Domain knowledge informed multitask learning for landslide induced seismic classification

Jiangfeng Li, *Student Member,* Minxiang Ye, Lina Stankovic, *Senior Member,*
Vladimir Stankovic, *Senior Member*, Stella Pytharouli

***Abstract*—Automatic seismic signal classification methods are extensively investigated to reduce or replace manual interpretation, with great potential in previous research. Discriminative seismic wave propagation physical characteristics, such as velocities and accelerations, are rarely considered for classification. A multitask learning scheme is proposed that utilises the seismic wave equation and three-dimensional (3D) P-wave velocity *Vp* model for signal representation learning. The classifier uses the obtained latent feature maps on a convolutional neural network architecture for classification of rockfall, slide quake, earthquake, and natural/anthropogenic noise events, recorded at an ongoing landslide. Our experimental results show that our approach outperforms state-of-the-art methods.**

***Index Terms*—Seismic wave equation, P-wave velocity, landslide-induced seismic classification, multitask learning**

I. Introduction

Seismic monitoring systems have become a viable option for slope stability monitoring due to their high sensitivity relative to other more conventional slope monitoring technologies (e.g. geodetic monitoring), high network density, low cost, and relatively low power supply, and multi-channel seismic time series data analysis enables better understanding of the origins of various seismic events, to predict major geological disturbances, and to further minimise fatalities and infrastructure damage [1]. However, the challenges of (micro)seismic signal classification, include: 1) the signal-to-noise ratio (SNR) of the recorded signals is low [2], especially for microseismic signals with very low magnitude, typically between M-3 and M0, 2) limited datasets of events, which are time-consuming to label and often produced with uncertainty [3], 3) heterogeneity and complexity of recordings, that include anthropogenic and meteorological events [4]. Conventional (micro)seismic analysis, by visually inspecting the signal, is a tedious process that necessitates detailed domain knowledge [5]. Thus, recently, seismic signal analysis has evolved into a trend that depends on machine learning assisted with little or no domain knowledge.

Conventional machine learning approaches, such as, e.g., Support Vector Machine [2]; Random Forest [4], [3], and graph-based classifiers [2], [6], rely on handcrafted feature construction, extraction, and selection, commonly performed requiring extensive domain knowledge for physical characterisation of events. However, deep learning-based classifiers perform automatic feature engineering, requiring little or no expert input, and as such, have become attractive for seismic analysis. Some examples include Convolutional Neural Network (CNN) architectures to distinguish types of earthquakes [7], and quarry blasts and earthquakes [8]. [9] use global weighted average pooling, as the attention module, to distinguish amongst various types of seismic sources. For multi-sensor recorded signals, [10] propose a CNN and Graph Convolution Network-based (GCN) architecture to classify seismic events with information exchange. Deep learning-based approaches have been shown to learn a compact representation of the target signal domain without expert-designed handcrafted features [7], making the inference process autonomous, practical, and fast. However, the main issue with the current models, is their lack of interpretability [1], which limits their application, due to not only trust issues but also difficulties in debugging, improving performance, making fair comparisons across approaches and datasets, and improving understanding of underlying seismic processes.

Motivated by recent studies [11], [12], a domain-knowledge informed multitask learning scheme comprising signal representation learning and classification is proposed, which uses the seismic wave equation and the P-wave velocity *Vp* model. [11] classifies seismic events into volcanic tremors, earthquakes and ambient noise by evaluating the spatial coherence of recorded signals, concluding that the spectral differences of the seismic signals with high spatial coherence could enhance the classification performance. In this letter, instead, we exploit the high spatial coherency among signal recordings at different sensors to denoise the signal using the second derivative of the seismic wave propagation in the spatial domain, estimated from the recordings of the deployed sensors. [12] shows that, with a limited number of sensors, the seismic wave equation can be an alternative pathway to extract the spatial wave representations (second derivative of the seismic wave propagation in the spatial domain) from the acquired signals in the time domain. Rather than solving the seismic wave equation numerically, with a few sensor recordings, [12] use a neural network to simultaneously conduct the propagation velocity inversion and displacement prediction. As in [12], the signal representation learning task proposed in this paper exploits the target signal seismic wave propagation characteristics within a neural network. However, while [12] solves

This work was supported in part by the EPSRC Prosperity Partnership research and innovation programme under grant agreement No EP/S005560/1.

J. Li, L. Stankovic, V. Stankovic and S. Pytharouli are at University of Strathclyde, Glasgow G1 1XW. M. Ye is with Research Center for Intelligent Robot, Zhejiang Lab, Hangzhou 311121. e-mail: {jiangfeng.li, lina.stankovic, vladimir.stankovic, stella.pytharouli}@strath.ac.uk, yemx@zhejianglab.com.

the seismic wave equation with a small amounts of time and spatial domain data samples triggered by artificial shots, we utilise a CNN architecture to estimate the propagation of temporal characteristics of the real-field seismic recordings. Additionally, we leverage on the CNN architecture's ability to operate on patch-based input, which allows it to consider local seismic reflection patterns when defining and learning features of target structures [7], to identify and exclude random or coherent seismic noise and processing artifacts of distinct patterns, resulting in reduced misclassification. The proposed multitask learning scheme includes three key contributions: 1) a novel signal representation learning/reconstruction with two CNN architectures to estimate temporal and spatial characteristics of the signal propagation from the real-field seismic recordings; 2) classifying four classes of recorded landslide-induced seismic signals with a CNN architecture; 3) experimental results demonstrating competitiveness of the proposed system w.r.t state-of-the-art in seismic signal classification.

II. Preliminaries

A. Seismic wave equation

The location of the signal source is an important feature to distinguish seismic events occurring in different critical zones of the monitored slope deformation. For instance, rockfalls emit signals from locations concentrating on the surface of the slopes, while slide quakes are identified as shallow-depth local seismic events [13]. To estimate source location, it is necessary to model the signal propagation pathway from their origin or source to the deployed sensors.

In general, seismic signals involve the superimposition of several complex pulses traveling at speeds dictated by the elastic characteristics and density of the medium [14]. The propagation of these signals through an isotropic and homogeneous medium is usually modeled by a 1-dimensional time-dependent seismic wave equation [15]:

$$\frac{1}{V^2}\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial^2 u(x,t)}{\partial x^2}, \tag{1}$$

where $u(x,t)$ is the displacement in point x at time t. The solution to this differential equation can be expressed as $u(x,t) = E\cos[2\pi(t - x/V) - \varphi]$, where E and φ are the amplitude and phase of the wave, respectively, V is the P/S-wave velocity; for typical sensor arrays with vertical channels (as in our dataset) where P-waves are dominant, V is the P-wave velocity, *Vp*, determined by the elastic properties and density of the medium along the wave propagation pathway.

The solution of (1) depends on the *Vp* model. Generally, signal propagation characterisation and localisation require a three-dimensional (3D) *Vp* model of the crust, determined by the density and elastic properties of the material along the seismic wave pathway [16]. Specifically, each cube of this 3D *Vp* model represents *Vp* of the wave propagating through the material inside the small cube area, which is assumed to be homogeneous. *Vp* of the crust varies across layers and depends on the geological composition (e.g., rock, clay *etc.*); thus, it can be exploited to distinguish the distance that wave travels and determine its source [13].

III. Proposed method

The main idea behind the proposed method, as illustrated in Fig. 1, is to use the recorded signals and a *Vp* model to estimate the location of the seismic event source via domain-knowledge informed multitask learning relying on wave equation to optimise the model, and then use this information to classify the signal based on the fact that different classes of seismic events propagate differently through the medium. In the following, first, we provide an overview of the proposed method and then provide implementation details.

A. Overview

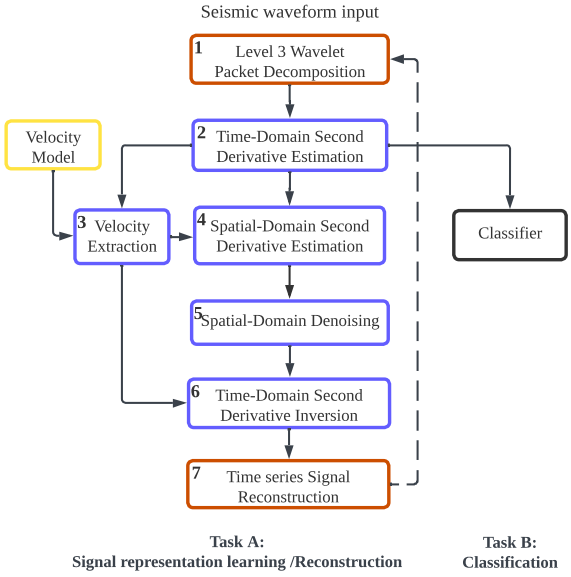
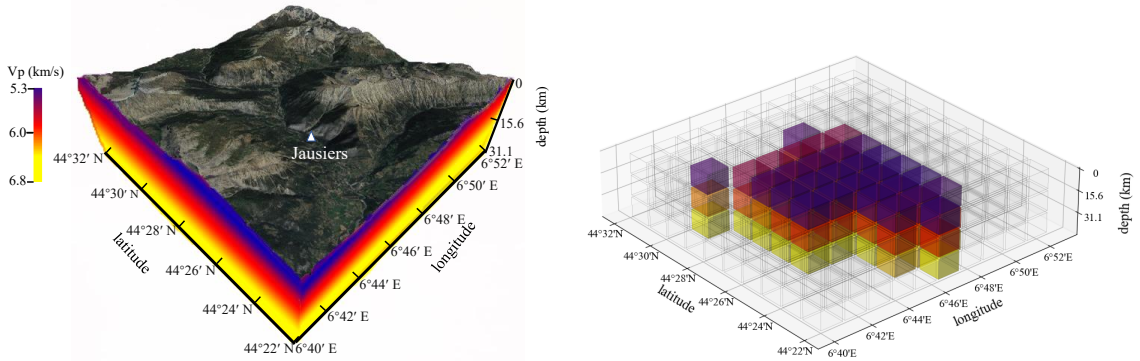


Figure 1. Proposed domain-knowledge informed multitask learning scheme.

First, the recorded seismic waveform is decomposed with a level three wavelet packet decomposition (Block 1), to identify its frequency distribution over time with variable-width windows using a bank of low-pass and high-pass filters, trading off time and frequency resolution [2]. This results in eight signal representations, each corresponding to a frequency band. Block 2 estimates the second derivative w.r.t time, of the displacement vector, $\frac{\partial^2 u(x,t)}{\partial t^2}$, from the latter wavelet packet-decomposed signals, which is the latent features for classification, since, the second derivative w.r.t. time carries information on wave curvature and depth, discriminating seismic events, particularly those originating from distinct locations.

Block 3 extracts one *Vp* value for each event, from each of the 3D *Vp* models, where the size of the grid cells can vary depending on the resolution required, trading off the resolution of the seismic velocity model and the complexity of the multi-tasking learning model. Here, the grid size of 10 × 10 × 3 velocity values was chosen to balance resolution and complexity. Block 3 comprises two Fully Connected



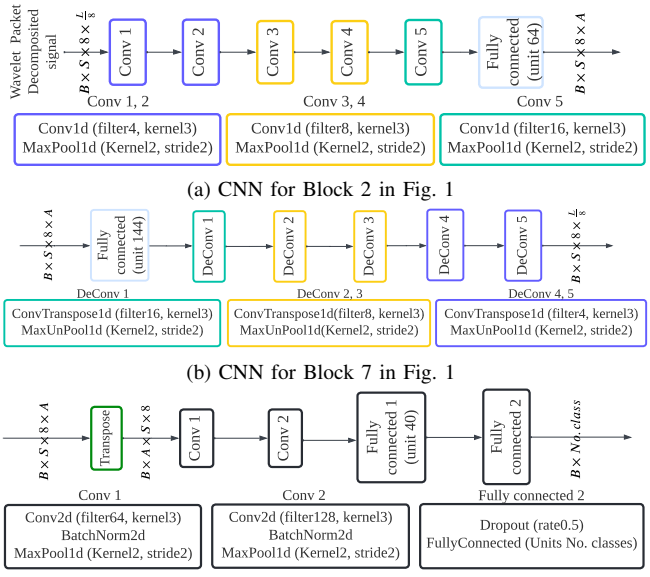
(a) Earthquake topographic *Vp* model (b) Model mapped to grid

Figure 2. The modified *Vp* model for earthquake signal [17].

(FC) layers, to estimate the sampling probabilities in the *Vp* model, which represent the likelihood of a velocity value being selected as the most likely propagation velocity of corresponding events, that in turn weighs the contribution of each velocity value to the overall proposed multi-task learning model fitting process. The *Vp* model is computed over a certain area with inversion techniques and interpolation of the seismic signals recorded by a dense network. Fig. 2 (a) shows a seismic topographical area of interest using, as an example, recorded earthquakes originating in South Alps near Jausiers. Using the *Vp* model in [18], the *Vp* values are extracted to form a tensor $\mathbf{G} \in \mathbb{R}^{M \times N \times C}$, as shown in Fig. 2 (b). To simplify the model, we only consider the areas (cubes) most likely to contain seismic sources (limited to 10 cubes in each direction, i.e., $M = N = 10$ and $C = 3$), where the blue, red, and yellow cubes show *Vp* values of $5.3\,, 6.0\,, 6.7\,\mathrm{km/s}$, respectively.

Seismic waves in earth materials are subject to attenuation and dispersion in a broad range of frequencies; in addition, the field-observed seismic signals are usually obscured by a significant amount of noise, which affects the precision of estimating the second derivative w.r.t time using the seismic wave equation Eq. (1) [19]. Hence, we perform signal enhancement on derivatives w.r.t the spatial coordinate via Block 4, which uses the seismic wave equation Eq. (1) to convert the second derivative w.r.t time $\frac{\partial^2 u(x,t)}{\partial t^2}$ (Block 2) to the second derivative w.r.t the spatial coordinate $\frac{\partial^2 u(x,t)}{\partial x^2}$, by performing element-wise division by the estimated V^2 (Block 3) $\frac{\partial^2 u(x,t)}{\partial x^2} = \frac{\partial^2 u(x,t)}{\partial t^2} \oslash V^2$, where $\oslash$ stands for element-wise division. Motivated by the high spatial coherency among seismic recordings observed in [11], Block 5 performs denoising of the spatial domain second derivative, by replacing the three highest-energy spectral components with their mean value. Block 6 performs the inverse operation to Block 4 on the denoised signal (Block 5) recovering the second derivative w.r.t time with Eq. (1). Finally, Block 7 reconstructs the decomposed time series signal using the recovered second derivative w.r.t time $\frac{\partial^2 u(x,t)}{\partial t^2}$ obtained by Block 6 after spatial-domain denoising in Block 5. The output of Block 7 updates the loss function of the reconstruction task by measuring the difference between the output of Block 1 (dashed line in Fig. 1), as explained in the following section. The CNN classifier takes the labeled second derivative w.r.t time (output of Block 2) as input to predict the event label.

B. Implementation Details



(a) CNN for Block 2 in Fig. 1

(b) CNN for Block 7 in Fig. 1

(c) CNN for classification

Figure 3. CNN architectures implemented

Fig. 3 shows our CNN architectures, where B is the batch size, S is the number of data recording channels, L is the signal length, and A represents the length of the second derivative of the displacement vector u w.r.t time. Since Blocks 4 and 6 provide a simple application of the seismic wave equation Eq. (1) and Block 5 performs denoising via mean averaging, the focus is on Blocks 2, 3, 7, and the classification block, which are described next. The proposed multi-task (signal representation learning and classification) scheme is an end-to-end learning process. CNN models, Figs. 3 (a) and (b), used for feature learning, perform signal reconstruction, taking as input the wavelet decomposed signal and the estimated time domain acceleration, respectively. The output of Fig. 3

(a) is the learnt latent feature set, fed to Fig. 3 (c) for classification. Additionally, the loss function of the proposed scheme contains two terms, reconstruction and classification loss, which are utilised to train the whole network. Thus, CNN models, Figs. 3 (a) and (b), are embedded into the signal representation learning task, whose weights are updated synchronously minimising Eq. (2).

1) Time-Domain Second Derivative Estimation (Block 2 in Fig. 1): The input signal is first decomposed with a db4 wavelet packet transform. Let $\mathbf{T} \in \mathbb{R}^{S\times 8\times \frac{L}{8}}$ be the wavelet-decomposed signal from Block 1, where 8 comes from the number of signal representations at Level 3, each corresponding to one frequency band. Taking signal $\mathbf{T}$ as the input, the proposed CNN architecture, illustrated in Fig. 3 (a) is used to estimate the second derivative w.r.t time, $\mathbf{D_t} = \frac{\partial^2 u(x,t)}{\partial t^2} \in \mathbb{R}^{S\times 8\times A}$ for each input window using ReLU activation function. Heuristically, we set A, the length of the second derivative signal, to 64.

2) Velocity Extraction (Block 3 in Fig. 1): Given initial *Vp* model $\mathbf{G} \in \mathbb{R}^{M\times N\times C}$ from [13], [17] and the estimated time-domain second derivative $\mathbf{D_t}$, we pose the *random sample consensus* as a voting strategy to extract velocities that optimally describe the underlying seismic wave attributes to pick one of the $M \times N \times C$ values in $\mathbf{G}$ for each given *Vp* model and each input window. As in [20], the sampling indices are generated by a set of *trainable* probability values $\mathbf{P} \in \mathbb{R}^{M\times N\times C}$ in order to select the optimal hypothesis of candidate *Vp* values that optimises the objective. A two-FC layer unit (with input (output) sizes of A (128) and 128 (300) for the first and the second layer, respectively) is employed to estimate the probability values $\mathbf{P}$ for generating random H hypotheses ($H = 12$ in our experiments) based on the multinomial distribution with $M = N = 10$ and $C = 3$, such that the dimension of $\mathbf{P}$ matches the dimensions of the initial *Vp* model $\mathbf{G}$ (see Subsection III-A). The network takes as an input the output of Block 2 ($\mathbf{D_t}$) and predicts the probability of signal originating in each of $M\times N\times C$ cube. At the output of Block 3, we extract the *Vp* that corresponds to the second derivative w.r.t time estimated in Block 2.

3) Time-Domain Signal Reconstruction (Block 7 in Fig. 1): Block 7 performs time series signal reconstruction by setting the network objective to be a reconstruction task (Task A in Fig. 1) that optimises the mean square error between the recovered time domain signal $\hat{\mathbf{T}} \in \mathbb{R}^{S\times 8\times \frac{L}{8}}$ after spatial domain denoising (output of Block 6) and the input wavelet-decomposed signal $\mathbf{T}$, defined as

$$L_r = \frac{1}{S}\sum_{s=1}^{S}\frac{1}{8}\sum_{ch=1}^{8}\Big(\frac{1}{\frac{L}{8}}\sum_{l=1}^{\frac{L}{8}}\Big(T_{s,ch,l} - \hat{T}_{s,ch,l}\Big)^2\Big), \qquad (2)$$

where ch and l represent the index of wavelet-decomposed frequency bands and the data samples of the event window. Specifically, we propose a CNN architecture that contains five deconvolution layers, ReLU activation function, and maxpooling operations, as illustrated in Fig. 3 (b) and performs the reverse operation to Fig. 3 (a).

4) Classification: With the learned optimal latent feature (output of Block 2, $\mathbf{D_t}$), a CNN architecture is employed for classification (Task B in Fig. 1), containing two 2-dimension convolution layers and two FC layers, with the Cross-Entropy loss function L_c. The classifier is trained with the labelled latent features.

IV. Experimental results

In this section, we evaluate the accuracy and precision of the proposed CNN-based multitask learning to classify the seismic signals recorded at an ongoing landslide. All experiments are conducted on NVIDIA RTX 3090 with PyTorch v1.12. All CNN-based models are trained over 450 epochs with 1.8 second per epoch using adaptive moment estimation (ADAM) optimiser with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and 7.7ms per sample for prediction. Note that the overall loss function of the proposed domain knowledge informed multitask learning is defined as $Loss = L_r + L_c$.

A. Dataset

The dataset is obtained from publicly available raw seismic recordings[1], sampled at $Fs = 250$ Hz, measures the ground shock caused by seismic signals, as described in [4]. The array comprises one 3-component center sensor and three vertical 1-component sensors organised as an equilateral triangle. A catalog of manually verified events, detected with spectrogram-based Short Time Average over Long Time Average (STA/LTA), contains 401 "Rockfall", 234 "Slide Quake", 388 "Earthquake", and 351 "Anthropogenic/Natural noise" events, totaling 1374 labeled events [4]. We filter 6-channel raw signals with a $5-100$ Hz bandpass filter, and then segment a 10 s window for each event from 2 s before the event starts (P wave pick), to form the input as $\mathbf{X} \in \mathbb{R}^{S\times L}$, where $S = 6$ channels, and $L = 2500$ samples. For benchmarking classification, we employ stratified sampling with 60% (training), 10% (validation), and 30% (testing). The validation set is used to evaluate the performance of the model on data that the model has not seen during training, which helps in determining the best set of hyperparameters for the model. At the testing stage, the best model corresponds to the one with the highest validation accuracy. To eliminate the impact of random data splitting, we repeat the experiment 50 times with a randomly split training, validation, and testing set.

B. Vp models

Table I
Seismic topographic areas

Event	Latitude (N)	Longitude (E)
Rockfall	$44°20'41''$-$44°21'8''$	$6°40'24''$-$6°40'47''$
Slide Quake	$44°20'51''$-$44°21'58''$	$6°40'30''$-$6°40'41''$
Earthquake	$44°24'10''$-$44°31'27''$	$6°40'44''$-$6°52'16''$
Noise	$44°20'41''$-$44°21'8''$	$6°40'24''$-$6°40'47''$

The *Vp* models are obtained from the seismic topographic areas (Table I) and depth from [13]. 3D *Vp* models for rockfall,

[1]https://seismology.resif.fr/networks/#/MT

slide quake and noise events are obtained by cubic interpolation of the 2D profiles, which are computed by inverting the anthropogenic shots recorded by the sensors deployed in the preset geophones using the Quasi-Newton inversion approach [13]. The 3D *Vp* model[2] for the earthquake is obtained by inverting the dense network recorded teleseismic data in [18].

Table II
EXPERIMENTAL RESULTS AVERAGED WITH 50 RUNS (MEAN (STD))

Classifier	Acc	Recall			
		Rockfall	Slide Quake	Earthquake	Noise
CNN + GCN [10]	0.8576 (0.03)	0.8715 (0.04)	0.7662 (0.07)	0.8442 (0.06)	0.9176 (0.04)
DeepQuake [21]	0.8650 (0.03)	0.9041 (0.04)	0.7560 (0.06)	0.8769 (0.05)	0.8796 (0.05)
ConNetQuake [7]	0.8470 (0.03)	0.8901 (0.04)	0.7254 (0.08)	0.8494 (0.05)	0.8756 (0.05)
$GWAP_1$ [9]	0.9140 (0.01)	0.9210 (0.03)	0.8249 (0.07)	0.9379 (0.03)	0.9389 (0.02)
$GWAP_2$ [9]	0.9345 (0.02)	0.9405 (0.03)	0.8520 (0.05)	**0.9554 (0.02)**	0.9592 (0.02)
Proposed	**0.9422 (0.02)**	**0.9413 (0.03)**	**0.9083 (0.04)**	0.9474 (0.02)	**0.9602 (0.02)**
Replace Task A (Fig. 1)	0.9312 (0.01)	0.9393 (0.02)	0.8571 (0.05)	0.9403 (0.03)	0.9491 (0.02)

C. Results

To evaluate the classification performance of the proposed multitask learning scheme, we compare the classification performance to the state-of-the-art in terms of Accuracy (Acc) and Recall (as in [9]). The benchmarks are based on recent CNN [7], [21], GCN [10] and attention-module based [9] architectures. Table II shows that the proposed method outperforms the benchmarks in terms of Acc averaged over all four classes. Furthermore, it provides the highest recall rate for all classes except Earthquake, where the schemes optimised for the detection and classification of earthquakes, such as, [9] perform slightly better. However, the limitation of [9] is that the size of the training set required for good performances of seismic classification is large. To evaluate the performance of the intermediate CNN models (Fig. 3 (a) and (b)), the state-of-the-art feature extraction architecture of [22] is implemented to replace task A in Fig. 1. The optimised latent feature maps are fed into the proposed classifier (Fig. 1), and the resulting classification performance is shown in the last row of Table II. Comparing the last two rows which use the same proposed classifier, the proposed intermediate feature learning architectures (Fig. 3 (a) and (b)) outperform that of [22] for all classes. The largest improvement is observed for slide quakes.

V. CONCLUSION

A novel domain knowledge informed multitask learning scheme is proposed with the seismic wave equation and *Vp* models for ongoing landslide-induced signal classification. CNN architectures are utilised as an alternative technique to estimate the signal temporal and spatial attributes. Experimental results demonstrate that the proposed multitask learning scheme is effective and outperforms state-of-the-art methods. The proposed model can also be applied to other applications, such as volcano active monitoring systems and earthquake early warning systems. As future work, it is worth investigating the effect of the 3D velocity model accuracy comparing homogeneous, 2-layer and gradient velocity models.

[2]https://dataservices.gfz-potsdam.de/panmetaworks/showshort.php?id=b259c149-19dd-11ec-9603-497c92695674

REFERENCES

[1] Z. Ma and G. Mei, "Deep learning for geological hazards analysis: Data, models, applications, and opportunities," *Earth Sci Rev*, vol. 223, p. 103858, 2021.

[2] J. Li, L. Stankovic, S. Pytharouli, and V. Stankovic, "Automated platform for microseismic signal analysis: Denoising, detection, and classification in slope stability studies," *IEEE Trans Geosci Remote Sens*, vol. 59, no. 9, pp. 7996–8006, Nov. 2020.

[3] M. Wenner, C. Hibert, A. van Herwijnen, L. Meier, and F. Walter, "Near-real-time automated classification of seismic signals of slope failures with continuous random forests," *Nat. Hazards Earth Syst. Sci.*, vol. 21, no. 1, pp. 339–361, 2021.

[4] F. Provost, C. Hibert, and J.-P. Malet, "Automatic classification of endogenous landslide seismicity using the random forest supervised classifier," *Geophys. Res. Lett.*, vol. 44, no. 1, pp. 113–120, Jan. 2017.

[5] C. Levy, D. Jongmans, and L. Baillet, "Analysis of seismic signals recorded on a prone-to-fall rock column (vercors massif, french alps)," *Geophys. J. Int.*, vol. 186, no. 1, pp. 296–310, 2011.

[6] J. Li, C. Yang, V. Stankovic, L. Stankovic, and S. Pytharouli, "Graph-based micro-seismic signal classification with an optimised feature space," in *2020 IEEE Dig Int Geosci Remote Sens Symp (IGARSS)*, 2020.

[7] T. Perol, M. Gharbi, and M. Denolle, "Convolutional neural network for earthquake detection and location," *Sci. Adv.*, vol. 4, no. 2, p. e1700578, 2018.

[8] L. Linville, R. C. Brogan, C. Young, and K. A. Aur, "Global-and local-scale high-resolution event catalogs for algorithm testing," *Seism. Research Let.*, vol. 90, no. 5, pp. 1987–1993, 2019.

[9] B. Ku, G. Kim, J.-K. Ahn, J. Lee, and H. Ko, "Attention-based convolutional neural network for earthquake event classification," *IEEE Geosci. Remote. Sens.*, vol. 18, no. 12, pp. 2057–2061, 2020.

[10] G. Kim, B. Ku, J.-K. Ahn, and H. Ko, "Graph convolution networks for seismic events classification using raw waveform data from multiple stations," *IEEE Geosci. Remote. Sens.*, vol. 19, pp. 1–5, 2021.

[11] T. Permana, T. Nishimura, H. Nakahara, and N. Shapiro, "Classification of volcanic tremors and earthquakes based on seismic correlation: application at sakurajima volcano, japan," *Geophys. J. Int.*, vol. 229, no. 2, pp. 1077–1097, 2022.

[12] S. Karimpouli and P. Tahmasebi, "Physics informed machine learning: Seismic wave equation," *Geosci. Front.*, vol. 11, no. 6, pp. 1993–2001, 2020.

[13] F. Provost, J. Malet, J. Gance, A. Helmstetter, and C. Doubre, "Automatic approach for increasing the location accuracy of slow-moving landslide endogenous seismicity: the apoloc method," *Geophys. J. Int.*, vol. 215, no. 2, pp. 1455–1473, 2018.

[14] S. Stein and M. Wysession, *An introduction to seismology, earthquakes, and earth structure*. John Wiley & Sons, 2009.

[15] R. E. Sheriff and L. P. Geldart, *Exploration seismology*. Cambridge university press, 1995.

[16] A. AlAli and F. Anifowose, "Seismic velocity modeling in the digital transformation era: a review of the role of machine learning," *J Pet Explor Prod Technol.*, pp. 1–14, 2021.

[17] L. Jenatton, R. Guiguet, F. Thouvenot, and N. Daix, "The 16,000-event 2003–2004 earthquake swarm in ubaye (french alps)," *J. Geophys. Res.*, vol. 112, no. B11, 2007.

[18] M. Paffrath, W. Friederich, S. M. Schmid, M. R. Handy, AlpArray, A.-S. D. W. Group *et al.*, "Imaging structure and geometry of slabs in the greater alpine area–a p-wave travel-time tomography using alparray seismic network data," *Solid Earth*, vol. 12, no. 11, pp. 2671–2702, 2021.

[19] T. M. Müller, B. Gurevich, and M. Lebedev, "Seismic wave attenuation and dispersion resulting from wave-induced flow in porous rocks—a review," *Geophysics*, vol. 75, no. 5, pp. 75A147–75A164, 2010.

[20] E. Brachmann, A. Krull, S. Nowozin, J. Shotton, F. Michel, S. Gumhold, and C. Rother, "Dsac-differentiable ransac for camera localization," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.(CVPR)*, 2017, pp. 6684–6692.

[21] L. Trani, G. A. Pagani, J. P. P. Zanetti, C. Chapeland, and L. Evers, "Deepquake—an application of cnn for seismo-acoustic event classification in the netherlands," *Comput Geosci.*, vol. 159, p. 104980, 2022.

[22] B. Ku, J. Min, J.-K. Ahn, J. Lee, and H. Ko, "Earthquake event classification using multitasking deep learning," *IEEE Geosci. Remote. Sens. Lett.*, vol. 18, no. 7, pp. 1149–1153, 2020.