Bidirectional Diffeomorphic Temporal Alignment Network Application to Distributed Acoustic Sensing

Konstantinos Vassalos[1], Cédric Richard[1]
[1]Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, France
Email: kon.vassalos@gmail.com

***Abstract*—Temporal alignment remains a key issue in multisensor time-series data analysis, particularly when signals are affected by local temporal distortions and mixed shifts of opposite signs. In this paper, we propose a bidirectional time-warping framework that extends regularization-free Diffeomorphic Temporal Alignment Networks to signals that contains both forward and backward shifts. The method introduces a bidirectional formulation that evaluates alignment in both directions and uses a soft gating mechanism for selecting the most consistent alignment locally. This enables the model to handle signals with varying temporal shifts within a same a same temporal segment. We illustrate the approach on Distributed Acoustic Sensing data collected for traffic monitoring on two-way highways, where overlapping vehicle-induced signatures from opposite directions introduce significant alignment challenges. Experimental results show improved alignment consistency and clearer separation of opposing traffic patterns.**

I. Introduction

Temporal alignment is a fundamental problem in time-series data analysis, aiming to compensate for temporal variations so that signals with similar structure can be aligned in a common temporal reference frame. This arises in applications such as handwriting recognition [1], speech recognition [2], and human action recognition [3]. In these settings, alignment methods must handle both global shifts and local distortions, including stretching and compression. Time warping is a standard approach that uses monotonic transformations to match corresponding events across signals [4]. Recent learning-based methods [5] provide greater flexibility by estimating warps from data but generally assume preserved temporal order, which limits their effectiveness when signals contain patterns with local shifts of opposite signs [6].

The aim of this paper is to address bidirectional temporal alignment by formulating it as a time-warping problem. To this end, we introduce a bidirectional extension of the Regularization-Free Temporal Alignment Net (RF-DTAN) [7]. The method evaluates forward and time-reversed warping hypotheses and uses a soft gating mechanism to select the most consistent alignment locally, enabling the handling of mixed temporal orientations.

To illustrate the proposed approach, we conduct experiments on Distributed Acoustic Sensing (DAS) data, obtained from a fiber-optic sensing technology that turns a standard optical fiber into a continuous array of sensors. A laser pulse is transmitted along the fiber, and small portions of light are backscattered due to microscopic imperfections in the fiber. The time delay localizes each measurement, while changes in the backscattered signal reflect local vibrations or strain, enabling continuous monitoring over long distances. DAS has been successfully applied to seismic monitoring [8], pipeline leak detection [9], power grid monitoring [10] and traffic speed estimation [11]. In two way highway scenarios, overlapping vehicle signatures from opposite directions make it difficult to separate individual events and align them consistently, motivating the use of bidirectional time-warping; see Figure 1.

The rest of the paper is organized as follows. The formulation of the problem and the proposed method are described in Section II. Section III describes the experimental setup. Section IV reports the experimental results, and Section V concludes the paper.

II. Methodology

To highlight the limitations of existing temporal alignment methods, we first analyze the regularization-free approach proposed in [7], which performs time warping through monotonic transformations. Although effective, this formulation cannot handle signals containing bidirectional temporal patterns with opposite-sign time shifts due to its order-preserving nature. To address this limitation, we extend the family of admissible transformations to include reversed warps, enabling the alignment of signals with opposite temporal alignment directions. We further introduce a gating mechanism that selects, at each time point, the most consistent alignment hypothesis. The final alignment is obtained by combining the selected warps across time.

A. RF-DTAN background

RF-DTAN is a learning-based framework for joint temporal alignment and time series averaging. Let $x_k = (x_k(t))_{t=1}^n$ denote the k-th input signal of length n defined on a temporal domain $\Omega = \{1, \dots, n\}$. In the multi-class setting, let $y_k \in \{1, \dots, M\}$ be the class label of x_k, and let N_m denote the number of signals in class m. The goal is to estimate a temporal warp $T^{\theta_k} \in \mathcal{T}$, where $\mathcal{T}$ is a family of diffeomorphic transformations. Every $T \in \mathcal{T}$ is a warp function $T : \Omega \to \mathbb{R}$. In our implementation, the transformation family $\mathcal{T}$ is chosen as the Continuous Piecewise Affine Basis (CPAB) class [12], ensuring that each T^{θ_k} is smooth and invertible. Each warp is parameterized by a warping parameter θ_k which is predicted by a Temporal Transformation Network (TTN) which consists of three modules. The first module is the *localization network* which takes as input a batch of sequences $(x_k(t))_{t=1}^n$ and

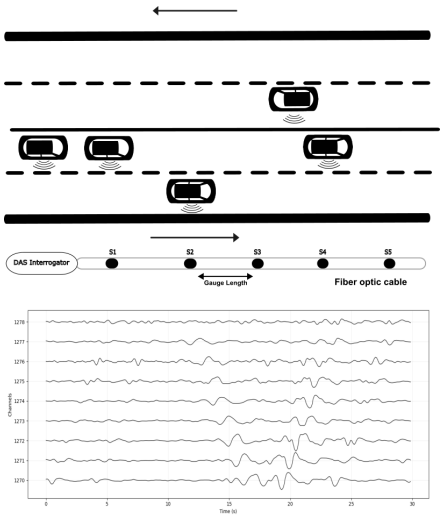

Fig. 1: At the top is the schematic overview of the problem geometry. It consists of the two way highway, the fiber optic cable which is parallel to the road and the DAS interrogator connected to the fiber. DAS channels are denoted by S_1, S_2, etc. At the bottom is depicted an example of DAS data with 9 DAS channels and 30 seconds duration. Each channel corresponds to a sensing location along the fiber, as shown in the schematic above.

predicts the corresponding warping parameters $(\theta_k)_{k=1}^n$ by defining a piecewise-affine velocity field over a fixed partition of the time axis. The second is a *grid generator* which creates an evenly spaced grid $G \subset \Omega$ of points which are then warped by T^{θ_k}. Lastly, a *sampler* computes the warped signal $\hat{x}_k = x_k \circ T^{\theta_k}$ by interpolation. Because the transformation is differentiable with respect to θ, the whole alignment module can be trained end to end with backpropagation learning algorithm.

Unlike earlier approaches [13], RF-DTAN does not include an explicit regularization term on the warps to penalize lack of smoothness or large deviations. Instead, it relies on the loss of the Inverse Consistency Averaging Error (ICAE). For each class m, a centroid signal μ_m is first computed from the aligned signals:

$$\mu_m = \frac{1}{N_m} \sum_{k:y_k=m} x_k \circ T^{\theta_k}. \tag{1}$$

Consistency is then enforced by mapping the centroid back to each original signal using the reverse CPAB warp, leading to the loss:

$$\mathcal{L}_{\text{ICAE}}(\theta) = \sum_{m=1}^{M} \frac{1}{N_m} \sum_{k:y_k=m} \left\| \mu_m \circ T^{-\theta_k} - x_k \right\|_{\ell_2}^2. \tag{2}$$

The loss is minimized when the average sequence is both a minimizer of the variance and consistent with its class. This formulation encourages coherent alignment across signals without requiring dataset-specific regularization. However, the solution is restricted to order-preserving transformations.

B. Pairwise RF-DTAN

In our setting, we use a *pairwise* version of RF-DTAN rather than the original class-wise formulation. Let $x_k = (x_k(t))_{t=1}^n$ denote the k-th input signal of length n defined over the temporal domain $\Omega = \{1, \ldots, n\}$. Let $\mathcal{T}$ be the family of warping CPAB transformations used by RF-DTAN. We adopt the single-class setting and train the model on adjacent pairs (x_k, x_{k+1}). For each pair, the localization network estimates a warp parameter $\theta_k = f_{\text{loc}}(x_k)$, yielding the forward-warped signal

$$\hat{x}_k = x_k \circ T^{\theta_k}. \tag{3}$$

The pairwise centroid is then computed as

$$\mu_k = \frac{1}{2} \left(\hat{x}_k + x_{k+1} \right). \tag{4}$$

To enforce inverse consistency, we also define the backward-warped centroid

$$\hat{\mu}_k = \mu_k \circ T^{-\theta_k}. \tag{5}$$

The pairwise loss is given by

$$\mathcal{L}(\theta) = \sum_{k=1}^{N_c-1} \left[(\hat{\mu}_k - x_k)^2 + (\mu_k - x_{k+1})^2 \right], \tag{6}$$

where N_c denotes the total number of channels. The batch loss is obtained by accumulating the pairwise loss over all consecutive channel pairs in the batch. By learning a separate warp for each input signal, this formulation avoids the need for a global class centroid and is therefore better suited to weakly correlated signals. Nevertheless, it remains unable to handle cases in which events with opposite temporal directions occur within the same signal, because a single monotonic warp cannot align both directions at once.

C. Limitation in bidirectional temporal data and extension

In this section, we briefly explain why RF-DTAN cannot directly handle opposite-sign time shifts and describe how the admissible CPAB warp family can be extended. Each warp is parameterized by θ and is order-preserving, i.e.,

$$t_1 < t_2 \;\Rightarrow\; T^\theta(t_1) < T^\theta(t_2), \qquad T^\theta \in \mathcal{T}. \tag{7}$$

Thus, corresponding events preserve their temporal order after warping. To represent a signal with the opposite temporal shifts, we introduce the time-reversal operator $\rho : \Omega \to \Omega$, defined as

$$\rho(t) = n + 1 - t, \qquad t \in \Omega. \tag{8}$$

If a signal is transformed by the time-reversal operator ρ, then two landmarks occurring at times $a < b$ in the original signal appear at $\rho(b) < \rho(a)$ in the reversed signal, thus reversing their temporal order. Because every transformation $T^\theta \in \mathcal{T}$ is order-preserving, the same warp family cannot restore a common temporal ordering when forward and reversed patterns coexist within the same temporal window. Consequently, a single monotonic warp can align one temporal orientation, but cannot simultaneously align events whose ordering is reversed.

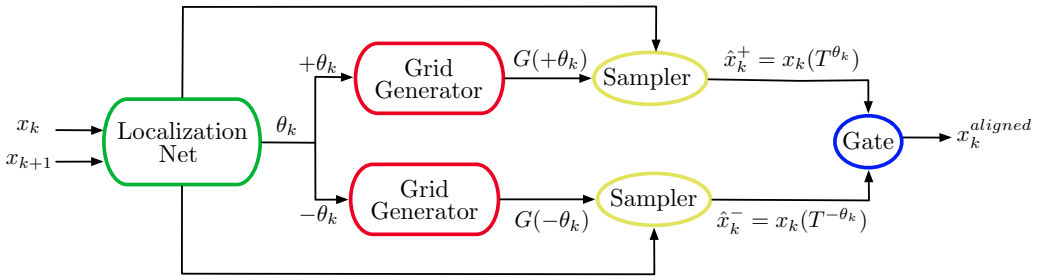

Fig. 2: Architecture of the proposed method.

To address this limitation, we extend the family of admissible transformations to include reversed warps and introduce a bidirectional formulation; see Figure 2:

$$\mathcal{T}^{\pm} = \mathcal{T} \cup (\rho \circ \mathcal{T}). \tag{9}$$

Thus, each admissible transformation is either a forward warp

$$T^{+\theta}(t) = T^{\theta}(t), \tag{10}$$

or a backward warp

$$T^{-\theta}(t) = (\rho \circ T^{\theta})(t), \qquad T^{\theta} \in \mathcal{T}. \tag{11}$$

With this extended warp family, a signal that differs from another only by a global time reversal can be aligned by composition with ρ, while the unreversed signal is handled by the original warp family. Although this extension can account for globally reversed patterns, it remains insufficient when opposite-sign temporal shifts coexist within the same analysis window. In this setting, no single global transformation from $\mathcal{T}^{\pm}$ can minimize the alignment error across the entire window, which motivates a local bidirectional alignment formulation.

D. Proposed bidirectional RF-DTAN (BRF-DTAN) method

To handle locally varying temporal directions, we adopt a pairwise formulation shown in Figure 2. Let $x_k = (x_k(t))_{t=1}^{n}$ be defined on $\Omega = \{1, \ldots, n\}$ with $k \in \{1, \ldots, N_{\text{channels}}\}$. For each pair (x_k, x_{k+1}), the localization network estimates θ_k, defining a warp $T^{\theta_k} \in \mathcal{T}$.

We construct two alignment hypotheses using the extended family $\mathcal{T}^{\pm}$:

$$\hat{x}_k^+ = x_k \circ T^{\theta_k}, \qquad \hat{x}_k^- = x_k \circ T^{-\theta_k}. \tag{12}$$

Each hypothesis defines a pairwise centroid:

$$\mu_k^+(t) = \frac{\hat{x}_k^+(t) + x_{k+1}(t)}{2}, \qquad \mu_k^-(t) = \frac{\hat{x}_k^-(t) + x_{k+1}(t)}{2}. \tag{13}$$

The corresponding alignment errors are defined as

$$\begin{aligned} e_k^+(t) &= \left[(\mu_k^+ \circ T^{-\theta_k})(t) - x_k(t)\right]^2 + \left[\mu_k^+(t) - x_{k+1}(t)\right]^2 \\ e_k^-(t) &= \left[(\mu_k^- \circ T^{+\theta_k})(t) - x_k(t)\right]^2 + \left[\mu_k^-(t) - x_{k+1}(t)\right]^2 \end{aligned} \tag{14}$$

In these error expressions, the first term applies the estimated warp to the centroid and evaluates its reconstruction of x_k, whereas the second term measures the discrepancy between the centroid and the adjacent time series x_{k+1}.

To select the most consistent hypothesis locally, we introduce a soft gating function:

$$\begin{aligned} W_k^+(t) &= \frac{\exp\left(-e_k^+(t)/\tau\right)}{\exp\left(-e_k^+(t)/\tau\right) + \exp\left(-e_k^-(t)/\tau\right)}, \\ W_k^-(t) &= 1 - W_k^+(t). \end{aligned} \tag{15}$$

The temperature parameter τ controls the sharpness of the soft assignment, with smaller values producing a directional selection that is close to winner-takes-all and larger values producing smoother mixtures.

The final loss is defined as

$$\mathcal{L}_{\text{B-ICAE}}(\theta) = \sum_{k=1}^{N_{\text{ch}}-1} \sum_{t \in \Omega} \left[W_k^+(t) e_k^+(t) + W_k^-(t) e_k^-(t)\right] \tag{16}$$

This loss is minimized in an end-to-end manner with backpropagation. For each mini-batch, the localization network first estimates the warp parameters θ. The forward and backward alignment hypotheses are then computed, followed by the evaluation of the soft gating function. The resulting loss is differentiated with respect to the network parameters, and all trainable parameters are optimized using Adam.

The aligned signal is reconstructed as

$$x_k^{\text{aligned}}(t) = W_k^+(t)\, \hat{x}_k^+(t) + W_k^-(t)\, \hat{x}_k^-(t). \tag{17}$$

This approach, named Bidirectional RF-DTAN (BRF-DTAN) provides a local mechanism for selecting between forward and reverse temporal alignment hypotheses, enabling the model to represent bidirectional patterns within the same analysis window while degenerating to standard RF-DTAN in purely unidirectional settings.

III. Experimental Setup

For model training and evaluation, we recorded DAS data in the city of Nice, France, on January 29, 2026, between 09:05 and 17:05. The data were acquired using an 88.2 km optical fiber installed along the highway from Nice to Carros.

DAS data were recorded using an hDAS interrogator from Aragon Photonics. The temporal sampling frequency was set to 125 Hz, with a gauge length of 15.3 m. As a preprocessing step, the data were bandpass filtered between 0.3 Hz and 2 Hz and subsequently downsampled to 10.4 Hz. The dataset used in this study corresponds to a 14.4 km section in which the fiber is deployed along a two-way highway connecting Nice and Carros, with two lanes in each direction. Because the fiber is located on one side of the highway, it predominantly captures vehicles traveling in the nearest direction, while heavier vehicles traveling in the opposite direction can also be observed.

Alignment quality was assessed over the full 8-hour DAS dataset by comparing RF-DTAN and BRF-DTAN in terms of residual temporal misalignment and vehicle-detection performance. Residual misalignment was quantified using a lag-based metric derived from the Pearson correlation coefficient. Specifically, for each window and each channel, we estimated the lag that maximized its Pearson correlation with a reference channel. The mean absolute value of these optimal lags was then used as a measure of the remaining temporal offset after alignment, where lower values indicate better temporal alignment. As an additional evaluation criterion, alignment performance was assessed by counting the number of vehicles detected in both travel directions using the method in [11].

For two adjacent raw signals x_k and x_{k+1}, the Pearson correlation coefficient is given by

$$\rho_k^{\text{raw}} = \text{corr}(x_k, x_{k+1}) = \frac{\sum_{t=1}^{T}(x_k(t)-\bar{x}_k)(x_{k+1}(t)-\bar{x}_{k+1})}{\sqrt{\sum_{t=1}^{T}(x_k(t)-\bar{x}_k)^2}\sqrt{\sum_{t=1}^{T}(x_{k+1}(t)-\bar{x}_{k+1})^2}},$$

where

$$\bar{x}_k = \frac{1}{T}\sum_{t=1}^{T} x_k(t), \qquad \bar{x}_{k+1} = \frac{1}{T}\sum_{t=1}^{T} x_{k+1}(t) \tag{18}$$

are the temporal means of the two signals. The same measure is computed after alignment as

$$\rho_k^{\text{aligned}} = \text{corr}(\tilde{x}_k, \tilde{x}_{k+1}), \qquad k = 1, \ldots, N_c - 1. \tag{19}$$

To quantify residual temporal misalignment, let $x_{w,k}^{(\ell)}$ denote channel k in window w shifted by ℓ samples, and let r denote the reference channel. For each window and channel, the optimal lag before alignment is defined as

$$\ell_{w,k}^{\text{raw}} = \arg\max_{\ell \in [-K,K]} \text{corr}\left(x_{w,k}^{(\ell)}, x_{w,r}\right), \tag{20}$$

and the corresponding optimal lag after alignment is

$$\ell_{w,k}^{\text{aligned}} = \arg\max_{\ell \in [-K,K]} \text{corr}\left(\tilde{x}_{w,k}^{(\ell)}, \tilde{x}_{w,r}\right). \tag{21}$$

The mean absolute residual lag before and after alignment is then computed as

$$\mu_{\text{raw}} = \frac{1}{N}\sum_{(w,k)} \left|\ell_{w,k}^{\text{raw}}\right|, \quad \mu_{\text{aligned}} = \frac{1}{N}\sum_{(w,k)} \left|\ell_{w,k}^{\text{aligned}}\right|, \tag{22}$$

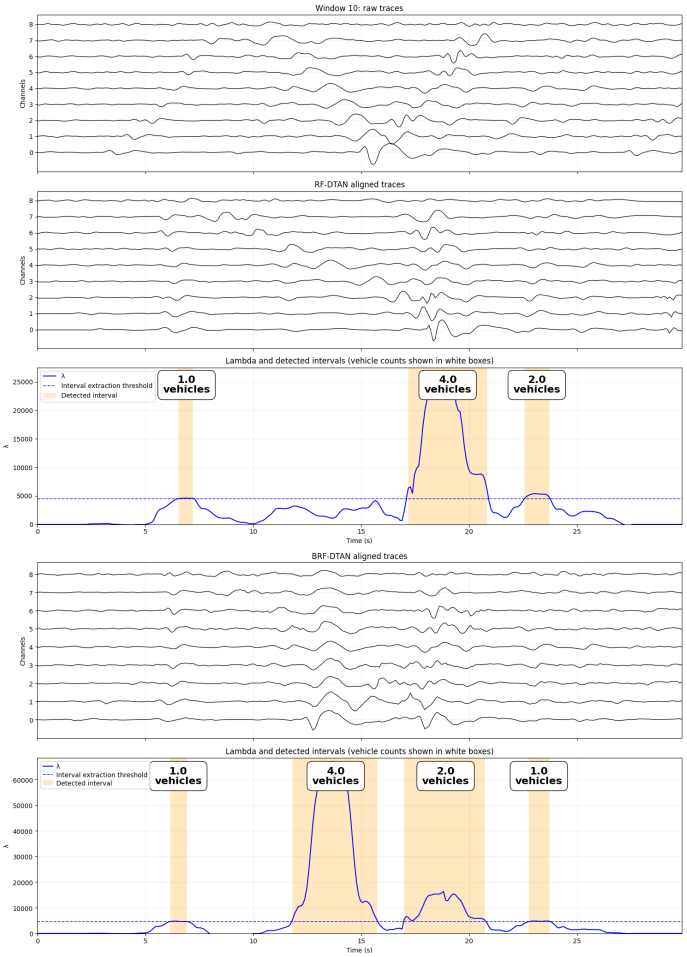

Fig. 3: Vehicle detections after alignment with RF-DTAN (top) and BRF-DTAN (bottom) over a 30 s window.

with N the total number of evaluated window–channel pairs. These quantities measure the average residual temporal misalignment before and after correction.

Finally, the relative improvement in alignment is defined as

$$\text{Improvement}(\%) = \frac{\mu_{\text{raw}} - \mu_{\text{aligned}}}{\mu_{\text{raw}}} \times 100. \tag{23}$$

Positive values indicate a reduction in residual temporal misalignment after alignment.

IV. Results

In this section, we present the results obtained by RF-DTAN and BRF-DTAN. For each model, we used analysis windows comprising 9 DAS channels and spanning 30 seconds. Both methods were trained for 300 epochs on the same 8-hour dataset and over the same range of channels. The temperature parameter τ, introduced in (15), was initialized at 0.5 and linearly annealed to 0.1 during training. This annealing schedule encourages the model to first explore soft combinations of the alignment hypotheses and then progressively favor the most consistent hypothesis.

Method	# detected intervals	# vehicles per detected interval
RF-DTAN	3	[1, 4, 1]
BRF-DTAN	4	[1, 4, 2, 1]

TABLE I: Detected vehicle intervals obtained with RF-DTAN and BRF-DTAN over the 30 s analysis window in Figure 3.

As shown in Figure 3, the two methods produce noticeably different results. This 30 s window is selected from the full 8-hour dataset for illustration purposes, while the correlation-based analysis reported below evaluates performance over the complete recording. Table I summarizes the corresponding vehicle counts for this example. RF-DTAN does not properly align the two groups of vehicles traveling in opposite directions. Instead, it aligns one group while overshifting the other. Consequently, the detection model cannot reliably separate vehicles moving in the opposite direction. By contrast, the proposed BRF-DTAN method separates the two opposing vehicle groups more effectively and provides improved vehicle-count estimates. Based on the detected counts, BRF-DTAN identifies two more vehicles than RF-DTAN.

Metric	RF-DTAN	BRF-DTAN
raw mean lag	16.99	16.99
aligned mean lag	9.72	8.54
mean absolute lag	0.4276	0.4970
lag improvement (%)	42.76	**49.70**

TABLE II: Comparison of correlation statistics after alignment between RF-DTAN and BRF-DTAN.

We further validate these observations by computing the Pearson-correlation-based lag statistic for both methods over the full 8-hour dataset. As reported in Table II, the proposed method improves the residual lag statistic by 7% relative to RF-DTAN. This gain should be interpreted in light of the challenging bidirectional setting. Since the fiber is deployed on one side of the highway, vehicles traveling in the opposite direction typically exhibit a lower signal-to-noise ratio. Moreover, crossing and overlapping trajectories further complicate the separation problem. While BRF-DTAN is able to separate such bidirectional patterns, as shown in Fig. 3, the local selection mechanism may introduce residual reconstruction noise in overlapping regions, which reduces the improvement measured by the correlation-based metric.

V. Conclusions and Future Work

In this work, we addressed bidirectional temporal alignment from a time-warping perspective and showed that standard RF-DTAN is limited by its reliance on order-preserving transformations, which cannot represent opposite temporal directions within the same analysis window. To overcome this limitation, we extended the admissible warp family and introduced a bidirectional formulation with a soft gating mechanism that locally selects the most consistent alignment hypothesis. The resulting framework can handle both unidirectional and bidirectional temporal patterns. While demonstrated on DAS traffic-monitoring data, the formulation is general and can be applied to other multidirectional time-series alignment problems. Future work will focus on reducing reconstruction noise in overlapping regions and extending the method to more complex geometries, such as curved road segments.

References

[1] Y. Qiao and M. Yasuhara, "Affine invariant dynamic time warping and its application to online rotated handwriting recognition," in *18th International Conference on Pattern Recognition (ICPR'06)*, vol. 2, 2006, pp. 905–908.

[2] M. Cuturi, J.-P. Vert, O. Birkenes, and T. Matsui, "A kernel for time series based on global alignments," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'07)*. IEEE, 2007.

[3] S. Sempena, N. U. Maulidevi, and P. R. Aryan, "Human action recognition using dynamic time warping," in *Proceedings of the 2011 International Conference on Electrical Engineering and Informatics*, 2011, pp. 1–5.

[4] *Dynamic Time Warping*. Springer Berlin Heidelberg, 2007, pp. 69–84.

[5] M. Cuturi and M. Blondel, "Soft-dtw: a differentiable loss function for time-series," in *Proceedings of the 34th International Conference on Machine Learning*, vol. 70. JMLR, 2017, p. 894–903.

[6] H. Janati, M. Cuturi, and A. Gramfort, "Spatio-temporal alignments: Optimal transport through space and time," *Proceedings of Machine Learning Research*, vol. 108, pp. 1695–1704, 2020.

[7] R. Shapira Weber and O. Freifeld, "Regularization-free diffeomorphic temporal alignment nets," in *Proceedings of the 40th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 202. PMLR, 2023, pp. 30 794–30 826.

[8] J. Azzola, K. Thiemann, and E. Gaucher, "Integration of distributed acoustic sensing for real-time seismic monitoring of a geothermal field," *Geothermal Energy*, vol. 11, no. 1, p. 30, 2023.

[9] J. Zuo, Y. Zhang, H. Xu, X. Zhu, Z. Zhao, X. Wei, and X. Wang, "Pipeline leak detection technology based on distributed optical fiber acoustic sensing system," *IEEE Access*, vol. 8, pp. 30 789–30 796, 2020.

[10] X. Zhang, J. Qi, X. Liang, Z. Guan, Z. Liu, C. Zhang, D. Chen, W. Deng, C. Xu, X. Wang, and H. Liu, "Fiber-optic distributed acoustic sensing for smart grid application," *Photonics*, vol. 12, no. 1, 2025.

[11] Y. Khacef, M. van den Ende, C. Richard, A. Ferrari, and A. Sladen, "Precision traffic monitoring: Leveraging distributed acoustic sensing and deep neural networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 6, pp. 7678–7689, 2025.

[12] O. Freifeld, S. Hauberg, K. Batmanghelich, and J. W. Fisher, "Transformations based on continuous piecewise-affine velocity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2496–2509, 2017.

[13] R. Shapira Weber, M. Eyal, N. Skafte, O. Shriki, and O. Freifeld, "Diffeomorphic temporal alignment nets," in *Advances in Neural Information Processing Systems*, vol. 32, 2019, p. 6348–6358.