Spatial Deep Deconvolution U-Net for Traffic Analyses With Distributed Acoustic Sensing

Siyuan Yuan, Martijn van Den Ende, Jingxiao Liu, Hae Young Noh, Robert Clapp, Cédric Richard, Biondo Biondi

Spatial Deep Deconvolution U-Net for Traffic Analyses with DAS

Siyuan Yuan, Martijn van den Ende, Jingxiao Liu, Hae Young Noh, Robert Clapp, Cédric Richard, Biondo Biondi

***Abstract*—Road-side Distributed Acoustic Sensing (DAS), being capable of capturing car-induced subsurface strain at high spatial-temporal resolution, has shown potential to revolutionize urban traffic monitoring by providing instantaneous updates on traffic flow, and possibly traffic accidents to optimize traffic management. However, due to noise contamination and interference among closely traveling cars, car detection and tracking accuracy is limited. To address these challenges, we design and train a self-supervised U-Net model that compresses car-induced DAS signals into sharp pulses through a spatial deconvolution. The localized and narrow outputs from our model lead to accurate and highly resolved car position and speed tracking, which can help identify driving behaviors and detect sudden stops due to accidents. We evaluate the effectiveness and robustness of our method through field recordings under different traffic conditions and various driving speeds. Our results show that our method can enhance the spatial-temporal resolution and better resolve closely traveling cars, which is beneficial for detecting and tracking cars under heavy traffic conditions and enables the characterization of large-size vehicles such as retrieving the number of bus axles and train bogies.**

***Index Terms*—Traffic monitoring, intelligent transportation, Distributed Acoustic Sensing, deconvolution, U-Net.**

I. Introduction

Traffic monitoring systems, which automatically and continuously detect, track, and characterize vehicles in moving traffic, provide valuable information for urban management, maintenance, and planning. Conventional monitoring systems include vision-based [1]–[3] and pavement sensing technologies (e.g., inductive loops [4]–[6] and piezoelectric sensors [4], [7], [8]). These approaches are well-developed but have several drawbacks. For example, camera systems bring individual-privacy concerns and are sensitive to weather conditions; point pavement sensing systems provide spatially sparse sampling and are challenging to maintain.

An emerging geophysical technology, Distributed Acoustic Sensing (DAS), which turns (pre-existing) optical fibers into dense seismic recording arrays, has the potential to revolutionize urban traffic monitoring by providing instantaneous updates on traffic flow and accidents. A DAS experiment is performed by connecting an optoelectronic DAS instrument called interrogator to one end of a standard telecommunications-grade optical fiber. The interrogator sends short laser pulses into the optical fiber and measures the subtle phase shifts of Rayleigh scattered light returning to the detector at a predicted two-way travel time [9], [10]. In this way, the strain field induced by urban activities (e.g. moving vehicles, construction, pumps) acting on the fiber coupled to the Earth can be sampled at a meter-scale spatial resolution over tens of linear fiber kilometers. Thanks to their broadband nature, DAS recordings comprise rich traffic information, including high-frequency (> 3 Hz) surface waves due to vehicle-road dynamic interaction and the low-frequency quasi-static deformation caused by vehicle loading [11]–[13]. The surface-wave components have been used to image the subsurface velocity structure down to hundreds of meters beneath the road [11]. The quasi-static components reflecting the vehicle trajectories are much more compact and have a simpler waveform than the surface waves. [12], [14], [15] demonstrate the feasibility of using the quasi-static signals to estimate vehicle count and speeds, and illustrate several advantages for a DAS traffic monitoring system: first, a DAS signal is fully anonymous that cannot be tied to any individual; second, DAS has been proven effective even on pre-existing fiber infrastructure ("dark fiber"), which is promising for cost-efficient city-scale traffic monitoring; last, DAS has low day-to-day operation and maintenance costs with only a single interrogator being deployed in a secure and easily accessible location, as opposed to having numerous individually powered instruments deployed across a city (exposed to meteorological conditions and detrimental interactions with humans, plants, and animals).

Despite the aforementioned advantages, analyzing data from a DAS array located in an urban environment is challenging because it records a complex mixture of inherently unlabeled signals [16]. A robust and accurate car-signal detection algorithm is essential to enable the real-world usage of a DAS-based traffic monitoring system. To detect car-induced quasi-static signals, [12] applied a common seismological method, the short-time-average through long-time-average trigger (STA/LTA). To exploit the array geometry of DAS, beamforming algorithms have been applied to detect cars and measure their speed [14], [15]. These simple methods performed well on roads with relatively light traffic and without complicated traffic patterns. But they are likely to become inaccurate if many vehicles transit simultaneously close to the same segment of the fiber cable, as signals from different cars start overlapping leading to complex patterns.

To reduce the interference among closely traveling cars, [17] proposed a self-supervised deep deconvolution Auto-Encoder (DAE) to deconvolve the cars' impulse response from the quasi-static recordings. Compared to a conventional channel-wise deconvolution algorithm, the DAE model has the benefits of incorporating the spatial-temporal characteristics of car signals, leading to much sharper and localized outputs.

This paper was produced by the IEEE Publication Technology Group. They are in Piscataway, NJ.

Manuscript received xxx; revised xxx.

Applying a beamforming algorithm to the localized outputs rather than the original inputs shows significant improvements in terms of the resolution in car-speed estimation, and the detection accuracy. Furthermore, as shown in [17], the DAE model, once trained, can deconvolve 24-hour recordings in less than 30 seconds, achieving $>$ 400 times speedup compared to a conventional iterative approach, which makes the method promising for real-time processing. However, the DAE model assumes a stationary Ricker wavelet in the time domain as the cars' impulse response. As the following sections of the paper will show, cars' temporal wavelet vary with car speeds, which makes the original DAE model not ideal for speed-varying traffic. Additionally, the DAE model appears to be less effective for large-size and multi-axle vehicles, possibly because of the wavelet mismatch between the assumed stationary time-domain wavelet and the actual wavelets.

This paper characterizes car-induced quasi-static signals and proposes a new deep deconvolution U-Net model assuming a stationary car wavelet in space-domain. We show our model is effective to compress signals of speed-varying cars into compact pulses while suppressing background noises with multiple real-data examples. Our model is also robust to map large-size vehicle signals to compact pulses corresponding to their axles, which is useful to vehicle type identification. In the following sections, we first characterize cars' impulse response using synthetic simulation and observations from field experiments. We then describe the method of our spatial DAE model in details, and evaluate the performance of our space-domain DAE benchmarked with the original time-domain DAE model and a traditional iterative deconvolution approach. Specifically, we test their performance for single-car tracking, followed by a local slant stack algorithm to estimate car speed. Additionally, we investigate their performance to monitor traffic patterns and identify the axles of large-size and heavy vehicles. Finally, we discuss that the spatial DAE model could be sub-optimal where car impulse response varies spatially due to the near-surface heterogeneity. To address the challenge, we estimate the location-dependent car impulse responses by statistically averaging and training separate DAE models for different parts of the fiber.

II. Vehicle-Induced DAS Response

Permanently deployed fibre-optic cables are often buried (trenched) or placed within underground conduits. We assume here that the DAS fiber is deployed alongside a road at some depth below the surface. When a vehicle passes over the virtual sensors of the roadside telecom fiber cable, the interaction between the vehicle and the road structure induces deformation of the telecom fiber cable. The signal pattern of vehicle-induced telecom fiber deformation is a function of the vehicle characteristics, fiber conduit properties, the ambient conditions, etc. There are mainly two components of signals produced by moving vehicles: quasi-static signals ($<$ 1 Hz) resulting from the ground deformation due to the vehicle's weight, and surface waves (3 to 20 Hz) caused by the dynamic vehicle-road interaction resulting from the roughness of the road (e.g., bumps). Previous studies [11], [12] have found that the quasi-static component dominates the energy of vehicle-induced telecom fiber vibration and is theoretically described by the Flamant-Boussinesq approximation [18]. As a vehicle approaches the virtual sensor, ground deformation above the sensor increases, and the fiber coupled to the earth is stretched, resulting in increased tension in the fiber. As the vehicle moves away, ground deformation near the virtual sensor and the fiber tension decreases. As a result, the vehicle motion creates a bell-shaped response when it passes a virtual sensor. Due to the relatively strong energy, simplicity and compactness compared to the surface-wave component, quasi-static signals have been used for car-tracking and detection tasks [17]. For the same reason, our car tracking method is also based on the quasi-static response. As the car impulse response serves as an input to the DAE model, achieving a better understanding of the quasi-static signals can guide the design of the DAE model.

We propose a new numerical simulation approach that generalizes the homogeneous subsurface medium assumed in the Flamant-Boussinesq approximation to a vertically heterogeneous medium. We adapt the modelling method from [19] that study the response of heterogeneous (layered) medium to wind pressure. We define the z axis such that the top of the lowermost elastic half-space is at $z = 0$ and the surface of the Earth is at $z = H$. Vertical displacement (u_z), vertical normal stress (σ_{zz}), horizontal displacement (u_x), and shear stress (σ_{xz}) are continuous throughout the medium and satisfy an ODE equation [19]. We model a car as uniform forces, $P(x,t)$ applied at the surface under each wheel as shown in Figure 1. Using $\sigma_{zz} = -P(x,t)$ as a boundary condition, we can solve the ODE equation through numerical integration and obtain the strain projected on the fiber. Assuming gradually increasing transverse and longitudinal wave speeds (V_S and V_P, respectively) and density model from the free surface down to $z = 0$, we simulate recordings of a two-axle car with a wheelbase of 2.8 m traveling with speeds of 10, 20, 30, and 40 mph recorded by one channel in Figure 2 (a). We can see that as the car speed increases, the wavelet becomes more compact (bandwidth becomes broader). In Figure 2 (b), we plot signals along the fiber at a constant time, we can see that waveforms in the space-domain is invariant to the speed change. Instead, the waveforms in the space domain depend on the vehicle size and the number of axles. In Figure 2 (c), we simulate signals for small cars with one and two axles (spaced by 2.8 m), and a large-size vehicle with three axles with an axle spacing of 8 m. We see that due to the long wavelength of the quasi-static signal, the one-axle car generates a similar waveform as the two-axle car with a wheelbase of 2.8 m. Waveform differences can be seen between the large-size vehicle (with an 8 m axle spacing) and small cars, indicating the sensitivity of DAS measurements for large-vehicle (e.g. bus or truck) identification.

Our observation that the space-domain rather than the time-domain signals is speed invariant is confirmed using field experiments at Sand Hill road monitoring by the Stanford DAS2 array [11] (See the fiber map in Figure 3). We drove a test car along the Sand Hill road to understand the car impulse responses for various driving speeds, including 10, 20, 30, and 40 mph. Figure 4 shows the quasi-static signals of our

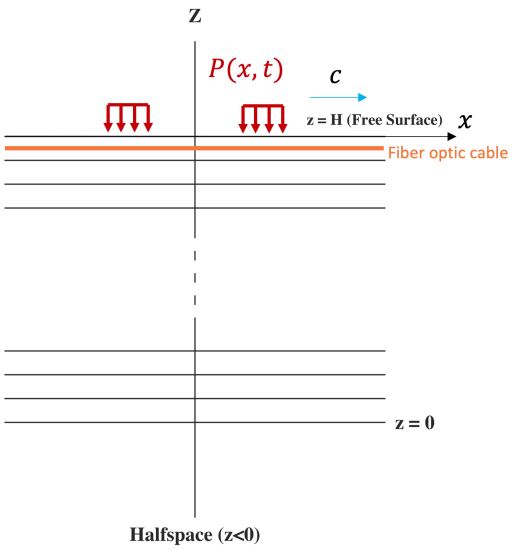


Fig. 1. Vertically heterogeneous medium used for car impulse response simulation (Figure adapted and revised from [19]. $P(x,t)$ represents the forces from the car applied to the ground surface under the wheels. Density, V_P and V_S change with depth between $z = 0$ and the free surface). The lowermost half-space is assumed to be homogeneous.

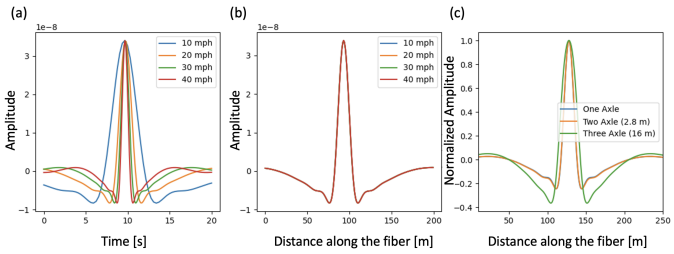


Fig. 2. Numerical simulation of quasi-static signals of two-axle car with constant speeds of 10, 20, 30 and 40 mph. (a) Signal in the time domain; (b) signal in the space domain for the four speeds. (c) Comparisons of signals for vehicles with different axle numbers and lengths.

car indicated from DAS recordings using orange arrows. The bottom panels show the corresponding F-K spectra. It can be seen that the frequency components vary with speeds, i.e., a higher car speed leads to a broader bandwidth (narrower temporal wavelet). We can also observe that the wavenumber components remain relatively invariant to car speeds, implying that the spatial wavelet for different speeds is stationary.

III. Method

A. Space-domain DAE model

The original DAE model performs deconvolution along the time axis with a stationary wavelet in time. The model is effective when car speed is approximately constant in time. However, as was shown in the previous section, the car's time-domain impulse response changes with speed, indicating the method could be suboptimal when the speed varies (e.g. stopping or accelerating). In contrast, the impulse response in the space domain is speed-invariant and can be tied to wheelbase and axle numbers. Therefore, we introduce a space-

Fig. 3. Stanford DAS-2 experiment Array (Adopted from [12])

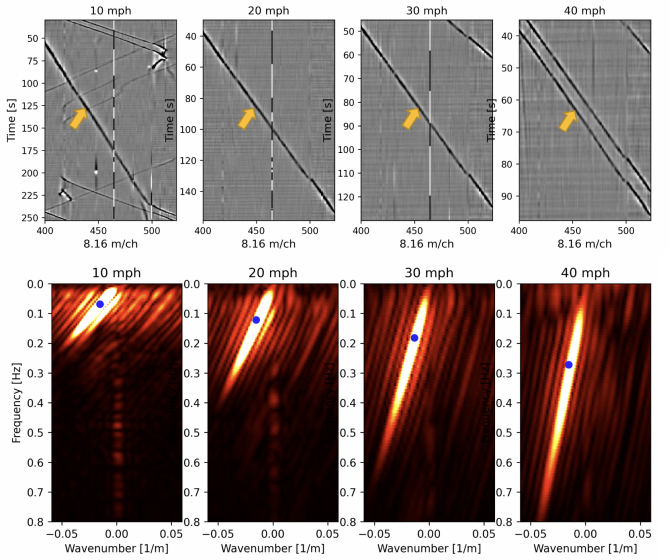


Fig. 4. (Top) DAS recordings of our car driving with constant speeds of 10, 20, 30 and 40 mph. (Bottom) The corresponding F-K spectra of the DAS recordings above. We can see that the frequency range becomes broader with increasing speed, whereas the wavenumber components are invariant to car speeds.

domain DAE model to perform deconvolution along the spatial axis with a space-domain wavelet.

1) Architecture: Our space-domain DAE model shown in Figure 5 is a 2-D fully convolutional U-Net adpated from [17]. The input is a set of $N_x = 256$ (256 meters) consecutive waveforms of $N_t = 1024$ time samples (20.48 s) in duration, organised in an $N_x \times N_t$ matrix. The U-Net model comprises 3 convolutional layers, followed by 3 encoder blocks containing a downsampling (max pooling) layer and 3 convolutional layers. The kernel sizes for the convolution layers are 3×5. The number of convolutional filters is initialized at 8 and get doubled after each downsampling operation. The maxpooling operation downsamples the data by a factor 2 along the DAS sensor axis and by a factor 4 along the time axis (i.e, the maxpooling kernel and strides are of size 2×4). The decoder reverses the encoding operations with 3 blocks of bilinear upsampling. The U-Net contains skip-connections, which directly connect the output of one encoder block with

the corresponding decoder block. Lastly, the output layer is a single convolutional layer with 1 output channel and ReLU activation, which enforces positivity in the model output.

Our model is a semi-supervised algorithm, in a sense that no ground truth deconvolution is required as labels to train the model. A weak supervision comes from the spatial car impulse response kernel shown as the red bell-shaped curve in Figure 5. The key difference between our DAE model and the original time-domain DAE model is that we obtain the reconstructed input by spatial convolution using the impulse response and the network output. The loss function is defined as a combination of the L-2 norm of the difference between the reconstructed input and the original input, and the L-1 norm of the outputs:

$$\mathcal{L} = \frac{1}{N_b}\sum_{i=1}^{N_b}(||[k * x_i]_d - y_i||_2^2 + \rho||x_i||_1), \tag{1}$$

where x_i and y_i denote the i-th output and input of the U-Net model, respectively. ρ is a weighting term that promotes sparseness in the deconvolved results. d refers to convolution along the sensor axis. As we will see in the following sections, unlike conventional channel-independent linear filters, the 2-D deconvolution operations incorporate spatial-temporal features in the DAS recordings. The non-linear nature of the U-Net introduces high frequencies that are not present in the inputs, producing sharp and localized outputs.

2) Spatial kernel estimation: Our DAE model assumes the spatial kernel of the car impulse response is stationary in space. In practice, the spatial kernel can be estimated either through numerical simulation described in the last section or by performing statistical averaging of responses of multiple passing cars assuming that the spatial impulse response is constant in time at each fiber location. The statistical averaging approach requires detecting several passing cars in an interested subsection of the fiber. The detection can be achieved through manual inspection. Herein, we apply a find local maximum algorithm from the SciPy library [20] to the recordings at quiet midnight to detect and average waveform from isolated cars.

3) Training procedure: To show the effectiveness of our approach, we trained two models, our proposed space-domain DAE and the original time-domain DAE, using simulated spatial and temporal wavelets respectively. The dataset we used are 2-hours' worth of traffic recordings of the Sand Hill DAS fiber. We split the 2-hour recordings into a training and evaluation set with a ratio of 80% to 20%. For fair comparisons, we use the same U-Net architecture for the two models.

B. FISTA deconvolution algorithm

We benchmark our spatial DAE model using a conventional deconvolution algorithm with an objective function:

$$\hat{x}_q = \underset{x}{\operatorname{argmin}}_q\{\frac{1}{2}||[k * x_q]_d - y_q||_2^2 + \rho||x_q||_1\} \tag{2}$$

Notations have same meaning as in equation (1). Note that $[k*x]_d$ stands for convolution in space between a known spatial impulse response kernel and the underlying impulse model, x. One commonly used algorithm to solve this optimisation problem is the Iterative Shrinkage Thresholding Algorithm (ISTA; [21], [22]). For this study, we adopt an accelerated version of ISTA (Fast-ISTA or FISTA) due to [23], which exhibits faster convergence guarantees. With FISTA performing spatial deconvolution, each signals at each time index is processed independently.

IV. Results

A. Car tracking

To test the performance of the proposed spatial DAE model, we conducted controlled driving experiments where we drove a car equipped with a speed sensor and a GPS receiver along a subsection of the Sand Hill road monitored by the Stanford DAS-2 array. In Figure 6, we focus on a case where our car is first speeding up and then slowing down. (a) shows the quasi-static signal of our car. The car speed is inversely proportional to the slope in the time-space coordinate. We can see in (a) that when the car speed is low (< 10 s and > 30 s), the time-domain wavelet is "stretched". When the car speed is relatively higher, the time-domain wavelet is "compressed", agreeing with our observation in Section II. (b), (c) and (d) show the deconvolution results from the proposed space-domain DAE model, the original time-domain DAE model and the space-domain FISTA algorithm. We can see that the proposed space-domain DAE model yields the sharpest and the most localized results regardless car-speed variation. Meanwhile, we can see that the background noise is suppressed by the DAE model. In contrast, the time-domain DAE model yields results that are dependent on the speed, e.g. the output is more compact in space with a higher speed but becomes stretched out when the speed is lower. The spatial deconvolution via the FISTA algorithm yields results that are not as compact as the spatial-DAE results, and signals < 5 seconds are oversuppressed by the L1 norm term in the equation 2.

Car speed estimates can be obtained through a local slant stack algorithm applied to either the quasi-static signals or the deconvolution results (Figure 7). The spectrum is computed by conducting a slant stack using a moving time window of 4 s with a stride of 0.5 s. In Fig. 7 a brighter color indicates higher stack energy in each panel. The red curves represent speed measurements from an onboard speed sensor. Our speed estimates, indicated with the black curves, are obtained from picking the maximum amplitude of the slant-stack spectrum at each time step. We can see that our slant-stack estimates match the red curve well in all four cases. Using the speed-sensor measurement as the ground truth, the root-mean-squared errors (RMSE) for the original recording, space-domain DAE, time-domain DAE, and FISTA are 3.52, 3.54, 3.94, and 3.99 km/h, respectively, indicating the effectiveness of using DAS and slant-stack for car-speed monitoring. It is noteworthy that the space-domain DAE slant stack spectrum in (b) shows an evidently higher resolution around the sensor measurement, which can be attributed to the sharp deconvolution results as shown in Figure 6 (b).

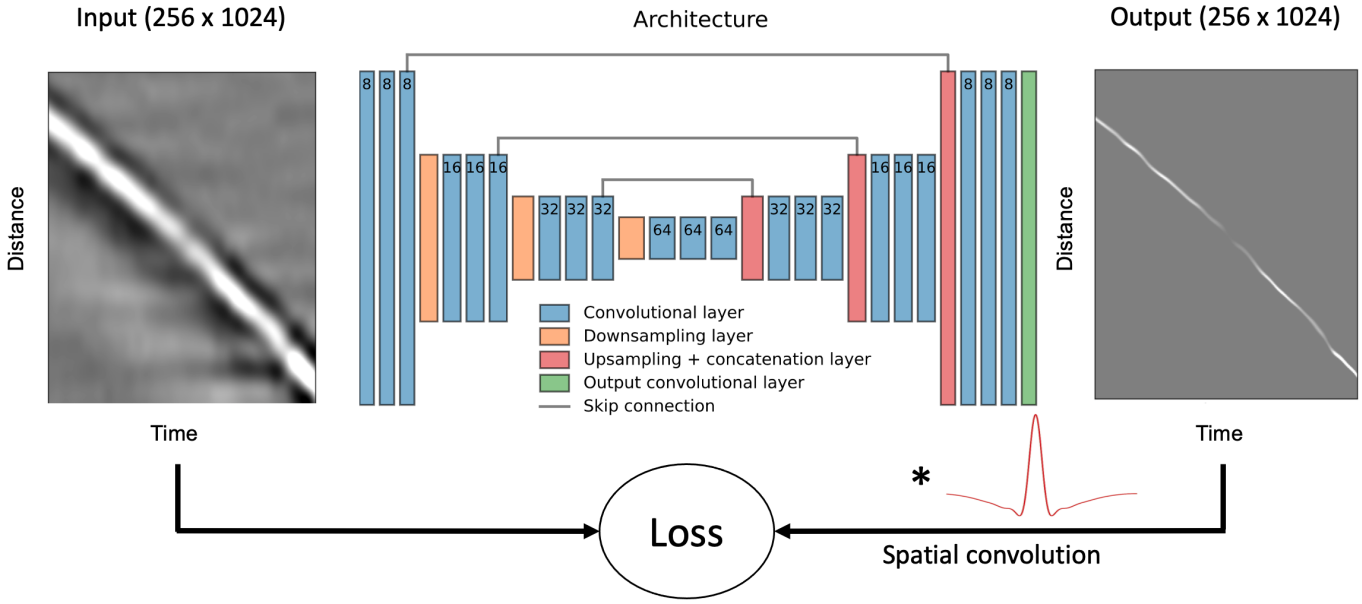


Fig. 5. Conceptual overview of the spatial DAE model. The input is the quasi-static response of DAS to cars, which can be viewed as a matrix, 256 channels (256 meters) x 1024 time steps (20.48 seconds). The output is the deconvolution results. The loss is computed with the input and the reconstructed input obtained through a spatial convolution of the output with a known car impulse response in the spatial domain.

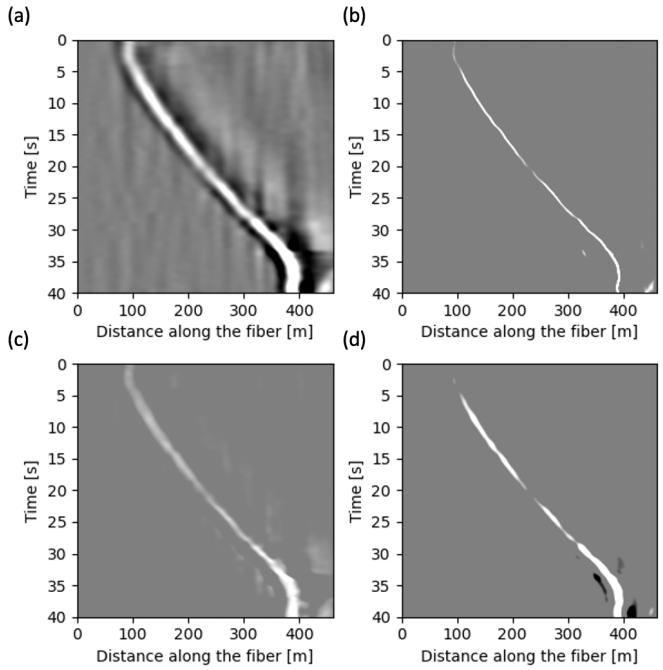


Fig. 6. (a) DAS recording of the quasi-static signal of a controlled driving experiment where we first speed up and then slow down; Deconvolution results: (b) the proposed space-domain DAE model, (c) the time-domain DAE model; (d) spatial deconvolution via the FISTA algorithm

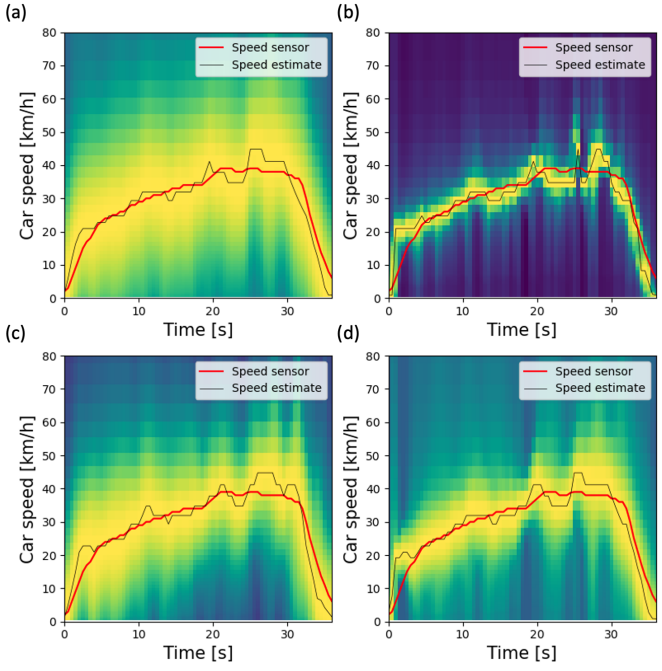


Fig. 7. Local slant stack spectra using the original and the deconvolved data shown in Figure 6. The red curve in each panel shows the speed measurements of an onboard speed sensor. The black curves indicate the speed estimates picked from the amplitude of the spectra at each time step.

B. Traffic pattern monitoring

This section investigates the performance of the space-domain DAE model under dense traffic conditions. A challenging case would be a faster car passing a slower, causing signal interference. Figure 8 (a) shows a 50-second quasi-static recording of a subsection of the Sand Hill DAS array (bandpass filtered between 0.1 - 1 Hz). Each car is represented by a relatively simple and compact pattern similar to the one shown in Figure 6. (b), (c), and (d) show the deconvolution results of the space-domain DAE model, time-domain DAE model and FISTA, respectively. We can see that the two DAE

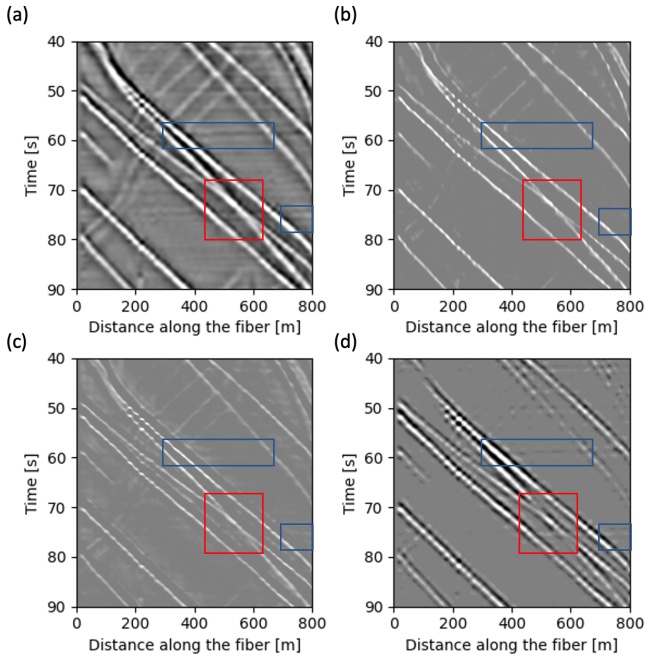


Fig. 8. (a) Car signals in a 50-second recording. Deconvolution results of (b) space-domain DAE model, (c) time-domain DAE model, and (d) FISTA, respectively. The red boxes indicate two-car passing. The blue boxes indicate ocean-wave noises in the (a) and the corresponding residues in (b), (c) and (d).

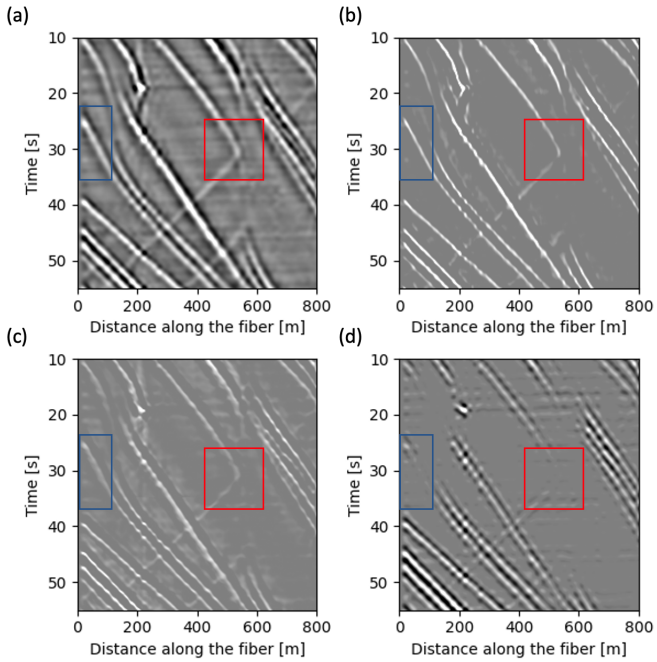


Fig. 9. (a) Car signals containing a car performing U-turn indicated with a red box. Deconvolution results of (b) space-domain DAE model, (c) time-domain DAE model, and (d) FISTA, respectively. The blue boxes indicate a signal of an accelerating vehicle.

models yield much sharper results than FISTA by introducing high frequencies that are not existing in the original recordings, owing to the highly non-linear nature of the U-Net model. The red boxes indicate a two-car passing scenario where quasi-static signals of the two cars overlap. We can see that both DAE models handle the crossing point well. By constrast, the FISTA model oversuppresses parts of signals of the slower car, which can be seen at the bottom right corner of the red box in (d). Additionally, we observe that the space-domain DAE model yields outputs containing the lowest level of background noise energy. The blue boxes in (a) point out a common type of non-car signal, which manifests itself as horizontal stripes. We interpret these signals as ocean waves. The ocean wave signals can have a similar time-domain wavelet as the car impulse response, which could explain the artefacts in the time-domain DAE results. We also observe residuals of ocean waves in the FISTA results indicating the lack of robustness of FISTA to noise.

Cars performing U-turns can also be captured in DAS recordings, as indicated with the red box in Figure 9 (a). (b), (c), and (d) show the deconvolution results, respectively. The spatial-DAE model again yields the sharpest results with the lowest energy level of ocean-wave imprints. The blue boxes indicate a signal of an accelerating vehicle. When the car speed is low, the wavelet in the time-domain is "stretched", which explains why the low-speed signal is less well resolved in the time-domain DAE results than in the space-domain DAE results.

Our space-domain DAE model is especially important in cases where multiple cars trailing closely with their quasi-

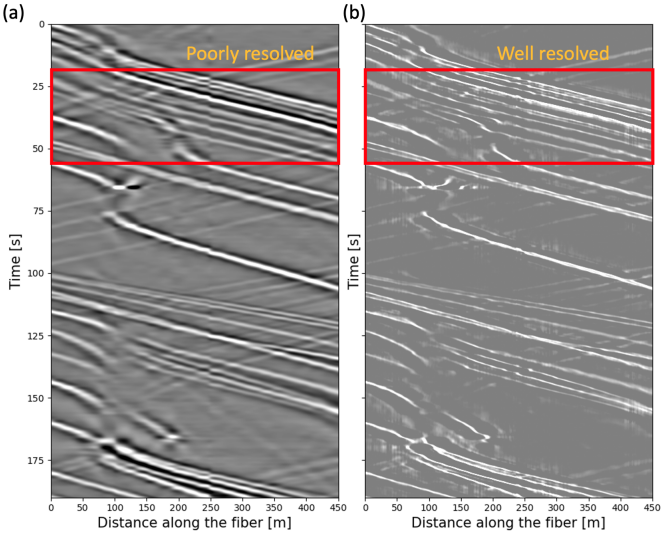


Fig. 10. (a) Quasi-static recordings under a heavy traffic conditions where signals of closely traveling cars start overlapping producing complicated patterns; (b) Deconvolution results from the space-domain DAE model.

static signals overlapping on each other. As we can see in the red box of Figure 10 (a), the overlaps produce complicated patterns posing ambiguity for individual car identification. On the other hand, we can see that results from the space-domain DAE model in (b), cars are much better resolved, which is critical to achieve a high-accuracy traffic counting.

C. Large-size vehicle monitoring

Heavier and longer vehicles, e.g. buses, trucks, and trains, generate quasi-static signals with much larger amplitudes and

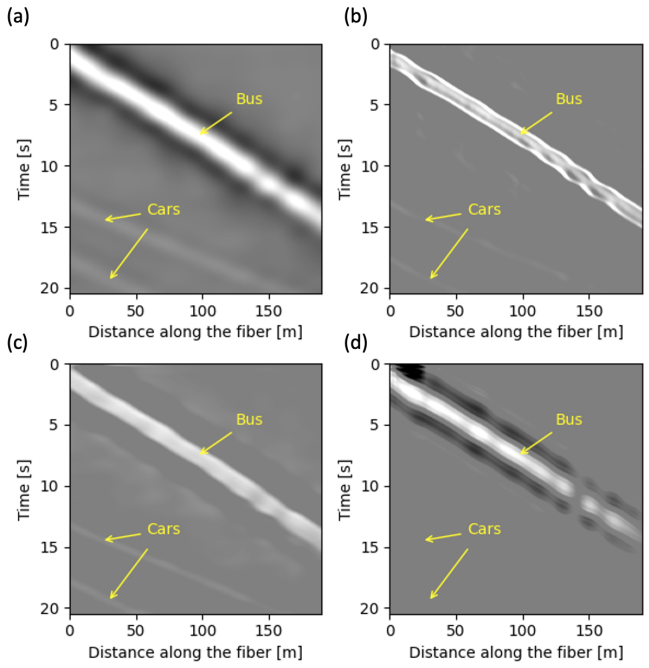


Fig. 11. (a) Signals of an 18-meter bus and two regular sized cars. Deconvolution results of (b) space-domain DAE model, (c) time-domain DAE model, and (d) FISTA, respectively.

wider spatial wavelets than personal vehicles (see Figure 11 (a) for visual comparisons between car and bus signals). The wider spatial wavelet can be viewed as the superposition of quasi-static signals of wheels at axles that are farther apart. Because of the wavelet difference, deconvolution of the data generated by such large vehicles along the time dimension with either the time-domain DAE model or FISTA yields undesirable artifacts as shown in [17]. Herein, we show that deconvolution along the spatial axis with the proposed space-domain DAE model properly resolves the axles. Figure 11 shows the recordings and deconvolution of an 18-meter three-axle bus. Performing spatial deconvolution with the proposed DAE model is more robust to vehicle sizes and axle numbers as we can see from (b). The bus can be recognized from the spatial-domain DAE model as three strong-energy trajectories corresponding to the three axles. The distance between the two most substantial peaks is about 18 m agreeing with the bus length, indicating the usage of our model for car axle and length characterization. We can see from (c) and (d) that the time-domain DAE model and the FISTA cannot resolve individual axles.

DAE models trained using only a vehicular traffic dataset could also be used to recover train axles at different bogies, which can be potentially useful to studies (e.g. [24]) on rail track health monitoring with the dynamic response of the train. Figure 12 (a) shows DAS recordings of a passing train recorded by a railroad side DAS array at somewhere in England. The train data displays multiple main lobes corresponding to train bogies because train axles at different bogies are much farther apart than car or bus axles. But it's challenging to resolve the two axles from a same bogie, as the length of a bogie is much shorter than the quasi-static wavelength. From the quasi-static signals in (a), we can see that the train first slowed down and then accelerated. The spatial DAE model in (b) yields the results with the highest resolution regardless of the train speed. In Figure 13, we compare the performance of the three methods by overlaying the untreated data and the three deconvolution results along the spatial dimension at a constant time step. Identified train axles at different bogies are numbered in the plot. We can see more clearly that the pulse from the space-domain DAE model is the most localized. Besides, we observe that the DAE models are robust to low-amplitude peaks (e.g. indicated as #10 in Fig. 13). Note the DAE models used here are trained on traffic dataset with a car wavelet. The future work would be to improve the results by training a train specific DAE model with wavelets of train axles.

D. Spatial non-stationarity

Our spatial model assumes that the car impulse response is stationary in space. However, the impulse response is a function of the recording system and the near-surface conditions surrounding the fiber optic cable. Thus, deconvolution of a DAS array covering heterogeneous near-surface conditions using a single stationary spatial impulse response could be suboptimal. Figure 14 (a) shows DAS recordings of traffic in downtown San Jose City (SJC). (b) shows deconvolution results of the proposed space-domain DAE model with a simulated car impulse response kernel. The deconvolution results look encouraging for most of the fiber, which can be contributed to the consistency of the assumed wavelet to the real wavelet. However, we can see that the results in the red box are relatively poorly resolved. The low resolution is more noticeable from the zoomed-in view of data and deconvolution results in the red box in Figure 15 (a) and (b). The under-performance could be explained as the wavelet inconsistency due to heterogeneity of the near-surface properties. Assuming that the near-surface properties, and in turn the car response, at each channel location is stationary in time. A possible solution would be to estimate the spatial car response at different parts of the fiber by averaging responses of multiple passing cars around each location. To verify this, we employ a local maximum finding algorithm as a simple car detector for the quasi-static signals. We average the responses of six identified cars passing the area in the red box of Figure 14 as the impulse response input to the U-net model. We retrain the U-net using this estimated kernel, which produces a sharper results in Figure 15, indicating the effectiveness of our approach.

V. Discussion and Conclusions

This paper focuses on the application of traffic monitoring with car-induced quasi-static signals recorded by an urban DAS array. To denoise the data and to reduce the interference among closely traveling cars, we propose a self-supervised convolutional U-Net model (space-domain DAE model) that can compress the bell-shaped quasi-static signals into sharp pulses and remove the background noises. The goal is achieved through spatial deconvolution with an assumed

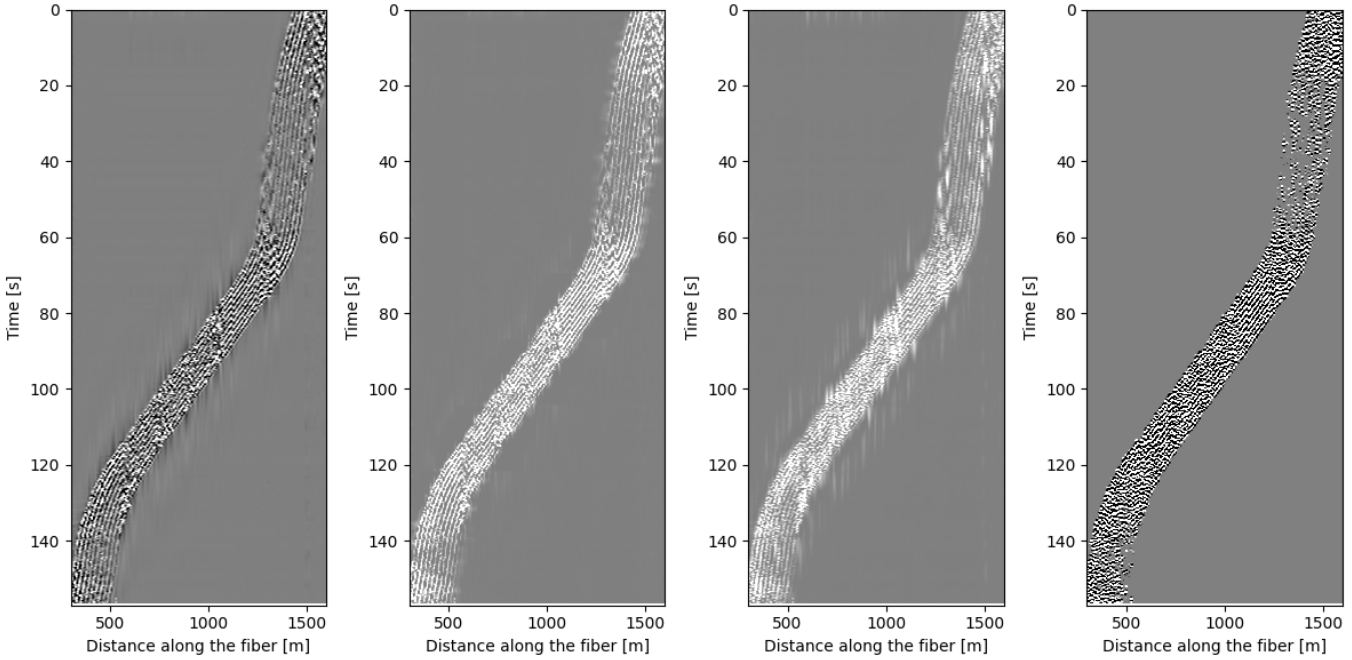


Fig. 12. (a) Signals of a passing train recorded by a DAS array along a railroad in England. Deconvolution results of (b) space-domain DAE model, (c) time-domain DAE model, and (d) FISTA, respectively.

Fig. 13. Constant-time train deconvolution results and the corresponding input taken from Figure 12. Numbers indicate the identified train axles at different train bogies.

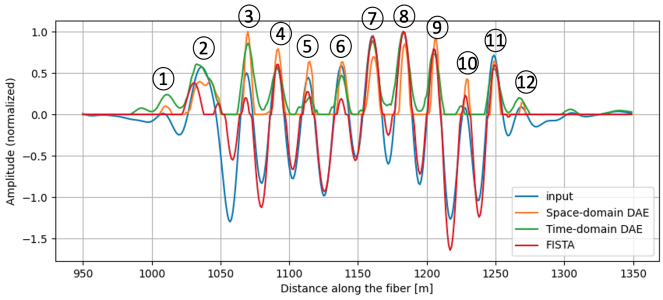
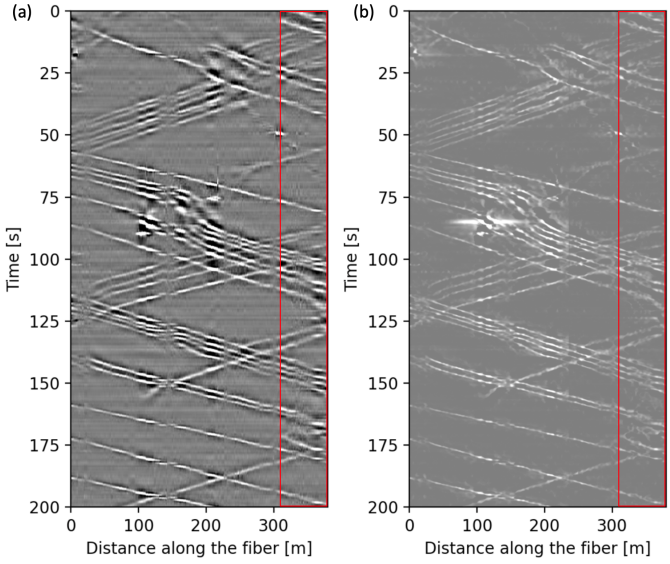


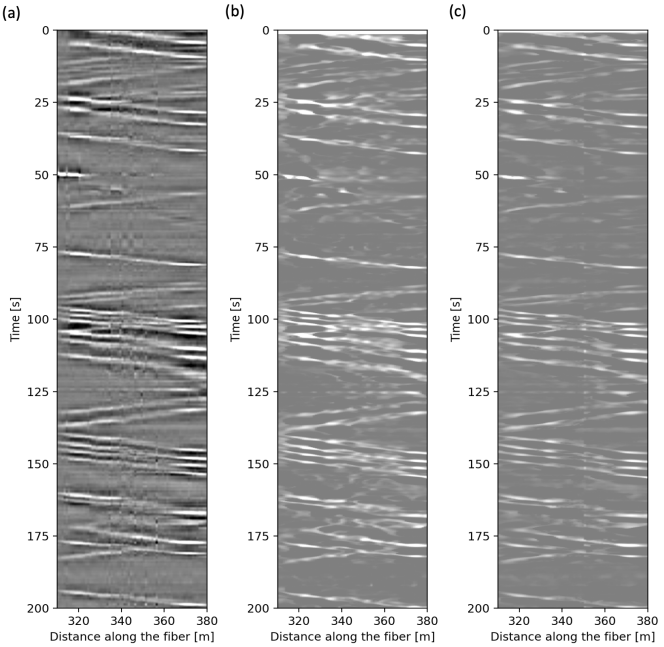
Fig. 14. Deconvolution of traffic recording in downtown SJC via a space-domain DAE model with a simulated impulse response stationary in space. (a) Input data; (b) deconvolution results. The red box indicates results that is poorly resolved due to the spatial heterogeneity of the car wavelets.



spatial wavelet of the quasi-static signals, which is a major difference from the previously proposed time-domain DAE model.

This paper shows that using the spatial instead of temporal kernel is advantageous because it is invariant to car speed variation. This leads to an improved precision robustness of tracking cars with varying speed, which is essential for driving behavior identification and accident detection. The benefits of our DAE model are more obvious for heavy-traffic conditions where car signals overlap. With our DAE model, compressing signals of individual cars to sharp pulses, traffic patterns are much better revealed. Besides, we show that our space-domain DAE model is robust to deconvolve signals of large-size and heavy vehicles. Being able to identify large vehicles can be helpful to extract low-frequency surface waves to image the velocity structure down to hundreds of meters beneath the fiber. Additionally, we demonstrate the application of the DAE model to identify rail train axles, which can be useful to rail track health monitoring using train's dynamic response. Lastly, we point out that a limitation of our spatial-domain DAE model arises when car impulse responses at different locations along the fiber vary due to the soil heterogeneity and/or fiber properties. We address the issue by extracting

Fig. 15. (a) and (b) are respectively zoomed-in view of the data and deconvolution results in the red box shown in Figure 14. (c) shows the deconvolution results with a statistical estimate of spatial wavelet at the distance along the fiber of 340 m.



car wavelets using statistical averaging, and train separated networks. Future work would be adapting the DAE model to different parts of the fiber using location-dependent spatial wavelets.

ACKNOWLEDGMENTS

This research was supported financially by the affiliates of the Stanford Exploration Project and the Stanford Sustainability Initiative and the UPS Foundation Endowment Fund. Jingxiao Liu is supported by Leavell Fellowship on SustainableBuilt Environment at Stanford University. Martijn van den Ende was supported by the French government through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002. The interrogator unit was loaned to us by OptaSense Inc. We thank Martin Karrenbach, Victor Yartsev, and Lisa LaFlame from Optasense, as well as the Stanford ITS fiber team, and in particular Erich Snow, for crucial help with the Stanford DAS-2 experiment. We also thank the Stanford School of Earth IT team for hosting the interrogator in the Scholl computer room.

REFERENCES

[1] K. Robert, "Video-based traffic monitoring at day and night vehicle features detection tracking," in *2009 12th International IEEE Conference on Intelligent Transportation Systems*, 2009, pp. 1–6.

[2] S. R. E. Datondji, Y. Dupuis, P. Subirats, and P. Vasseur, "A survey of vision-based traffic monitoring of road intersections," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 10, pp. 2681–2698, 2016.

[3] 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, pp. 149–158, 2006, theme Issue: Airborne and Spaceborne Traffic Monitoring. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271606001146

[4] N. K. Jain, R. Saini, and P. Mittal, "A review on traffic monitoring system techniques," *Soft Computing: Theories and Applications*, pp. 569–577, 2019.

[5] S.-T. Jeng and L. Chu, "A high-definition traffic performance monitoring system with the inductive loop detector signature technology," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2014, pp. 1820–1825.

[6] T. Cherrett, H. Bell, and M. McDonald, "Traffic management parameters from single inductive loop detectors," *Transportation Research Record*, vol. 1719, no. 1, pp. 112–120, 2000. [Online]. Available: https://doi.org/10.3141/1719-14

[7] J. Zhang, Y. Lu, Z. Lu, C. Liu, G. Sun, and Z. Li, "A new smart traffic monitoring method using embedded cement-based piezoelectric sensors," *Smart Materials and Structures*, vol. 24, no. 2, p. 025023, 2015.

[8] Z.-X. Li, X.-M. Yang, and Z. Li, "Application of cement-based piezoelectric sensors for monitoring traffic flows," *Journal of transportation engineering*, vol. 132, no. 7, pp. 565–573, 2006.

[9] R. Posey, "Rayleigh scattering based distributed sensing system for structural monitoring," in *Fourteenth International Conference on Optical Fiber Sensors*, A. G. Mignani and H. C. Lefèvre, Eds., vol. 4185, International Society for Optics and Photonics. SPIE, 2000, p. 41850E. [Online]. Available: https://doi.org/10.1117/12.2302157

[10] A. Masoudi and T. P. Newson, "Contributed review: Distributed optical fibre dynamic strain sensing," *Review of Scientific Instruments*, vol. 87, no. 1, p. 011501, 2016. [Online]. Available: https://doi.org/10.1063/1.4939482

[11] S. Yuan, A. Lellouch, R. G. Clapp, and B. Biondi, "Near-surface characterization using a roadside distributed acoustic sensing array," *The Leading Edge*, vol. 39, no. 9, pp. 646–653, 09 2020. [Online]. Available: https://doi.org/10.1190/tle39090646.1

[12] N. J. Lindsey, S. Yuan, A. Lellouch, L. Gualtieri, T. Lecocq, and B. Biondi, "City-scale dark fiber das measurements of infrastructure use during the covid-19 pandemic," *Geophysical Research Letters*, vol. 47, no. 16, p. e2020GL089931, 2020, e2020GL089931 2020GL089931. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL089931

[13] S. Yuan, J. Liu, H. Young Noh, and B. Biondi, "Urban system monitoring using combined vehicle onboard sensing and roadside distributed acoustic sensing," in *SEG International Exposition and Annual Meeting*, vol. Day 1 Sun, September 26, 2021, 09 2021, d011S137R003. [Online]. Available: https://doi.org/10.1190/segam2021-3584136.1

[14] X. Wang, Z. Zhan, E. F. Williams, M. G. Herráez, H. F. Martins, and M. Karrenbach, "Ground vibrations recorded by fiber-optic cables reveal traffic response to covid-19 lockdown measures in pasadena, california," *Communications Earth Environment*, vol. 2, 2021.

[15] 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*, 2021, pp. 1104–1108.

[16] F. Huot, E. R. Martin, and B. Biondi, *Automated ambient-noise processing applied to fiber-optic seismic acquisition (DAS)*, 2018, pp. 4688–4692. [Online]. Available: https://library.seg.org/doi/abs/10.1190/segam2018-2997880.1

[17] M. van den Ende, A. Ferrari, A. Sladen, and C. Richard, "Deep deconvolution for traffic analysis with distributed acoustic sensing data," 2021.

[18] Y. Fung, "Foundations of solid mechanics prentice-hall," *Inc, New Jersey*, 1965.

[19] T. Tanimoto and J. Wang, "Theory for deriving shallow elasticity structure from colocated seismic and pressure data," *Journal of Geophysical Research: Solid Earth*, vol. 124, no. 6, pp. 5811–5835, 2019. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JB017132

[20] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0:

Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[21] A. Chambolle, R. De Vore, N.-Y. Lee, and B. Lucier, "Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 319–335, 1998.

[22] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.20042

[23] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm with application to wavelet-based image deblurring," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 693–696.

[24] G. Lederman, S. Chen, J. Garrett, J. Kovačević, H. Y. Noh, and J. Bielak, "Track-monitoring from the dynamic response of an operational train," *Mechanical Systems and Signal Processing*, vol. 87, pp. 1–16, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327016302230

VI. Simple References

You can manually copy in the resultant .bbl file and set second argument of `\begin` to the number of references (used to reserve space for the reference number labels box).

VII. Biography Section

If you have an EPS/PDF photo (graphicx package needed), extra braces are needed around the contents of the optional argument to biography to prevent the LaTeX parser from getting confused when it sees the complicated `\includegraphics` command within an optional argument. (You can create your own custom macro containing the `\includegraphics` command to make things simpler here.)

If you include a photo:

Michael Shell Use `\begin{IEEEbiography}` and then for the 1st argument use `\includegraphics` to declare and link the author photo. Use the author name as the 3rd argument followed by the biography text.

If you will not include a photo:

John Doe Use `\begin{IEEEbiographynophoto}` and the author name as the argument followed by the biography text.