Early Safety Warnings for Long-Distance Pipelines: A Distributed Optical Fiber Sensor Machine Learning Approach

Tsinghua University

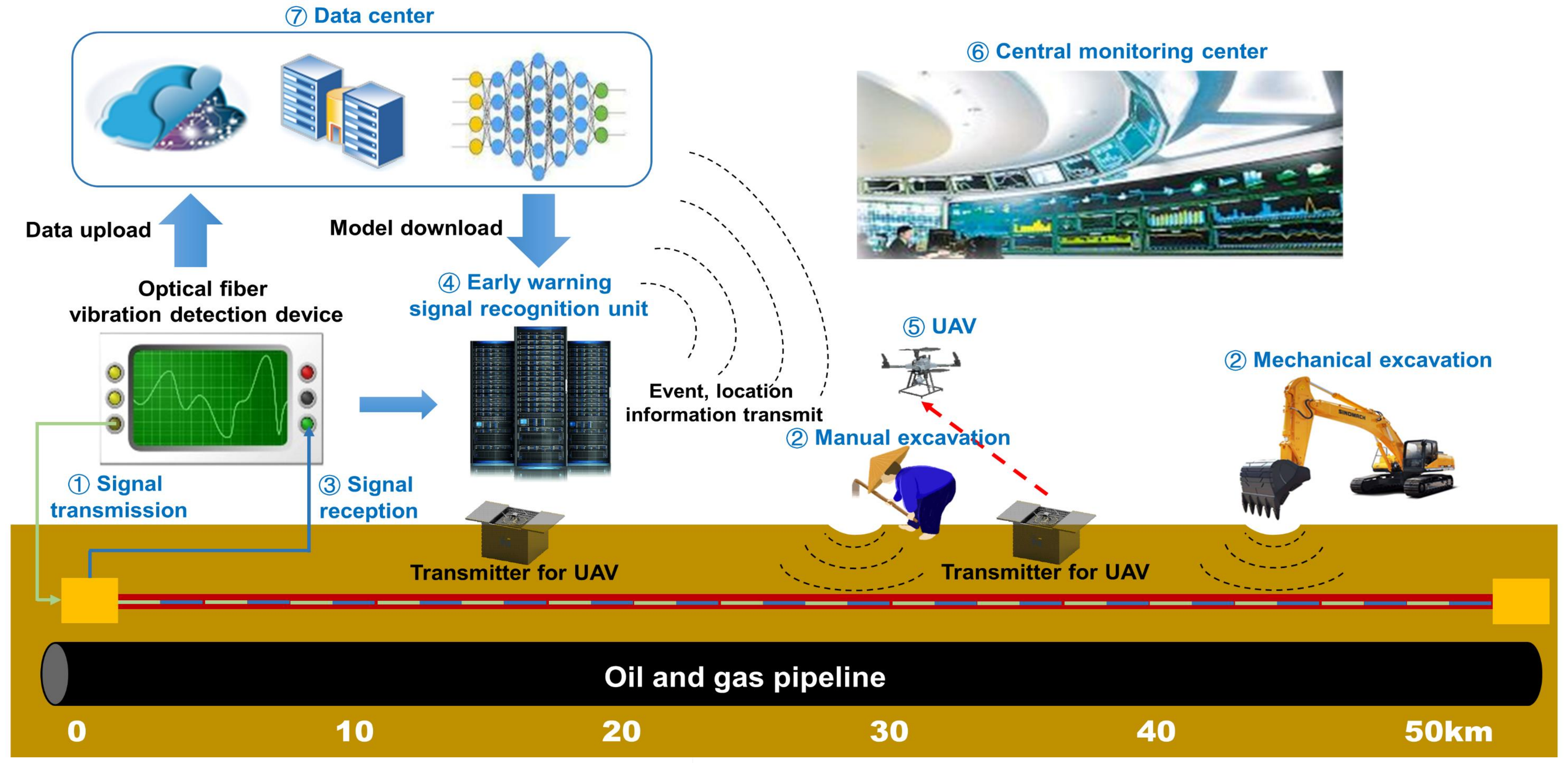
35th AAAI Conference on Artificial Intelligence, 2021

Association for the Advancement of Artificial Intelligence

Yiyuan Yang[1], Yi Li[1], Taojia Zhang[2], Yan Zhou[2], Haifeng Zhang[3*]
[1]International Graduate School at Shenzhen, Tsinghua University, Shenzhen, China
[2]PetroChina Pipeline Company, Langfang, China, [3]Research Institute of Tsinghua, Pearl River Delta, Guangzhou, China
yangyy19@mails.tsinghua.edu.cn, liyi@sz.tsinghua.edu.cn, {zhangtj01, kjzhouyan}@petrochina.com.cn, zhanghf@tsinghua-gd.org

Problem

Ensuring the safety of energy pipelines is related to the energy supply, environmental protection and the stability of the economy and society. Pipeline safety early warning (PSEW) systems aim to automatically identify and locate damage events on energy pipelines and replace traditional, inefficient manual inspection methods. However, existing systems cannot achieve universality for various complex environments because they are sensitive to the spatiotemporal stability of the signal obtained by distributed sensors at various locations and times. Our research aims to improve the identify and location algorithm through ML algorithm based on our novel PSEW system, as shown in the figure below.

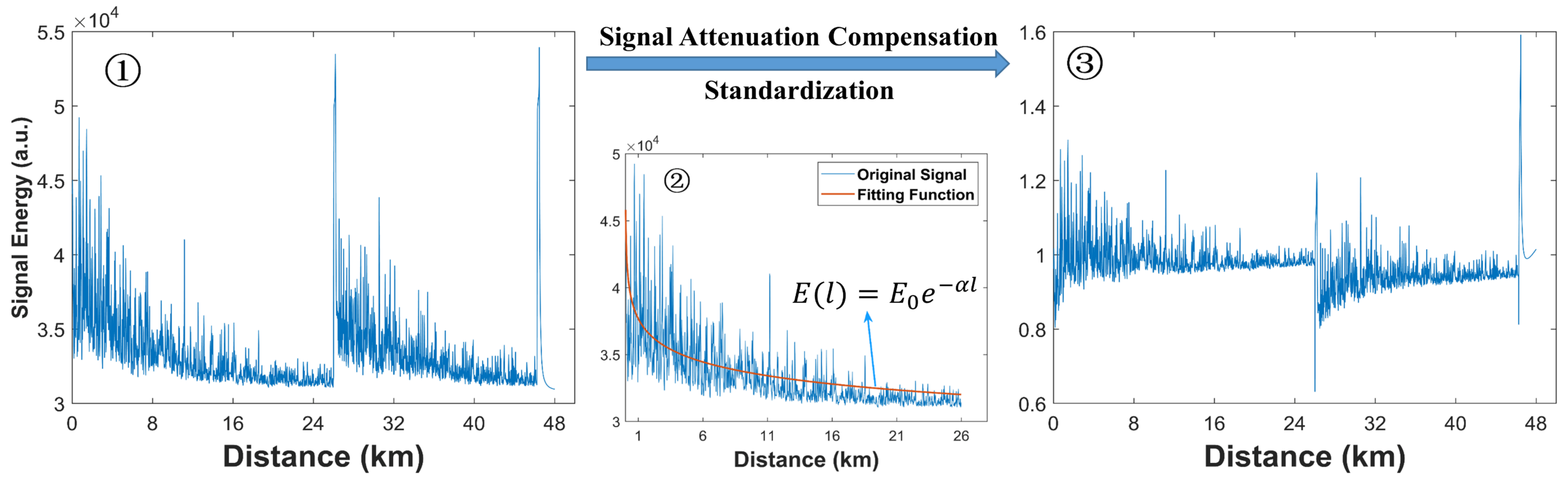


Contributions

- Two complementary features calculation methods based on the spatiotemporal information of distributed signals
- A novel Deep Learning model for real-time action recognition and spatiotemporal localization of damage events
- Our method has better real-time environmental adaptability, wider deployment, greater extendibility for various hardware, and model performance than SOTA based on real sites experiments with strong noise, weak signals, and signal fluctuations.

Methodology

• Attenuation Compensation

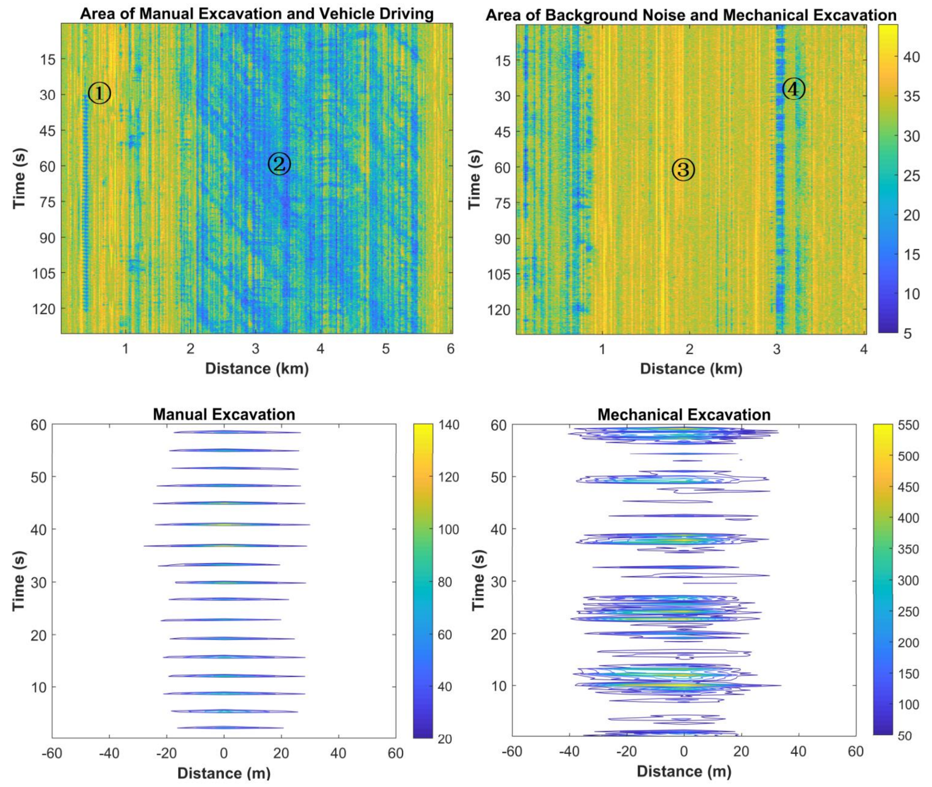


• Feature Generator

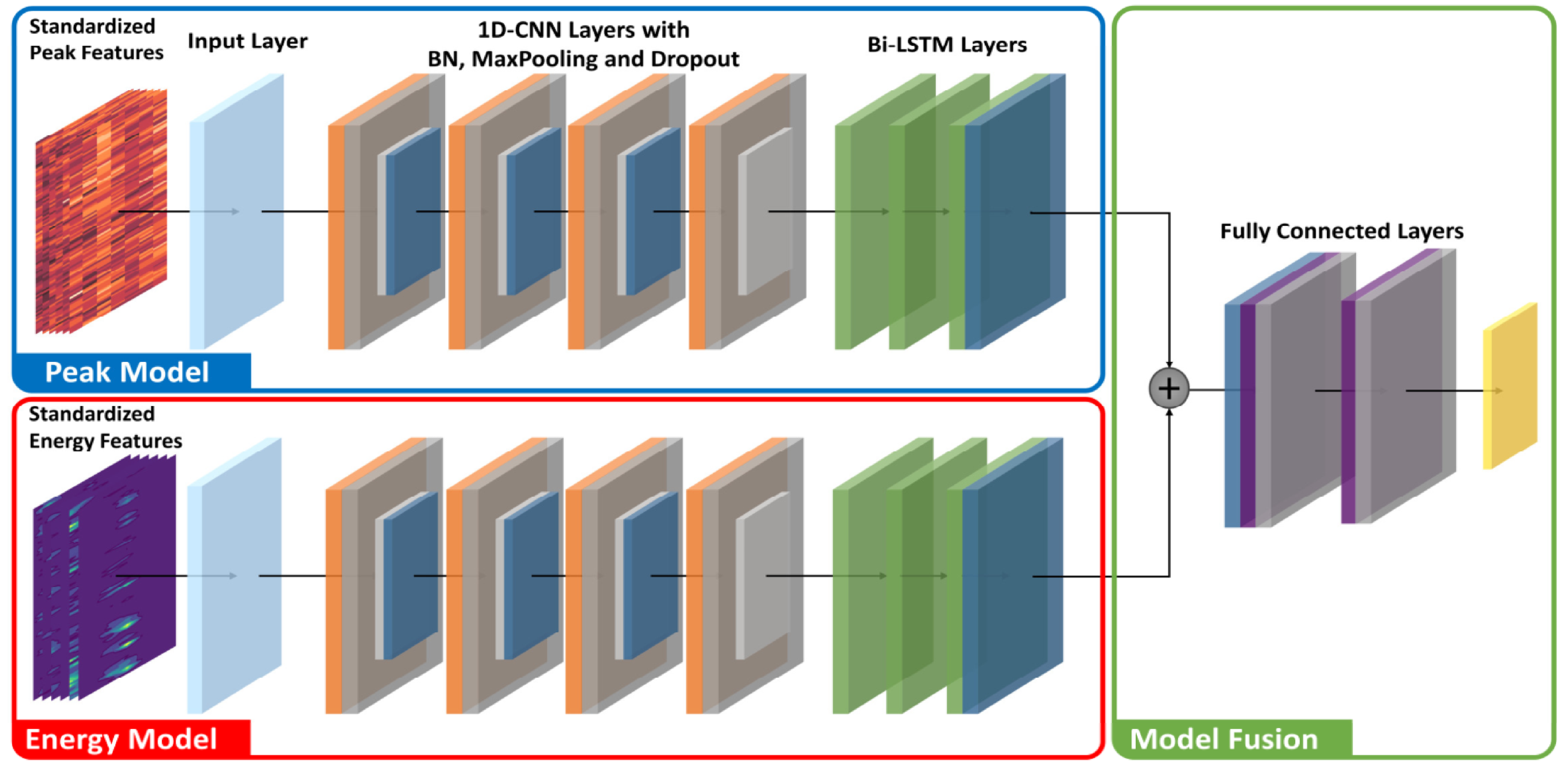
Algorithm I Matrix of Peak and Energy Features M_{peak}, M_{energy}

Input: Origin data X, Background noise data X_{base}.
Output: Matrix of Peak and Energy Features M_{peak}, M_{energy}
variable: Length of window and step N_{win}, N_{step}, Number of observation points L, Number of data in time dimension T, Number of observation points and windows to be considered $N_{d-point}$, N_{d-win}, Threshold α and β.

Attenuation compensation and standardization of X and X_{base}.
for each $i = 1, \ldots, L$ **do**
  for each $j = 1, \ldots, \frac{T}{N_{win}}$ **do**
    $F_{peak}[i,j] \leftarrow$ **Count peak**$(X[i, j * N_{step} : j * N_{step} + N_{win}])$
    for each $k = 1, \ldots, N_{win} - 2$ **do**
      $F_{\text{energy}}[i,j] \leftarrow \frac{\sum(\text{data}[k+\alpha] - \text{data}[k])^2}{T * X_{base}}$
    end for
    Set $F_{\text{energy}}[i,j] \leftarrow 1$ **if** $F_{peak}[i,j] > \beta$
  end for
end for
for each $m = \frac{N_{d-point}}{2}, \ldots, L - \frac{N_{d-point}}{2}$ **do**
  for each $n = 1, \ldots, \frac{T}{N_{win} * N_{d-win}}$ **do**
    $M_{peak}, M_{energy} \leftarrow F_{peak}, F_{energy}[m - \frac{N_{d-point}}{2} : \frac{N_{d-point}}{2} + \frac{m+1}{2}, n * \frac{N_{d-win}}{2} : n * \frac{N_{d-win}}{2} + N_{d-win}]$
  end for
end for



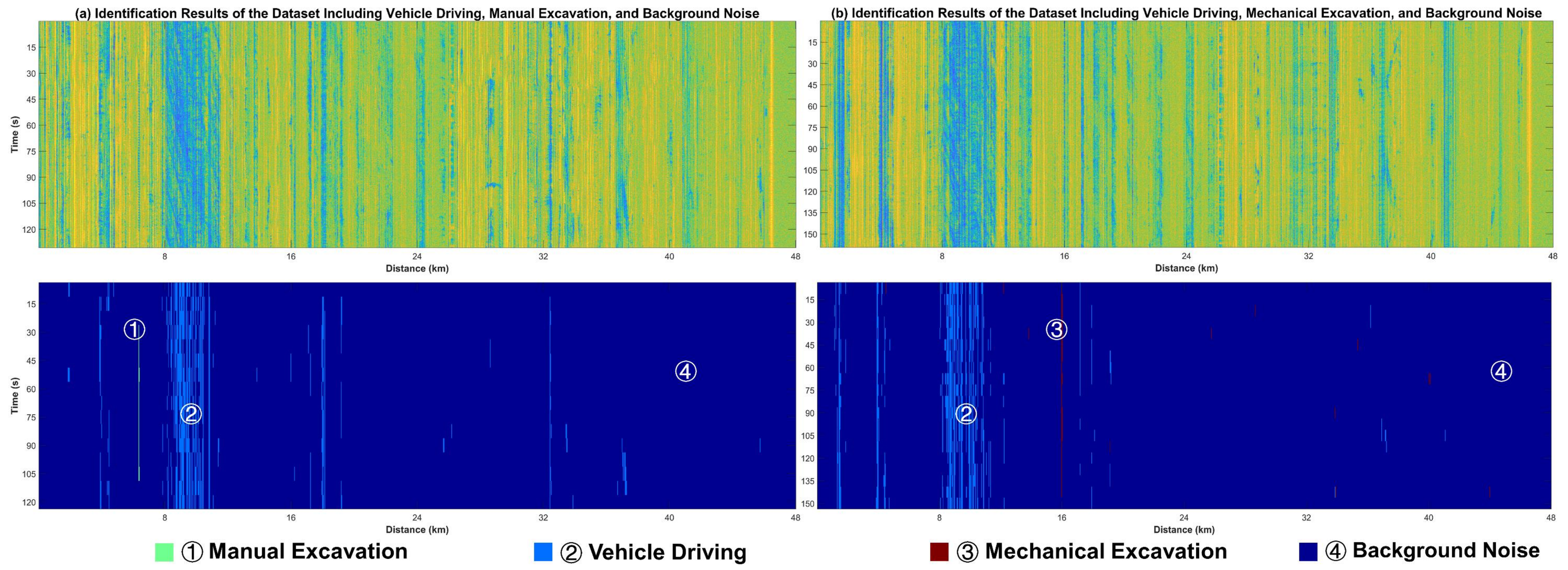
• Action Recognizer



Datasets

The data were gathered at a China National Petroleum Corporation real pipeline from May to June and from November to December. The total data size is approximately 494 GB. The test pipeline is approximately 48 km with complex environment under several types of strong noise and weak valuable signal information.

Experiment and Deployment



	2D CNN	1D CNN+LSTM	1D CNN+Bi-LSTM		
	Fusion model	Fusion model	Peak feature model	Energy feature model	Fusion model
Background noise					
Precision (%)	**100.0/100.0**	**100.0/100.0**	99.86/99.26	98.18/98.13	**100.0/100.0**
Recall (%)	**100.0/100.0**	**100.0/100.0**	99.86/99.16	99.23/98.13	**100.0/100.0**
F1-score (%)	**100.0/100.0**	**100.0/100.0**	99.86/99.21	99.08/98.13	**100.0/100.0**
AUC	**1.00/1.00**	**1.00/1.00**	0.999/0.993	0.998/0.991	**1.00/1.00**
Manual excavation					
Precision (%)	**100.0/100.0**	**100.0/100.0**	**100.0/100.0**	98.84/94.39	**100.0/100.0**
Recall (%)	91.38/98.06	98.83/**100.0**	94.25/88.35	97.70/98.06	**98.85/100.0**
F1-score (%)	95.50/99.02	99.40/**100.0**	97.04/93.81	98.27/96.19	**99.42/100.0**
AUC	0.957/0.990	**0.994/1.00**	0.971/0.942	0.986/0.977	**0.994/1.00**
Mechanical excavation					
Precision (%)	83.51/75.19	95.51/81.82	82.28/76.60	90.59/70.31	**96.55/83.33**
Recall (%)	95.29/**100.0**	**100.0/100.0**	76.47/80.01	90.59/**100.0**	98.82/**100.0**
F1-score (%)	89.01/85.71	**97.70**/90.00	79.27/78.26	90.59/82.57	97.67/**90.91**
AUC	0.959/0. 973	**0.996**/0.982	0.867/0.880	0.944/0.966	0.991/**0.984**
Vehicle driving					
Precision (%)	97.70/**100.0**	**100.0/100.0**	89.30/86.67	95.32/**100.0**	99.42/**100.0**
Recall (%)	**99.42**/80.05	98.83/84.62	97.66/**100.0**	95.32/64.62	**99.42**/86.15
F1-score (%)	98.55/88.89	99.41/91.67	93.30/**92.86**	95.32/78.50	**99.42**/92.56
AUC	0.992/0.900	0.994/0.923	0.961/**0.981**	0.966/0.823	**0.996**/0.931
Total					
Accuracy (%)	96.28/95.33	99.22/96.89	93.68/93.44	96.28/92.05	**99.26/97.20**

	Feature time	Model time	Total time	Model size
500 Hz	13.23 s	3.715 s	17.22 s	18.7 MB
100 Hz	3.028 s	3.489 s	6.597 s	

Supplement

Our method has been deployed and used at other real pipelines in North China and has been validated to maintain a high-level of real-time and performance with adverse weather conditions during the long field test.

Besides, we are interested in exploring the applications of distributed signal early warning in other areas, such as early warnings of undersea and land quakes, traffic flow statistics for urban road networks, and illegal cross-border behavior monitoring.