Self-Supervised Velocity Field Learning for High-Resolution Traffic Monitoring with Distributed Acoustic Sensing

Yacine Khacef, Martijn P.A. van den Ende, André Ferrari, Cédric Richard, Anthony Sladen

Self-Supervised Velocity Field Learning for High-Resolution Traffic Monitoring with Distributed Acoustic Sensing

Yacine Khacef, Martijn van den Ende, André Ferrari, Cédric Richard, Anthony Sladen
Université Côte d'Azur, CNRS, OCA, France
Email: yacine.khacef@oca.eu

***Abstract*—Distributed Acoustic Sensing (DAS) is a technology that can be employed to record vibrations along fiber optic (telecommunication) cables, including those generated by human activities. Since the optical fiber cables are often deployed along existing traffic infrastructures, DAS has the potential to record vehicular traffic flows, which permits high-resolution traffic analysis and long-term monitoring. In this work, we propose a Machine Learning (ML) model for estimating the speed of vehicles using DAS data. A major component of the proposed model is based on Continuous Piecewise Affine (CPA) transformations, which allows us to extract the speed as a function of space and time. We demonstrate the efficiency of our approach, which is significantly faster than non-ML solutions in estimating the vehicle speed.**

I. Introduction

Owing to increasing urbanization, traffic management has long been a major challenge. To optimize the flow of traffic in cities, policy makers and operators require traffic data such as the number and type of vehicles on the road, and their speed. Depending on the granularity of the data, they may inform decisions such as infrastructure expansion, traffic jam intervention [1], smart traffic light management [2] and optimal route planing [3]. Commonly used instruments for traffic data collection include roadside cameras [4], GPS data [5] and inductive loop detectors [6]. These methods suffer from several limitations, like the prohibitive cost of large-scale deployment and maintenance, the low spatial resolution, and privacy concerns in the case of cameras and GPS. To overcome these issues and perform the traffic monitoring task in a cost-efficient way, we propose to use fiber-optic Distributed Acoustic Sensing (DAS).

Distributed Acoustic Sensing represents a subcategory of Distributed Fiber Optic Sensing techniques. When connected to a conventional optical fiber cable (like those used for telecommunication), a DAS interrogator performs laser-pulse interferometry to infer local perturbations of strain around the fiber, measured at fixed locations along the cable. As such, DAS effectively converts a telecom cable into an array of equally-spaced vibration sensors. Commercially available DAS interrogators now provide a sensing range of over 100 km at a minimum spatial resolution of around 1 metre, with temporal sampling rates of several kHz or more. These sensing characteristics, along with the robustness of fiber optic cables, make DAS an attractive measurement technique for numerous applications, like earthquake seismology [7], structural integrity monitoring [8], pipeline surveillance [9], and road traffic monitoring.

When a fiber optic cable is deployed alongside a road (as is often the case for commercial telecom cables), vehicles that drive along are recorded at a given location as short-duration signals that exist within a narrow frequency band (Fig. 1). These characteristic signals (or "signatures") do not depend on the details of the vehicle, and are recorded by multiple DAS sensors ("channels") as the vehicle drives along the cable. The translation of these characteristic signals from one DAS channel to the next is a simple function of the vehicle's velocity, which can be estimated with an appropriate analysis of the data [10]. In typical urban traffic scenarios, multiple vehicles may be closely-trailing with potentially different speeds, which is a challenge for techniques that employ a fixed (space)time-window to estimate the vehicle speed.

In this paper, we propose a data-driven model to obtain high-resolution detections and speed estimations of vehicles using DAS data. Our approach is inspired by the model proposed by [11], which is based on Continuous Piecewise Affine (CPA) transformations [12]. To greatly accelerate the analysis, we leverage an Artificial Neural Network (ANN) to extract the parameters of the transformations, trained in a self-supervised manner. The detail the model architecture in Section 2, and the experimental setup in Section 3 provides. Subsequently, the obtained speed estimates are presented in section 4 along with a comparison of these results with data generated by conventional traffic sensors. Section 5 concludes the paper and provides some perspectives for the use of DAS in traffic monitoring.

II. Methodology

Three main steps are required to achieve the detection and speed estimation tasks. After splitting the data into several windows, we first use a model that applies a non-uniform time warping in order to align the vehicle's signatures. We used two architectures, one including an ANN and one without. The second step consists of using the resulting data windows to detect the vehicles. The speed estimation is achieved using the aligned data and detection results. All these steps are detailed in the following sections.

A. Data alignment

1) CPA-based transformations model without ANN: This section is an application of the example provided by [12] to

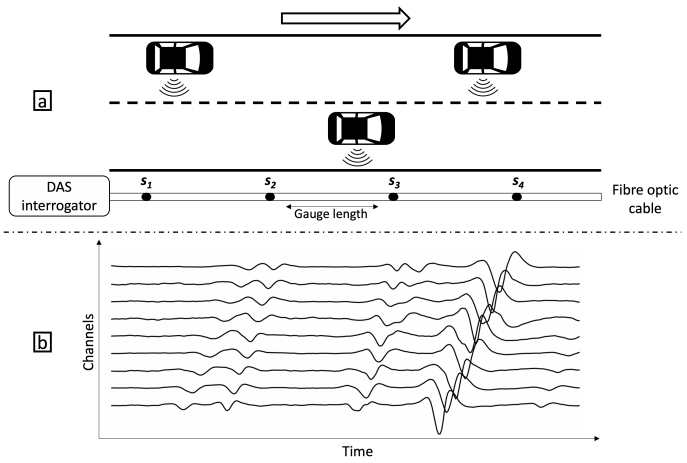

Fig. 1. a) Schematic overview of the problem geometry, which features a fiber optic cable parallel to a road. The DAS channels are denoted by S_1, S_2, etc. b) Example of DAS recordings with 9 DAS channels and 30 seconds duration. Each channel corresponds to a consecutive virtual sensor along the fiber, as indicated in panel a.

DAS data, with tuned parameters that fit well our application. Consider any pair of time series recorded by consecutive DAS channels I_n and I_{n+1}, with $I_n \in \mathbb{R}^{N_t}$, for all $n \in [1, 2, \cdots, N_{ch} - 1]$. N_{ch} represents the number of channels in the data window, and N_t its time duration. Our objective is to apply a transformation to I_n that minimises the difference between the transformed time series and I_{n+1}, i.e.:

$$\theta_n = \arg\min_{\theta'_n} L, \quad L = \|E^{\theta'_n}(I_n) - I_{n+1}\|^2 \tag{1}$$

where $E^{\theta'_n}(I_n)$ is the transformation that warps I_n and $\theta_n = [\vartheta_1^n, \vartheta_2^n, \cdots, \vartheta_{N_p+1}^n]^T \in \mathbb{R}^{N_p+1}$ is a parameter vector defining the transformation $E^{\theta'_n}$ for N_p time intervals (as defined momentarily). The latter can be subdivided into two components: a grid generator and a sampler (Fig. 2.a, without the localisation network module).

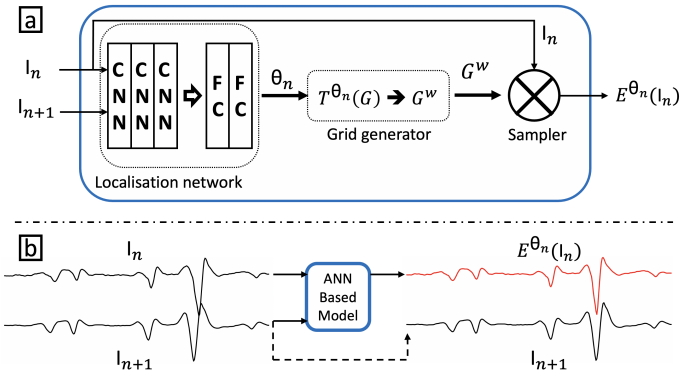

Fig. 2. a) Architecture of the ANN-based model. b) Left: Two consecutive channels I_n and I_{n+1}, which are misaligned in time. Middle: Model as shown in panel a. Right: Stack of the transformed version of I_n and the original signal I_{n+1}. We can clearly see that $E^{\theta_n}(I_n)$ and I_{n+1} are well aligned.

Grid generator:

The grid generator applies a CPA-based transformation T^{θ_n} to the time domain $G = [x_0, x_1, \cdots, x_{N_t}]$ of the input, which integrates a CPA velocity field v^{θ_n}.

$$T^{\theta_n}(x) = x + \int_0^1 v^{\theta_n}(\phi^{\theta_n}(x, \tau))\mathrm{d}\tau, x \in G$$

in which $\phi^{\theta_n}(x, \tau)$ is the position of the transformed version of x at integration time τ. v^{θ_n} is obtained by first defining a Piecewise Affine (PA) velocity field consisting of N_p intervals, each one representing a one-dimensional affine transformation (equivalent to translation in time). To get a continuous velocity field from the PA one, it is sufficient to add continuity conditions on the borders of each interval. The obtained CPA velocity field is defined as:

$$v_n^{\theta}(x) = x \sum_{j=1}^{N_p+1} \vartheta_j^n A_{x,j} + \sum_{j=1}^{N_p+1} \vartheta_j^n B_{x,j}$$

where A and B are two-dimensional matrices resulting from the continuity conditions.

Now that we defined the transformation, we can compute the warped version of G called G^w:

$$G_n^w = [T^{\theta_n}(x_0), T^{\theta_n}(x_1), \cdots, T^{\theta_n}(x_{Nt})]$$

The next step is to map the input signal I_n into the warped domain G_n^w.

Sampler:

The sampler takes two inputs, being the time series I_n to transform and the warped domain G_n^w. Then, the signal is projected onto G_n^w using a linear interpolation. This last step generates the warped signal $E^{\theta_n}(I_n)$. CPA-based transformations are continuous and differentiable, which allows us to do backpropagation of the loss L defined in Eqn. (1) and optimise with respect to the parameters θ_n that define the transformation.

2) CPA-based transformations model with ANN: Although the previous model performs well in term of warping precision, the optimal parameters θ_n are obtained through an iterative inversion procedure, which is too computationally costly for (real-time) traffic analysis. For this reason, an ANN is employed to extract θ directly from the time series. This ANN is referred to as the localisation network, and we adopt a similar architecture as the one proposed by [11], which consists of three stacked 1D convolutional layers followed by two fully connected layers. For the present study, we provide the model with two time series simultaneously, instead of one (as was done by [11]). This allows the localisation network to generate the suitable parameters for each pair of consecutive channels. The training objective of the localisation network $\mathcal{N}_\omega : (I_n, I_{n+1}) \rightarrow \theta_n$ for one data window is given by:

$$\omega = \arg\min_{\omega'} \frac{1}{N_{ch}-1} \sum_{n=1}^{N_{ch}-1} (||E^{\theta_n}(I_n) - I_{n+1}||^2 + \alpha||\theta_n||_{\Sigma^{-1}_{CPA}})$$
$$\theta_n = \mathcal{N}_{\omega'}(I_n, I_{n+1})$$

where ω' represents the parameters of the localisation network, $\alpha||\theta_n||_{\Sigma^{-1}_{CPA}}$ is a regularization term defined by [11]. It is used to avoid getting too large θ values (which leads to large transformations) and controls the smoothness of the transformation.

To align a full data window using either the model with or without the ANN, it is sufficient to align all consecutive DAS channels I_n and I_{n+1}. Then, select a reference channel (I_4 in our case, corresponding to the mid-channel) to align all the remaining ones towards it. This is done by stacking the transformations from all channels to the middle one (e.g. to transform the first channel I_1 to the fourth one I_4, we apply $E^{\theta_3}(E^{\theta_2}(E^{\theta_1}(I_1)))$). An example of aligned data window is shown in Fig. 5.b).

B. Vehicles detection

Using the aligned data windows instead of the original ones, makes the process of detecting vehicles easier and more efficient. Since the signatures of every vehicle are almost similar in all channels, taking the mean of the aligned data window along the spatial axis generates a time series $\overline{I}$, containing the average signature of each vehicle. Noise level in $\overline{I}$ is significantly reduced comparing to the individual channels I_n, making the detection more accurate.

We compute the envelope (magnitude of the analytic signal) of $\overline{I}$, and use a threshold to detect and locate in time the vehicles. This generates a binary signal $\overline{I}_B$, with ones where the envelope is above the threshold, and zeros elsewhere.

C. Speed estimation

Using either the first or the second model, we convert the extracted parameters θ_n into velocity estimations v_n as a function of space and time using the offset between the original grid G and the warped version G^w_n, $v_n \propto \Delta_{G_n}, \Delta_{G_n} = G^w_n - G$. This offset represents a time shift required to align the data of two consecutive DAS channels. Given the spatial distance between each pair of DAS channels, the velocity vector is computed as:

$$v_n = f_s \times \frac{\delta}{\Delta_{G_n}}, n \in [1, 2, \cdots, N_{ch}-1]$$

with f_s the temporal sampling frequency, and δ the spacing between the DAS channels (or "gauge length").

This generates a speed map (Fig. 5.c) defined between all channels and all along the time axis. This map can not be used as it is for extracting the vehicles speeds, since it includes parts of data where only noise is present. In order to filter the speed map and keep only vehicles signatures sections, we multiply it by a binary mask (Fig. 5.d). It is generated using inverse stacked transformations of $\overline{I}_B$ (e.g. to filter v_1, we multiply it by $E^{-\theta_1}(E^{-\theta_2}(E^{-\theta_3}(\overline{I}_B)))$). The resulting filtered speed map is shown in Fig. 5.e . We can then extract the individual/group of vehicles speeds, and compute the average velocity for every data window.

III. Experimental setup

We recorded DAS data in the city of Nice, France, between November 28, 2021 and January 5, 2022. We used a 20 km fiber that goes around the city (see Fig. 3). The fiber was already installed and used for telecommunication purposes.

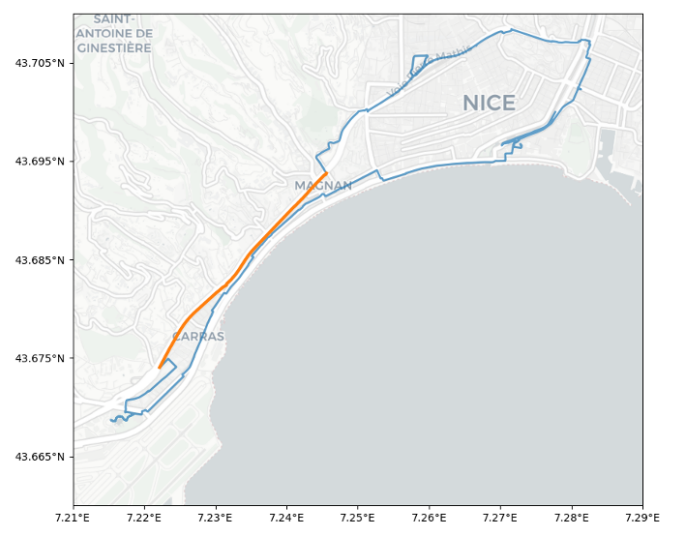

Fig. 3. Sensed fiber path going around the city of Nice, France. Orange section represents portion of the fiber from where we collected DAS data used in this work.

We used a hDAS interrogator (Aragon Photonics) to record the DAS data. The temporal sampling frequency was set to 250 Hz and the gauge length (as well as the channel spacing) was set to 10 m. Before processing, data were filtered between 0.5 Hz and 2 Hz then downsampled to 10 Hz. All data we used in this work came from one section where the fiber is deployed along the Voie Pierre Mathis, a three lanes highway going through part of the city (orange section in Fig. 3). Since the fiber is located on the side of the highway, it only catches the vibrations of the vehicles driving on one direction (from South West to North East in Fig. 3).

IV. Results and discussion

For each model, we used data windows consisting of 9 DAS channels and 30 seconds duration. The localisation network was trained over 150 epochs using 138 hours of data. To get similar results as the ANN-based model, 20 optimization iterations per pair of channels are required for the model without ANN. This leads to significantly different execution times. While the one without localisation network takes up to 10 seconds to process one data window, the ANN-based model performs the same task in around 0.01 seconds, which yields

a performance gain of a factor 1000 without loss of accuracy. This speed-up is critical for real-time traffic analysis.

To have a reference to compare the obtained results with, we used speed estimates from loop detectors placed in one section of the monitored highway. Fig. 4 shows the results of applying our model to one day of data and the velocities provided by the loop detectors for the same period of time. We can see high velocities during early morning and evening times, this is due to the low traffic density, so vehicles can drive faster. During the daytime, vehicles drive around 70 $\mathrm{km\ h^{-1}}$, which corresponds to the speed limit. Velocities are equal to zero before 5 am and after 11:30 pm, which corresponds to no vehicles, because the monitored highway is closed in this time interval. We can see more fluctuations in our model speed estimates than the loop detectors ones, and this due to several reasons. Among them, one can highlight the high sensitivity of our model to heavy vehicles and vehicles driving close to the fiber. These two cases generate high amplitude signals, where the ANN gives more importance to align, which may affect the speed estimates of the data window. We can also notice around 5 am that the loop detectors are not detecting any vehicle, unlike our model, which gives fairly high speeds. We checked in the DAS data, and we found that most of our speed estimates were correct around this time. We suspect that the loop detectors were not enabled yet, or need more vehicles to provide speed estimates.

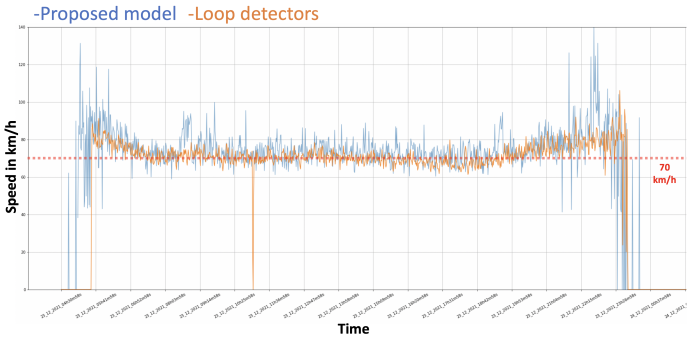

Fig. 4. Speed estimates provided by our model (blue curve) and the loop detectors (orange curve). The dashed red line indicates the speed limit (70 $\mathrm{km\ h^{-1}}$).

V. Conclusions and perspectives

In this paper we introduced a novel solution for high-resolution vehicles detection and speed estimation using DAS. We showed the efficiency of this technology, combined with the proposed models, achieving these two tasks. Although both models, with and without the localisation network, generate similar transformations, their execution times are significantly different. Making the ANN-based model a clear choice for our application, especially since we plan to upgrade the system to real-time. The speed estimates provided by the loop detectors constitute a good reference to compare our results with. Despite some differences between the two techniques results, the overall speed trends are similar. The self-supervised aspect of the model, makes it suitable for a large variety of time series alignment applications, even when no labels are available.

Our proposed method opens the door to several other traffic monitoring applications. Starting by a real-time version of the system, where DAS data will be processed right after acquisition. Taking advantage of the fact that vehicles generate signatures in DAS data that are proportional to their characteristics (i.e. weight, length, number of wheels, etc.), we can extract the types of vehicles from the DAS recordings. Traffic anomaly detection, like traffic jams or vehicles accidents, can also be achieved using, for example, the speed estimates we obtain using our model combined with a tracking algorithm.

Acknowledgements

This research was supported by the French government through the 3IA Côte d'Azur Investments in the Future project with the reference number ANR-19-P3IA-0002. We thank the Métropole Nice Côte d'Azur (MNCA) for providing access to the fibre.

References

[1] P. P. Dubey and P. Borkar, "Review on techniques for traffic jam detection and congestion avoidance," in *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*. IEEE, 2015, pp. 434–440.

[2] B. Płaczek, "A self-organizing system for urban traffic control based on predictive interval microscopic model," *Engineering applications of artificial intelligence*, vol. 34, pp. 75–84, 2014.

[3] W.-C. Hu, H.-T. Wu, H.-H. Cho, and F.-H. Tseng, "Optimal route planning system for logistics vehicles based on artificial intelligence," *Journal of Internet Technology*, vol. 21, no. 3, pp. 757–764, 2020.

[4] P. Reinartz, M. Lachaise, E. Schmeer, T. Krauss, and H. Runge, "Traffic monitoring with serial images from airborne cameras," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 61, no. 3-4, pp. 149–158, 2006.

[5] L. X. Pang, S. Chawla, W. Liu, and Y. Zheng, "On detection of emerging anomalous traffic patterns using gps data," *Data & Knowledge Engineering*, vol. 87, pp. 357–373, 2013.

[6] L. Bhaskar, A. Sahai, D. Sinha, G. Varshney, and T. Jain, "Intelligent traffic light controller using inductive loops for vehicle detection," in *2015 1st International Conference on Next Generation Computing Technologies (NGCT)*. IEEE, 2015, pp. 518–522.

[7] M. van den Ende and J.-P. Ampuero, "Evaluating seismic beamforming capabilities of distributed acoustic sensing arrays," *Solid Earth*, vol. 12, no. 4, pp. 915–934, 2021.

[8] H.-N. Li, D.-S. Li, and G.-B. Song, "Recent applications of fiber optic sensors to health monitoring in civil engineering," *Engineering structures*, vol. 26, no. 11, pp. 1647–1657, 2004.

[9] J. Tejedor, J. Macias-Guarasa, H. F. Martins, J. Pastor-Graells, P. Corredera, and S. Martin-Lopez, "Machine learning methods for pipeline surveillance systems based on distributed acoustic sensing: A review," *Applied Sciences*, vol. 7, no. 8, p. 841, 2017.

[10] M. van den Ende, A. Ferrari, A. Sladen, and C. Richard, "Next-generation traffic monitoring with distributed acoustic sensing arrays and optimum array processing," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2021, pp. 1104–1108.

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

[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.

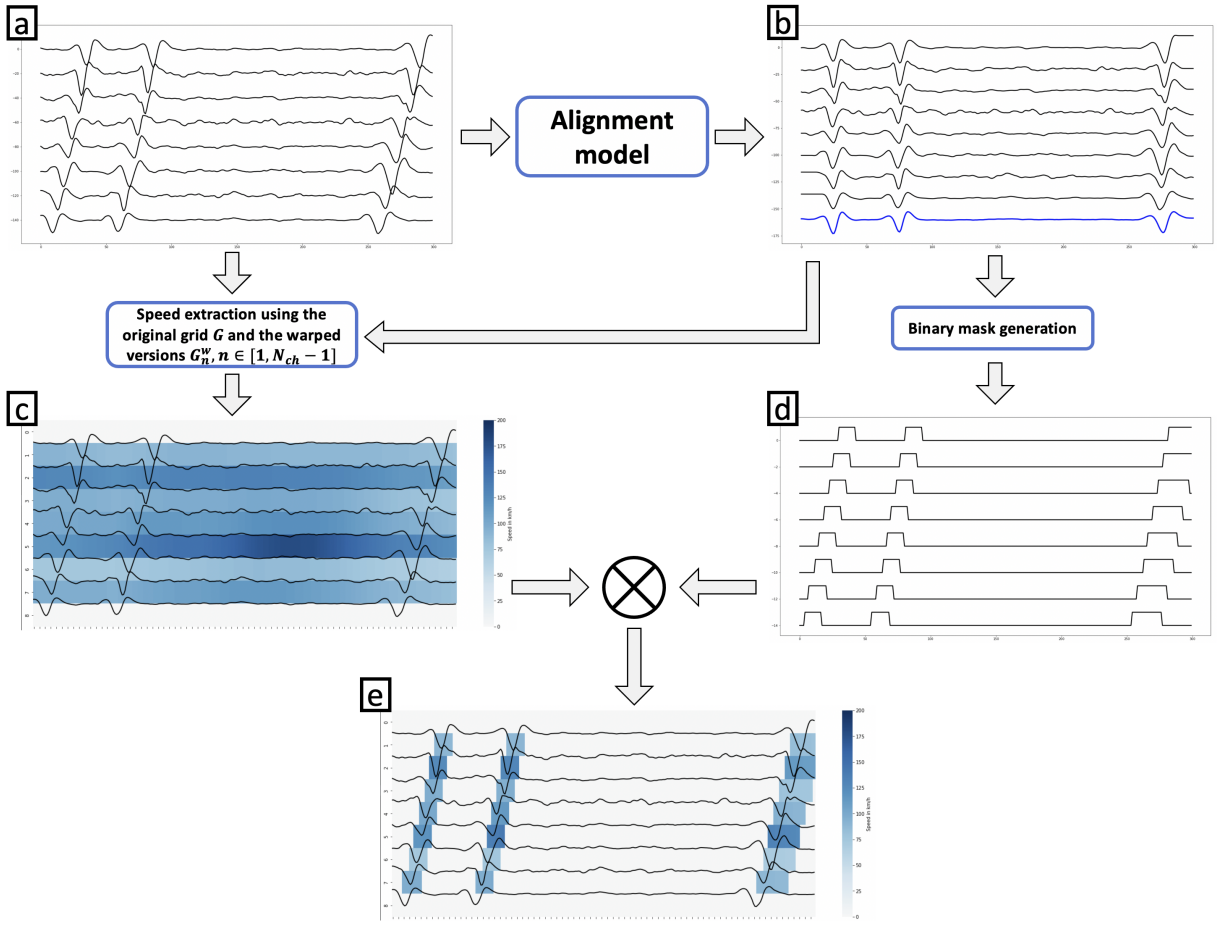

Fig. 5. Full procedure of vehicles detection and speed estimation for one DAS data window. a) DAS data window example with 3 vehicles. b) Aligned version of the data window in panel a using the ANN-based model. The bottom time series, in blue color, represents the average of all aligned channels along the spatial axis. c) Obtained speed map for the data window in panel a. d) Binary mask generated following the procedure in section II-C. e) Filtered speed map obtained by multiplying the one in panel c by the binary mask in panel d. For this example, since the vehicles signatures are not close to each other, we can extract the individual velocities for each of them.