

EVENT CLASSIFICATION WITH CLASS-LEVEL GATED UNIT USING LARGE-SCALE PRETRAINED MODEL FOR OPTICAL FIBER SENSING

Noriyuki Tonami¹, Sakiko Mishima¹, Reishi Kondo¹, Keisuke Imoto², Tomoyuki Hino¹

¹NEC Corporation, Japan, ²Doshisha University, Japan

ABSTRACT

Optical fiber sensing is a technology in which sounds, vibrations, and temperature are detected using an optical fiber; especially the sounds/vibrations-aware sensing is called distributed acoustic sensing (DAS). DAS has the potential to capture various types of sounds and/or vibrations in wide areas, e.g., the ground, the sea, and a city area, in our everyday life. To precisely recognize the various types of events, e.g., whale calls, car horns, and wind, by DAS, therefore two problems. First, there is little publicly available data and few pretrained models for the various types of events. Second, the signal-to-noise ratio (SNR) of DAS data is lower than that of other sensor data, such as microphone data, because of optical noise and low sensitivity of DAS. To tackle the lack of DAS data, we first demonstrate a DAS simulation method where DAS observations are simulated by exploiting a microphone simulation. We then propose a method of event classification for DAS utilizing a pretrained audio recognition model, where none of the DAS data are used for training. Moreover, we advocate a class-level gated unit with the pretrained model to overcome the poor classification performance caused by the low SNR of the DAS data. In the proposed method, class probabilities, which are the output of the pretrained model, are employed for controlling priors of DAS, such as events of interest or optical noise. Directly controlling the class probabilities, which are non-black-box values, as priors enables us to utilize not only a pretrained model but also powerful human knowledge. To verify the performance of the proposed method, we conduct event classification, where we simulate observed signals by DAS with the ESC-50 dataset. Experimental results show that the accuracy of the proposed method is improved by 36.75 percentage points compared with that of conventional methods.

Index Terms— Optical fiber sensing, distributed acoustic sensing, deep neural network, sound event classification

1. INTRODUCTION

Optical fiber sensing is the detection of sounds and/or vibrations using an optical fiber [1, 2], which is known as distributed acoustic sensing (DAS) or phase-sensitive optical time domain reflectometry (ϕ -OTDR). Optical fiber sensing, including DAS, is superior to other methods in terms of the scalability of the sensing area, the electromagnetic resistance, and the usability of existing optical fibers. Because of these features, DAS is widely used for various applications, especially detecting vibrations, such as whale call detection [3], structural health monitoring [4], seismic activity monitoring [5], border monitoring [6], and pole localization [7]. Owen *et al.* [6] introduced the DAS-based system for distinguishing people, vehicles, and flying objects. Waagaard *et al.* [8] proposed large-scale sensing over 171 km using DAS. Ip *et al.* [1, 2] discussed the use of telecom cables for sensing sounds and/or vibrations.

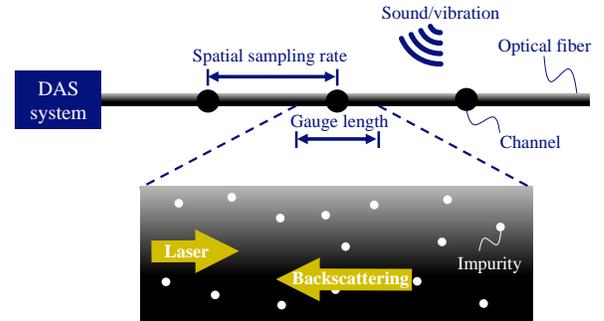


Figure 1: Overview of DAS system

Recently, deep neural networks (DNNs) have been in the spotlight in the field of DAS [7, 9, 10, 11]. DNN-based methods require a large amount of training data to achieve the high performance of DAS. To alleviate this limitation, Zhao *et al.* [10] proposed a data augmentation of DAS and its application to seismic data. Generative adversarial network (GAN)-based methods [11] have been studied to produce training data and are designed for seismic applications.

DAS has two problems in the precise recognition of various types of events, such as whale calls, dog barking, and footsteps. First, there is little publicly available data and few pretrained models for analyzing various types of event classes in the field of DAS. On the other hand, in communities where acoustic signal processing and statistical methods have been studied for various types of sounds [12], various types of dataset or pretrained model [13, 14, 15] are available. Second, the signal-to-noise ratio (SNR) is lower in DAS compared with methods using other sensors, such as microphones. The lower SNR of DAS is caused mainly by optical noise and the low sensitivity of DAS.

To address the lack of DAS data, we first demonstrate a DAS simulation method where a simulator of microphones is utilized for simulating DAS observations. We then propose an event classification method of DAS using a pretrained audio recognition model trained by microphone data. Moreover, we introduce a class-level gated unit with the pretrained model to tackle the problem of the low SNR of DAS data. In the proposed method, posteriors of the pretrained model are employed to control prior information, i.e., events of interest or optical noise, which can be directly manipulated by humans in the inference stage. The proposed gated unit that directly controls the probabilities of event classes, which are non-black-box values, enables us to utilize not only the pretrained model but also human knowledge.

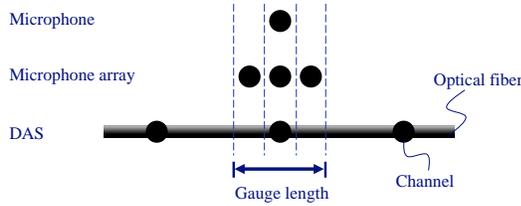


Figure 2: Comparison of DAS and microphone

2. DISTRIBUTED ACOUSTIC SENSING

2.1. Principle of DAS

In DAS, a sensor captures the phase change of a Rayleigh backscattering light wave. Figure 1 shows an overview of the DAS system. The backscattering is triggered by a coherent laser that collides with impurities in the optical fiber. The phase change of the backscattering is proportional to the optical fiber stretching [16], that is, sounds or vibrations that propagate through the optical fiber.

In DAS, the stretching of the optical fiber is measured over the gauge length (GL) L . The total phase change $\Delta\phi$ in L [17] is

$$\Delta\phi = \int_{-\frac{L}{2}}^{\frac{L}{2}} \epsilon(x) dx, \quad (1)$$

where $\epsilon(x)$ indicates the strain, that is, the observed sounds and/or vibrations along the optical fiber, at position x of the optical fiber. The point is that the observed signals depend on L . In general, larger L suppresses optical noise, although it distorts the observed signal of sounds or vibrations. An optical fiber sensor with GL of L is also interpreted as a linear sensor array in Eq. 1 where the directivity of the angle of a source signal and its distortion [18]. As can be seen in Fig. 1, arbitrary multiple sensing points can be set along the optical fiber in accordance with a predefined spatial sampling rate.

2.2. Difference between DAS and acoustical microphones

There are two main differences between DAS and microphones: optical noise and GL. The differences make the SNR of DAS data lower compared with that of microphone data. In the first difference, optical noise [19, 20], shot noise [2] is dominant because of the randomness of photons. The second difference is the idea of the gauge. A larger GL distorts signals observed by DAS. On the other hand, a smaller GL, i.e., $\lim_{L \rightarrow 0} \Delta\phi$, approximates a point sensor such as a microphone. As can be seen in Fig. 2, DAS data within the GL is thus regarded as a microphone array where channels are densely distributed. Note that the effects of the GL are evident in a single channel of DAS data, unlike a microphone array.

3. PROPOSED METHOD

In this section, we first introduce a simulation method of DAS observation to address the lack of DAS evaluation data for analyzing various types of events. Second, the event classification method for DAS utilizing the pretrained audio recognition model and the class-level gated unit are proposed.

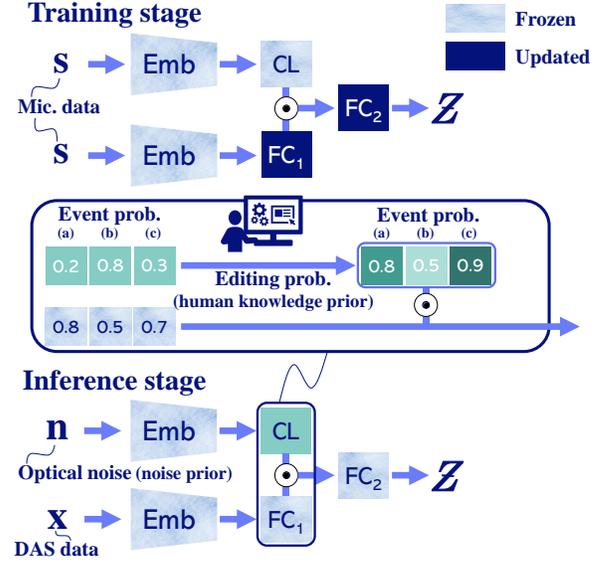


Figure 3: Overview of proposed method

3.1. DAS simulation utilizing microphone simulation

To simulate DAS, we employ multiple simulated microphones, i.e., a microphone array. This is because the sound propagation and the GL are easily implemented using tools of a microphone array simulation, e.g., pyroomacoustics [21] where distance attenuation, reverberation, or directivity is easily simulated. We simply take an average of signals observed by microphones of M channels.

$$\tilde{s}_{m'} = \frac{1}{M} \sum_{m=0}^{M-1} s_m. \quad (2)$$

Here, $\tilde{s}_{m'} \in \mathbb{R}^T$ indicates a DAS signal of the m' -th channel with T temporal frames. $s_m \in \mathbb{R}^T$ denotes a signal captured by a microphone m of the array. The average of the multiple channels corresponds to L in Eq. 1. $\tilde{s}_{m'}$ is distorted by the GL, i.e., no-delay-and-sum operation. As an example referring to Fig. 2, there is a single channel data value of DAS for each observed data of three microphones ($M = 3$).

The signal detected by DAS is reportedly expressed as [22]

$$\mathbf{x} = \tilde{s}_{m'} + \mathbf{n}, \quad (3)$$

where $\mathbf{n} \in \mathbb{R}^T$ represents the noise signal. $\mathbf{x} \in \mathbb{R}^T$ is the noisy signal that is corrupted by the GL and the noise signal. When \mathbf{n} is the shot noise of optical noise, it follows the Gaussian distribution $\mathcal{N}(\mu, \sigma)$.

3.2. Event classification of DAS with pretrained audio recognition model and class-level gated unit

To precisely classify various types of events under a lower SNR condition of DAS data, we propose the event classification method of DAS with the class-level gated unit utilizing the pretrained audio recognition model.

[Training stage] In the proposed method, only the data and pretrained model trained with the microphone data are used for the training to tackle the lack of DAS data. Figure 3 shows an overview

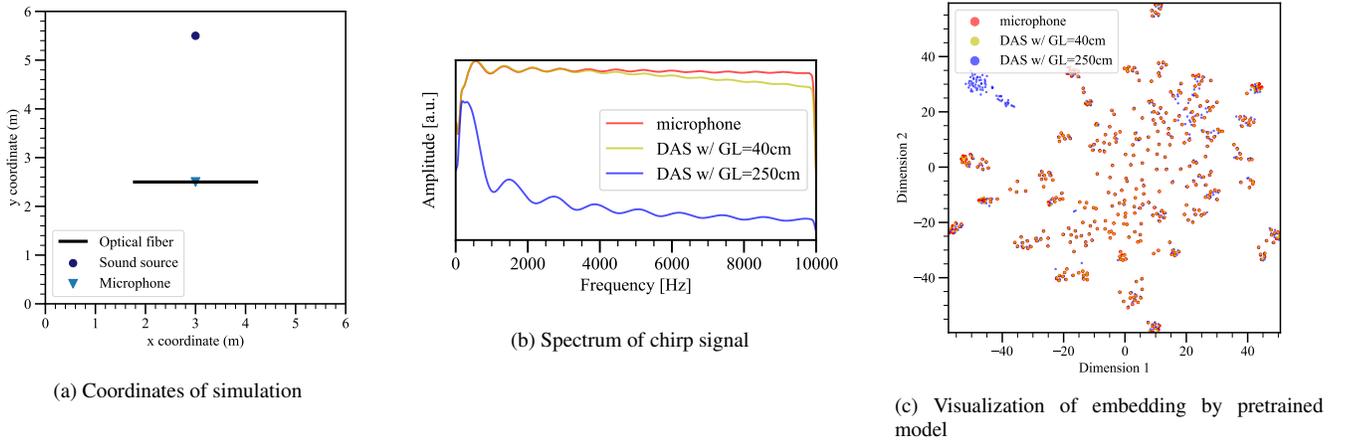


Figure 4: (a) Coordinates of simulation and (b), (c) results of preliminary experiments

of the proposed method. Given a pretrained audio recognition model of the microphone, its deep feature extractor $\text{Emb} : \mathbb{R}^T \rightarrow \mathbb{R}^D$ and the classification layer $\text{CL} : \mathbb{R}^D \rightarrow [0, 1]^E$ are defined. The posterior for event classes is then

$$\mathcal{P} = \text{CL}(\text{Emb}(\mathcal{A})) \big|_{\mathcal{A}=\mathbf{s}}, \quad (4)$$

where T , D , and E represent the temporal length of an audio clip, the number of dimensions of embedding, and the number of event classes of the pretrained model, respectively. \mathcal{A} denotes an arbitrary signal. When the model is trained with the microphone data, $\mathbf{s}_m \mid \exists m$ is used as \mathcal{A} . Given two fully connected layers (FCs) $\text{FC}_1 : \mathbb{R}^D \rightarrow \mathbb{R}^{E'}$ and $\text{FC}_2 : \mathbb{R}^{E'} \rightarrow \mathbb{R}^{E'}$, the class-level gated unit is

$$\mathcal{Z} = \text{FC}_2(\text{FC}_1(\text{Emb}(\mathcal{B})) \odot \mathcal{P}) \big|_{\mathcal{B}=\mathbf{s}}. \quad (5)$$

\odot and E' indicate elementwise multiplication and the number of target event classes of DAS, respectively. \mathcal{B} denotes an arbitrary signal. Equation 5 is similar to that of a gated linear unit (GLU) [23]. In our method, gating is conducted on the probabilities of event classes of the pretrained audio recognition model, which are non-black-box values.

[Inference stage] In inference stages, DAS data \mathbf{x} are used as \mathcal{A} and \mathcal{B} in Eqs. 4 and 5, respectively. \mathcal{P} can then be edited by human interaction. Directly controlling the interpretable values, i.e., the probabilities of event classes, enables us to utilize human knowledge as an additional prior. Human knowledge can enhance the classification performance with low SNR of DAS data since humans have rich knowledge compared with pretrained audio recognition models. For example, values of event classes of \mathcal{P} unrelated to a recording situation and/or events of interest can be directly masked with zero or a small value:

$$\mathcal{P} = (p_0, \dots, p_e, \dots, p_{E-1}), \quad (6)$$

where $p_e \in [0, 1]$ is the value of event e of the predefined event classes for the pretrained model and can be directly set by a human. Secondly, $\mathcal{A} \big|_{\mathcal{A}=\mathbf{n}}$ in Eq. 4 can be used for mitigating the optical

noise in the level of the class of the pretrained model in the inference stages:

$$\mathcal{P} = \max(\text{FC}_1(\text{Emb}(\mathcal{B})) - \text{CL}(\text{Emb}(\mathcal{A})), \mathbf{0}) \big|_{\mathcal{B}=\mathbf{x}, \mathcal{A}=\mathbf{n}}, \quad (7)$$

where $\max(\mathbf{a}, \mathbf{b})$ is a function that returns the larger value element of vectors \mathbf{a} and \mathbf{b} in an elementwise manner. $\mathbf{0}$ represents the E dimensional vector where all elements are zero. In Eq. 7, only denoised probabilities of event classes are expected to be passed through. In the inference stages, the softmax function is applied to \mathcal{Z} for obtaining the maximum value of the posteriors of the event classes.

4. EXPERIMENT

4.1. Experimental conditions

[Simulation procedure] To simulate DAS observations, we followed the procedure described in Sec. 3.1. We first simulated a linear microphone array and a sound source, as shown in Fig. 4a, using the pyroomacoustics toolbox [21]. By using pyroomacoustics, we can easily simulate the sound propagation and the idea of the GL. The microphone array consists of 250 channels at intervals of 1 cm. The observed signals of the channels are then averaged using Eq. 2. M was set to 40 ($L=40\text{cm}$) or 250 ($L=250\text{cm}$) in our experiment. Here, the center of the gauge with $M=40$ matches those with $M=250$. We finally obtained single channel data of DAS from the signals observed by M microphones using Eqs. 2 and 3. The signal of the sound source was omnidirectionally propagated. In our experiments, we did not simulate any reverberations or reflections of the optical fiber. Moreover, for the shot noise of optical noise, we use Gaussian noise $\sim \mathcal{N}(0, 1)$ with variable SNRs.

[Dataset, classification model, and acoustic feature] We used the ESC-50 dataset [13] to evaluate the performance of our methods. ESC-50 comprises 4,000 5-second audio clips with 50 event classes. For the classification model, we used pretrained CNN14 in PANNs [15]. PANNs were trained using AudioSet [14] where event classes are organized in a hierarchy, i.e., ontology. In our experiment, Emb and CL are those of CNN14 with the frozen parameters, where D and E are set to 2,048 and 527, respectively. E' is set to 50,

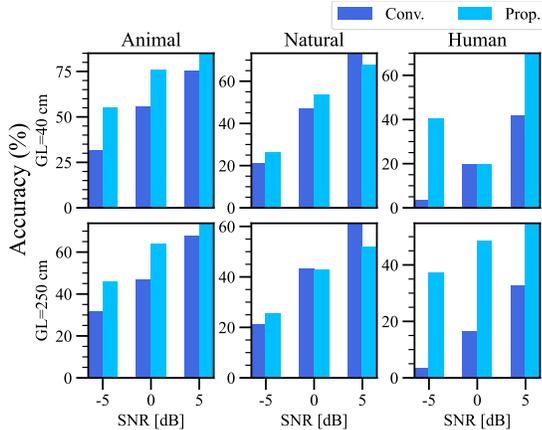


Figure 5: Accuracy (%) of event classification for each target major category with directly controlling gate

which is the number of classes in ESC-50. All clips are downsampled to 32kHz in our experiment. As acoustic features, we used 64-dimensional log-mel energies calculated for every 1,024 sample window and 320 sample hop sizes.

[Training and evaluation] We conducted the 5-fold cross-validation with the ESC-50 dataset. In all experiments, we trained the models using $(M/2)$ -th-channel signals of the microphone array before all of the channels were averaged. Note that the $(M/2)$ -th channel is located at the center of the microphone array. This is because we aim to reproduce the situation of the lack of DAS data. In the inference stages, we used only the DAS simulation data, which were averaged with Eqs. 2 and 3. In the training stages, the parameters of only FC_1 and FC_2 were updated by backpropagation using Adam [24]. In all FCs, except for FC_2 , ReLU activation was used.

4.2. Experimental results

[Comparison between DAS and microphone data] We first confirm the amplitude spectrum of a chirp signal in the DAS and microphone simulations. Figure 4b shows the 50-to-10,000Hz spectrum. As shown in the figure, $GL=40$ cm does not considerably distort the signal compared with $GL=250$ cm. Furthermore, in Fig. 4c, the embedding vectors of DAS and microphone data obtained by Emb of PANNs [15] are visualized by t-SNE [25]. As can be seen in the figure, there is little difference between the distributions of the microphone and DAS data of $GL=40$ cm. Even when GL is set to 250cm, most of the embeddings of DAS are overlapped with those of the microphone. The results prove that microphone data are similar to DAS data except for optical noise, obtained using large-scale pretrained audio recognition models.

[Event classification with controlling gate] In this experiment, we directly control \mathcal{P} to evaluate the performance of targeting an event of interest. To conduct this experiment, we utilize the overlap between the major categories of ESC-50 and the ontology of AudioSet. In the major categories of ESC-50, we focus on “Animal,” “Natural,” and “Human,” which correspond to “Animal,” “Natural sounds,” and “Human sounds” of the AudioSet ontology. When events in a major category of ESC-50 are targeted, p_e in Eq. 6 is set to 1; otherwise, 0. For example, when event classes of the major category “Human” of ESC-50 are targeted, p_e in Eq. 6 corresponding to “Human sounds” of the AudioSet ontology is set to 1; otherwise, 0. Figure 5 indicates the results of event classification with

Table 1: Accuracy (%) of event classification with denoising optical noise

	SNR [dB]					
	-5		0		5	
	GL [cm]					
Conv.	40	250	40	250	40	250
w/ spectral subtract. [26]	18.90	15.65	39.00	32.00	59.00	48.45
w/ Wiener filter [27]	23.60	24.25	54.65	37.45	69.65	39.50
Prop. w/ Eq. 7	21.35	16.65	34.15	23.90	47.35	32.20
	32.45	27.55	51.95	40.05	64.50	51.50

Eq. 6 in terms of each major category. “Conv.” indicates CNN14 [15] fine-tuned with ESC-50 of microphone data, where the last two FCs were trained, as described in [15]. “Prop.” represents the proposed method with directly controlled \mathcal{P} . The results show that the proposed method outperformed the conventional method in terms of classification accuracy. In particular, our method improved the accuracy of “Human” by 36.75 percentage points compared with that of the conventional method under the condition of $SNR = -5$ dB and $GL = 40$ cm. On the other hand, the “Natural” class is misclassified when using the proposed method under some conditions. This is because the “Natural” class, e.g., wind or rain, possibly co-occurs with other classes. In other words, the proposed method with Eq. 6 might discard the information of co-occurrence among event classes.

[Event classification with denoising] In this experiment, we evaluated the denoising performance for event classification of DAS data with optical noise signals. Note that none of the DAS data were used for training models to simulate the lack of DAS data. We thus employed non-machine-learning-based denoising methods for the comparison. “Prop.” represents the proposed method where \mathcal{P} was produced by Eq. 7 with the optical noise \mathbf{n} . Table 1 shows the results of event classification with denoising optical noise. The results reveal that the classification performance is improved when using the proposed method compared with the conventional methods. In particular, the more degraded signals, i.e., lower SNR and/or larger GL, receive greater benefits from the proposed method with Eq. 7. This is because the proposed method does not further distort the signals distorted by the GL, unlike the conventional methods. Moreover, the proposed denoising method can utilize the statistical information of the pretrained model, unlike the conventional methods.

5. CONCLUSION

In this paper, we proposed the event classification of DAS data utilizing the pretrained audio recognition model with the class-level gated unit for accurately classifying various types of events under low SNR conditions without DAS training data. In the proposed method, the class-level outputs of the pretrained model, which are non-black-box values, are employed for controlling priors of DAS data, that is, the optical noise and/or events of interest. This enables us to exploit not only the statistical information of the pretrained model but also human knowledge. To evaluate the performance of the proposed method, we conducted event classification where signals observed by DAS were simulated with the ESC-50 dataset. Experimental results show that the accuracy of event classification by the proposed method is improved by 36.75 percentage points compared with that of the conventional methods.

6. REFERENCES

- [1] E. Ip, Y. Huang, M. Huang, M. Salemi, Y. Li, T. Wang, Y. Aono, G. Wellbrock, and T. Xia, “Distributed fiber sensor network using telecom cables as sensing media: Applications,” *Proc. Optical Fiber Communications Conference and Exhibition (OFC)*, pp. 1–3, 2021.
- [2] E. Ip, J. Fang, Y. Li, Q. Wang, M. Huang, M. Salemi, and Y. Huang, “Distributed fiber sensor network using telecom cables as sensing media: technology advancements and applications,” *Journal of Optical Communications and Networking (JOCN)*, vol. 14, no. 1, pp. 61–68, 2022.
- [3] L. Bouffaut, K. Taweasantanon, H. J. Kriesell, R. A. Rorstadbotnen, J. R. Potter, M. Landro, S. E. Johansen, J. K. Brenne, A. Haukanes, O. Schjelderup, and F. Storvik, “Eavesdropping at the speed of light: Distributed acoustic sensing of baleen whales in the arctic,” *Frontiers in Marine Science*, vol. 9, pp. 1–13, 2022.
- [4] H. Peter, X. James, Z. Shenghan, D. Matthew, L. Linqing, S. Kenichi, P. Carlo, Z. Christian, M. Demetrio, F. Fabio, L. Francisco, and M. Chris, “Dynamic structural health monitoring of a model wind turbine tower using distributed acoustic sensing (DAS),” *Journal of Civil Structural Health Monitoring*, vol. 11, pp. 833–849, 2021.
- [5] T. Parker, S. Shatalin, and M. Farhadiroushan, “Distributed acoustic sensing a new tool for seismic applications,” *First Break*, vol. 32, no. 2, pp. 61–69, 2014.
- [6] A. Owen, G. Duckworth, and J. Worsley, “Optasense: Fibre optic distributed acoustic sensing for border monitoring,” *Proc. European Intelligence and Security Informatics Conference*, pp. 362–364, 2012.
- [7] Y. Lu, Y. Tian, S. Han, E. Cosatto, S. Ozharar, and Y. Ding, “Automatic fine-grained localization of utility pole landmarks on distributed acoustic sensing traces based on bilinear resnets,” *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4675–4679, 2021.
- [8] O. Waagaard, E. Rønnekleiv, A. Haukanes, F. Stabo-Eeg, D. Thingbø, S. Forbord, S. Aasen, and J. Brenne, “Real-time low noise distributed acoustic sensing in 171km low loss fiber,” *OSA Continuum*, vol. 4, no. 2, pp. 688–701, 2021.
- [9] L. Shiloh, A. Eyal, and R. Giryas, “Deep learning approach for processing fiber-optic das seismic data,” *Proc. International Conference on Optical Fiber Sensors*, pp. 1–4, 2018.
- [10] Y. Zhao, Y. Li, and N. Wu, “Data augmentation and its application in distributed acoustic sensing data denoising,” *Geophysical Journal International*, vol. 228, no. 1, pp. 119–133, 2021.
- [11] A. Venketeswaran, N. Lalam, J. Wuenschell, P. Ohodnicki, M. Badar, K. Chen, P. Lu, Y. Duan, B. Chorpeningand, and M. Buric, “Recent advances in machine learning for fiber optic sensor applications,” *Advanced Intelligent System*, vol. 4, no. 1, pp. 1–24, 2022.
- [12] <https://dcase.community/>.
- [13] J. Piczak, “ESC: Dataset for environmental sound classification,” *Proc. the 23rd Annual ACM Conference on Multimedia (ACMM)*, pp. 1015–1018, 2015.
- [14] J. Gemmeke, D. Ellis, D. Freedman, A. Jansen, W. Lawrence, R. Moore, M. Plakal, and M. Ritter, “Audio set: An ontology and human-labeled dataset for audio events,” *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 776–780, 2017.
- [15] Q. Kong, Y. Cao, T. Iqbal, Y. Wang, W. Wang, and M. Plumbley, “PANNs: Large-scale pretrained audio neural networks for audio pattern recognition,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)*, vol. 28, pp. 2880–2894, 2020.
- [16] Z. Zhong, F. Wang, M. Zong, Y. Zhang, and X. Zhang, “Dynamic measurement based on the linear characteristic of phase change in ϕ -otdr,” *IEEE Photonics Technology Letters*, vol. 31, no. 14, pp. 1191–1194, 2019.
- [17] T. Dean, T. Cuny, and A. Hartog, “The effect of gauge length on axially incident p-waves measured using fibre optic distributed vibration sensing,” *Geophysical Prospecting*, vol. 65, pp. 184–193, 2016.
- [18] W. Li, Y. Chen, Y. Liang, Y. Lu, and Z. Meng, “Directivity dependence of a distributed fiber optic hydrophone on array structure,” *Sensors*, vol. 22, no. 6297, pp. 1–12, 2022.
- [19] J. Zhou, Z. Pan, Q. Ye, H. Cai, R. Qu, and Z. Fang, “Characteristics and explanations of interference fading of a ϕ -otdr with a multi-frequency source,” *Journal of Lightwave Technology*, vol. 31, no. 17, pp. 2947–2954, 2013.
- [20] A. Hartog, O. Kotov, and L. Liokumovich, “The optics of distributed vibration sensing,” *Proc. Second EAGE Workshop on Permanent Reservoir Monitoring 2013–Current and Future Trends*, pp. 1–5, 2013.
- [21] R. Scheibler, E. Bezzam, and I. Dokmanić, “Pyroomacoustics: A python package for audio room simulations and array processing algorithms,” *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 351–355, 2018.
- [22] L. Jiajing, W. Zhaoyong, L. Bin, W. Xiao, L. Luchuan, Y. Qing, Q. Ronghui, and C. Haiwen, “Distributed acoustic sensing for 2d and 3d acoustic source localization,” *Optics Letters*, vol. 44, no. 7, pp. 1690–1693, 2019.
- [23] Y. Dauphin, A. Fan, M. Auli, and D. Grangier, “Language modeling with gated convolutional networks,” *Proc. International conference on machine learning (ICML)*, pp. 933–941, 2017.
- [24] D. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *Proc. International Conference on Learning Representations (ICLR)*, 2015.
- [25] L. Maaten and G. Hinton, “Visualizing data using t-SNE,” *Journal of machine learning research*, vol. 9, no. 11, pp. 1–27, 2008.
- [26] M. Berouti, R. Schwartz, and J. Makhoul, “Enhancement of speech corrupted by acoustic noise,” *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 208–211, 1979.
- [27] J. Lim and A. Oppenheim, “All-pole modeling of degraded speech,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)*, vol. 26, no. 3, pp. 197–210, 1978.