parallelize these computations, thereby keeping the additional training time within acceptable bounds for practical applications. In general, most semantic segmentation tasks involving seismic data comprise a relatively low number of classes, typically ranging from three to seven. The few classes portray the nature of geological interpretation, where the objective generally consists of identifying a limited set of key structural features. This characteristic restricts the computational overhead of class-dependent operations and makes the integration of more advanced loss formulations computationally feasible in practice.

4. Experimental Results

This section provides a comprehensive analysis of the performance of the proposed Distance Transform Loss (DTL), benchmarking it against state-of-the-art and widely adopted cost functions in the semantic segmentation domain. To rigorously assess the statistical significance of the results, we employ the dependent-sample *t-Test* statistic, ensuring that observed improvements are not due to random variation but reflect meaningful advancements in performance. To enable a deeper and more nuanced interpretation of the obtained results, we direct readers to the supplementary materials accompanying this work. The additional content provides an exhaustive breakdown of performance under a variety of quantitative metrics, such as class accuracy, intersection over union, pixel accuracy, and F1 score, enabling a holistic evaluation of the segmentation capabilities of DTL.

4.1. Assessment Details

The implementation of our approach is carried out using the PyTorch framework (Paszke et al., 2019), a widely recognized and versatile deep learning library. Experiments were conducted on a high-performance computational server configured with cutting-edge hardware: the server is equipped with an AMD EPYC 7742 CPU, offering 64 cores and 128 threads for efficient data processing and parallel computation. The system also includes an impressive 1.97 TB Random Access Memory (RAM) for intensive tasks such as seismic volume processing. In addition, the server contains a NVIDIA A100-SXM4 GPU with 80 GB of dedicated memory optimized for accelerated training of deep convolutional neural networks. Despite all available resources, no more than 60 GB of RAM and GPU memory were required to train the chosen architectures, considering entire sections of the chosen seismic volumes. Training with whole sections rather than fragments or patches requires more memory and computational resources, as the network must process larger amounts of data in each training step. Besides, it also increases the training time due to the higher computational burden involved in training higher-resolution seismic volumes.

4.2. Segmentation Metrics

The Mean Intersection over Union, also referred to as mIoU, is a widely recognized and extensively used evaluation metric for semantic segmentation tasks. It quantifies the degree of overlap between the segmented regions predicted

by a model and the corresponding ground-truth masks, offering a straightforward measure of segmentation accuracy. Predominantly, the Intersection over Union is commonly defined as $\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{|P \cap G|}{|P \cup G|}$, where P denotes the pixels in the predicted region and G the ground-truth region. As a consequence, the mIoU is obtained by averaging the IoU values across each class $k \in K$ in such a way that $\text{mIoU} = \frac{1}{|K|} \sum_k \text{IoU}_k = \frac{1}{|K|} \sum_k \frac{|P_k \cap G_k|}{|P_k \cup G_k|}$, where $|K|$ is the total number of classes. Even though mIoU is particularly informative in the presence of class imbalance in virtue of preventing dominant classes from overshadowing minority ones, it is insensitive to boundary precision, which is a key limitation in seismic segmentation tasks.

To address the lack of boundary sensitivity in mIoU, we also consider the Boundary F1 Score (BF1S) (Csurka et al., 2013), which evaluates segmentation accuracy in terms of contour alignment. More precisely, BF1S evaluates the quality of predicted boundary delineation by calculating the harmonic mean of the precision and recall of frontier pixels, thereby taking into account the proximity of class contours. This metric is defined as $\text{BF1S} = 2 \cdot \frac{\text{Precision}_b \cdot \text{Recall}_b}{\text{Precision}_b + \text{Recall}_b}$, where Precision_b is the fraction of predicted boundary pixels correctly overlapping the ground-truth boundaries within a tolerance distance, and Recall_b is the fraction of ground-truth boundaries correctly detected by the prediction. A higher BF1S indicates that the predicted boundaries are closer to the ground truth boundaries. In summary, this metric is analogous to the conventional F1 score but applied to boundary pixels, emphasizing the ability of a model to capture fine structural details, which is especially important for seismic horizons and fault delineation.

We also include two pixel-level classification metrics: Pixel Accuracy (PA) and Mean Class Accuracy (MCA). PA measures the proportion of correctly classified pixels across all classes with respect to the total number of pixels in the ground truth. Formally, it is defined as $\text{PA} = \frac{\sum_k |P_k \cap G_k|}{\sum_k |G_k|}$ where P_k and G_k denote the predicted and ground-truth masks of class k . In plain English, PA corresponds to the ratio between the total number of correctly predicted pixels (summed over all classes) and the total number of ground-truth pixels. Despite its simplicity, PA can be misleading in highly imbalanced datasets, as models may achieve high PA by primarily predicting the majority class. In contrast, MCA computes the accuracy for each class individually and then averages these values, ensuring that each class contributes equally regardless of its frequency in the dataset. It is defined as $\text{MCA} = \frac{1}{|K|} \sum_k \text{CA}_i = \frac{1}{|K|} \sum_k \frac{|P_k \cap G_k|}{|G_k|}$, representing the arithmetic mean of the per-class accuracies, mitigating the bias toward majority classes that affects PA. This aspect makes MCA particularly valuable in scenarios with severe class imbalance, such as seismic interpretation, where minority classes often carry crucial structural information.

4.2.1. Evaluation Protocol

To ensure a reliable and robust evaluation of the proposed method and its baselines, we adopt a k -fold cross-validation scheme with $k = 10$ partitions. This widely used protocol offers a statistically grounded alternative to a simple train-test split by averaging performance across multiple folds, thus mitigating variance due to dataset partitioning. It also reduces the risk of *overfitting* to a specific data subset and provides a more generalizable estimate of model performance. Each model is trained and validated ten times ensuring that every inline and crossline instance appears in the validation subset exactly once and in the training subset the remaining $k - 1$ times. This repetition improves the robustness of the evaluation and helps account for variability across different sections of the seismic volumes. Given the high spatial correlation between adjacent seismic slices, as previously noted by Alaudah et al. (2019), we perform the selection of inlines and crosslines in a continuous and deterministic fashion across all folds. This methodology prevents data leakage and preserves the spatial structure inherent to the seismic volumes.

Tables 2 and 3 report the average performance across all folds, highlighting statistically significant improvements using asterisks derived from dependent-sample *t-Test*. This statistical test evaluates whether the mean differences in performance between DTL and each baseline loss function are statistically significant, assuming a paired design since the same data splits are used across all methods. The test yields p -values that quantify the likelihood that the observed improvements are not merely coincidental. We highlight the results using asterisks as follows: *** for p -values ≤ 0.001 , ** for ≤ 0.01 , and * for ≤ 0.05 , where more marks indicate increasing levels of statistical confidence and superiority. These results validate the consistency and reliability of the observed gains, particularly for metrics sensitive to boundary accuracy such as BF1s. In addition, Table 1 presents the mean and 95% confidence intervals for the hyperparameters selected for DTL during model optimization. These intervals provide a statistical measure of variability across folds and serve as a reference for future studies aiming to reproduce or extend the findings.

4.2.2. Seismic Datasets

We conducted a comprehensive evaluation of our proposed method across three publicly available seismic datasets, each presenting unique geological and structural characteristics. This diversity allows for a well-rounded assessment of segmentation performance under varying conditions and degrees of annotation availability.

The F3 BLOCK dataset (Alaudah et al., 2019), acquired from the Dutch North Sea sector, represents the smallest but most thoroughly annotated volume in our study. It includes clearly defined training and test partitions, making it suitable for rigorous benchmarking. To assess model robustness, we employ k -fold cross-validation by partitioning the training set into k disjoint folds. The final results are reported as the average performance across all folds, evaluated on the two official test sets, ensuring fair comparison and reproducibility.

Table 1

bF1s results concerning optimal values for w and e .

$w \backslash e$	F3 BLOCK + DECONVNET + DTL		
	1/1	3/4	1/2
0.05	62.14 \pm 0.56	63.55 \pm 0.62	63.45 \pm 0.55
0.10	50.85 \pm 0.99	63.59 \pm 0.54	64.06 \pm 0.53
0.50	13.33 \pm 0.36	64.01 \pm 0.55	64.89 \pm 0.59
1.00	03.45 \pm 0.58	62.67 \pm 0.56	64.97 \pm 0.54

For the PARIHAKA dataset (Inc, 2020) from New Zealand's Taranaki Basin, we worked with the annotated training volume containing six facies classes. Since the companion test sets lack ground truth annotations, we adopted a conservative evaluation approach using k -fold cross-validation exclusively on the labeled training data. This strategy ensures all reported metrics derive from verified annotations while maintaining methodological consistency with the other experiments. The dataset's particular challenge lies in its complex depositional environments, including mass transport deposits and submarine canyon systems.

Lastly, we included the PENOBSCOT dataset (Baroni et al., 2019), collected from the Scotian Shelf in Nova Scotia, Canada. This dataset comprises a single seismic volume with annotations spanning seven seismic facies. Given its structure, we employed a strict k -fold cross-validation protocol, rotating the validation set across k equally sized folds while using the remaining $k - 1$ folds for training. This procedure enables the utilization of the entire annotated volume while preventing data leakage, offering a more robust and statistically sound evaluation of model generalization. Together, these three datasets cover a wide range of geological scenarios, annotation levels, and acquisition conditions, providing a strong foundation for benchmarking the proposed segmentation approach across realistic and challenging seismic interpretation tasks.

4.2.3. Training Settings

To ensure consistent and fair training across all datasets, we adopted a unified training framework while adapting specific hyperparameters to the characteristics of each seismic volume. Our approach aimed to maximize model performance while maintaining generalizability and robustness across diverse geological settings. For the F3 BLOCK dataset (Alaudah et al., 2019), we employed a batch size of eight samples per iteration, which was well-suited given the moderate volume dimensions and memory constraints. Training was conducted using the *Adam* optimizer, widely recognized for its adaptive learning capabilities and effectiveness in various deep learning applications. The learning rate was set to 0.0003, a value determined empirically to offer a reliable trade-off between convergence speed and training stability. Each model was trained for a maximum of 60 epochs, with early stopping applied to terminate training when no improvement was observed in validation performance for a predefined number of epochs. Data augmentation techniques such as

Table 2

Results exposing average **bF1s** and **mIoU** results for Active Boundary Loss (ABL), Boundary Loss (BFL), Cross-Entropy Loss (CEL), and the proposed DTL. The results are marked with asterisks to indicate statistical significance: *** for $p \leq 0.001$, ** for $p \leq 0.01$, and * for $p \leq 0.05$, with more asterisks denoting higher confidence levels.

Datasets	Arch Loss	Boundary F1 Score (bF1s)			Mean Intersection Over Union (mIoU)		
		DECONVNET	SEGNET	U-NET	DECONVNET	SEGNET	U-NET
F3 BLOCK	ABL	63.48	64.21	65.23	88.80	89.34*	89.31
	BFL	63.35	62.37	66.25	86.60	84.69	89.93
	CEL	63.75	64.12	66.31	88.68	88.55	90.70
	DTL	64.97***	65.05*	67.17*	90.06***	89.26*	91.50**
PENOBSCOT	ABL	72.58	63.16	57.77	90.08	83.94	81.88
	BFL	71.46	65.52	61.19	85.16	79.63	74.25
	CEL	75.08***	68.03	60.21	91.50	87.57	84.91
	DTL	75.33***	73.57***	65.96***	92.70**	91.33***	88.73***
PARIHAKA	ABL	69.88	70.51	71.25	87.36	88.19	87.92
	BFL	73.91***	69.59	74.68***	88.73***	84.57	88.67
	CEL	69.50	69.90	71.36	86.76	87.66	88.00
	DTL	72.62	72.90***	73.66	88.54***	89.82**	89.36**

random rotations, vertical flipping, and additive Gaussian noise were applied to promote generalization and reduce overfitting.

In contrast, the PARIHAKA dataset (Inc, 2020) features significantly larger input volumes, necessitating a reduction in batch size to six samples per iteration to accommodate GPU memory limitations. Despite the change in batch size, the same optimization strategy, learning rate, and stopping criteria were employed to ensure consistency in training dynamics. The use of data augmentation remained critical due to the geological complexity of the Taranaki Basin, enabling the network to learn more invariant and generalized features across facies boundaries. As mentioned before, we followed a k -fold cross-validation scheme exclusively on the annotated training portion of the data, given the absence of ground truth for the provided test volumes. Similarly, the PENOBSCOT dataset (Baroni et al., 2019), which includes a single annotated volume with considerable spatial extent and stratigraphic variability, also required a reduced batch size of six. We adhered to the same optimizer, learning rate, and early-stopping policy used for the other datasets. Due to its single-volume nature, a k -fold cross-validation protocol was crucial to properly evaluate model generalization and prevent overfitting to specific regions of the dataset. The same set of data augmentation techniques was employed to simulate plausible variability in seismic patterns, thereby improving robustness.

Across all datasets, we avoided the use of small patches during training and instead processed full 2D seismic sections (e.g., entire inlines or crosslines). This design choice was motivated by findings from Alaudah et al. (2019), which indicate that using full sections reduces the risk of context loss and misclassification errors that often occur near geological boundaries when training on small patches. This strategy

allowed the models to exploit broader spatial context and better capture large-scale geological structures critical for accurate seismic segmentation.

4.3. Parameters Selection

To determine the optimal hyperparameters for the proposed DTL, we conduct an extensive grid-search heuristic as described in the evaluation protocol, focusing on the two key components introduced in Equation 4b: the weighting factor w and the exponentiation parameter e . This experiment is carried out using a DECONVNET architecture trained on the F3 BLOCK dataset, which allows us to assess the robustness and adaptability of DTL in a new context, distinct from the original model-dataset combinations. The results, summarized in Table 1, present the mean bF1s (Boundary F1 Score) for each configuration, offering empirical insights into how different settings influence the boundary-aware segmentation performance.

The analysis reveals that configuring DTL with an exponent $e = 0.50$ consistently leads to superior boundary delineation, regardless of the value chosen for w . Notably, when $e = 0.50$, even a neutral weight factor $w = 1.00$ yields competitive results, indicating that the exponentiation operation alone sufficiently modulates the loss function to prioritize relevant boundary regions. This outcome suggests that the influence of w becomes marginal when e is adequately set, effectively eliminating the need for tuning w in practical applications. Consequently, DTL can be simplified to a single-hyperparameter formulation under this configuration, which not only reduces the computational burden of hyperparameter tuning but also facilitates its integration into existing pipelines with minimal manual intervention. These findings further highlight DTL practicality and ease of deployment in real-world seismic interpretation tasks.

Table 3

Results exposing average **mCA** and **PA** results for Active Boundary Loss (ABL), Boundary Loss (BFL), Cross-Entropy Loss (CEL), and the proposed DTL. The results are marked with asterisks to indicate statistical significance: *** for $p \leq 0.001$, ** for $p \leq 0.01$, and * for $p \leq 0.05$, with more asterisks denoting higher confidence levels.

Datasets	Arch Loss	Mean Class Accuracy (mCA)			Pixel Accuracy (PA)		
		DECONVNET	SEGNET	U-NET	DECONVNET	SEGNET	U-NET
F3 BLOCK	ABL	94.89	95.14	95.63	97.99	98.13	98.21
	BFL	94.34	93.33	96.03	97.62	97.28	98.28
	CEL	95.40**	95.33*	96.54**	98.07	98.12	98.33
	DTL	95.45**	95.56*	96.60**	98.19***	98.22*	98.43*
PENOBSCOT	ABL	94.60	91.11	89.73	96.48	94.79	92.90
	BFL	89.96	84.60	82.48	94.99	96.37	89.66
	CEL	95.58	93.15	91.55	97.07	96.53	94.62
	DTL	96.05**	95.44***	94.07***	98.10***	97.24**	96.19***
PARIHAKA	ABL	92.97	93.40	93.38	97.03	97.11	97.19
	BFL	93.92	90.00	93.21	97.44***	96.68	97.38
	CEL	92.98	93.84	93.78	96.87	97.03	97.11
	DTL	94.56***	94.41***	94.47***	97.43***	97.51***	97.57**

4.4. Literature Comparison

In contrast to the previous experiment, which focused on optimizing the Distance Transform Loss parameters, this section investigates the influence of various cost functions on the performance of deep neural networks for the semantic segmentation of seismic data. Specifically, we trained three well-known architectures – DECONVNET (Noh et al., 2015), SEGNET (Badrinarayanan et al., 2017), and U-NET (Ronneberger et al., 2015) – using a supervised learning paradigm and evaluated their performance in labeling inline and crossline sections. The results, summarized in Tables 2 and 3, showcase the impact of employing four different cost functions across three geological volumes, resulting in a total of 36 unique performance measurements per evaluation metric. All cost functions were evaluated based on their ability to produce more accurate boundaries and enhance overall segmentation accuracy. Additionally, the variations in performance across different architectures and cost functions provide deeper insights into their relative capability.

This investigation is conducted to compare the performance improvements achieved by DTL when compared to the traditional CEL (Wang et al., 2021), as well as two domain-specific cost functions: BFL (Bokhovkin and Burnaev, 2019) and ABL (Wang et al., 2022). CEL, a pixel-wise variant of the long-established *log* loss, has long been the standard loss function applied in semantic segmentation tasks due to its simplicity and effectiveness in classifying individual pixels. In contrast, ABL and BFL are more recently developed loss functions that are specifically designed to improve boundary-aware segmentation. These methods were originally designed to penalize inaccuracies in the delineation of object boundaries in domains such as remote sensing and urban/indoor segmentation. Both ABL and BFL have demonstrated signifi-

cant success in these domains, achieving promising results by incorporating boundary-aware penalties that drive better performance on boundary-sensitive segmentation tasks. In this study, we systematically analyze these three baseline methods, as they represent both widely adopted approaches (like CEL) and more advanced, boundary-sensitive techniques (such as BFL and ABL).

Table 2 contrasts the performance of DTL with the baselines across multiple datasets and architectures using BFIS and MIOU metrics. In regard to BFIS, which emphasizes boundary precision, DTL exhibits clear improvements in nearly all settings. For the F3 BLOCK dataset, DTL achieves the highest BFIS scores across the three architectures with statistically significant gains, particularly in U-NET where it reaches 67.17. In the more heterogeneous PENOBSCOT, DTL again surpasses all baselines, most notably with SEGNET and U-NET, where the improvements are substantial: 73.57 for SEGNET and 65.96 for U-NET, both with high statistical confidence ($p \leq 0.001$). Similarly, in PARIHAKA, DTL maintains competitive performance, matching or exceeding baselines, with remarkable boundary improvements for SEGNET (72.90, $p \leq 0.001$). The advantages of DTL are further corroborated by the MIOU results, providing a global measure of segmentation accuracy across entire images. For F3 BLOCK, DTL again demonstrates superior performance across all architectures, achieving the highest values and reaching 91.50 with U-NET, which represents a significant improvement over CEL (+0.80) and BFL (+1.57). In the PENOBSCOT dataset, DTL achieves the strongest results overall, delivering 92.70 with DECONVNET, 91.33 with SEGNET, and 88.73 with U-NET, all statistically significant improvements over the competing baselines. Even though BFL achieves a slightly superior score on the PARIHAKA

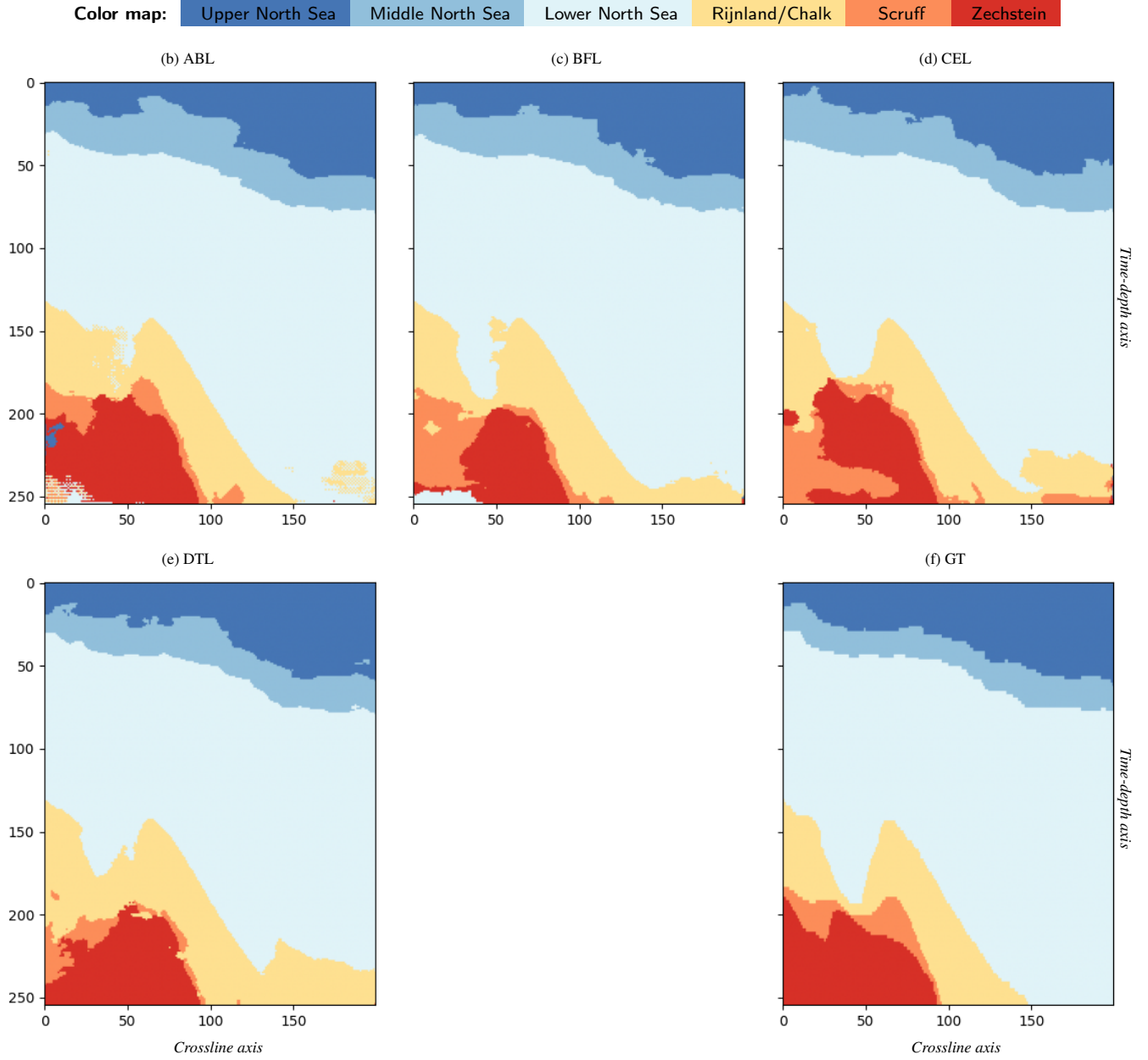


Figure 2: F3 BLOCK + U-NET: Qualitative results regarding inline#99 from test set #2 displaying its respective ground truth (GT) alongside the network reconstructions obtained with Active Boundary Loss (ABL), Boundary F1 Loss (BFL), Cross-Entropy Loss (CEL), and the proposed Distance Transform Loss (DTL).

dataset when coupled with DECONVNET, DTL still provides consistently higher or comparable MIOU, particularly with SEGNET (89.82, $p \leq 0.01$) and U-NET (89.36, $p \leq 0.01$).

Table 3 compares the four cost functions under the perspective of class and pixel-level accuracy. In terms of MCA, which is particularly relevant in scenarios of class imbalance, DTL consistently achieves the best or near-best performance across datasets and architectures. On the F3 BLOCK dataset, DTL reaches the highest MCA in all architectures, most notably 96.60 with U-NET as it is statistically tied with CEL (96.54) and markedly higher than ABL and BFL. In the more challenging PENOBSCOT dataset, where class im-

balance is more pronounced, DTL exhibits a clear superiority, reaching 96.05 with DECONVNET, 95.44 with SEGNET, and 94.07 with U-NET. Similarly, in the PARIHAKA dataset, DTL achieves the highest MCA across all architectures, with values above 94.40 in every case, outperforming both CEL and ABL. The proposed loss also delivers substantial improvements over BFL, which shows weaker performance, particularly when paired with SEGNET (90.00). PA further reinforces these trends as DTL reaches the highest PA with all architectures on F3 BLOCK, peaking at 98.43 for U-NET. DTL again provides a noticeable boost on PENOBSCOT, especially with U-NET, where it achieves 96.19 against CEL's

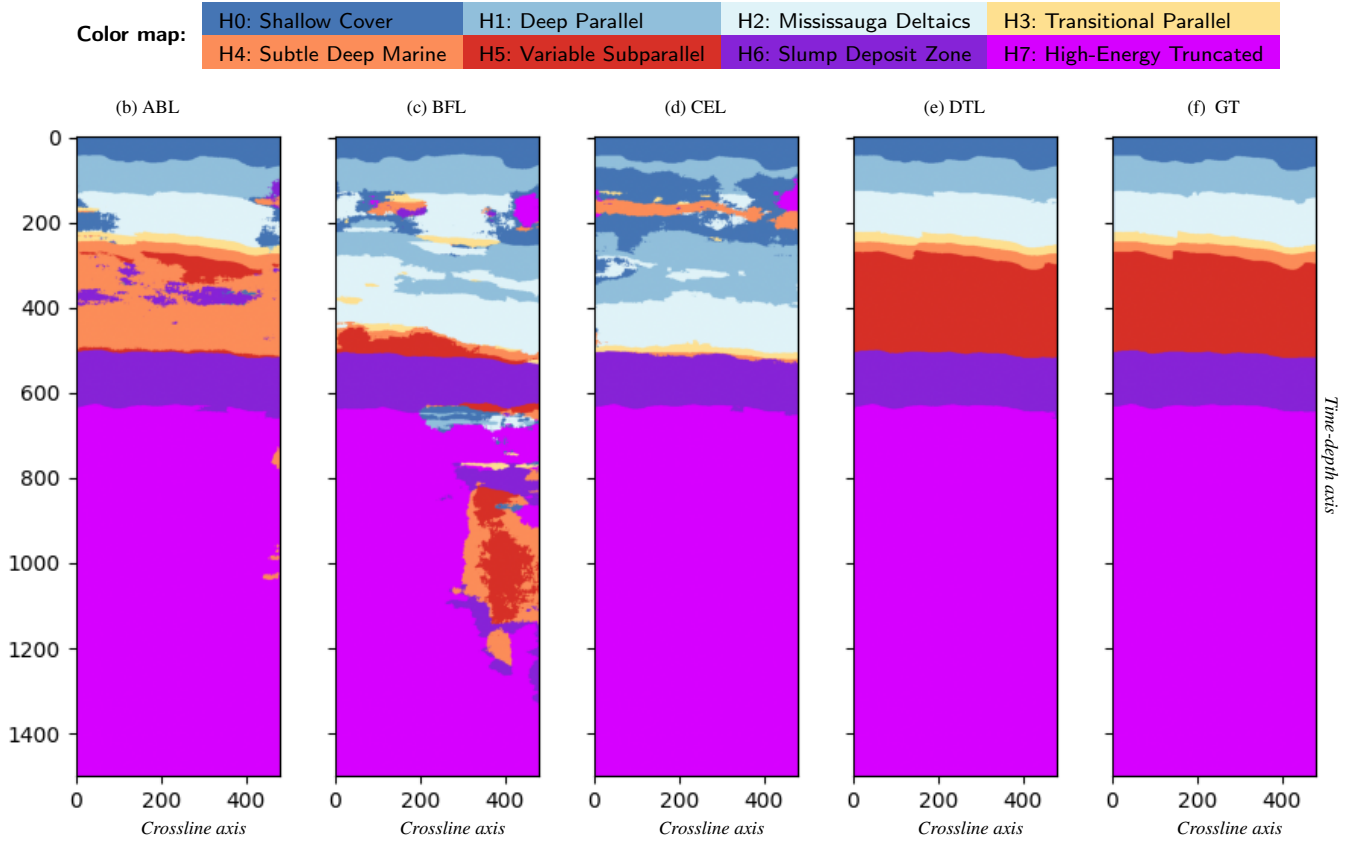


Figure 3: PENOBSCOT + SEGNET: Qualitative results regarding inLine #387 displaying its respective ground truth (GT) alongside network reconstructions obtained with Active Boundary Loss (ABL), Boundary F1 Loss (BFL), Cross-Entropy Loss (CEL), and the proposed Distance Transform Loss (DTL).

94.62 and ABL's 92.90, reflecting gains of up to +6.53 points in pixel-level correctness. Even on PARIHAKA, where the baselines achieve strong PA values, DTL holds statistically significant improvements with SEGNET with 97.57 in U-NET and 97.51 in SEGNET. These results suggest that DTL not only ensures class-level fairness but also enhances overall pixel-level correctness, even in difficult settings.

In a previous work, Alaudah et al. (2019) utilized DECONVNET in conjunction with the CEL to perform semantic segmentation on the Dutch F3 BLOCK dataset, presenting results with a reported Pixel Accuracy of 90.5% and Mean Class Accuracy of 87.7%. Even though their work demonstrated the potential of deep learning in seismic segmentation, their results were outperformed by the proposed DTL, which achieved a remarkable Pixel Accuracy of 98.19% and Mean Class Accuracy of 95.45%. This substantial improvement brings light to the robustness of DTL in addressing the inherent challenges of boundary delineation and spatial localization, particularly in scenarios with complex subsurface structures. We refer readers to the additional quantitative metrics for semantic segmentation and further detailed information provided in the supplementary material. The auxiliary content presents additional experiments and evaluation metrics that are typically employed in semantic segmentation research, enabling a rigorous comparison of our findings with

those reported in the literature.

Curiously, ABL and BFL functions were not able to match the performance of both CEL and DTL on the PENOBSCOT benchmark. We believe that ABL and BFL struggled to assign high probability scores to pixels located on class boundaries due to the lack of continuity in many horizons, a prevalent issue also observed in the training data of the Nova Scotia dataset (Chevitarese et al., 2018). This discontinuity possibly undermines the ability of ABL and BFL to effectively delineate class boundaries, especially in datasets where horizon structures are fragmented or poorly defined. Apparently, this limitation did not hinder the performance of DTL, which outperformed all three baselines. DTL also achieved competitive results on the PARIHAKA dataset, surpassing ABL and CEL under both evaluation metrics. This is the only domain where BFL marginally outperformed DTL in terms of the boundary F1 score (BF1s), although DTL still maintains superior performance in the mean Intersection over Union (mIoU) assessment.

Unlike more complex approaches such as ABL and BFL, DTL achieves these results with notable simplicity, emphasizing the balance between efficiency and performance. Its ability to generalize across networks and benchmarks highlights its potential as a reliable loss function for semantic segmentation in seismic interpretation tasks. It is worth em-

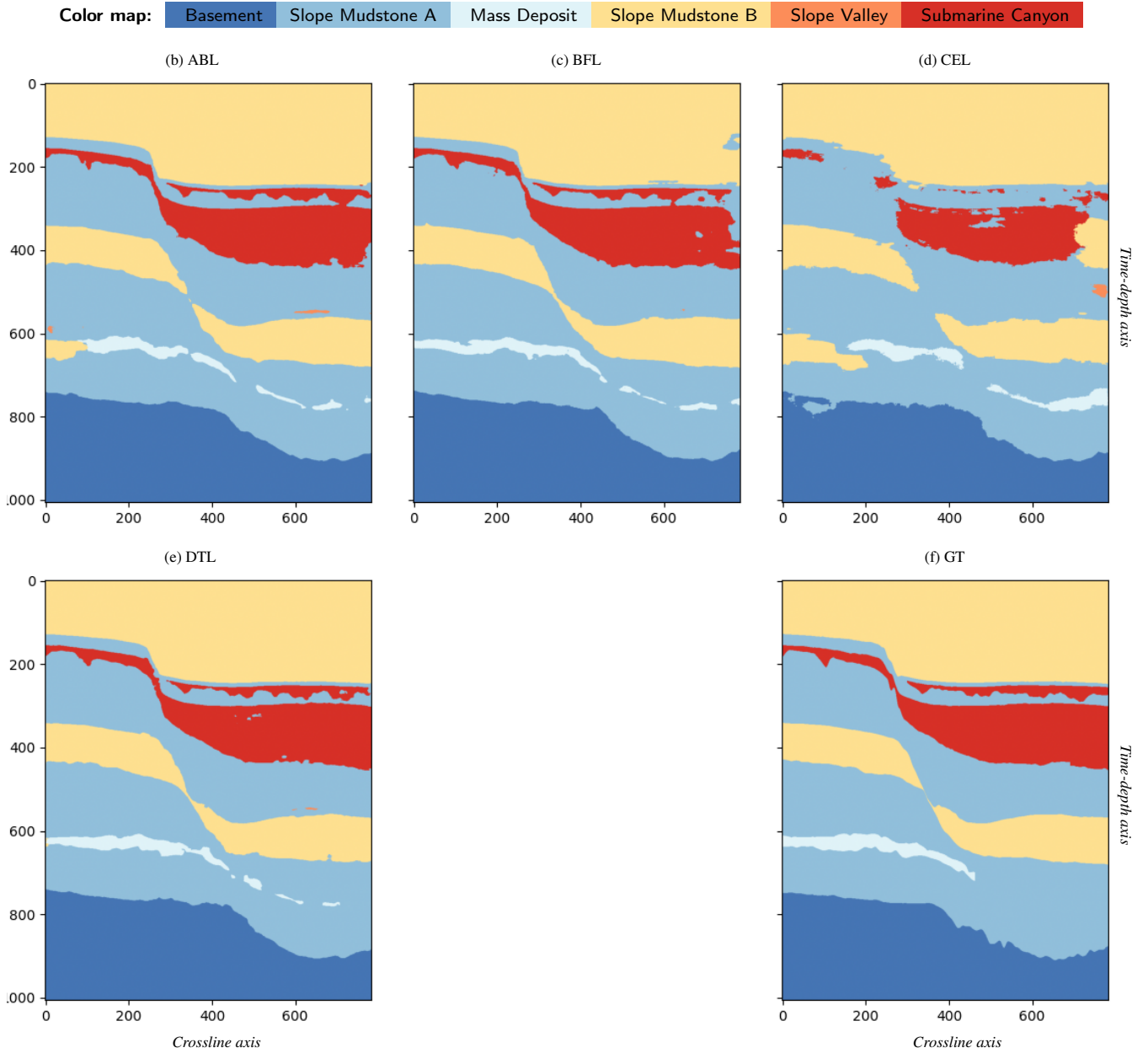


Figure 4: PARIHAKA + DECONVNET: Qualitative results regarding inline #413 displaying its respective ground truth (GT) alongside four network reconstructions obtained with Active Boundary Loss (ABL), Boundary F1 Loss (BFL), Cross-Entropy Loss (CEL), and the proposed Distance Transform Loss (DTL).

phasizing that we do not proclaim DTL as a universal replacement for traditional losses, but rather as a complementary strategy – particularly beneficial in delineating stratigraphic structures where boundary precision and shape continuity are critical. To ensure a fair and transparent comparison, all experiments have been conducted under controlled settings with identical network architectures and data splits, isolating the effect of the loss function itself. Moreover, we believe that this consistent quantitative improvement across multiple datasets is not only a reflection of DTL’s robustness but also a compelling argument for its broader applicability in seismic interpretation tasks.

4.5. Qualitative Evaluation

Figure 2 illustrates the segmentation results obtained with a U-NET trained on the F3 BLOCK dataset. CEL produces reasonable predictions but suffers from scattered misclassifications within the Lower North Sea and Rijnland/Chalk intervals, where transitions appear blurred. BFL improves boundary delineation between Upper and Middle North Sea units but introduces inconsistencies in the Scruff and Zechstein formations, fragmenting their continuity. ABL better preserves the slope geometries of the Rijnland/Chalk region and shows improved facies transitions, yet still fails to capture the structural coherence of the Zechstein facies in the south-

ern section. In contrast, DTL produces cleaner and more geologically consistent results across all units: the wedge-shaped Rijnland/Chalk interval is sharply defined, Upper and Middle North Sea layers remain continuous, and the Scruff–Zechstein boundary is more faithfully reconstructed. Overall, DTL outperforms other losses by simultaneously enhancing facies coherence and maintaining sharper boundary transitions across structurally complex zones of the F3 BLOCK volume.

Figure 3 presents segmentation outcomes from a SEGNET-based model trained on the PENOBSCOT dataset. CEL achieves good results in the Shallow Cover (H0) and High-Energy Truncated (H7) facies but struggles in the Subtle Deep Marine (H4) and Variable Subparallel (H5) zones, where transitions become diffuse and vertically smeared. BFL sharpens horizon boundaries, particularly between Deep Parallel (H1) and Mississauga Deltaic (H2) facies, but generates noisy artifacts within the Slump Deposit Zone (H6), breaking structural continuity. Similarly, ABL improves boundary localization in transitional facies (H3) yet underrepresents smooth layering within deeper regions, leading to geologically implausible segmentation. DTL, however, preserves both vertical layering and lateral continuity across facies: it cleanly separates Shallow Cover from Deep Parallel zones, maintains coherent H4 and H5 intervals, and delivers consistent reconstruction of the H6 slump region without spurious fragmentation. These results confirm that DTL produces the most geologically coherent and stratigraphically reliable segmentation in the PENOBSCOT volume.

Lastly, Figure 4 shows qualitative results from a DECONVNET applied to the PARIHAKA dataset. CEL generates plausible facies assignments but produces poor transitions between Slope Mudstone A, Slope Mudstone B, and Mass Deposit units, leading to blurred contacts and over-segmentation. BFL sharpens the separation of Submarine Canyon deposits from adjacent mudstone facies but introduces boundary artifacts along slope geometries. ABL delivers precise delineation of curved and dipping features, especially at the interface between Slope Valley and Submarine Canyon, but misses some stratigraphic contrasts within Mass Deposits. DTL achieves the most balanced performance: boundaries between Slope Mudstone facies are well preserved, the Submarine Canyon is sharply defined, and reflector continuity is maintained across Mass Deposit and Basement intervals. These results highlight that while boundary-aware methods consistently outperform CEL, DTL stands out by combining sharp boundary delineation with structural continuity, yielding the most geologically consistent segmentation for the PARIHAKA dataset.

Although DTL may not always yield the visually sharpest boundaries in every instance, it regularly delivers robust segmentation performance across diverse geological regions by effectively capturing subtle structural variations and maintaining coherence in class assignments. This balance between boundary delineation and global consistency highlights its strength in producing geologically plausible segmentations, even in challenging and heterogeneous subsurface scenarios.

Furthermore, the adaptability and resilience of DTL across varying stratigraphic settings underscore its suitability for tasks that demand both precise boundary reconstruction and reliable facies classification. The qualitative results thus reinforce the importance of carefully selecting loss functions that align with the specific objectives of seismic segmentation, ultimately contributing to more accurate geological interpretations and improved efficiency in resource exploration and decision-making workflows.

4.6. Advantages and Limitations

DTL introduces some advantages tailored to the challenges of seismic data interpretation. Unlike standard pixel-wise cost functions, it leverages a distance transform map to impose greater penalties on misclassifications occurring near class boundaries. DTL makes the neural network “more sensitive” to structural discontinuities and stratigraphic transitions, which are essential for the accurate interpretation of seismic facies and geological features such as faults, salt bodies, and channel systems. Additionally, DTL is easily integrated with traditional loss functions, such as CEL, requiring only minimal adjustments to existing training pipelines. Its modular nature allows it to complement boundary-agnostic supervision strategies, leading to enhanced segmentation performance with minimal architectural modifications.

Experiments across multiple seismic datasets, pointed out in Tables 2 and 3, demonstrate that DTL consistently improves boundary scores (BFIS), indicating better alignment with class borders of the ground-truth. This is particularly beneficial in marine and subsurface seismic surveys where continuous reflectors and subtle stratigraphic changes must be captured accurately. Furthermore, DTL contributes to more geologically plausible segmentation outputs by preserving reflector continuity and avoiding over-segmentation of layered structures. Its effectiveness remains consistent across distinct architectures (e.g., DECONVNET, SEGNET, U-NET), reinforcing its generalizability and adaptability to different model backbones. The grid-based use of the L_1 Manhattan distance within the distance transform also brings a computational advantage, leveraging efficient two-pass algorithms for runtime feasibility during training.

Despite these strengths, DTL has certain limitations that merit discussion. For instance, its reliance on accurate boundary information from the ground-truth annotations means that performance may degrade in settings where labels are noisy, incomplete, or ambiguous — conditions that are not uncommon in real-world seismic datasets. DTL introduces a computational overhead relative to simpler loss functions due to the need to compute distance maps at each training iteration, although this cost remains tractable with modern graphics processing unit resources. Another drawback is that while DTL enhances boundary precision, it may slightly compromise performance in homogeneous regions where boundary guidance is less critical. Lastly, as a supervised method, DTL is inherently limited by the availability of labeled data, which is often scarce in geoscience. While self-supervised and semi-supervised paradigms can reduce the

need for large labeled datasets during pretraining or representation learning, the supervision provided by DTL during the fine-tuning stage requires high-quality annotations to yield its intended boundary-aware benefits. Consequently, the annotation burden is not entirely eliminated but rather shifted toward ensuring precision in the labeled data used for problem specialization.

5. Conclusions

In this work, we introduce the Distance Transform Loss (DTL), a cost function specifically designed to guide deep networks into segmenting more accurate inter-class boundaries. DTL works in an intuitive and straightforward manner by assigning higher penalty scores the farther a predicted boundary lies from the ground truth.

The proposed approach contrasts with most existing methods in the literature, which often require complex calculations or hyper-parameter tuning tailored to specific datasets. In opposition to many segmentation methods that fine-tune hyper-parameters on each dataset, DTL parameters were selected using only a single domain, the F3 BLOCK dataset, and then applied directly to two other distinct seismic datasets, PENOBSCOT and PARIHAKA. Despite the cross-domain evaluation, experiments have demonstrated versatility and effectiveness of DTL, delivering either superior or competitive results across multiple benchmarks. In fact, the quantitative analysis reveals DTL standing out in 5 out of 6 scenarios, showcasing its ability to generalize well beyond the initial training domain. Chances are the proposed approach could have obtained even higher performance on PENOBSCOT and PARIHAKA datasets if we had considered their validation sets when picking the best parameters.

In conclusion, the qualitative results underscore the impact of incorporating DTL into the training process of deep networks. Not only does the use of DTL enhance the visual accuracy of boundary delineation in seismic data but it also improves the overall classification performance in semantic segmentation tasks. The findings suggest that DTL is a valuable contribution to the computer and earth science communities, producing artificial neural networks with robustness and adaptability in diverse geophysical contexts.

As a direction for future work, we propose exploring the integration of self-supervised learning pretraining with boundary-aware loss functions like DTL during downstream fine-tuning. This combined strategy holds potential to leverage large volumes of unlabeled seismic data for robust feature learning while enhancing boundary delineation through task-specific optimization. Such an approach could improve segmentation performance in scenarios with limited annotated data, further extending the practical applicability of DTL in real-world seismic interpretation tasks.

Acknowledgements

The authors would like to thank Brazilian Petroleum Corporation (Petróleo Brasileiro S.A., PETROBRAS) for the technical and financial support through its cooperation agreement

with the Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, Minas Gerais, Brazil.

Computer Code Availability

The source code containing the Distance Transform Loss has been developed by Mr. Rafael Henrique Vareto, currently pursuing a Ph.D. degree at the Federal University of Minas Gerais, Brazil (Email: rafael@vareto.com.br). The approach is publicly available as a Python Package through the PIP installer program (`pip install segloss`), an effortless way to integrate our method into new projects. It can also be accessed in the following GitHub repository: <https://github.com/rafaelvareto/segmentation-loss>. The module also encompasses cost functions used as baselines and evaluated in the experimental section, namely Active Boundary Loss, Boundary Loss and Cross-Entropy Loss, which have been previously published to GitHub by their respective authors Alaudah et al. (2019); Bokhovkin and Burnaev (2019); Wang et al. (2022).

References

- Alaudah, Y., Michałowicz, P., Alfarraj, M., AlRegib, G., 2019. A machine-learning benchmark for facies classification. *Interpretation* 7, SE175–SE187. <https://zenodo.org/record/3755060>.
- An, Y., Guo, J., Ye, Q., Childs, C., Walsh, J., Dong, R., 2021. Deep convolutional neural network for automatic fault recognition from 3d seismic datasets. *Computers & Geosciences* 153, 104776.
- Audebert, N., Boulch, A., Le Saux, B., Lefèvre, S., 2019. Distance transform regression for spatially-aware deep semantic segmentation. *Computer Vision and Image Understanding* 189, 102809.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 39, 2481–2495.
- Baroni, L., Silva, R.M., Ferreira, R.S., Civitarese, D., Szwarcman, D., Brazil, E.V., 2019. Penobscot dataset: Fostering machine learning development for seismic interpretation. *arXiv:1903.12060*. <https://zenodo.org/record/3924682>.
- Bokhovkin, A., Burnaev, E., 2019. Boundary loss for remote sensing imagery semantic segmentation, in: *International Symposium on Neural Networks (ISNN)*, Springer. pp. 388–401.
- Borse, S., Wang, Y., Zhang, Y., Porikli, F., 2021. Inverseform: A loss function for structured boundary-aware segmentation, in: *Computer Vision and Pattern Recognition (CVPR)*, IEEE/CVF. pp. 5901–5911.
- Caliva, F., Iriando, C., Martinez, A.M., Majumdar, S., Padoia, V., 2019. Distance map loss penalty term for semantic segmentation. *Medical Imaging with Deep Learning (MIDL)*.
- Chen, L., Wu, W., Fu, C., Han, X., Zhang, Y., 2020. Weakly supervised semantic segmentation with boundary exploration, in: *European Conference on Computer Vision (ECCV)*, Springer. pp. 347–362.
- Civitaresse, D.S., Szwarcman, D., e Silva, R.G., Brazil, E.V., 2018. Deep learning applied to seismic facies classification: A methodology for training, in: *Saint Petersburg International Conference and Exhibition, European Association of Geoscientists & Engineers*. pp. 1–5.
- Csurka, G., Larlus, D., Perronnin, F., Meylan, F., 2013. What is a good evaluation measure for semantic segmentation?, in: *British Machine Vision Conference (BMVC)*, IET. pp. 10–5244.
- Fabbri, R., Costa, L.D.F., Torelli, J.C., Bruno, O.M., 2008. 2d euclidean distance transform algorithms: A comparative survey. *ACM Computing Science Surveys (CSUR)* 40, 1–44.
- Felzenszwalb, P.F., Huttenlocher, D.P., 2012. Distance transforms of sampled functions. *Theory of Computing* 8, 415–428.
- Gerl, S., Paetzold, J.C., He, H., Ezhov, I., Shit, S., Kofler, F., Bayat, A., Tetteh, G., Ntziachristos, V., Menze, B., 2020. A distance-based loss for

- smooth and continuous skin layer segmentation in optoacoustic images, in: *Medical Image Computing and Computer Assisted Intervention*, Springer. pp. 309–319.
- Gonzalez, R.C., Woods, R.E., 2018. *Digital Image Processing*. 4 ed., Pearson Education Limited.
- Haggerty, R., Sun, J., Yu, H., Li, Y., 2023. Application of machine learning in groundwater quality modeling-a comprehensive review. *Water Research* 233, 119745.
- Inc, C.U., 2020. Annual meeting machine learning interpretation workshop. <https://public.3.basecamp.com/p/JyT276MM7krjYrMoLqLQ6xST>.
- Jadon, S., 2020. A survey of loss functions for semantic segmentation, in: *Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, IEEE. pp. 1–7.
- Kaur, H., Pham, N., Fomel, S., Geng, Z., Decker, L., Gremillion, B., Jervis, M., Abma, R., Gao, S., 2023. A deep learning framework for seismic facies classification. *Interpretation* 11, T107–T116.
- Lee, D., Ovanger, O., Eidsvik, J., Aune, E., Skauvold, J., Hauge, R., 2024. Latent diffusion model for conditional reservoir facies generation. *Computers & Geosciences*, 105750.
- Li, F., Zhou, H., Wang, Z., Wu, X., 2020. Addcnn: An attention-based deep dilated convolutional neural network for seismic facies analysis with interpretable spatial-spectral maps. *Transactions on Geoscience and Remote Sensing (TGRS)* 59, 1733–1744.
- Liang, Z., Wang, T., Zhang, X., Sun, J., Shen, J., 2022. Tree energy loss: Towards sparsely annotated semantic segmentation, in: *Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE/CVF. pp. 16907–16916.
- Martin, D.R., Fowlkes, C.C., Malik, J., 2004. Learning to detect natural image boundaries using local brightness, color, and texture cues. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 26, 530–549.
- Monteiro, B.A., Canguçu, G.L., Jorge, L.M., Vareto, R.H., Oliveira, B.S., Silva, T.H., Lima, L.A., Machado, A.M., Schwartz, W.R., Vaz-de Melo, P.O., 2024. Literature review on deep learning for the segmentation of seismic images. *Earth-Science Reviews*, 1–32.
- Monteiro, B.A., Oliveira, H., dos Santos, J.A., 2022. Self-supervised learning for seismic image segmentation from few-labeled samples. *Geoscience and Remote Sensing Letters (GRSL)* 19, 1–5.
- Noh, H., Hong, S., Han, B., 2015. Learning deconvolution network for semantic segmentation, in: *International Conference on Computer Vision (ICCV)*, IEEE/CVF. pp. 1520–1528.
- Paszke, A., Gross, S., Massa, F., Lerer, A., et al., 2019. Pytorch: An imperative style, high-performance deep learning library, in: *Advances in Neural Information Processing Systems (NeurIPS)*, Curran. pp. 1–12.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer. pp. 234–241.
- Saad, O.M., Chen, W., Zhang, F., Yang, L., Zhou, X., Chen, Y., 2022. Self-attention fully convolutional densenets for automatic salt segmentation. *Transactions on Neural Networks and Learning Systems* 34, 3415–3428.
- Sun, J., Liu, Y., Cui, J., He, H., 2022. Deep learning-based methods for natural hazard named entity recognition. *Scientific reports* 12, 4598.
- Taghanaki, S.A., Abhishek, K., Cohen, J.P., Cohen-Adad, J., Hamarneh, G., 2021. Deep semantic segmentation of natural and medical images: a review. *Artificial intelligence review* 54, 137–178.
- Tao, H., Hameed, M.M., Marhoon, H.A., Zounemat-Kermani, M., Heddam, S., Kim, S., Sulaiman, S.O., Tan, M.L., Sa'adi, Z., Mehr, A.D., et al., 2022. Groundwater level prediction using machine learning models: A comprehensive review. *Neurocomputing* 489, 271–308.
- Tian, J., Mithun, N.C., Seymour, Z., Chiu, H.P., Kira, Z., 2022. Striking the right balance: Recall loss for semantic segmentation, in: *International Conference on Robotics and Automation (ICRA)*, IEEE. pp. 5063–5069.
- Vareto, R.H., Günther, M., Schwartz, W.R., 2023. Open-set face recognition with neural ensemble, maximal entropy loss and feature augmentation, in: *Conference on Graphics, Patterns and Images (SIBGRAPI)*, IEEE. pp. 1–6.
- Vareto, R.H., Linghu, Y., Boulton, T.E., Schwartz, W.R., Günther, M., 2024. Open-set face recognition with maximal entropy and objectosphere loss. *Image and Vision Computing* 141, 104862.
- Wang, C., Zhang, Y., Cui, M., Ren, P., Yang, Y., Xie, X., Hua, X.S., Bao, H., Xu, W., 2022. Active boundary loss for semantic segmentation, in: *Conference on Artificial Intelligence (CAI)*, AAAI. pp. 2397–2405.
- Wang, W., Zhou, T., Yu, F., Dai, J., Konukoglu, E., Van Gool, L., 2021. Exploring cross-image pixel contrast for semantic segmentation, in: *International Conference on Computer Vision (ICCV)*, IEEE/CVF. pp. 7303–7313.
- Wang, X., Ma, H., Chen, X., You, S., 2017. Edge preserving and multi-scale contextual neural network for salient object detection. *Transactions on Image Processing (TIP)* 27, 121–134.
- Xu, Z., Ma, W., Lin, P., Shi, H., Pan, D., Liu, T., 2021. Deep learning of rock images for intelligent lithology identification. *Computers & Geosciences* 154, 104799.
- Xue, Y., Tang, H., Qiao, Z., Gong, G., Yin, Y., Qian, Z., Huang, C., Fan, W., Huang, X., 2020. Shape-aware organ segmentation by predicting signed distance maps, in: *Proceedings of the AAAI conference on artificial intelligence*, pp. 12565–12572.
- Yang, L., Fomel, S., Wang, S., Chen, X., Saad, O.M., Chen, Y., 2024. Salt3dnet: A self-supervised learning framework for 3d salt segmentation. *Transactions on Geoscience and Remote Sensing (TGRS)*.
- Zhang, H., Chen, T., Liu, Y., Zhang, Y., Liu, J., 2021. Automatic seismic facies interpretation using supervised deep learning. *Geophysics* 86, IM15–IM33.
- Zhao, S., Wang, Y., Yang, Z., Cai, D., 2019. Region mutual information loss for semantic segmentation. *Advances in Neural Information Processing Systems (NIPS)* 32.
- Zhu, X., Zhou, H., Yang, C., Shi, J., Lin, D., 2018. Penalizing top performers: Conservative loss for semantic segmentation adaptation, in: *European Conference on Computer Vision (ECCV)*, pp. 568–583.