Otto-von-Guericke-University Magdeburg
Faculty for Computer Science

A hybrid deep-learning approach for reliable real-time assessment of high magnitude earthquakes

Master Thesis

Author:
Viola Hauffe

Advisor:
Prof. Sebastian Stober

Other advisors:
Dr. Marius Kriegerowski
Jannes Münchmeyer

Magdeburg, 23rd August 2021

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG FAKULTÄT FÜR INFORMATIK

GFZ Helmholtz-Zentrum POTSDAM

Contents

Abstract

wie angekündigt hier die Paper, an die ich gedacht hatte, und nochmal eine Kurzbeschreibung davon, was ich bei Hybridsystem im Kopf hatte.

Das Gerät zeichnet kontinuierlich Wellenformen auf und auf denen müsste man dann Picking machen, also feststellen wann eine seismische Welle auftrifft. Dazu gibt es inzwischen einen ganzen Stapel Paper. Ein guter Einstieg ist Ross et al. (2018) "Generalized Seismic Phase Detection with Deep Learning". Entweder damit gekoppelt (indem man die Trainingsdaten geeignet zusammenstellt) oder als nächster Schritt, muss man unterscheiden, ob es tatsächlich ein Erdbeben oder nur impulsive noise ist. Hier gibt es zum Beispiel Li et al. (2018) "Machine learning seismic wave discrimination: Application to earthquake early warning".

Der übliche Deep Learning Ansatz wäre jetzt Magnitude und Lokalisation direkt zu schätzen, siehe z.B. Mousavi et al. (2019) "A Machine-Learning Approach for Earthquake Magnitude Estimation" und Mousavi et al. (2019) "Bayesian-Deep-Learning Estimation of Earthquake Location from Single-Station Observations". Mein Vorschlag war jetzt, Deep Learning nur für die Distanz zu nutzen und für die Magnitude eine parametrische Modellierung zu wählen. Das ist im early warning eine etablierte Methode, um Magnituden schnell zu schätzen (z.B. Zollo et al. (2006) "Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records"). Für eine genauere Schätzung könnte man wahrscheinlich Teile der Methode auf Nutzung in Echtzeit adaptieren, die ich hier vorgeschlagen habe (Münchmeyer et al. (2020) "Low uncertainty multifeature magnitude estimation with 3-D corrections and boosting tree regression: application to North Chile"). Eine Echtzeit Magnitudenschätzung wäre schon ein exzellentes Ergebnis und könnte zum Beispiel ziemlich direkt für early warning verwendet werden.

Todo list

Index of Notation

In a lot of cases it makes sense to give an overview over your mathematical notation.

Mathematical

x	Point in 3D space
$\overrightarrow{\mathrm{xy}}$	Normalized direction vector from x to y
$\mathbf{v}$	Direction vector in 3D space
$\mathrm{p}_x, \mathbf{v}_x$	x component of point / vector
$\mathbf{v} \cdot \mathbf{w}$	Dot product of vectors $\mathbf{v}$ and $\mathbf{w}$
$(\mathbf{v} \cdot \mathbf{w})^+$	Dot product of vectors $\mathbf{v}$ and $\mathbf{w}$ with negative values clamped to zero
$\mathbf{v} \times \mathbf{w}$	Cross product of vectors $\mathbf{v}$ and $\mathbf{w}$
$\|\|\mathbf{v}\|\|$	Euclidean length of vector $\mathbf{v}$
$\hat{\mathbf{v}}$	Normalized vector $\mathbf{v}$

Quantities & Functions

A	**Area**
ω	**Solid Angle**
ϕ	**Radiant Flux**, light power
I	**Radiant Intensity**, flux density per solid angle
E	**Irradiance**, flux density per area
L	**Radiance**, flux density per area per solid angle
ρ	**Reflectance**, ratio between incoming and outgoing flux
f_r	**BRDF**, function on the relation between irradiance and outgoing radiance

1 Introduction

Earthquakes are still an ongoing thread for humans, infrastructure and buildings. There were some big earthquakes in the last years. (include photo) As technology improves early warning systems get more elaborate each year. Much work is put into researching earthquake types, earthquake physics simulation and the prediction of shaking or magnitude of an arriving earthquake.

1.1 Background

1.1.1 Seismometers

Since the 1900s seismographs were developed. Today's seismometers can measure ground motion very precisely. In earthquake rich countries there exit big networks of seismometers. We will also use such a network of seismometers in our dataset.

1.1.2 Strength of an earthquake

While we will have a look at the magnitude of an earthquake its important to keep in mind, that, while the magnitude is based on the physical properties of an earthquake, it is not the sole factor how strong the shaking is, or how much damage is going to occur.

1.2 Motivation

1.3 Goals

Algorithm goals: Safety before accuracy, detect big earthqaukes, even though they are underrepresented in data, estimate distance, detect, and estimate magnitude Algorithmusentwurf, Ziele Schnelligkeit, Embeddedgeeignet, läuft auf schlechten Sensordaten, sagt Magnitude, Epizentrum mit einer gewissen Sicherheit richtig voraus

1.4 Tasks

Algorithm implementation, testing

1.5 Data

what kind of data, which structure, which dataset

1.6 structure and contents of thesis

explain how the thesis is going to explain the work

2 Prerequisites

2.1 Earthquake Physics

2.1.1 The Earthquake Event

An earthquake are caused by movements within the earth crust. These might result from earth plates moving under or against each other or artificial causes like mining works. After building up a lot of stress against friction, rock fractures along a fault line while the other tectonic layers can slip past each other. The fault line can even be visible on the surface.

The point where the fracturing starts is called hypocenter, while the epicenter is this point's projection onto the surface as seen in figure 2.1. The breakage can extend to other points along the fault line as well as cause more ruptures.

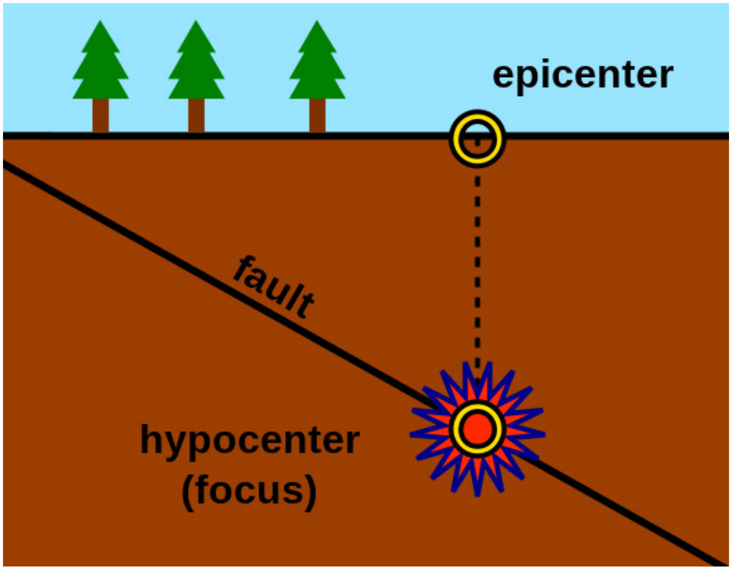

Figure 2.1: Hypocenter and Epicenter

2.1.2 Seismic Waves

The energy, which is emitted as the fracturing occurs, travels through the earth crust as seismic waves. We denote different types of waves, which travel at different velocities. Soil and material type (air, liquid or solid) also affect wave arrival times. Waves also reflect on certain materials, creating many sub-types of waves to arrive.

The first wave to arrive at a site is the p-wave or primary/pressure wave. The propagating

waves moves the material back and forth and is therefore also known as a compressional wave.
S-waves, second or shear waves move the material in a right angle to the movement direction. Contrary to p-waves, s-waves cannot move through liquid and are slightly slower than p-waves. Figure 2.2 shows this relationship. S-Wave and P-Waves are often

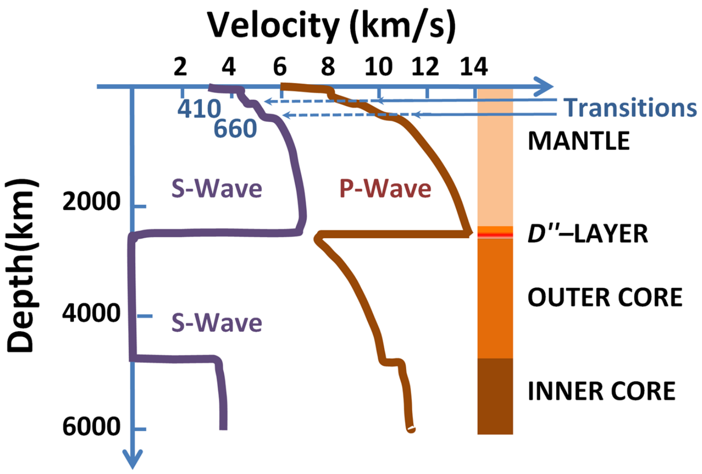

Figure 2.2: Velocity of S and p Wave

denoted as body waves. Surface waves arrive even later and are the most damaging ones. Two important waves are Rayleigh and Love waves. Rayleigh waves cause the surface to "ripple", they move the ground up and down while Love waves cause horizontal shearing. We can see all mentioned types of waves in figure 2.3.¸

2.2 Recording an Earthquake

2.2.1 Seismic Data

The seismic waves which arrive on a site can be recorded by a seismometer, which is the instrument recording ground motion as displacement, velocity or acceleration. A seismograph is the system build around the seismometer. Seismograms show the measured motions as a 1D signal for the three motion orthogonal axes, namely north-south, east-west and ground-up. The recordings depend on the sensitivity of the sensors, but also on earthquake and location. For further reading see Stein and Wysession [2003], page 398. An example for a recording of one axis at an earthquake arrival can be seen in figure 2.4. The change from noise to P-Wave to S-Wave arrival is clearly visible in the example.

2.2.2 Networks

Multiple seismometers might be organized in a network, where stations are linked together virtually or physically and their data can be processed nearly simultaneously. Network goals are often earthquake prediction with the according depth and distance, but network size, sensor type and network range can be drastically different depending on the type and magnitude of earthquake which must be detected. As the emitted waves get weaker depending on distance and soil type good coverage is crucial to get realistic predictions. This is especially important for location computation, as we need three or more sensors on the corners of a preferably equilateral triangle to get a precise location with the circle method. More information on the organisation of sensor networks can be found in Havskov and Alguacil [2016].

2.2.3 Early Warning

2.3 Differences between regression and classification

2.3.1 Typical networks and techniques

2.4 Estimation Representation

It is in important to keep in mind, that we do not take into account the uncertainty of our magnitude model and the representation of an earthquake with just ground motion sensors is itself a simplification.

2.4.1 Types of uncertainty

We divide between epistemic and aleatoric uncertainty.

- Epistemic uncertainty is our model uncertainty. The model can be unsure about certain parts of our dataset, because it has seen too few data points, but it can also be lacking the design to wholly learn and understand the dataset. The model can also be unsure, if we are unsure about the model's parameters for prediction or if we have multiple models to choose from when predicting.
  Epistemic uncertainty can be reduced by adding more training examples and choosing a good model architecture.

- Aleatoric uncertainty is uncertainty engrained in the data. This includes instrument noise and (over-)simplification of data information. We further divide between homoscedatic and heteroscedatic aleatoric noise. While homoscedatic uncertainty stays the same for all data point examples, as it would be with a constant sensor-offset noise, heteroscedatic uncertainty can be different for each input sample and is much harder to track and.

2.4.2 Uncertainties in our model

There are a few key points in the way our dataset is build which contribute to aleatoric uncertainty. While seismometers are extremely precise instruments, they are not completely free from sensor noise. We remove the sensor response in preprocessing, but they are still physical instruments, which cannot guarantee total precision, although it is more than enough for our model.
Even more, our learning goals, distance and magnitude are itself hard to determine. The distance to an earthquake has to be assumed from the emitted waves recorded by multiple seismometers and relies on certain simplifications of earthquake physics. As this is the same for the magnitude, the magnitude itself is a physical concept, that tries to capture the strength of an earthquake, but uses a simplified representation of the energy the earthquake emits, which again is only computed on the basis of our recorded data.
Epistemic uncertainty is far easier to detect. As also seen in /refdataset our data is not equally distributed and we are lacking examples for higher magnitudes and (naturally, as small earthquakes cannot be detected by remote sensors) higher distances.
As distance and magnitude are measurements of our sensor data, the model uncertainty also changes while the signal goes in and out, as earthquake waves arrive and depart. We do expect that, because both magnitude and distance are linked to the wavelength and amplitude of certain emitted waves.

2.4.3 Implementation

A very basic technique is taking the softmax output of a classification network as an uncertainty measure. While this is a good rule-of-thumb, most larger networks are overconfident in their predictions, meaning their accuracy is in general lower than the certainty that is suggested by the softmax output. Guo et al. [2017] propose a few methods for improvement.
For Bayesian neural networks instead of a single value a gaussian distribution is placed over the weights. Techniques similar to "Bayes by Backprop" Blundell et al. [2015] allow learning the weight distributions with an algorithm very similar to standard backpropagation. When predicting an example, the weights are sampled from the distribution every time, giving an uncertainty measure for this example after a reasonable number of runs.
In contrast to Bayesian neural networks, which require a change of loss functions and learning routines, "Monte Carlo Dropout" does not need this. Instead of just using dropout during training, Gal and Ghahramani [2016] apply dropout also when predicting examples and generate an uncertainty measure by aggregating the (all slightly different) outputs.
An even simpler method, that can be used for regression networks, is not learning a single valued output, but the mean and standard deviation of a gaussian distribution by changing the loss function and the last layer of the network. The mean can then be used as a output prediction, while the standard deviation gives an uncertainty.

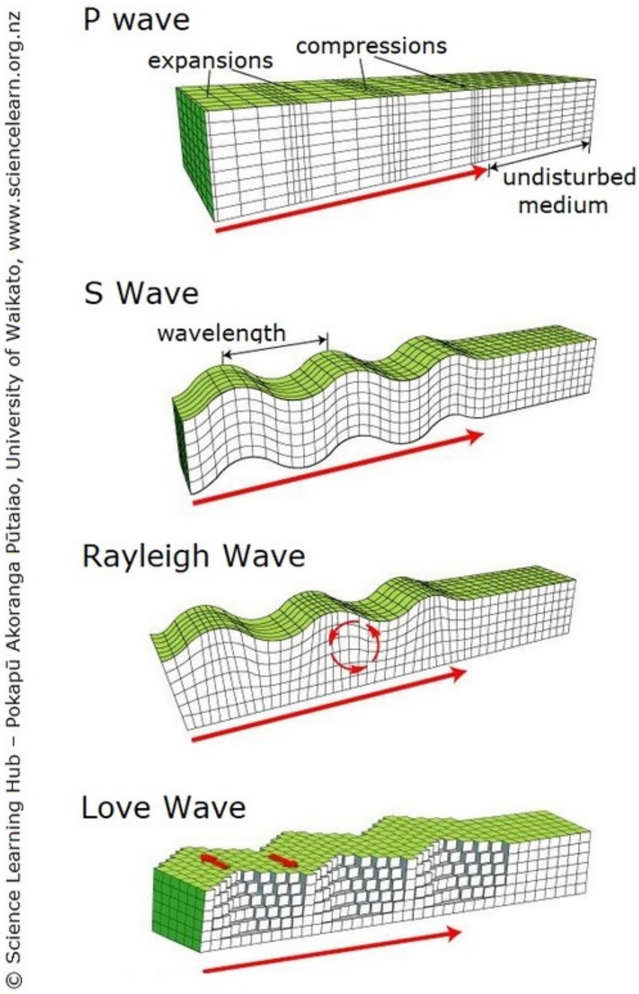

Figure 2.3: Overview of seismic waves

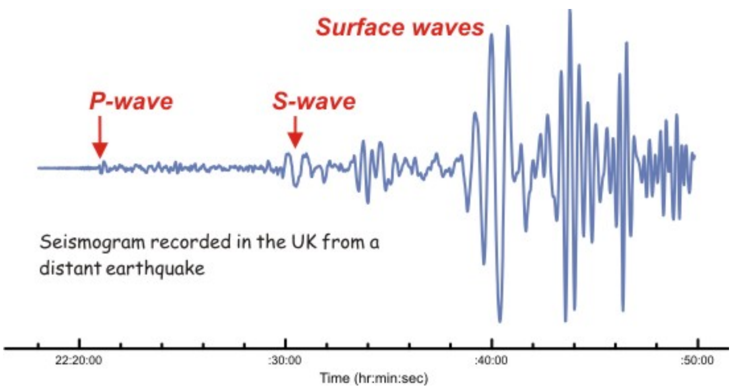

Figure 2.4: One axis recording of a seismograph

3 Related Work

Ross et al. (2018) "Generalized Seismic Phase Detection with Deep Learning"Ross et al. [2018]

short introduction about early warning in literature?

3.1 Earthquake Detection

3.2 Estimating magnitude and distance

3.3 Parametric Methods

4 Algorithm

4.1 Algorithm Design and requirements

Our aim is to design a framework, which is in general suitable for early warning. Therefore we decided to firstly implement an algorithm which detects an incoming earthquake, as this makes it easier to eventually use the whole project later and gives us good reason to continue computing the magnitude, if we are sure, that there is really an earthquake. As time is also an important factor, we just use data of a single station. While it is much more reliable to use a network of seismometers, we omitted this due to time and complexity overhead. Plus, the project was designed to work on small devices, which might not be able to process lots of data at once. As we wanted a reliable prediction, which would be able to include the uncertainty of the dataset, we want to at least capture the uncertainty of our deep neural network by not learning a distinctive value, but a Gauss function with an expected value and a variance, representing the uncertainty.

4.2 Data overview

4.2.1 The Dataset

4.2.2 Data preparation

Before using the seismometer data as an network input, we are preparing it for neural network use. This means we normalize it between 0 and 1, remove any trend by aligning start and end point at zero horizontally and filtering it with a two factor butterworth filter. We are doing that separately for every input window.

4.3 Detecting the earthquake

4.4 Ground-truth algorithm

The algorithm which we will use to compare our new technique to, is a simple CNN network, similar to the network we will use in our algorithm. It directly gives us a value for the magnitude from a 20 second seismometer input. Our proposed algorithm will be evaluated against a basic algorithm on the same dataset.

4.5 Proposed algorithm

The proposed algorithm consists of two parts: At first we take the whole 20 second input and compute a distance to the earthquake. Then we estimate the magnitude by using the formula proposed in cite

4.6 Bringing it all together

´

5 Evaluation

In evaluation it is important to not take into account any input, where the network already sees the s-wave. While we have included these examples into the learning, we now do not want to include those when evaluating the performance.

6 Conclusion

6.1 Summary

Sum up your work. Similar to the abstract but more technical.

6.2 Evaluation

Be hard with yourself, but not too hard. Stay scientific!

6.3 Future Work

What would be next? What did your thesis not touch?

Appendices

A. A section

Weird stuff that you didn't want to put into the main text but didn't want to leave it out either? You found the right place for it.

Bibliography

C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra. Weight uncertainty in neural network. In *International Conference on Machine Learning*, pages 1613–1622. PMLR, 2015.

Y. Gal and Z. Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059. PMLR, 2016.

C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger. On calibration of modern neural networks. In *International Conference on Machine Learning*, pages 1321–1330. PMLR, 2017.

J. Havskov and G. Alguacil. *Seismic Networks*, pages 261–308. Springer International Publishing, Cham, 2016. ISBN 978-3-319-21314-9. doi: 10.1007/978-3-319-21314-9_8. URL https://doi.org/10.1007/978-3-319-21314-9_8.

Z. E. Ross, M.-A. Meier, E. Hauksson, and T. H. Heaton. Generalized seismic phase detection with deep learning. *Bulletin of the Seismological Society of America*, 108 (5A):2894–2901, 2018.

S. Stein and M. Wysession. *Introduction to Seismology, Earthquakes, and Earth Structure*. Blackwell Publishing, 2003.

Statement of Authorship / Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbstständig und ausschließlich unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.
Die aus fremden Quellen direkt oder indirekt übernommenen Stellen sind als solche kenntlich gemacht.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form weder einer anderen Prüfungsbehörde vorgelegt oder noch anderweitig veröffentlicht.

Unterschrift	Datum