



## ORIGINAL PAPER

TYPE ANALYSIS OF LABORATORY SEISMIC EVENTS  
BY CONVOLUTIONAL NEURAL NETWORKSPetr KOLÁŘ<sup>1)</sup> \* and Matěj PETRUŽÁLEK<sup>2)</sup><sup>1)</sup> Institute of Geophysics of the Czech Academy of Sciences, Boční II 1401, 141 31 Prague, Czech Republic<sup>2)</sup> Institute of Geology of the Czech Academy of Sciences, Puškinovo nám. 44/79, 160 00 Prague 6, Czech Republic\*Corresponding author's e-mail: [kolar@ig.cas.cz](mailto:kolar@ig.cas.cz)

## ARTICLE INFO

## Article history:

Received 22 February 2021

Accepted 19 April 2021

Available online 4 May 2021

## Keywords:

Convolutional neural network

Machine learning

Earthquake identification

Acoustic emission

Seismic signal processing

Bayesian optimization

## ABSTRACT

In this work, we successfully identified seismic events (observations of earthquakes) in seismograms using a Convolutional Neural Network (CNN). In accordance with past (analogue) seismogram interpretations, we did not treat digital seismograms as a time series, as per the general method, but, rather, converted them into time snaps of continuous data flow. Multichannel seismograms were represented with a time-frequency domain in the form of multilayer images, with each signal channel forming one image layer. Images were then exposed to CNN (composed of three convolutional blocks whose architecture design was justified using Bayesian optimization). To improve reliability, we evaluated the posterior type function (PTP) as a combination of the probabilities of all of the considered classes of signal types (five in our case) which increased robustness of the identification. For data, we used records of acoustic emission (AE) events. The events were generated during laboratory loading experiments originally performed to study material/rock properties. As known, AE events may be studied in the same manner as natural earthquakes and treated in other ways as laboratory earthquake models. AE events are less complex compared to natural earthquakes where many of the physical parameters are known or may be controlled. Based on our results, we concluded that the successful identification of AE events is a necessary step prior to applying the proposed methodology for identifying natural earthquakes in seismograms.

## 1. INTRODUCTION

Over the last decade increasing quantities of recorded data for natural events consequently require sophisticated methods for data processing. Such is also the case for seismic events (earthquakes) and their records - seismograms. In this work, we propose an approach for data processing based on Convolutional Neural Networks (CNN). While CNN algorithms are currently available in standard software libraries, using CNN applications for seismic signals is a fairly new field as indicate recently published works, e.g. (Zhou et al., 2019; Perol et al., 2018; Curilem et al., 2018; Zhu and Beroza, 2018; Ross et al., 2018; Chen et al., 2019), an overview of various seismological application is given in (Kong et al., 2019).

The origin of the CNN formalism lays in computer vision and pattern recognition (Krizhevsky et al., 2012). However, it has appeared soon to be capable to efficiently solve the problems from other fields.

To perform such task successfully, the architecture of CNN has to be tuned and the data – seismograms in our case – must be converted into form suitable for CNN formalism and finally the obtained probabilities – the CNN outputs - have to interpret in seismological terms. Naturally, all the processing must

by tested and verified. In this number we have to deal with particular features of seismograms as e.g.: multi-channel signal inputs, appropriate signal representations, reasonable definitions of output classes, etc.

Formally, seismograms may be characterized as a time series (with a constant time sampling) for non-stationary processes. In the approach presented here, we, in some sense, suppress time dependent features of the seismic signal and treat it as an image in contrast to the older way of automatic seismogram processing, where the flow of signal or its changes are interpreted. We perform our computation in MATLAB platform and adopt their image processing/recognition library.

From a general perspective treating seismograms as images has traditionally been done within seismology: several decades ago, in the epoch of “analogue paper seismograms” and manual interpretations, operators were able to classify events using a type analysis, without the need of hard criteria, if they possessed long-term experience. At the beginning of our work, we considered this fact as a promising feature for suitable applications of the CNN approach. Details regarding problems related to analogue seismogram interpretations can be obtained, for example, from seismologist manuals printed

during the last decades of the 20<sup>th</sup> century (Willmore, 1979) (now also available in electronic format).

From a technical point of view, the problem may be solved using different approaches. In the past, a time-frequency signal analysis of quoted signals has been used, for example see (Plešinger et al., 2016), while in other studies, for example see (Truong et al., 2018), CNN algorithms have been applied.

Automatic identification of useful seismic signals is, obviously, not a new task. As an example of possible approach, it could be mentioned usage of the STA/LTA algorithm (the Short Time Average/Long Time Average of signal amplitudes) by (Blandford, 1982) or usage of the AR-AIC (autoregressive – Akaike Information Criteria) by (Kitagawa and Akaike, 1978). These algorithms are non-neural network algorithms and were not only used to identify an earthquake in the record but to trigger recording, especially during times when digital recording media capacity was heavily limited. Seismic event discrimination by classical/shallow neural networks has also been investigated, for example, by (Plešinger et al., 2000). These authors classified recorded events using two classes: natural events and artificial blasts. Even if the reliability obtained was fairly high (up to 98 % for testing data set), the algorithm was not used in routine data processing. A newer attempt for NN seismic event identification was performed e.g. in (Doubrovová et al., 2016). Here, a single recurrent NN was used with fairly high reliability. Implementation of the algorithm into routine use is part of the process (Doubrovová – personal communication). Obviously, earthquake phenomena form a fairly large class of events beginning with small induced events in mines, continuing with quarry blasts, and extending to local and regional events and/or to large worldwide earthquakes. Since they can be superposed in time, these types of events are possibly mixed in records. To simplify our task, we decided to begin with a particular type of event – acoustic emission (AE) events. Such events are observed during loading experiments using material or rock samples. Rock AE events are connected with the generation, growth, and interaction of microcracks (Paterson and Wong, 2005). In general, the laboratory scale AE experiments are performed to study the relation between the mechanical properties of rocks and the loading induced fracturing, e.g. (Petružálek et al., 2019). Note that the “AE events” term is only a convention based on our experience with mining events. Typical AE frequencies for our events ranged from 0.1 to 1.0 MHz and, obviously, could not be directly heard. At present, most AE processing methods originate from seismology (Grosse and Ohtsu, 2008). Alternatively, some AE studies are capable of helping us understand seismologic problems: e.g. (Thompson et al., 2009) or (Lockner et al., 1991). AE events may be considered as laboratory models of earthquakes. Over the past few decades, such events have been studied in the same manner as natural earthquakes, although a scale of similarity over

several orders of magnitude must be inferred (Lockner, 1993; Sellers et al., 2003). Our processed data possessed several particular features in comparison with real earthquakes. In contrast to real seismograms where various kinds of events such as local versus teleseismic, can occur, our data set only contained one kind of event. All events occurred inside a limited size rock sample. Due to preprocessing, the location of events was available. The experimental setup displays a fairly favorable directional coverage of the sample by sensors, especially when compared to the usual distribution of actual seismic stations. Experimental conditions were known or controlled, especially for the actual stress-strain state. On the other hand, limitations in laboratory equipment and/or experiment organization often led to challenges. For example, limitations occurred due to among other reasons: (i) special type of noise generated by a loading frame; (ii) the high frequencies considered during recording; (iii) the proportion of samples, sources, and sensors (point approximation is often used for natural earthquake studies and may not be valid for this type of work).

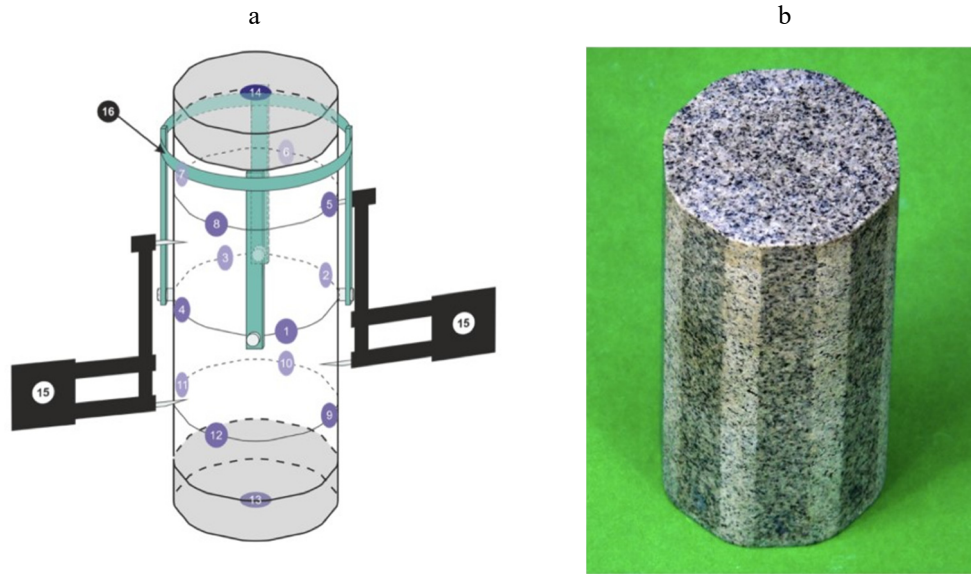
In general, preprocessed AE events seem to be a favorable choice of data for designing and verifying new identification method of seismic events in seismic signals.

## 2. DATA

### 2.1. EXPERIMENTAL SETUP AND MATERIAL

Westerly Granite, rock type often tested under laboratory conditions, was used in our experiment. The cylindrical rock sample was ground into a semi octagonal prism to enable flat contacts between the sample and the AE sensors (Fig. 1). The grinding improved sensitivity and homogeneity of registration in comparison to commonly used purely cylindrical shape. The sample diameter was 52 mm, the height 104 mm. Loading frame MTS 815 was used to generate uniaxial stress during the experiment; for details see a description given in (Petružálek et al., 2018, 2020).

We used 14 Fuji AE204A AE sensors with almost flat frequency characteristic within the range of 240–520 kHz and sensitivity 708 V/m/s. All the sensors were attached to the rock surface, but the two that were placed in the bottom and top loading platens. Each sensor could operate in two modes: (i) passive AE monitoring regime and (ii) active ultrasonic sounding (US) with use of high-voltage (200 V) input sine pulse. The sensors net covering the specimen is displayed in Figure 1a. Both AE and US waveforms were recorded using a multichannel transient recorder (AMSY-5 Vallen System, Germany) working in a triggering mode (if a defined signal level was reached at least on one channel, the 14 channel recording started). Waveforms length was set to 1,024 points including 512 points of pre-triggering interval; sampling frequency was 10 MHz (it corresponds to 0.1 ms record duration); 16-bits A/D converter was used.



**Fig. 1** Experimental setup. (a): points 1-14 stand for AE/US (acoustic emission/ultrasonic sounding) sensors; 15 for extensometers and 16 for a cantilever. Semi-octagonal Westerly Granite specimen before loading experiment (b).

## 2.2. DATA AND PREPROCESSING

About 21 000<sup>1</sup> AE events were recorded during the loading experiment. The two-step Akaike information criterion picker (Sedlak et al., 2009) was employed for automatic picking of first arrival times and amplitudes. Consequently, detected AE events were localized by a grid search method with an application of stress-strain dependent anisotropic velocity model in the form of three axis ellipsoid, that was determined from the periodical US (Petružálek et al., 2007). The localization error of  $\pm 2$  mm was estimated based on the localization of US events with the known position of transmitting sensors.

Three basic types of signals were distinguished in each trace (Fig. 2): (i) usual noise at the record without signal, (ii) arrival of the AE signal, the first onset is on the 512<sup>th</sup> sample (triggering setup) and the consecutive onsets follow in dependency of the mutual position of AE event locations and the sensor positions. (iii) the coda – i.e. the signal following AE event arrival, a superposition of all reflected phases. Of course, additional complication may arise. For example, if two AE events occur almost simultaneously or if a previous event coda is not fully attenuated etc., more complex signals are observed.

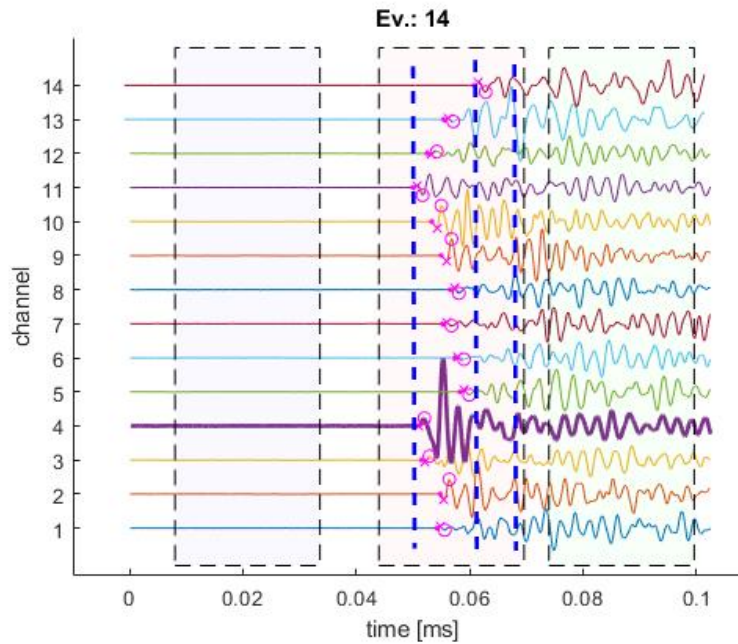
For our work, we exploited a previously processed pilot set of 38 selected AE events. These events were originally selected to develop and test the improved method for their mechanisms determination. Results of this interpretation may be found in (Petružálek et al., 2018). Considering presented experiment, the complete AE data set was used for a detail study of implementation of shear-tensile source model to AE (Petružálek et al., 2020) and

demonstration of source directivity in AE laboratory scale (Kolář et al., 2020).

## 2.3. IMAGES OF SEISMOGRAMS AND CLASS SETUP

In our approach, we treated multichannel seismograms as images. The images were converted from sub-records cut from the seismogram using a moving window, Figure 2. For our rock specimen (approximately 10 cm high and 5 cm in diameter) and for a seismic wave velocity of approximately 5 km/s the maximum difference in arrival times to the two most distant sensors (the two sensors located in the bottom and top loading platens) cannot exceed roughly 200 samples (regarding 10 MHz sampling frequency) for any event location. Therefore, we set the moving window width to 256 samples. Such a choice assured a sub-record that contained AE arrivals to all stations, including two consecutive signal extremes. The interval between signal arrival ( $T_0$ ) and the second pronounced extreme ( $T_2$ ) was taken as the definition of event duration. The times  $T_0$  and  $T_2$  (and  $T_1$  corresponding to the first pronounced maximum) are routinely interpreted during signal preprocessing; the signal amplitudes in  $T_1$  and  $T_2$  define the peak to peak amplitude, that is used for source mechanism determination (Petružálek et al., 2018). We primary were seeking the occurrence of an event in the investigated signal. In our analysis the event occurrence is defined as a portion of the record which include the interval  $\langle \min(T_{0i}), \max(T_{2i}) \rangle$ , where  $i=1-14$  are channel numbers. The definition also constitutes one of the CNN classes that should be distinguished as accurate.

<sup>1</sup> Note that 21.000 events is an upper number. Events or their records, respectively, are naturally of different quality and the number of events that can be used for a particular analysis is lower and depends on the type of the performed analysis, on the desired approach (qualitative or quantitative), on the required/acceptable level of reliability, etc.



**Fig. 2** An example of a 14 channel seismogram (event number 14). Channels are plotted using different colors/gray intensity. The triggering channel (i.e. the channel with the earliest occurrence of signal) is shown in bold, here as Channel 4. Dashed areas represent examples of the three basic types of signal; from left to right: (i) noise, (ii) event, and (iii) coda. The dashed boxes width is 256 samples (i.e. 26  $\mu$ s). Signal onsets are marked by dots, the first signal extremes by crosses, and the second extremes by circles. These values are routinely interpreted during signal preprocessing and are used for routine AE event analysis. The three vertical dashed lines indicate (from left to right): the earliest onset (min  $T_{0i}$ ), the latest onset (max  $T_{0i}$ ) and the latest second maximum (max  $T_{2i}$ ); these values are then used for classes' constitution.

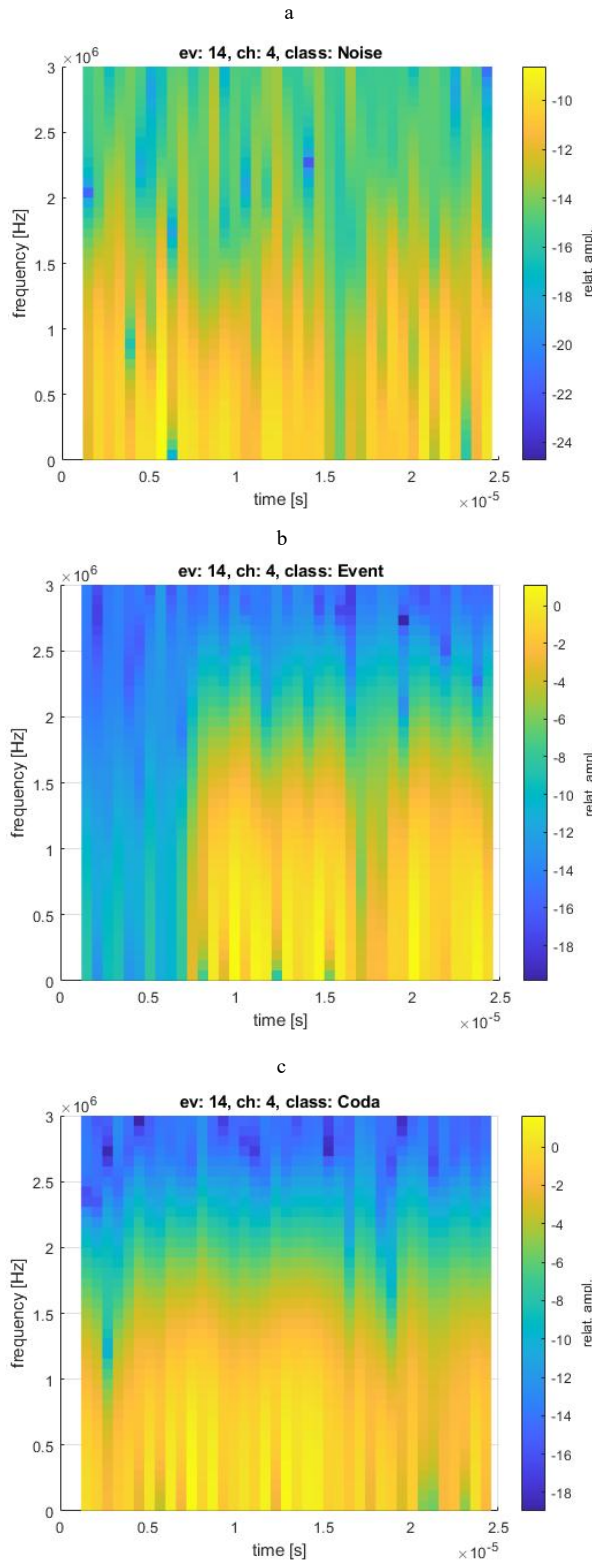
Sub-records, cut from record(s) with use of a moving window, were converted into image format using a time-frequency transformation. For our case, conversion was made possible using MATLAB function 'pspectrum' (<https://www.mathworks.com/help/signal/ref/pspectrum.html>, accessed 18 February 2021) based on (Welch, 1967; Harris, 1978). Image size was set to 40 x 40 points and the frequency range was set to 2 kHz – 3 MHz. We set up five (sub-)image classes. In addition to those previously mentioned (i.e. "noise" and AE arrivals/signals, labeled as an "event" and "coda"), we added two transitional classes: part-event and part-coda. The definitions for these classes are provided in Table 1. These images were input data for the CNN analysis. Image examples are provided in Figure 3. When the time shift for the cutting window was set to eight samples, we obtained classes, each with approximately 600 members; noise classes then had more than 1,000 members constructed from the considered set of 38 events. Successful detection of an "event" class would be a solution in our task. Others classes were non-event types of signals, however they represented different physical types of signal and could not be merged into a single class. Additionally, as it is show below, accurate identification of these non-event classes played an important role in final post-processing.

#### 2.4. CNN DESIGN AND TUNING, CLASS DEFINITION

CNNs are extensively used and have a large number of modifications, for example, see the review of (Schmidhuber, 2015). One of important field of CNN applications is image classification (Bishop, 2006). Since our computation was performed on the MATLAB platform (<https://www.mathworks.com/products/matlab.html>, accessed 18 February 2021), some technical details of our work are directly connected to MATLAB available features. An important point which has to be resolved is to form CNN input. Prior to MATLAB version R2018b, input images could either be 3D (most typically the RGB representation) or 1D (e.g. for B and W images). In MATLAB version R2018b, an option for multi-dimensional images was introduced (<https://www.mathworks.com/help/nnet/ref/nnet.cnn.layer.imageinputlayer.html>, accessed 18 February 2021). We took advantage of this feature and our input images consisted of 14 layers (i.e. one layer for each channel - Fig. 4). Our previous attempts using three- or one-layer images, composed of various combinations of seismograms or their time-frequency representations, did not yield satisfactory results. For processing by CNN, data were randomly divided for training, validating, and testing subsets using ratios of 60 %, 10 %, and 30 %, respectively.

To represent seismograms in multi-dimensional time-frequency domain, as the subject of a CNN

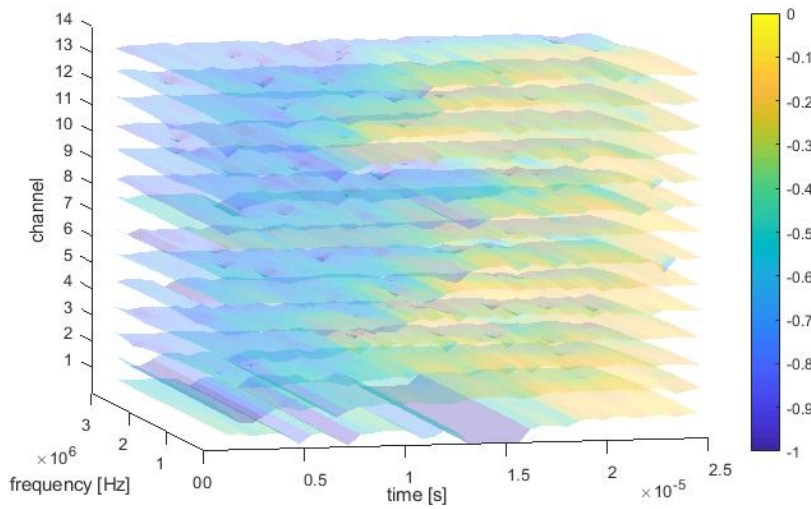




**Fig. 3** Examples of the time-frequency representation: event 14, channel 4 (triggering channel), class: Noise (a), Event (b), Coda (c). Amplitudes are converted into a color/intensity shadow scale. The figures are of different value scales.

analysis, is an innovative approach e.g. (Curilem et al., 2018; Nakano et al., 2019). On the other hand, the idea of the identification of events in seismograms using time frequency representations is not new and was used, for example, in (Joswig, 1990; Sick et al., 2015). Time frequency representations are also commonly used in other fields for time series processing, for example see (Plešinger et al., 2016).

Currently, only soft recommendations exist for CNN architecture design and their control parameters setup. We, therefore, followed, almost literary, MATLAB manual examples (<https://www.mathworks.com/help/nnet/ref/trainingoptions.html>, accessed 18 February 2021) and used a CNN composed of three convolutional cells, with control parameters set to the default or to the middle of the allowed intervals. Our CNN consists of three convolution blocks; its structure is described in Table 2. To justify our choice of CNN architecture and control parameters setup Bayesian optimization (BO) was used. The method is named after Thomas Bayes (1701?-1761). Among others, BO can be used for the optimization of hyper parameters of a task, here, the number of convolutional cells, their depth as well as values of CNN teaching control parameters. We again followed the MATLAB example ([https://www.mathworks.com/help/deeplearning/examples/deep-learning-using-bayesian-optimization.html?s\\_tid=srchtitle](https://www.mathworks.com/help/deeplearning/examples/deep-learning-using-bayesian-optimization.html?s_tid=srchtitle), accessed 18 February 2021), modified the necessary parts, and searched for optimal values of following CNN parameters: the number of convolutional cells and number of convolutions in each cell. The CNN learning parameters were the Initial Learning Rate, the Momentum, the Drop Out value and the L2 regularization value. We used CNN accuracy (the sum of successful class identifications given in the Confusion matrix trace) in order to compare the efficiency of the tested CNN configuration. For the initial CNN configuration the accuracy oscillated several percent above 90 % (Fig. 5a). For the fully optimized CNN, the accuracy was only some better (92.6 % vs. 93.9 %, Fig. 5b). Most false identifications occurred in the transition zones; i.e. between noises to part-event, from part-event to event, from event to part-coda, and from part-coda to coda classes. The fact that these classes were not clearly separated but mutually merged and that the sharp classifications could not be reliably setup is not a surprise as our definition (provided in Table 1) was partly artificial. Since results gained by the initial and optimized CNN are similar, we employed the initial CNN in future calculations. As the initial tests indicated, the pure classification of the data is not enough robust in our case. Hence, we created posterior type function (PTF) depending on combination individual class probabilities. As it is shown below, the values PTF or their changes indicate searched classes of seismograms.



**Fig. 4** Example of a 14 layer's image (event number 14, class Event). For better visibility, individual layers were normalized.

**Table 1** Classes definition.

$t_{winBeg}$  and  $t_{winEnd}$  are times of cut sub-records at the beginning or ending, respectively.  $T_{0i}$  are onset times and  $T_{2i}$  are times of the second pronounced extremes,  $i = 1-14$  are channel numbers; c.f. Figure 1.

The time of the first signal arrival is relatively clearly defined, even if possibly disturbed, for example, by higher noise level. In contrast, due to its interferential character, the coda phase has no exact beginning. For real earthquakes, the distances among the source and points of observation are relatively, as well as absolutely, longer than those in our laboratory experiment. Therefore, different seismic phases (i.e. the reflected and/or refracted signals) are often separated in time and can generally be distinguished in seismograms. Such is not the case for AE signals where individual phases are not separated and the signal of the first phase fluently merges into an interferential coda. As a result, for our study, we used the interval between signal onset and the second pronounced extreme as representative of an event class and the rest of the signal as the coda class.

Class	Definition
Noise	$t_{winEnd} \leq \min(T_{0i})$
Part Event	$\min(T_{0i}) < t_{winEnd} \ \&\& \ t_{winEnd} < \max(T_{2i})$
Event	$t_{winBeg} < \min(T_{0i}) \ \&\& \ \max(T_{2i}) < t_{winEnd}$
Part Coda	$\min(T_{0i}) < t_{winBeg} \ \&\& \ \max(T_{2i}) > t_{winEnd}$
Coda	$t_{winBeg} > \max(T_{2i})$

**Table 2** Architecture of the employed CNN. Three convolution blocks were used. The user's manual example for a type of CNN (<https://www.mathworks.com/help/nnet/ref/trainingoptions.html>, accessed 18 February 2021) was almost literary followed. However, necessary tuning for our data format and the required output was applied.

Convolution package	Layer No.	Name, function	Description, parameters
(input)	1	'Image Input'	(40 x 40 x 14) images with 'zerocenter' normalization
1 <sup>st</sup>	2	'Convolution 1'	14 (7 x 7) convolutions with stride [1 1] and padding 'same'
	3	'Batch Normalization 1'	Batch normalization
	4	'ReLU 1'	ReLU
	5	'Max Pooling 1'	(2 x 2) max pooling with stride [2 2] and padding [0 0 0 0]
2 <sup>nd</sup>	6	'Convolution 2'	28 (5 x 5) convolutions with stride [1 1] and padding 'same'
	7	'Batch Normalization 2'	Batch normalization
	8	'ReLU 2'	ReLU
	9	'Max Pooling 2'	(2 x 2) max pooling with stride [2 2] and padding [0 0 0 0]
3 <sup>rd</sup>	10	'Convolution 3'	56 (3 x 3) convolutions with stride [1 1] and padding 'same'
	11	'Batch Normalization 3'	Batch normalization
	12	'ReLU 3'	ReLU
(output and classification)	13	'Dropout'	50 % dropout
	14	'Fully Connected'	5 fully connected layer
	15	'Softmax'	Softmax
	16	'Classificaation output'	Crossentropyex

Confusion Matrix					
Output Class	Coda	Event	Noise	PartCoda	PartEvent
	146 13.7%	0 0.0%	0 0.0%	11 1.0%	0 0.0%
	0 0.0%	155 14.5%	0 0.0%	18 1.7%	4 0.4%
	0 0.0%	0 0.0%	343 32.1%	0 0.0%	9 0.8%
	16 1.5%	4 0.4%	0 0.0%	164 15.4%	0 0.0%
	0 0.0%	15 1.4%	2 0.2%	0 0.0%	180 16.9%
Target Class					
	Coda	Event	Noise	PartCoda	PartEvent
	90.1% 9.9%	89.1% 10.9%	99.4% 0.6%	85.0% 15.0%	93.3% 6.7%

**Fig. 5a** A confusion matrix for the initial CNN configuration and the five classes classifications. From the matrix it follows that most of false classifications (marked by circles) occur within two transient zones amongst the three basic classes/signal types (i.e. noise, event, and coda). The percentages in the individual cells are the portion of the identified class examples normalized as a function of the total number of examples.

### 2.5. POST-PROCESSING

The identification of classes using the CNN was slightly over the 90 % level (Fig. 5ab), a level that we did not consider as a fully acceptable result. We were not able to increase this percentage even with tuned CNN architecture and/or teaching control parameters. To emphasize the reliability of the algorithm we introduced a posterior type function (PTF) evaluated as the combination of individual probabilities for all considered classes for each classified subimage<sup>2</sup> The following formulas were optionally considered:

$$PTF1 = 0.5 \cdot pN + (1 - pPE) + pE + 0.5 \cdot pPE + (1 - pC), \quad (1)$$

$$PTF2 = (1 - pN) + 0.5 \cdot pPE + pE + 0.5 \cdot pPE + (1 - pC), \quad (2)$$

$$PTF3 = (-pN) + (pPE - 0.5) + pE + (pPC - 0.5) + (-pC), \quad (3)$$

where  $pN$ ,  $pPE$ ,  $pE$ ,  $pPC$ , and  $pC$  indicate the probability of individual classes: Noise, Part-Event, Event, Part-Coda, and Coda (the sum of these probabilities is always one). The  $PTFn$  values were evaluated for each CNN input image. The time sequence of those values or their changes respectively, can indicate an event occurrence (i.e. desirable information – Fig. 6). Such an approach increased the number of successfully identified events in a way that

Confusion Matrix					
Output Class	Coda	Event	Noise	PartCoda	PartEvent
	155 14.5%	0 0.0%	0 0.0%	22 2.1%	0 0.0%
	0 0.0%	167 15.7%	0 0.0%	12 1.1%	1 0.1%
	0 0.0%	0 0.0%	332 31.1%	0 0.0%	3 0.3%
	7 0.7%	3 0.3%	0 0.0%	159 14.9%	0 0.0%
	0 0.0%	4 0.4%	13 1.2%	0 0.0%	189 17.7%
Target Class					
	Coda	Event	Noise	PartCoda	PartEvent
	95.7% 4.3%	96.0% 4.0%	96.2% 3.8%	82.4% 17.6%	97.9% 2.1%

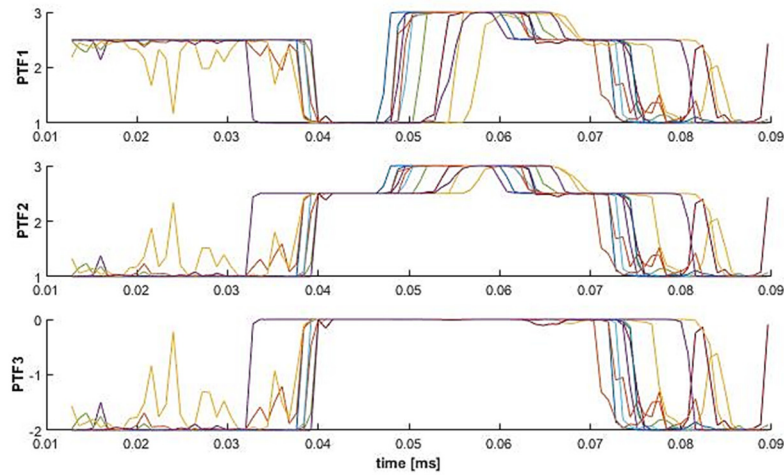
**Fig. 5b** A confusion matrix for the optimized CNN configuration (the same meaning as in Fig. 5a). It follows from the figure/numbers that the improvement of classification accuracy is only partial. This fact justifies usage of CNN with basic architecture and the default setting of control parameters.

all of the tested events were identified. Particular features of given  $PTFn$  formulas are discussed in the following paragraphs.

### 3. RESULTS

As described above we process set of 38 selected events. These events were cut into more than 3400 sub-images and this number represents also the real size of the input files submitted to CNN. Applying the method, we successfully identified events in all 11 randomly selected testing records from the processed set. Even three unexpected particular anomalies in records were also recognized (Fig. 6) and these events were studied in more details. In concordance with general CNN application rules, for training, validation, and data testing, subsets were randomly divided from the available data set. However, exceptional events number 8, 9, and 28, which exhibited some anomalies during method tuning, were forced to be members of the testing data to prove the method capabilities when loaded with anomaly types of records. Their description is, as follows: (i) for event number 8, the identification indicated a delay compared to the triggering time. In the seismogram (Fig. 7), it is evident that the majority of traces indicate a double event (here, we do not speculate as to whether it is really a double event, with a second stronger event, or some reflection that appears on the same

<sup>2</sup> In an ideal case the probability of identified class should be one, while the others should be zeros. However, the real data produced rather wide combination of considered class probabilities.



**Fig. 6** Superposition of PTF1, PTF2, and PTF3 (formula (1)-(3)) for a testing subset of 11 events randomly selected from an investigated set of 38 events (various colors/shadow intensities indicate separate events). The most deviating PTF behaviors are discussed individually in consequent Figures 7-9.

channels as a second event). The used CNN approach preferred to identify the second, more pronounced event. (ii) For event number 9, the identification was earlier than correspond to automatic trigger. In the seismogram, the trigger was activated not only by the first arrival but by the second arrival, and the event generally occurred sooner, see Figure 8. Such situations can exceptionally occur during the experiment. (iii) For event number 28, some unusual noise or false signal appeared during the pre-event time period. An analysis of seismograms indicated that for at least one sensor, a signal was not interpreted as noise. However, the signal was also not recognized as an event, see Figure 9. This result potentially indicates a part-event class of some other event, but not with a high probability.

Based on the data, we concluded that all of these incidences presented cases that proved the method's capability. With use of combination of CNN classification and posterior  $PTF_n$  function we were able to identify all events from testing data subset and excluded anomalies excited by other sources.

Based on the results provided in Figure 6, since their values significantly increased during the time of event occurrence, the posterior types PTF1 and PTF2 indicated rather "pure" event identification. Function PTF3 then seemed to be suitable for the identification of an event with a border (i.e. with a part-event and a part-coda parts), and the PTF2 seemed to be more robust for double events. PTF1 displayed the most figurative indication for part-event types (a decrease in the value). In general, differences, especially between PTF1 and PTF2, appeared to be small. A future analysis will indicate whether they bear significant information or whether they are, in practice, commutable.

We additionally processed an enlarged set of 660 events of which a subset of 200 events was used for

testing. Events were routinely preprocessed in the same manner as for the pilot set of 38 events, although a detailed analysis was not performed. The results are provided in Figure 10. The figure indicates good capability for the method in identifying event occurrence within the signal. Previously documented anomalies again appeared in the results, which indicate that such cases may occurred rarely but regularly during the recording.

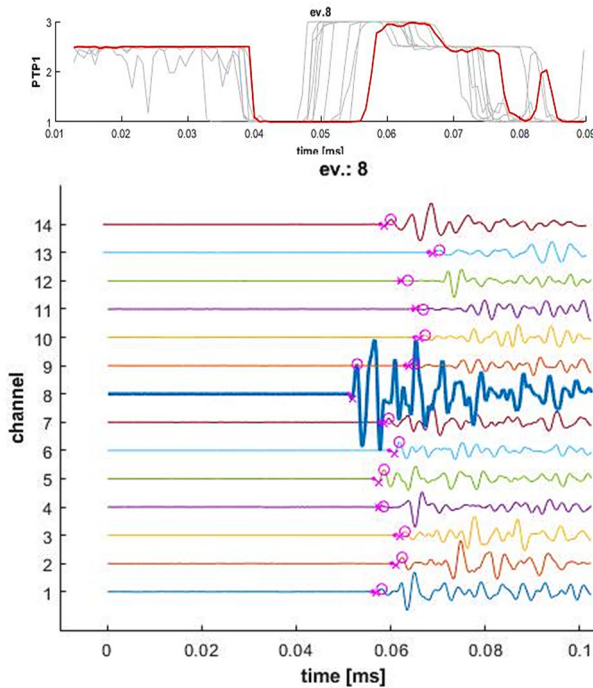
Some recently published works solved the simultaneous processing of multiple channels seismograms, as an example it could be mentioned e.g. in (Perol et al., 2018; Kriegerowski et al., 2019; Lomax et al., 2019). Also probability description of occurrence of searched feature (e.g. phase onset) is used e.g. in (Dokht et al., 2019). Our approach was developed independently on those works.

#### 4. DISCUSSION AND CONCLUSION

In this work, we introduced a new method for the identification of seismic events in multichannel seismograms using convolutional neural networks. The method consists of the following three principal steps:

1. Seismograms are represented in the time-frequency domain when each channel represents one layer of a multidimensional image that forms an input of the CNN.
2. The CNN employs three convolutional blocks. The CNN design employed standard architecture and learning parameters was set to default values. The performed optimization of CNN architecture as well as of the learning parameter values, only exhibited insignificant result improvements (i.e. our classification is not sensitive to CNN tuning and default setting can be employed).
3. To improve algorithm resolution, we suggest to use posterior type function instead of selecting

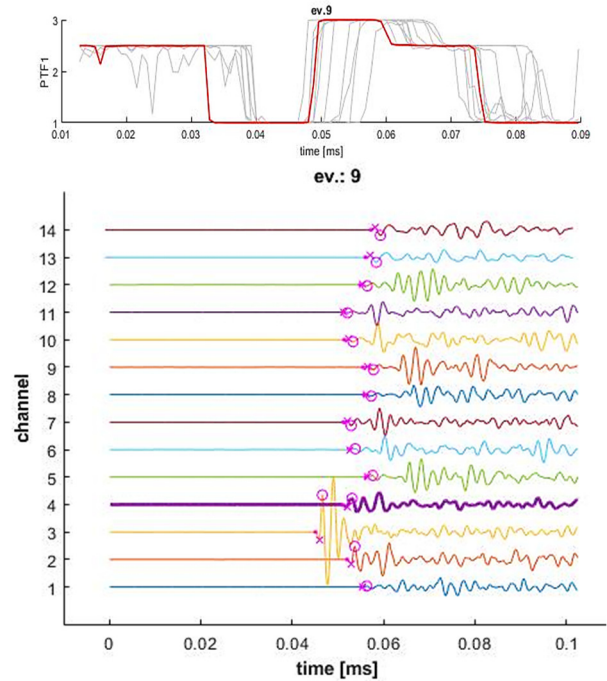




**Fig. 7** Example of a particular event identification for event number 8; PTF1 (above) and its corresponding seismogram (down). A significant PTF increase (red/sick line on the background for other event probabilities) was delayed. When seismograms were inspected, the first triggering event was clearly followed by the second event (namely on channels: 1, 5, 8, 12, and 14). The CNN preferred to identify the second, more pronounced, event. For our task, it was not (now) important whether or not the event was a real second event or only a reflection of the first event.

directly a class with the highest probability indicated directly by the CNN. The posterior type function is a combination of probabilities of all (five) considered classes. The functions' intervals with approximately constant values correspond to the ranges of individual classes, while the sudden changes indicate their limits. This step successfully solved the problem ambiguity for badly constrained classified signals.

The point (2) in our approach is formal up to certain point and the CNN architecture and control parameters setting follow more or less generally recommended schema. The point (1) – CNN input data in form of images multilayer spectrograms was adopted after numerous tests as the one yielding acceptable results. Within point (3) we convert series of “static” images back into time series. It can be directly matched to original signal time flow and the points of relevant signal occurrences can be identified. To our knowledge, such an identification of events in seismograms or other types of time series data have not been performed yet. In addition in our approach whole processing is delegated to Artificial Intelligence (to CNN in our case) and e.g. location of events is not performed subsequently in classical way from

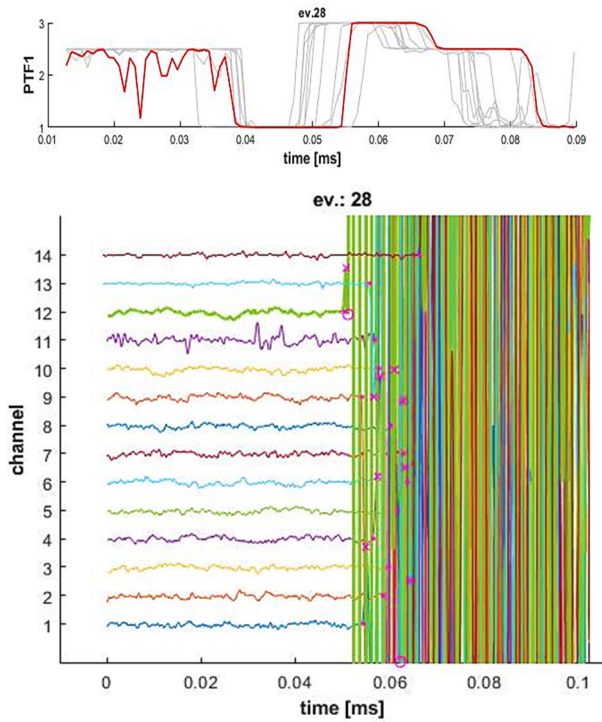


**Fig. 8** Example of a particular event identification for event number 9. Onset of the event was identified by an PTF that came earlier than expected (above). When the corresponding seismogram was inspected (down), it was evident that the record was triggered by the second arrival (channel 4) instead of the first occurrence (channel 3).

identified signal onsets as it is done e.g. in (Perol et al., 2018).

We achieved promising results for identifying seismic events in seismograms. We focused on AE events from loading experiments that, for our task, could be considered as laboratory earthquakes. The investigated data set was relatively simple and limited. That was advantageous for method development and testing and necessary before analyzing more complex input data. The successful identification of seismic signals within a seismogram opens the door to its more sophisticated (automatic) processing, such as onset interpretation, a crucial task that can more effectively be performed if event occurrence is successfully pre-indicated. Since our results for a limited set of AE events are promising, further development will include (i) the enlargement of the Acoustic Emission data sets and (ii) the inclusion of natural seismic events.

The method has a potential to deal with continuous AE signals that are becoming a standard with the increasing power of registration apparatus. Here, the reliable and effective AE event recognition is a crucial task necessary for any further processing and interpretation.



**Fig. 9** Example of a particular event identification for event number 28. Noise class indication by the PTF (above) was not constant for the record. When the seismogram (below) was inspected (here zoomed vertically), a signal mainly occurred in channel 11 (but also in channels 7, 12, and 13).

## ACKNOWLEDGEMENTS

This study was partially supported by Czech Science Foundation research grant 21-26542S, by the Czech Academy of Sciences project RVO 67985831, by the Institute of Geophysics (P.K. home institution) and by Institute of Geology (M.P. home institution).

Thanks to Humusoft co. ([www.humusoft.cz](http://www.humusoft.cz)), the international reseller of MATLAB, for organizing a Technical Computing Camp (2016, 2017, 2018) – an inspirational meeting of MATLAB users across various fields of science and technology.

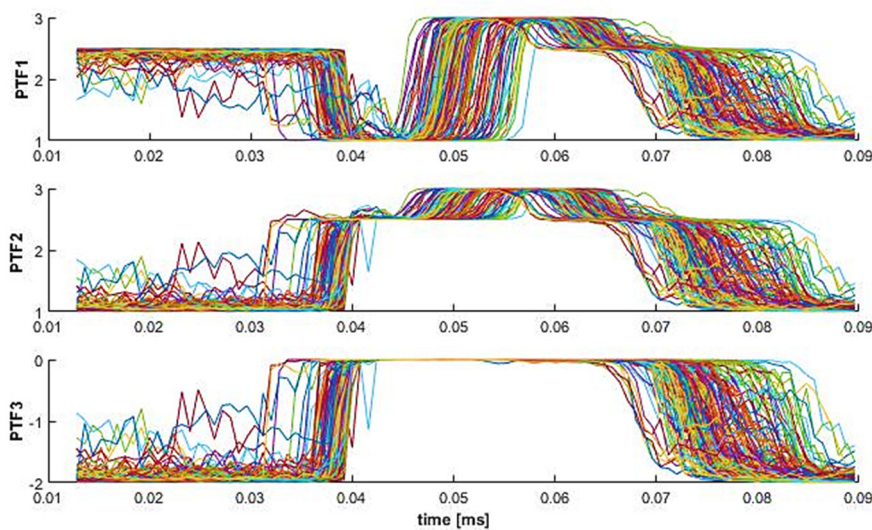
Loading experiments were performed in the Laboratory of Mechanical Properties of Rocks of the Institute of Geology of the Czech Academy of Science, located in Prague, Czech Republic.

Most of calculations were performed by LACoS (Laboratory for Advanced Computation in Seismology) of the Institute of Geophysics of the Czech Academy of Science, located in Prague, Czech Republic.

Thanks to two unknown reviewers for detailed and challenging comments.

## REFERENCES

- Bishop, C.M.: 2006, Pattern recognition and machine learning. Springer, New York.
- Blandford, R.R.: 1982, Seismic event discrimination. *Bull. Seismol. Soc. Am.*, 72, 569–587.
- Chen, Y., Zhang, G., Bai, M. et al.: 2019, Automatic waveform classification and arrival picking based on convolutional neural network. *Earth. Sp. Sci.*, 6, 1244–1261. DOI: 10.1029/2018EA000466
- Curilem, M., Canario, J.P., Franco, L. and Rios, R.A.: 2018, Using CNN to classify spectrograms of seismic events from Llaima Volcano (Chile). 2018 International Joint Conference on Neural Networks (IJCNN). IEEE, 1–8.



**Fig. 10** Superposition of posterior type functions for the 200 testing events (a subset) selected from the set of 660 events. The same as for Figure 6, but with a significantly enlarged data set. Behavior was similar to the originally investigated 11/38 events set.

- Dokht, R.M.H., Kao, H., Visser, R. and Smith, B.: 2019, Seismic event and phase detection using time-frequency representation and convolutional neural networks. *Seismol. Res. Lett.*, 90, 481–490. DOI: 10.1785/0220180308
- Doubravová, J., Wiszniowski, J. and Horálek, J.: 2016, Single layer recurrent neural network for detection of swarm-like earthquakes in W-Bohemia/Vogtland—the method. *Comput. Geosci.*, 93, 138–149. DOI: 10.1016/j.cageo.2016.05.011
- Grosse, C.U. and Ohtsu, M. (eds): 2008, *Acoustic emission testing*. Springer - Verlag Berlin Heidelberg. DOI: 10.1007/978-3-540-69972-9
- Harris, F.J.: 1978, On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE*, 66, 51–83. DOI: 10.1109/PROC.1978.10837
- Joswig, M.: 1990, Pattern recognition for earthquake detection. *Bull. Seismol. Soc. Am.*, 80, 170–186.
- Kitagawa, G. and Akaike, H.: 1978, A procedure for the modeling of non-stationary time series. *Ann. Inst. Stat. Math.*, 30, 351–363. DOI: 10.1007/BF02480225
- Kolář, P., Petružálek, M., Lokajíček, T. et al.: 2020, Acoustic emission events interpreted in terms of source directivity. *Pure Appl. Geophys.*, 177, 2, 1–8. DOI: 10.1007/s00024-020-02517-w
- Kong, Q., Trugman, D.T., Ross, Z.E. et al.: 2019, Machine learning in seismology: Turning data into insights. *Seismol. Res. Lett.*, 90, 3–14. DOI: 10.1785/0220180259
- Kriegerowski, M., Petersen, G.M., Vasyura-Bathke, H. and Ohrnberger, M.: 2019, A deep convolutional neural network for localization of clustered earthquakes based on multistation full waveforms. *Seismol. Res. Lett.*, 90, 510–516. DOI: 10.1785/0220180320
- Krizhevsky, A., Sutskever, I. and Hinton, G.E.: 2012, Image net classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.*, 25, 1097–1105.
- Lockner, D.A.: 1993, The role of acoustic emission in the study of rock fracture. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.*, 30, 883–899. DOI: 10.1016/0148-9062(93)90041-B
- Lockner, D.A., Byerlee, J.D., Kuksenko, V. et al.: 1991, Quasi-static fault growth and shear fracture energy in granite. *Nature*, 350, 39–42. DOI: 10.1038/350039a0
- Lomac, A., Michelini, A. and Jozinović, D.: 2019, An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network. *Seismol. Res. Lett.*, 90, 517–529. DOI: 10.1785/0220180311
- Nakano, M., Sugiyama, D., Hori, T. et al.: 2019, Discrimination of seismic signals from earthquakes and tectonic tremor by applying a convolutional neural network to running spectral images. *Seismol. Res. Lett.*, 90, 530–538. DOI: 10.1785/0220180279
- Paterson, M.S. and Wong, T.: 2005, *Experimental rock deformation—the brittle field*. Springer.
- Perol, T., Gharbi, M. and Denolle, M.: 2018, Convolutional neural network for earthquake detection and location. *Sci. Adv.*, 4, 2, e1700578. DOI: 10.1126/sciadv.1700578
- Petružálek, M., Jechumtálová, Z., Kolář, P. et al.: 2018, Acoustic emission in a laboratory: Mechanism of microearthquakes using alternative source models. *J. Geophys. Res., Solid Earth*, 123, 4. DOI: 10.1029/2017JB015393
- Petružálek, M., Jechumtálová, Z., Šílený, J. et al.: 2020, Application of the shear-tensile source model to acoustic emissions in Westerly granite. *Int. J. Rock Mech. Min. Sci.*, 128. DOI: 10.1016/j.ijrmms.2020.104246
- Petružálek, M., Lokajíček, T., Svitek, T. et al.: 2019, Fracturing of migmatite monitored by acoustic emission and ultrasonic sounding. *Rock Mech. Rock Eng.*, 52. DOI: 10.1007/s00603-018-1590-2
- Petružálek, M., Vilhelm, J., Lokajíček, T. and Rudajev, V.: 2007, Assessment of P-wave anisotropy by means of velocity ellipsoid. *Acta Geodyn. Geomater.*, 4, 23–31.
- Plesinger, A., Růžek, B. and Boušková, A.: 2000, Statistical interpretation of Webnet seismograms by artificial neural nets. *Stud. Geophys. Geod.*, 44, 251–271. DOI: 10.1023/A:1022119011057
- Plešinger, F., Klimes, P., Halamek, J. and Jurak, P.: 2016, Taming of the monitors: reducing false alarms in intensive care units. *Physiol. Meas.*, 37, 1313–1325. DOI: 10.1088/0967-3334/37/8/1313
- Ross, Z.E., Meier, M., Hauksson, E. and Heaton, T.H.: 2018, Generalized seismic phase detection with deep learning. *Bull. Seismol. Soc. Am.*, 108, 2894–2901. DOI: 10.1785/0120180080
- Schmidhuber, J.: 2015, Deep learning in neural networks: An overview. *Neural Netw.*, 61, 85–117. DOI: 10.1016/j.neunet.2014.09.003
- Sedlak, P., Hirose, Y., Khan, A.S. et al.: 2009, New automatic localization technique of acoustic emission signals in thin metal plates. *Ultrasonics*, 49, 254–262. DOI: 10.1016/J.ULTRAS.2008.09.005
- Sellers, E.J., Kataka, M.O. and Linzer, L.M.: 2003, Source parameters of acoustic emission events and scaling with mining-induced seismicity. *J. Geophys. Res., Solid Earth*, 108. DOI: 10.1029/2001JB000670
- Sick, B., Guggenmos, M. and Joswig, M.: 2015, Chances and limits of single-station seismic event clustering by unsupervised pattern recognition. *Geophys. J. Int.*, 201, 3, 1801–1813. DOI: 10.1093/gji/ggv126
- Thompson, B.D., Young, R.P. and Lockner, D.A.: 2009, Premonitory acoustic emissions and stick-slip in natural and smooth-faulted Westerly granite. *J. Geophys. Res., Solid Earth*, 114, B02205. DOI: 10.1029/2008JB005753
- Truong, N.D., Nguyen, A.D., Kuhlmann, L. et al.: 2018, Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Netw.*, 105, 104–111. DOI: 10.1016/J.NEUNET.2018.04.018
- Welch, P.D.: 1967, The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.*, 15, 70–73. DOI: 10.1109/TAU.1967.1161901
- Willmore, P.L. (ed): 1979, *Manual of seismological observatory practise*. World Data Center A for Solid Earth Geophysics, Boulder, Colorado.
- Zhou, Y., Yue, H., Kong, Q., Zhou, S.: 2019, Hybrid event detection and phase-picking algorithm using convolutional and recurrent neural networks, 90, 3, 1079–1087. DOI: 10.1785/0220180319
- Zhu, W. and Beroza, C.G.: 2018, PhaseNet: A deep-neural-network-based seismic arrival time picking method. *Geophys. J. Int.*, 216, 261–273. DOI: 10.1093/gji/ggy423