Classifying earthquakes and mining activity with deep neural networks

András Horváth[1,2], Máté Timkó[1,3], Márta Kiszely[1]
Tamás Bozóki[1] ,István Bozsó[1,3],Lukács Kuslits[1]

1, Institute of Earth Physics and Space Science, Sopron, Hungary
2, Faculty of Information Technology and Bionics, Péter Pázmány Catholic University, Budapest, Hungary
3, Doctoral School of Earth Sciences, Faculty of Science, Eötvös Loránd University, Budapest, Hungary

EGU 2022.05.26

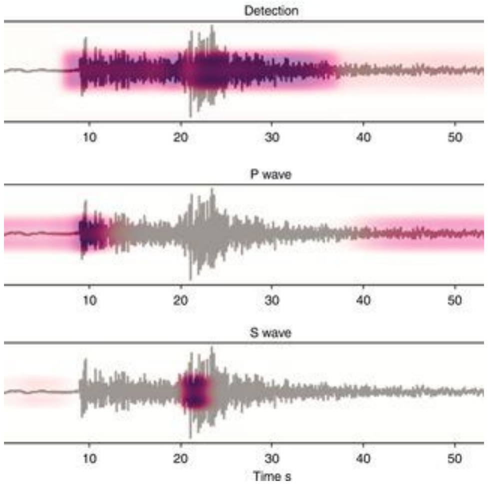
Application in seismology

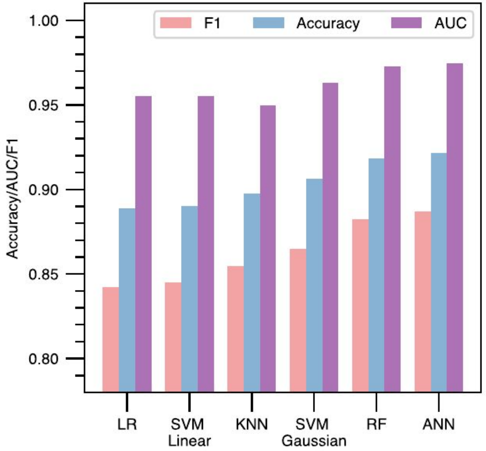
There are existing solutions for earthquake detection:

- manual labour (tedious)
- classical methods (not too accurate)
- machine learning (depends heavily on the dataset)

Chai, C et al. Automatic waveform quality control for surface waves using machine learning. Seismological Society of America (2022)

Mousavi et al. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature communications (2020)

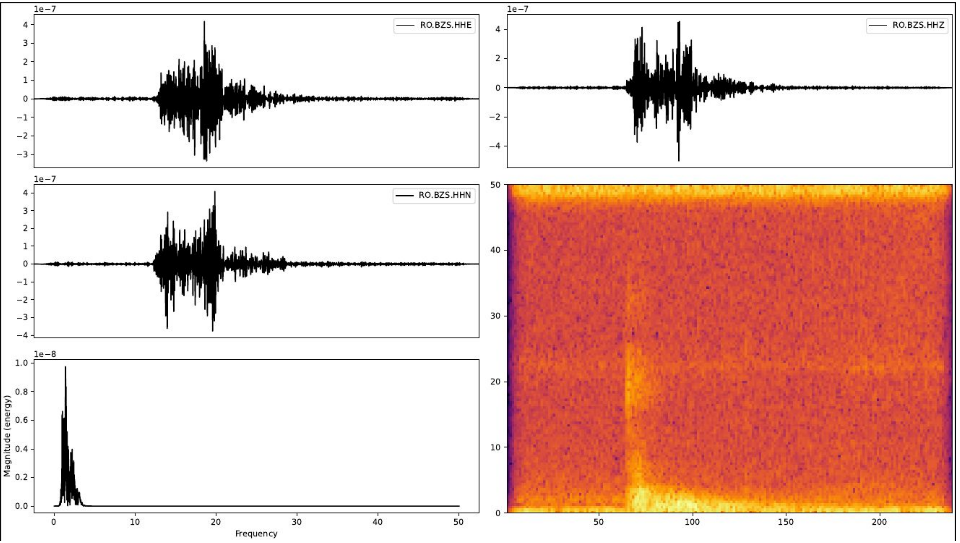


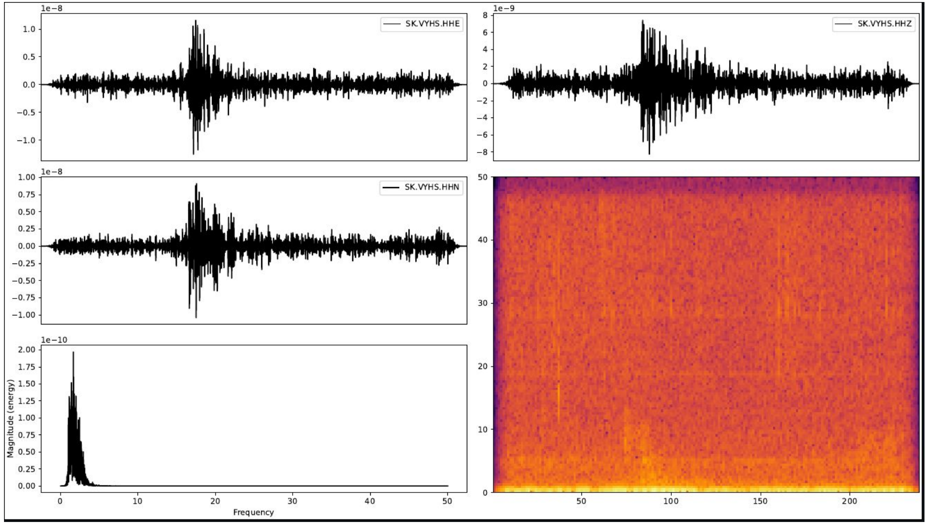


Soucre: https://www.nature.com/articles/s41467-020-17591-w

Manual identification

Classification using only the waveforms is a complex task even for an expert





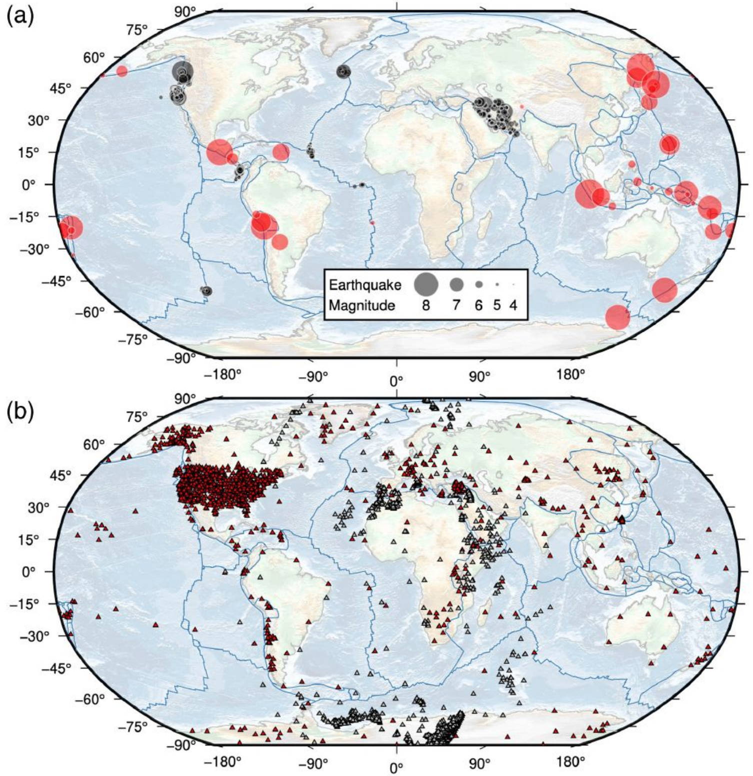
Application in seismology

Machine learning based solutions (especially deep neural networks) provide the best performance

These methods are trained on datasets containing „earthquakes” and „no activity” classes

The provide poor performance in case of other activities:
Typically mining activity can easily fool them (**98%** of the investigated samples have fooled EQ-transformer)

Not all mines report activities
Small amplitude signals → easily confused with other activities



Our Dataset

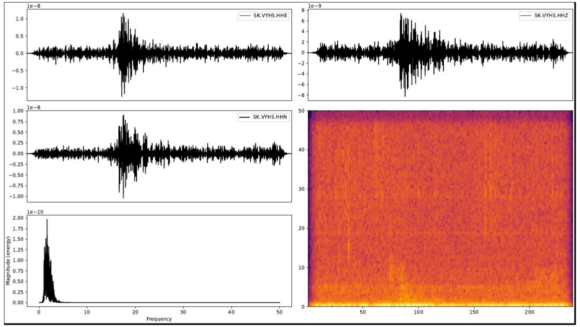
Created from the Hungarian Seismology Bulletin (1995-2020)

Input: Preprocessed temporal data

North, East, Z components
100 Hz sampling frequency

5 minute long timeseries which contains activity for all samples

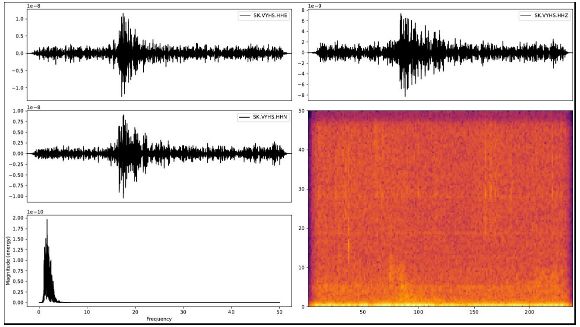
Two classes:
A, Seismic activity
B, Mining activity




```
Event HU20131018105830  Czech and Slovak Republics

   Date       Time        Err   RMS Latitude Longitude  Smaj  Smin  Az Depth   Err Ndef Nsta Gap  mdist  Mdist Qual   Author      OrigID    Rep   DPdep   Err Sgp
2013/10/18 10:58:30.56   1.38  0.51  48.3393   19.8547   6.0   5.3  56   0.0S         8    4 106   0.42   0.85 m   kx BUD       14092899   iLoc               212
2013/10/18 10:58:30.88f        0.52  48.3466   19.8590f                  0.0S         8      105   0.43   0.85 m   kx BUD_GT    14114118   BUD                210
 (#PRIME)
 (locality :  Velke Dravce)
 (GT info :  GT2 explosion)

Magnitude  Err Nsta Author      OrigID
ML     1.8          BUD       14114118

Sta     Dist  EvAz Phase        Time      TRes  Azim AzRes   Slow   SRes Def   SNR       Amp   Per Qual Magnitude    ArrID    Agy   Deploy   Ln Auth  Rep   PCh ACh L
PSZ     0.43 176.8 Pg       10:58:40.000  -0.3                           T__                     __e              14114120   FDSN  IR       -- BUD   BUD   ??? ??? _
PSZ     0.43 176.8 Lg       10:58:47.000  -0.4                           T__                     __e              14114121   FDSN  IR       -- BUD   BUD   ??? ??? _
KECS    0.44  71.6 Pg       10:58:39.500  -0.3                           T__                     __e              14114122   FDSN  IR       -- BUD   BUD   ??? ??? _
KECS    0.44  71.6 Lg       10:58:46.000  -0.5                           T__                     __e              14114123   FDSN  IR       -- BUD   BUD   ??? ??? _
VYHS    0.70 282.6 Pg       10:58:44.600  -0.3                           T__                     __e              14114124   FDSN  IR       -- BUD   BUD   ??? ??? _
VYHS    0.70 282.6 Lg       10:58:54.200  -1.0                           T__                     __e              14114125   FDSN  IR       -- BUD   BUD   ??? ??? _
LANS    0.85 342.3 Pg       10:58:48.000   0.5                           T__                     __e              14114126   FDSN  IR       -- BUD   BUD   ??? ??? _
LANS    0.85 342.3 Lg       10:58:59.500  -0.4                           T__                     __e              14114127   FDSN  IR       -- BUD   BUD   ??? ??? _
```

Our Dataset

Created from the Hungarian Seismology Bulletin (1995-2020)



Two classes:
A, Seismic activity
B, Mining activity

3400 events → labeled by human experts
(knowing the event times locations etc..., not just the waveforms)
66% of them are mining activity: 2000 samples were used to train the models

Each event was recorded
by multiple stations

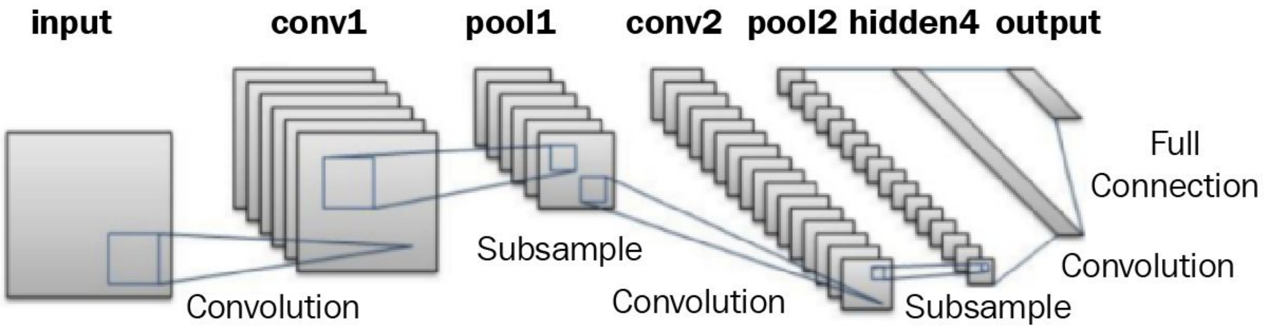
```
Event HU20131018105830  Czech and Slovak Republics

   Date       Time        Err   RMS Latitude Longitude  Smaj  Smin  Az Depth   Err Ndef Nsta Gap  mdist  Mdist Qual   Author      OrigID    Rep   DPdep   Err Sgp
2013/10/18 10:58:30.56   1.38  0.51  48.3393   19.8547   6.0   5.3  56   0.0S         8    4 106   0.42   0.85 m   kx BUD       14092899    iLoc              212
2013/10/18 10:58:30.88f        0.52  48.3466   19.8590f                  0.0S         8      105   0.43   0.85 m   kx BUD_GT    14114118    BUD               210
 (#PRIME)
 (locality :  Velke Dravce)
 (GT info :  GT2 explosion)

Magnitude  Err Nsta Author      OrigID
ML     1.8          BUD       14114118

Sta      Dist  EvAz Phase        Time      TRes  Azim AzRes   Slow   SRes Def   SNR       Amp   Per Qual Magnitude    ArrID    Agy   Deploy   Ln Auth  Rep   PCh ACh L
PSZ      0.43 176.8 Pg       10:58:40.000  -0.3                           T__                      __e              14114120   FDSN  IR       -- BUD   BUD   ??? ??? _
PSZ      0.43 176.8 Lg       10:58:47.000  -0.4                           T__                      __e              14114121   FDSN  IR       -- BUD   BUD   ??? ??? _
KECS     0.44  71.6 Pg       10:58:39.500  -0.3                           T__                      __e              14114122   FDSN  IR       -- BUD   BUD   ??? ??? _
KECS     0.44  71.6 Lg       10:58:46.000  -0.5                           T__                      __e              14114123   FDSN  IR       -- BUD   BUD   ??? ??? _
VYHS     0.70 282.6 Pg       10:58:44.600  -0.3                           T__                      __e              14114124   FDSN  IR       -- BUD   BUD   ??? ??? _
VYHS     0.70 282.6 Lg       10:58:54.200  -1.0                           T__                      __e              14114125   FDSN  IR       -- BUD   BUD   ??? ??? _
LANS     0.85 342.3 Pg       10:58:48.000   0.5                           T__                      __e              14114126   FDSN  IR       -- BUD   BUD   ??? ??? _
LANS     0.85 342.3 Lg       10:58:59.500  -0.4                           T__                      __e              14114127   FDSN  IR       -- BUD   BUD   ??? ??? _
```

Quantitative results

Five-fold cross validation with 80-20 split

	4 layered ConvNet	AlexNet 8 layers	4 layered Transformer
Accuracy	94.4 %	92.6%	91.3%

The decreasing accuracy could be caused by low amount of training data and increased number of parameters → overfitting



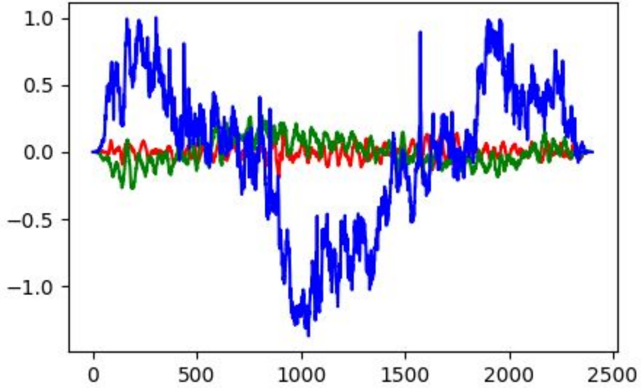
Explainable AI

Neural networks are considered black boxes

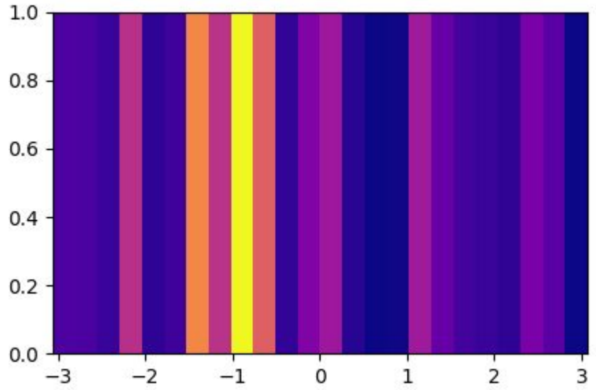
With saliency map generation approaches one can visualize the regions which were important for the network

We an ask experts about this features, whether they are meaningful or not

We can use these maps in data augmentation to preserve important regions



North, East, Z components



Importance map for the time series

Conclusions

- We collected a fairly large dataset containing seismic and mining activities, created a human baseline for this task

- Implemented data augmentation methods specially tailored for seismic activity

- We have demonstrated the the classification of seismic activities is solvable via deep neural networks

In case of any further questions please write us:

horvath.andras@itk.ppke.hu