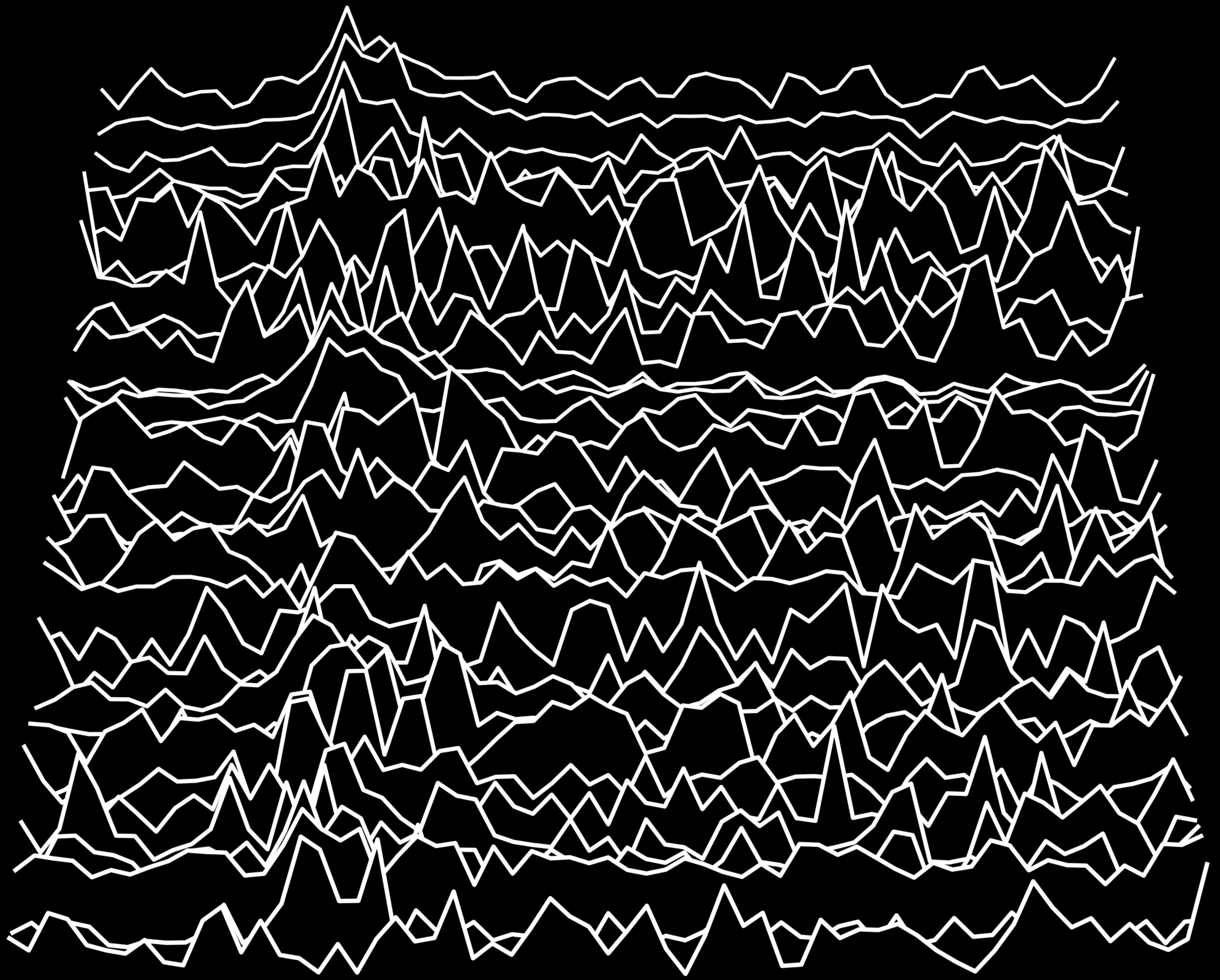
Tackling earthquake detection problem using 1D Convolutional Neural Networks

Josipa Majstorović, Sophie Giffard-Roisin and Piero Poli

UNKNOWN **FEATURES**



5èmes Rencontres Scientifiques et Techniques Résif, 17 November 2021, Obernai

josipa.majstorovic@univ-grenoble-alpes.fr

MONIFAULTS

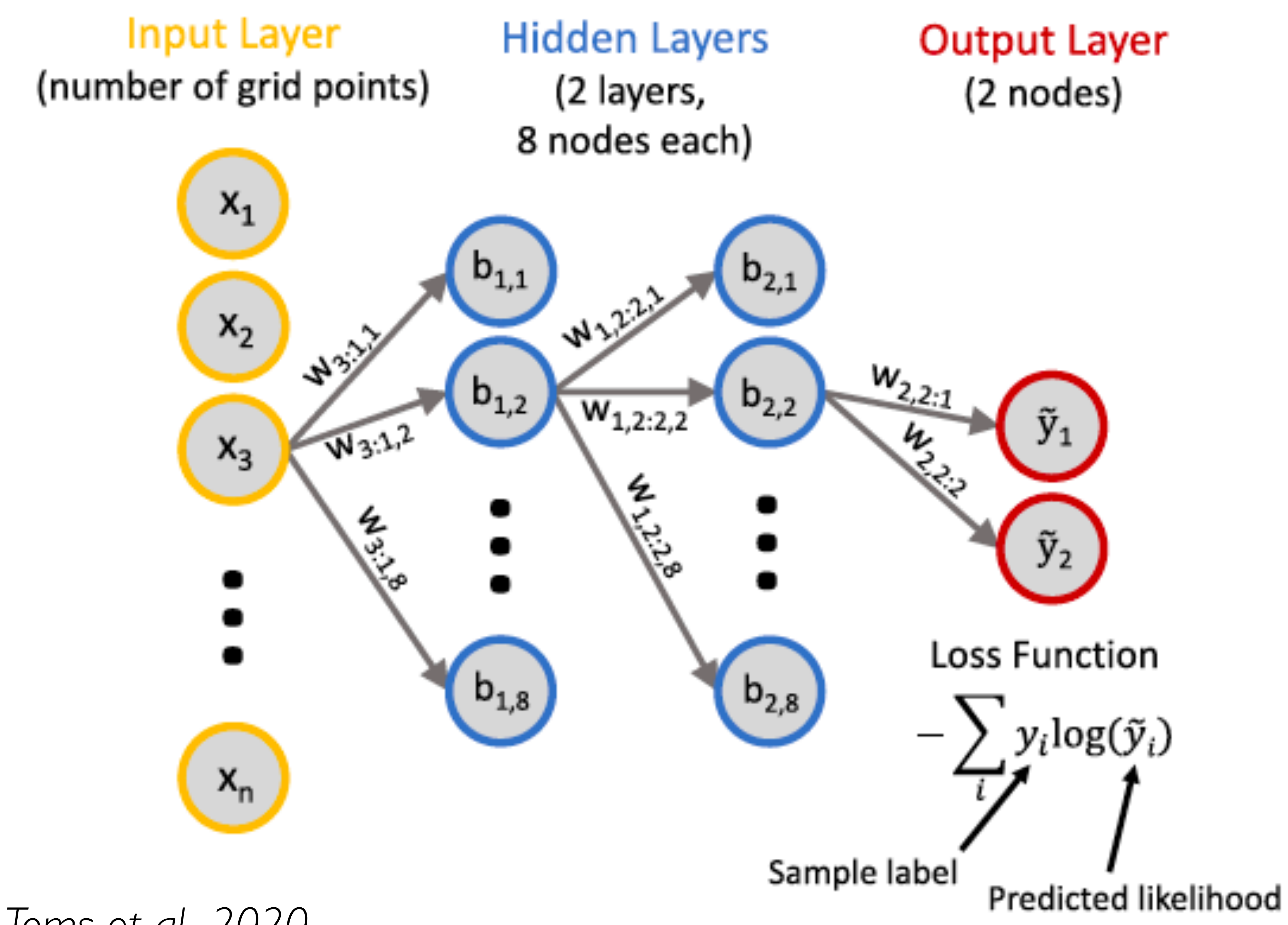
Basics of neural networks, deep learning models

Developing CNN detector

What does interpretation stand for in DL?

Goal: to further understand how we can exploit the existing DL models for the earthquake detection and whether we can find some other application.

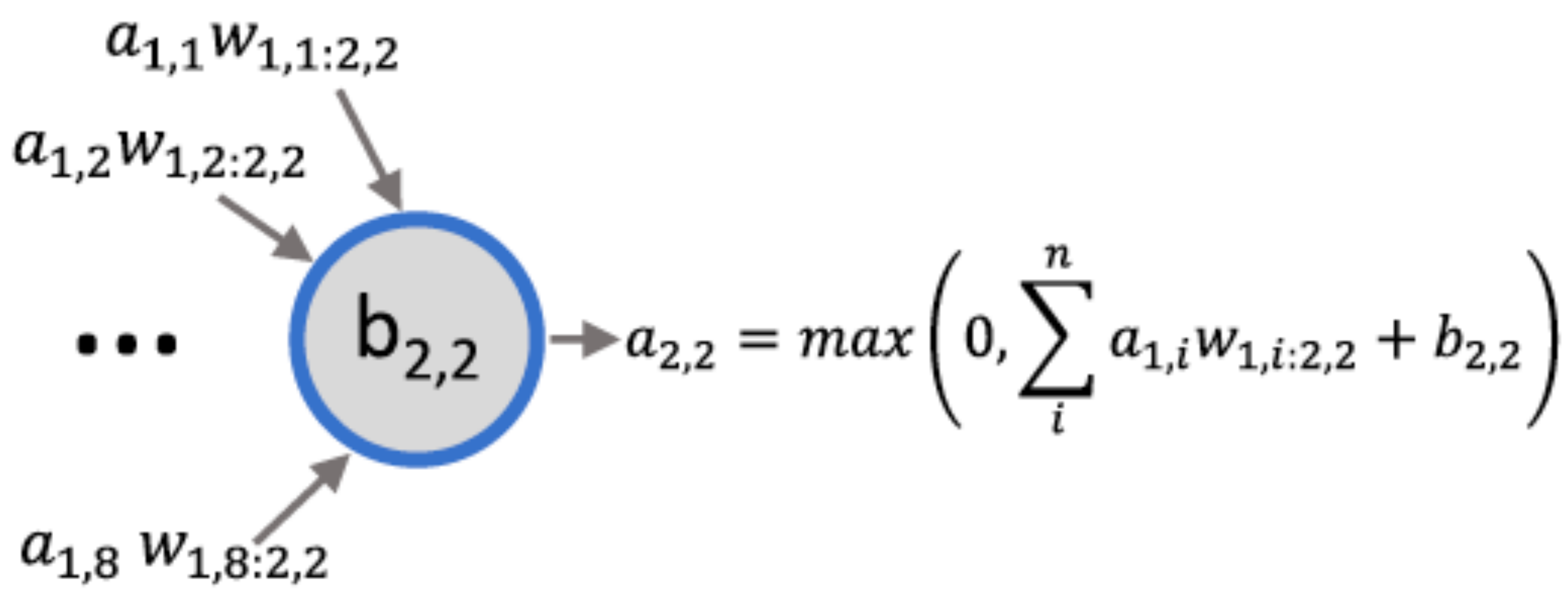
NEURAL NETWORK



Toms et al., 2020

Deep learning - Collection of neural networks, biologically inspired networks, that extract abstract features from the data in a hierarchical fashion.

BASIC UNIT - NEURON (NODE)



During the training process of the neural network we adjust the weights and biases.

The performance of the NN model highly depends on:

1	Training data	Datasets and labels!
2	Training process	Hyperparameters such as optimising algorithms, learning rates …
2	Modelling approach	Classification vs regression
3	NN architecture	Hyperparameters such as nb. of neurons, nb. of layers…

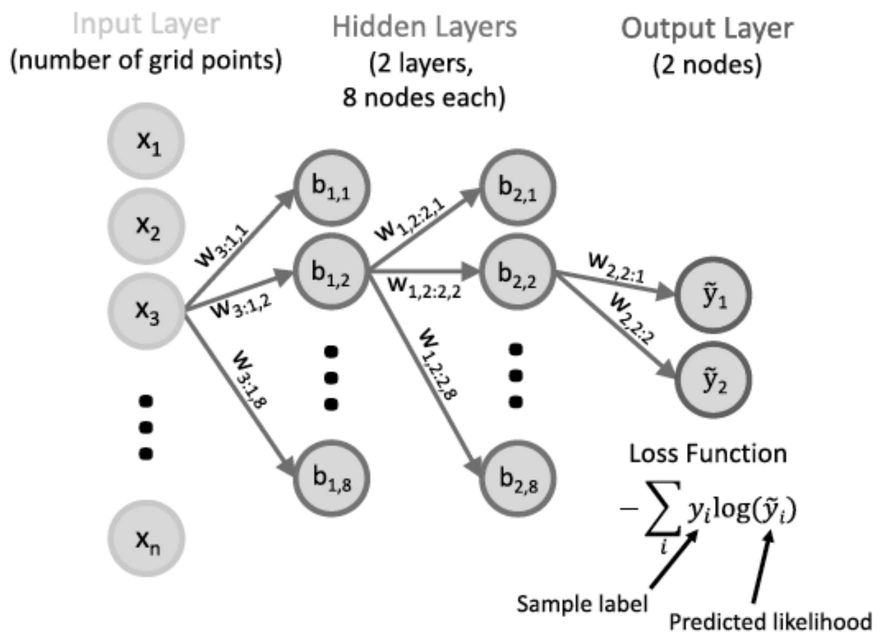
M. Mousavi presentation

Decision making process

Basics of neural networks, deep learning models

Developing CNN detector

Majstorović et. al., 2021, JGR



1 Training data

2 Training process

2 Modelling approach

3 NN architecture

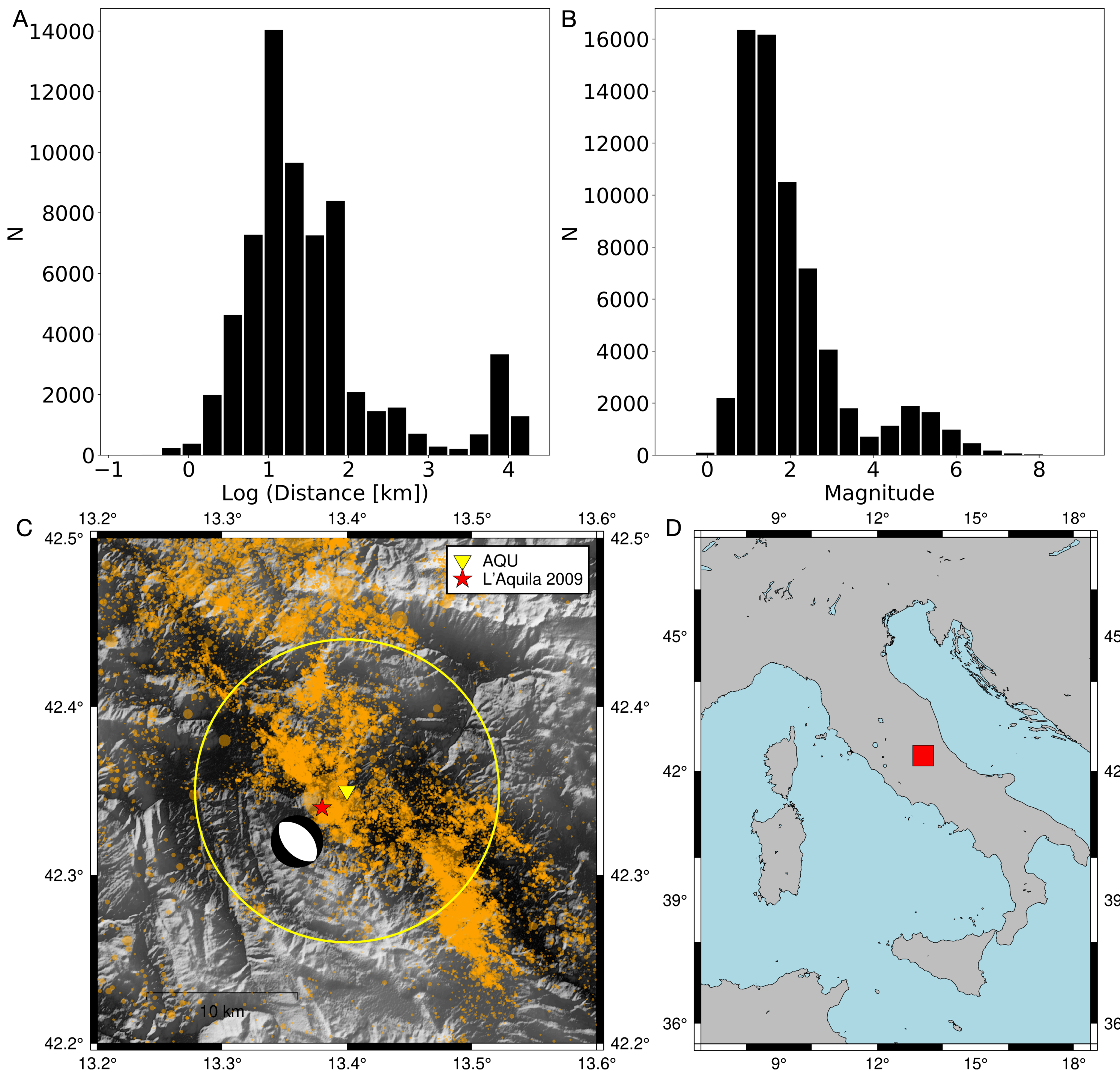
1 Training data

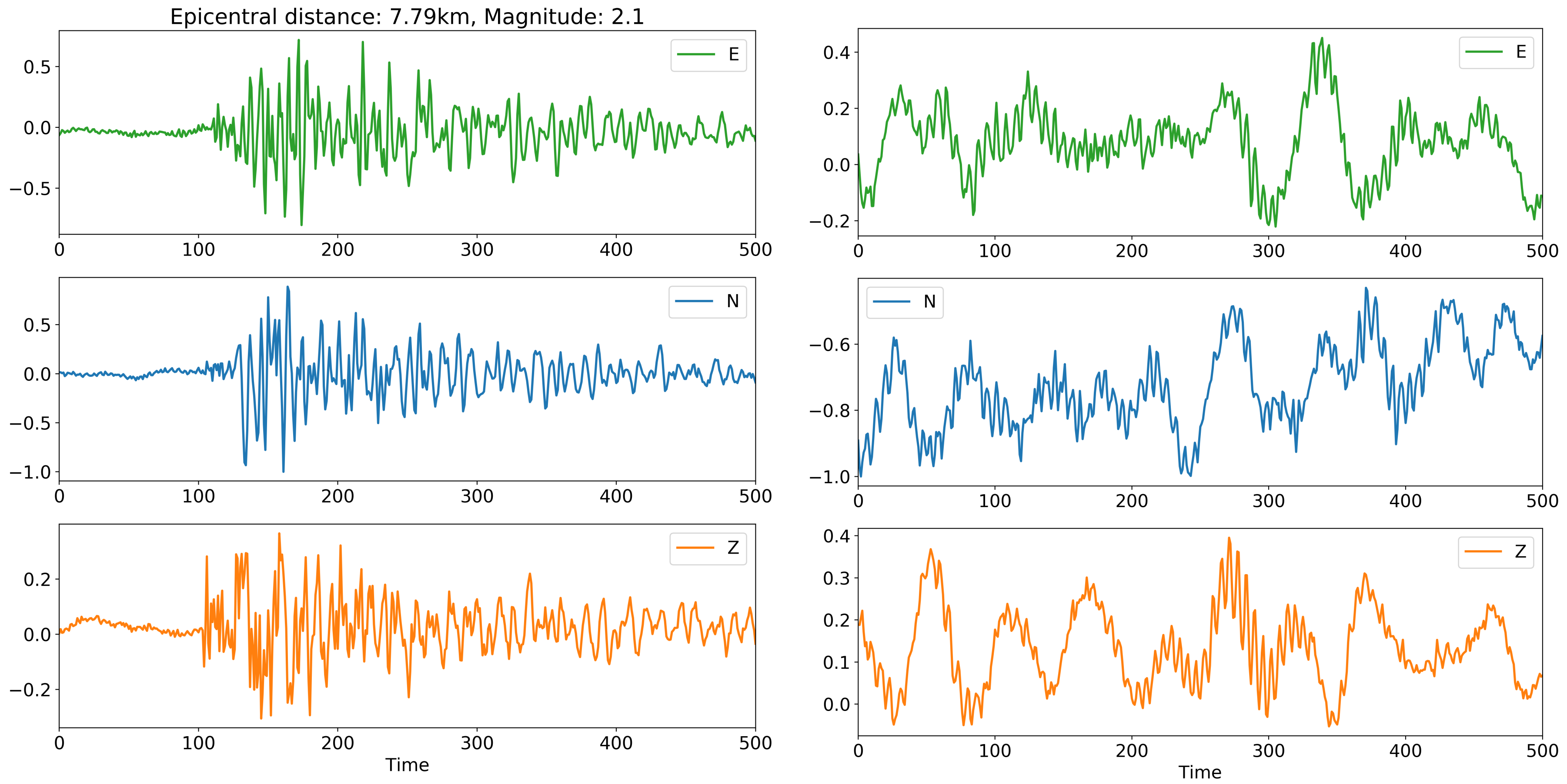
2 Training process

2 Modelling approach

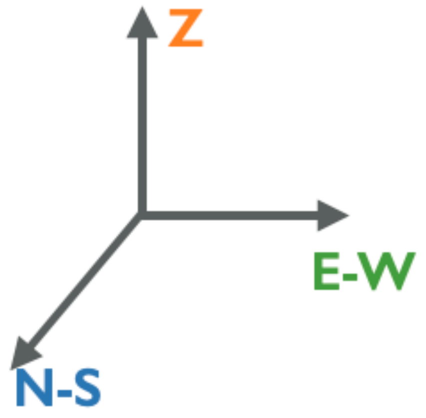
3 NN architecture

1 Creating a database using one station approach.





65k positive samples
+ 65k negative samples



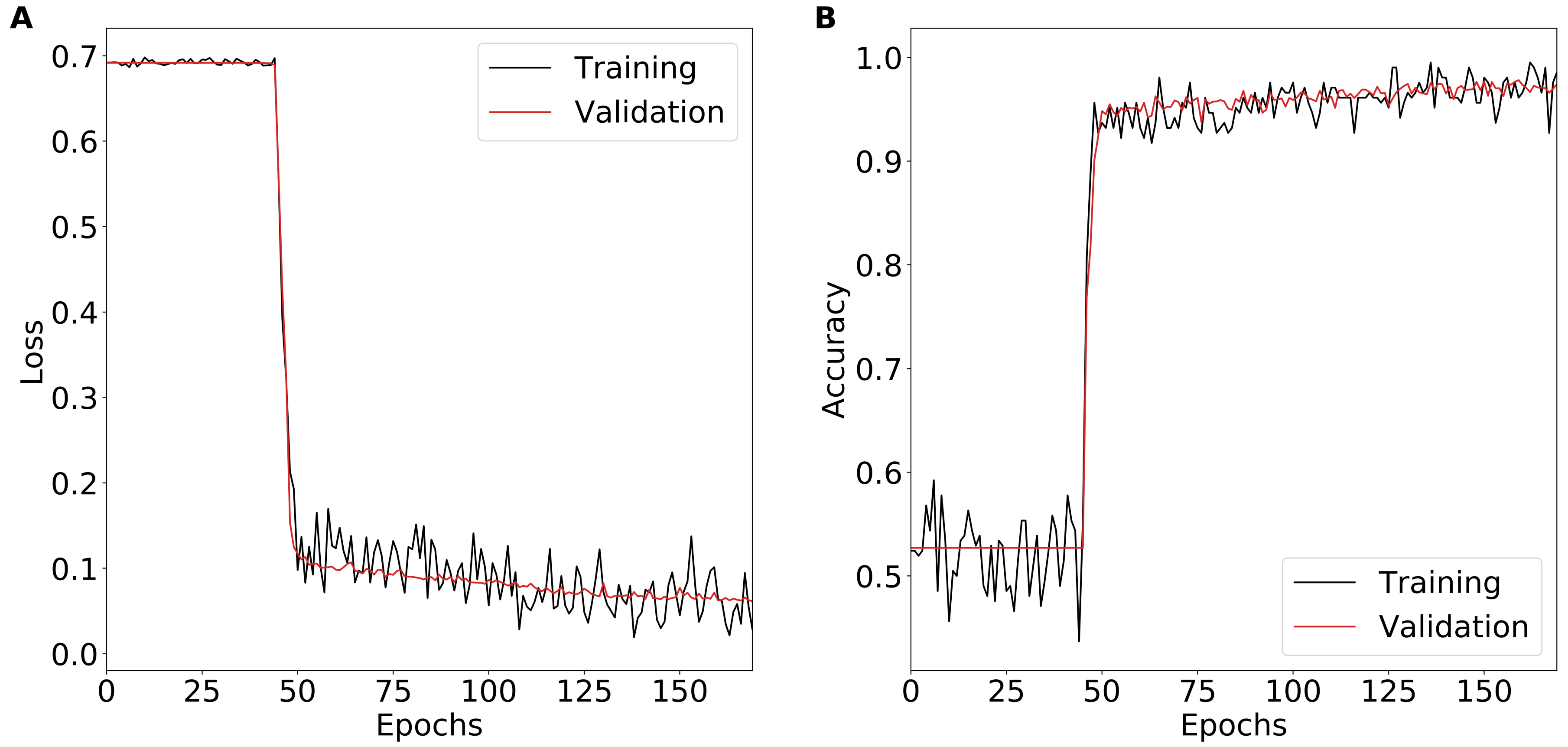
2 Classification task

Positive class (earthquake)

Negative class (noise)

Convolutional neural network (CNN)

After a grid search analysis, the best combination of the training hyperparameters are selected.



The training hyperparameters

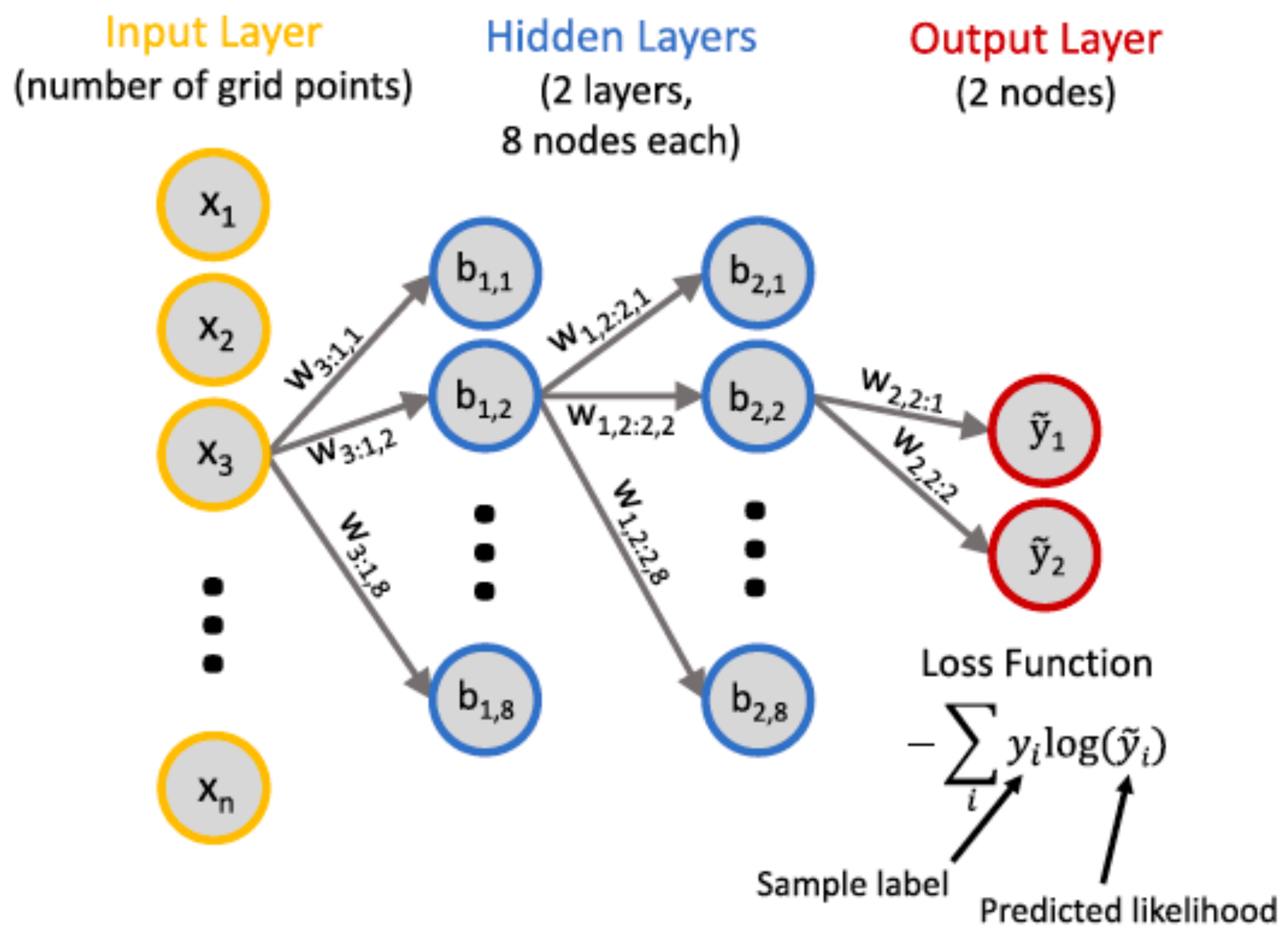
Optimising algorithm	SGD
Learning rate	1e-2
Momentum	0.9
Mini-batch	128

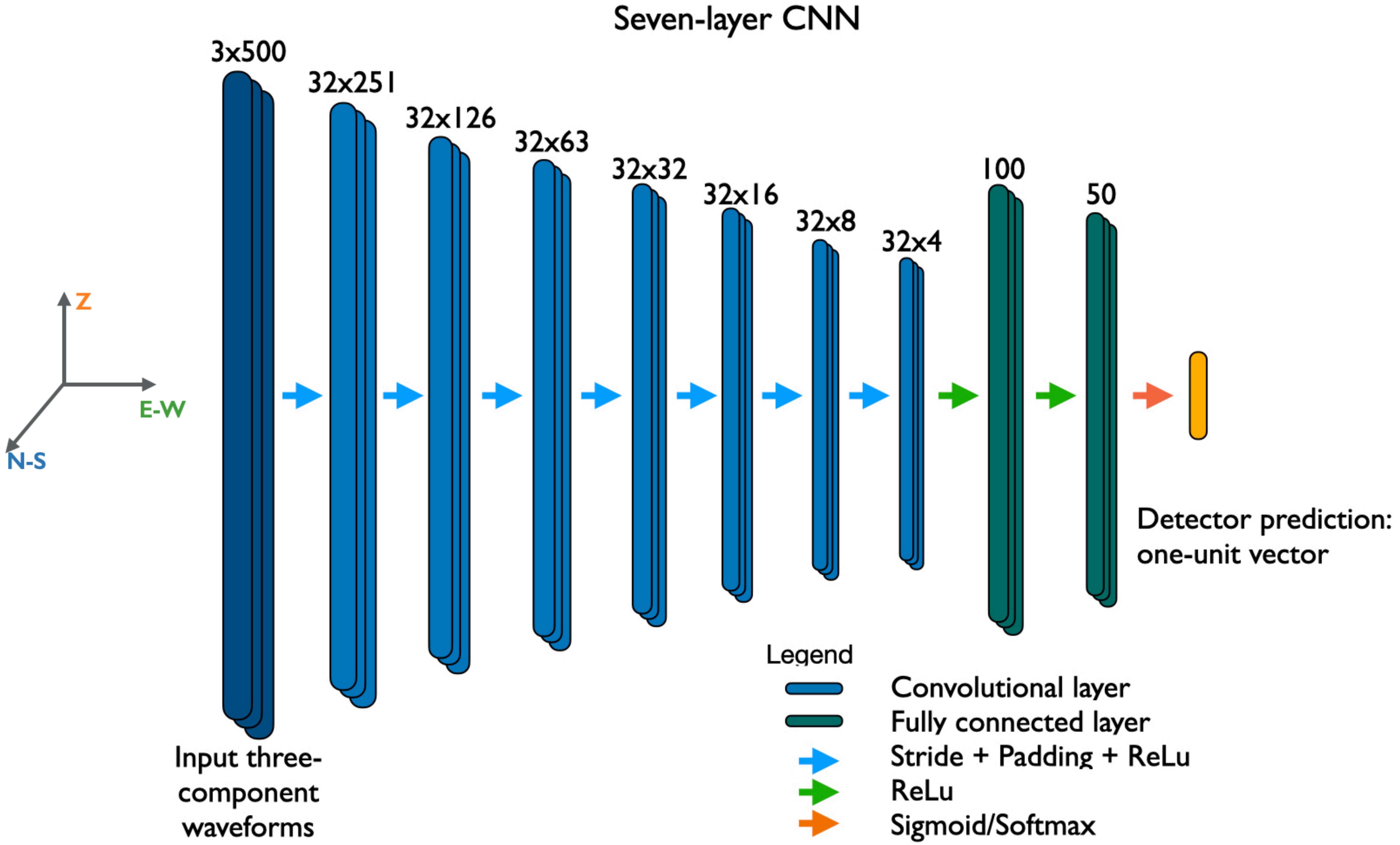
Majstorović et. al., 2021, JGR

- 1 Training data
- 2 Training process
- 2 Modelling approach
- 3 NN architecture

3 Convolutional neural network (deep learning) - supervised training process

- Convolutional layers
- Pooling layers
- Fully-connected layers





36k of trainable parameters

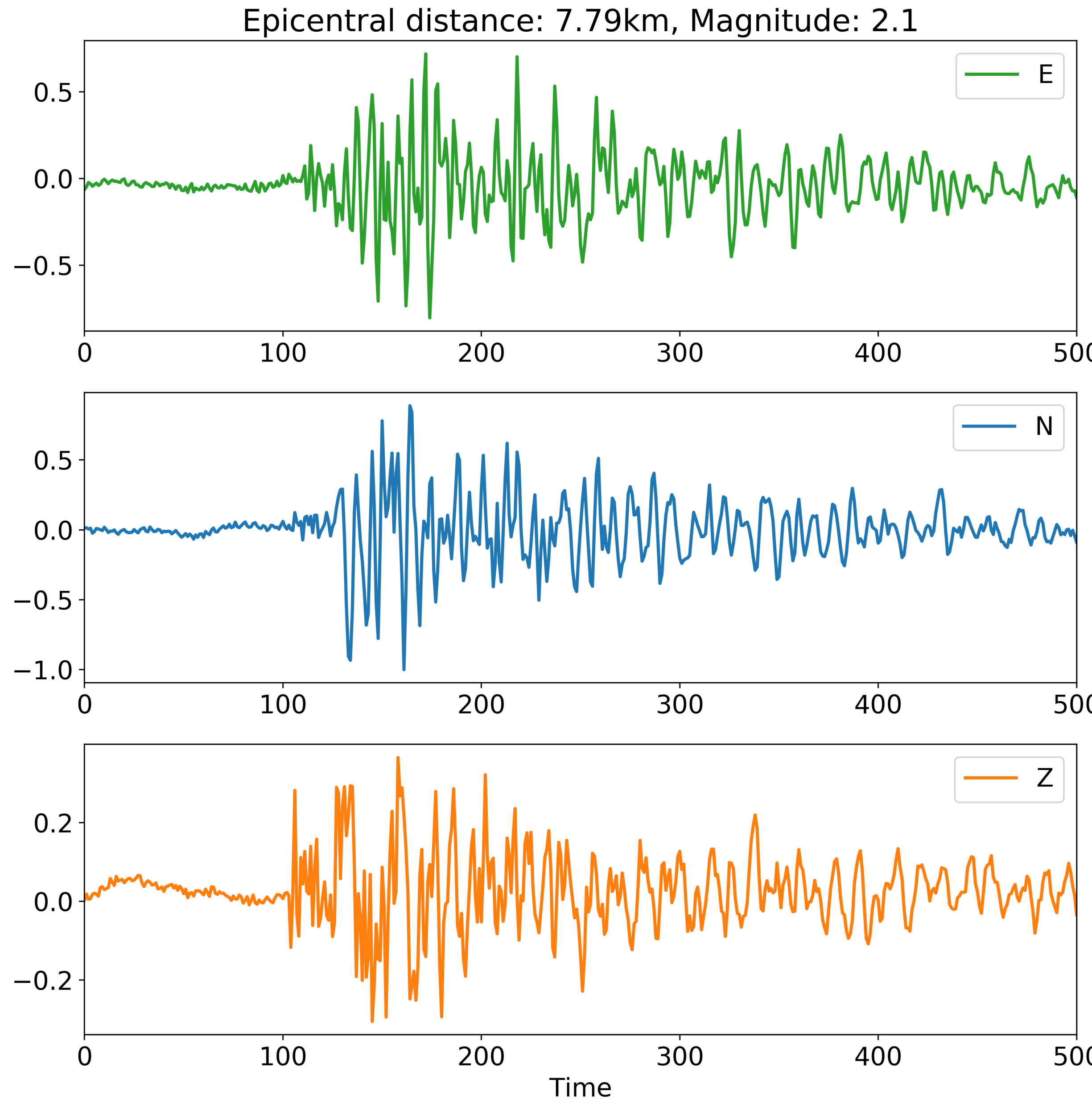
Majstorović et. al., 2021, JGR

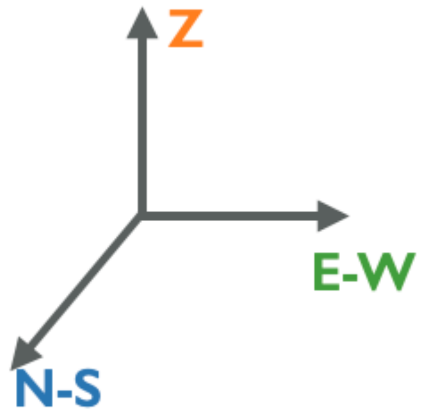
1 Training data
2 Training process
2 Modelling approach
3 NN architecture

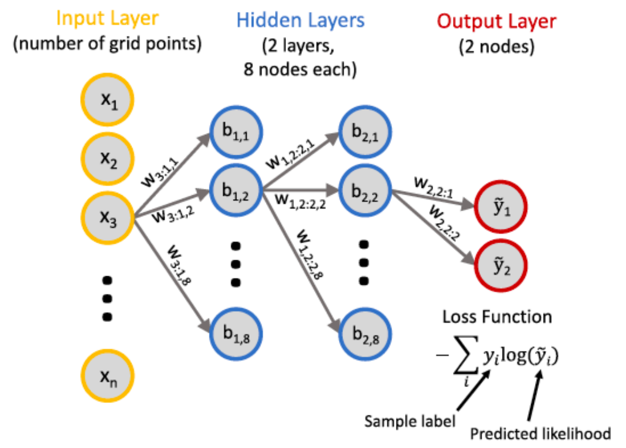
3 Convolutional neural network (deep learning) - supervised training process

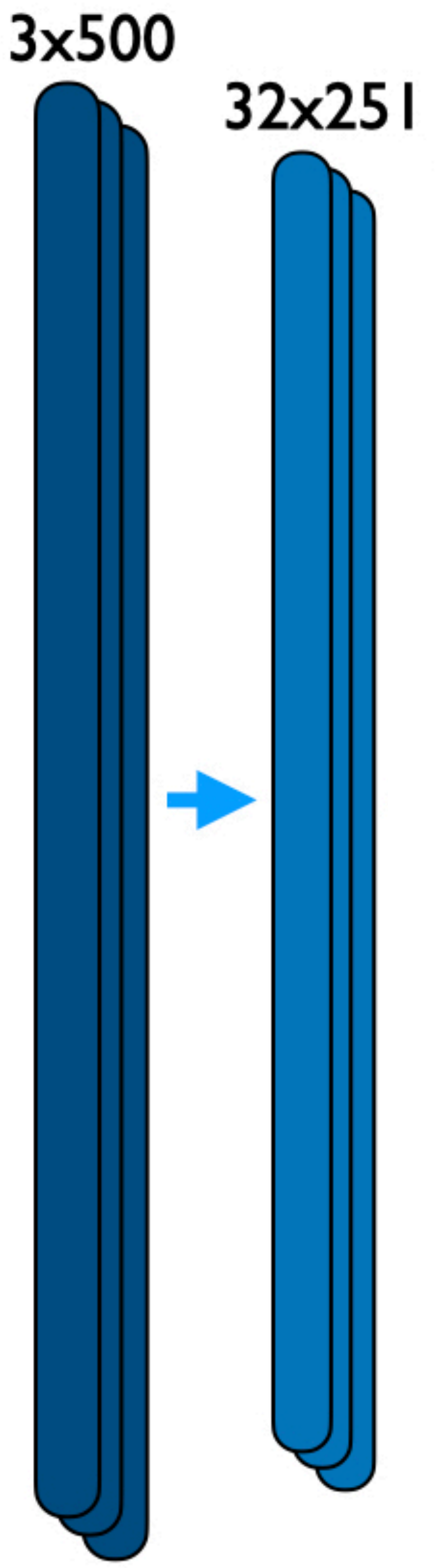
Convolutional layers

Legend
- Convolutional layer
- Fully connected layer
- Stride + Padding + ReLu
- ReLu
- Sigmoid/Softmax

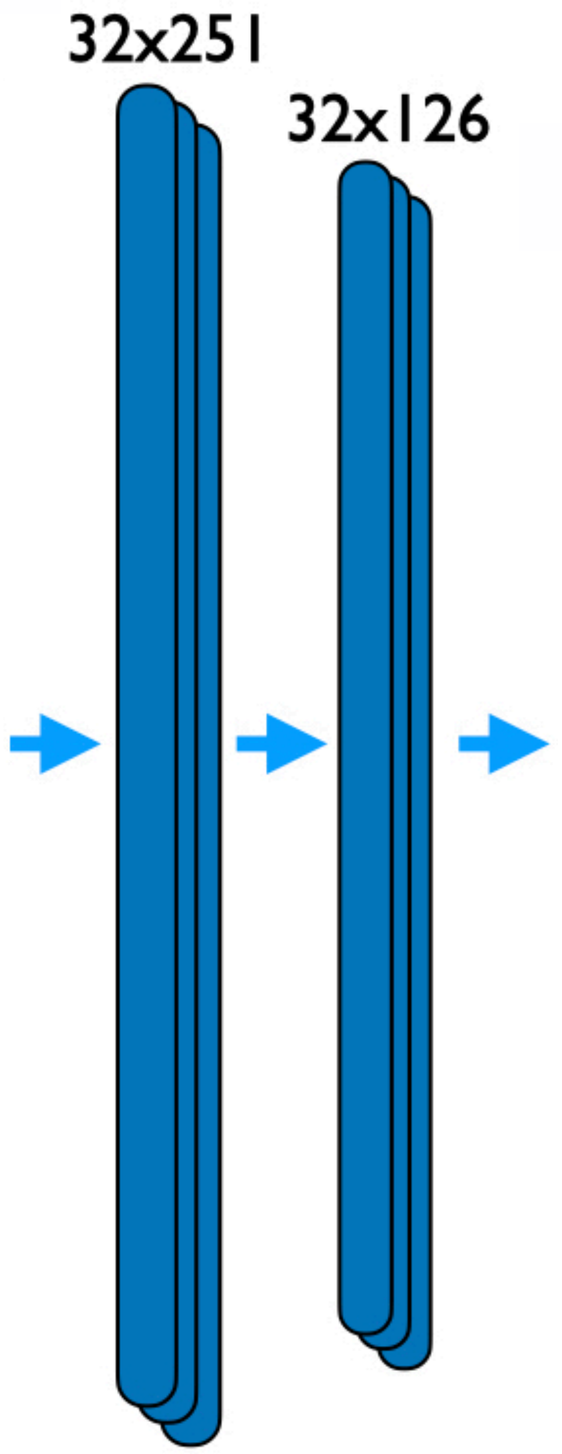








32 channels
Feature map
(251 features)



32 channels
Feature map
(126 features)

Input array
3 channels x 500 samples

Filter (kernel) size
3 x 3 x 32

Filter (kernel) size
3 x 32 x 32

$f \times c_{in} \times c_{out}$

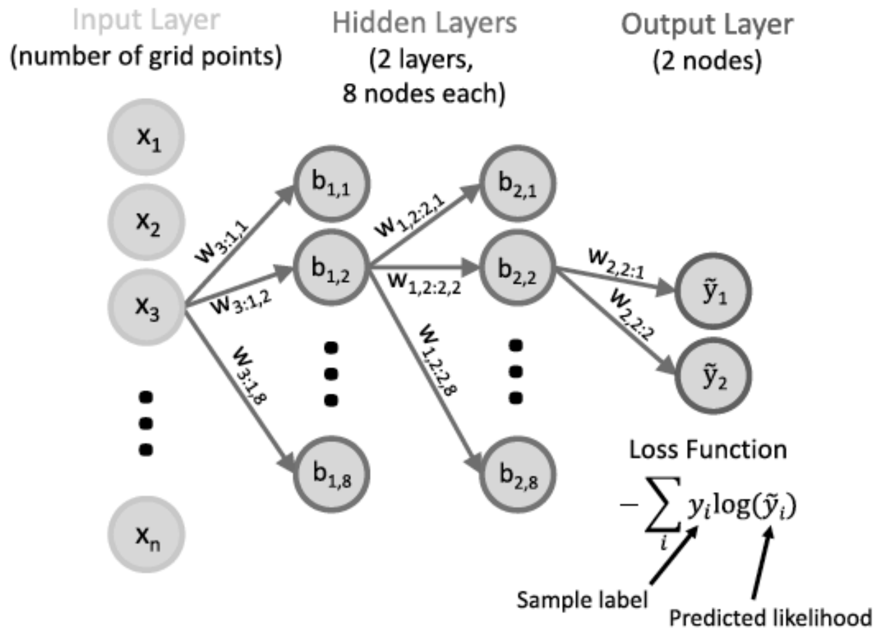
Majstorović et. al., 2021, JGR

Basics of neural networks, deep learning models

Developing CNN detector

What does interpretation stand for in DL?

Majstorović et. al., 2021, JGR

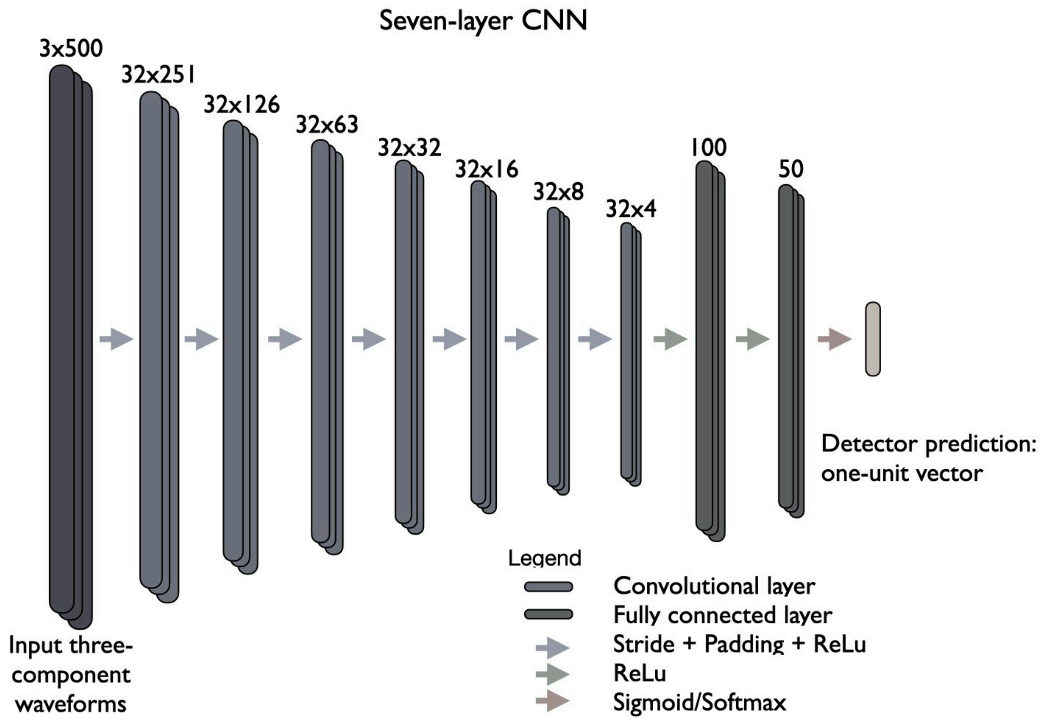


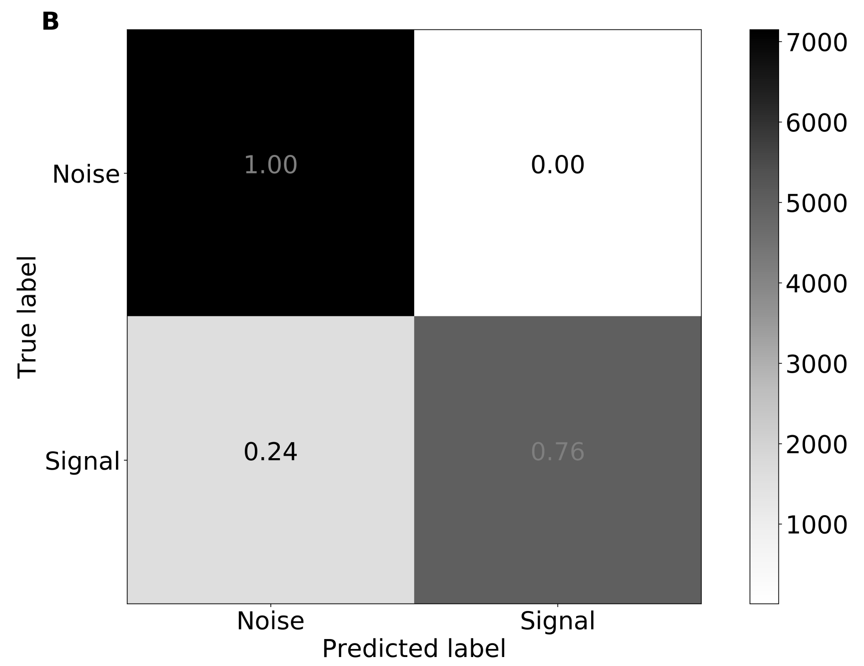
1 Training data

2 Training process

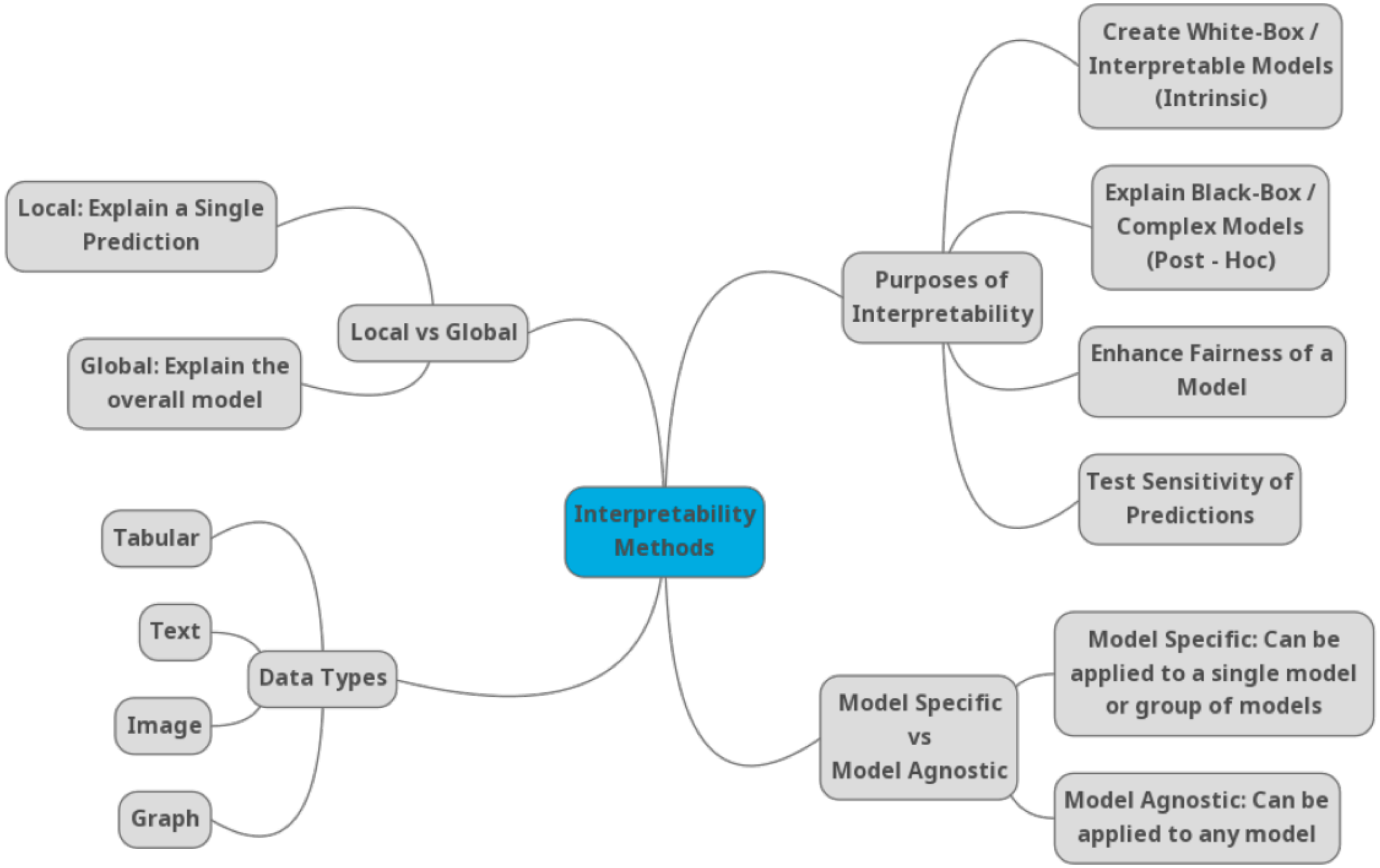
2 Modelling approach

3 NN architecture

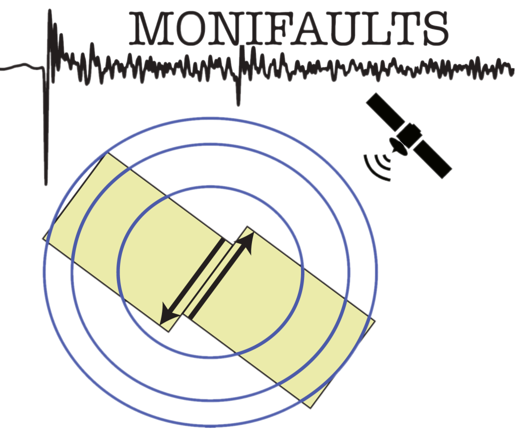


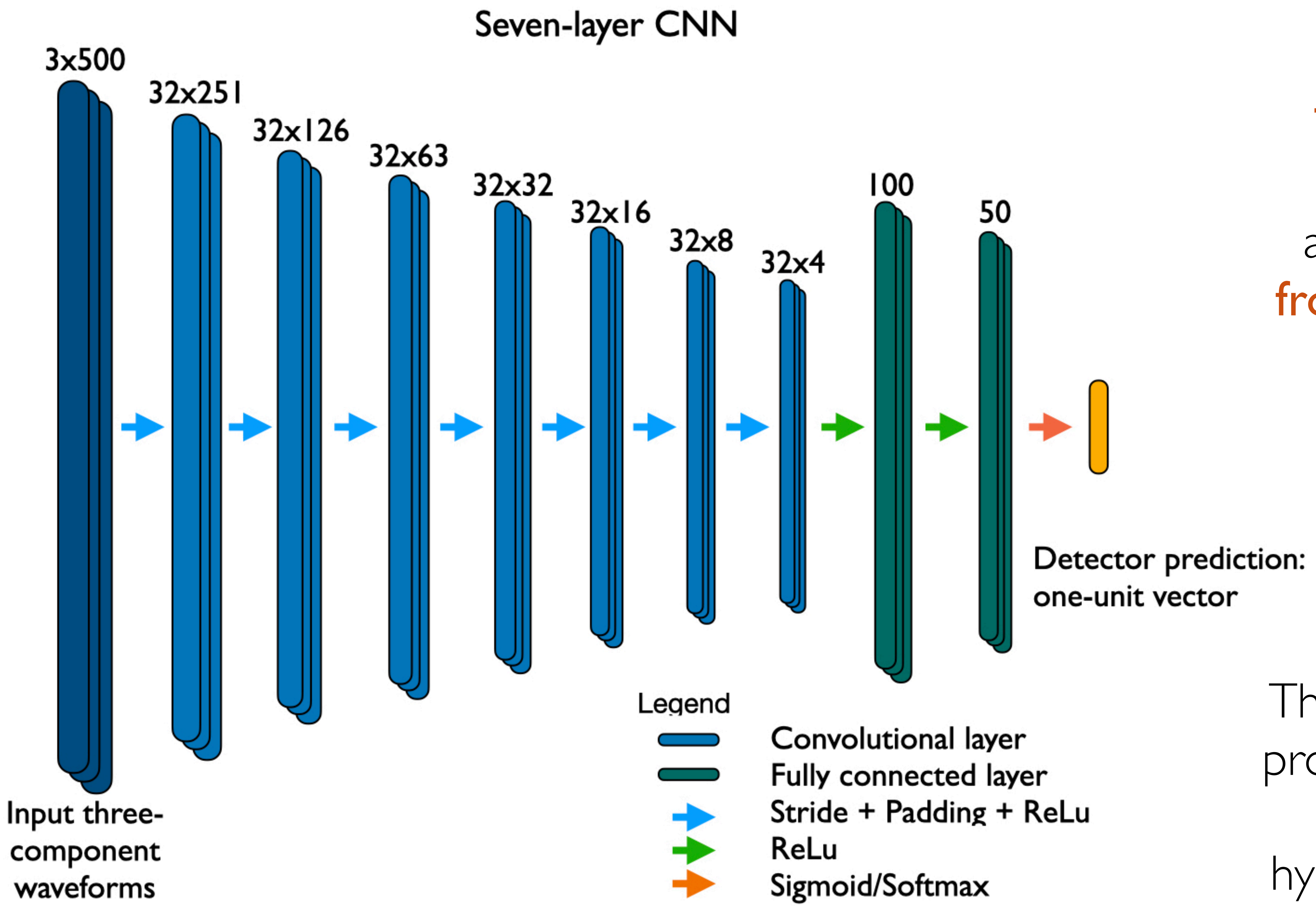


ML (NN, DL) models are often criticised by end users as being a **"black box"** because of the perceived inability to understand how ML makes its prediction.



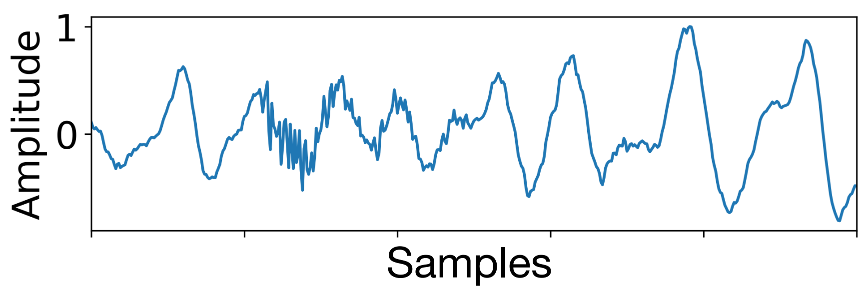
Linardatos, 2020, Entropy

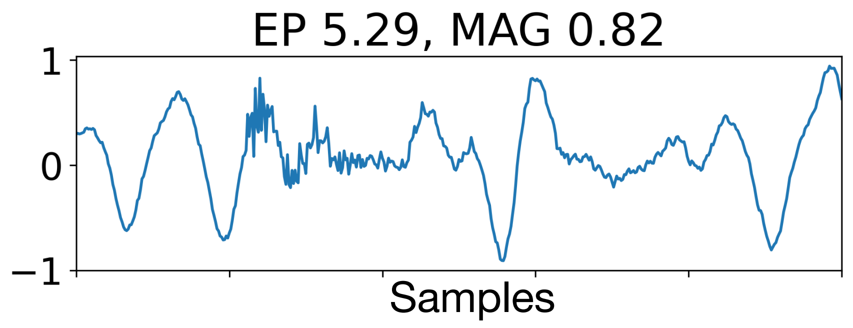


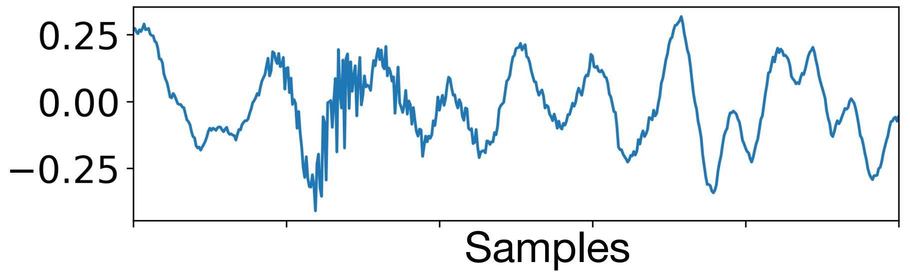


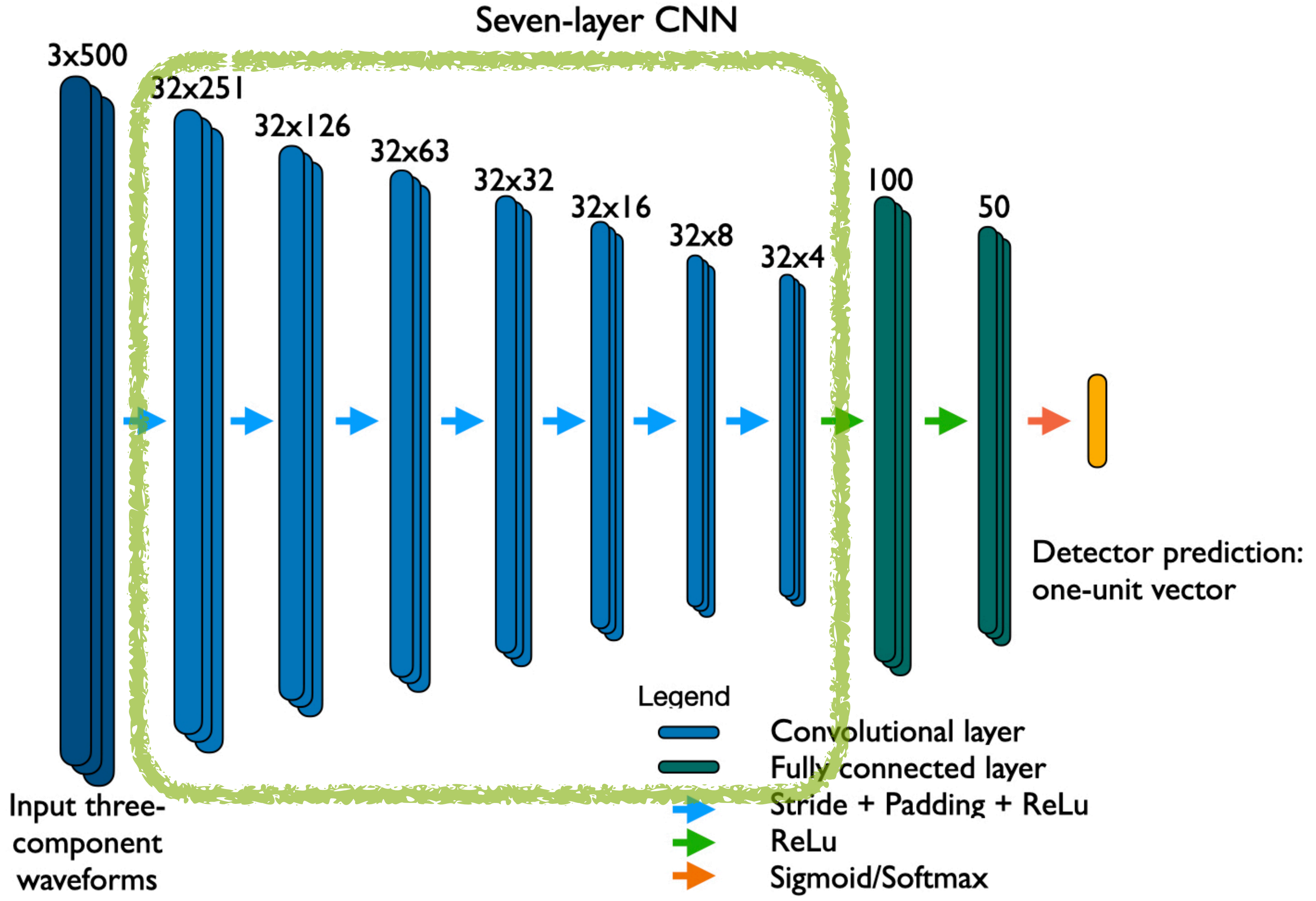
The weights and biases that define the kernels (the filters) and that are being adjusted in the training process **are now frozen** to investigate how those impact the input data.

The training of the DL model is a stochastic process - repeating the training process with the same dataset and the same hyperparamters yields different weights and biases.



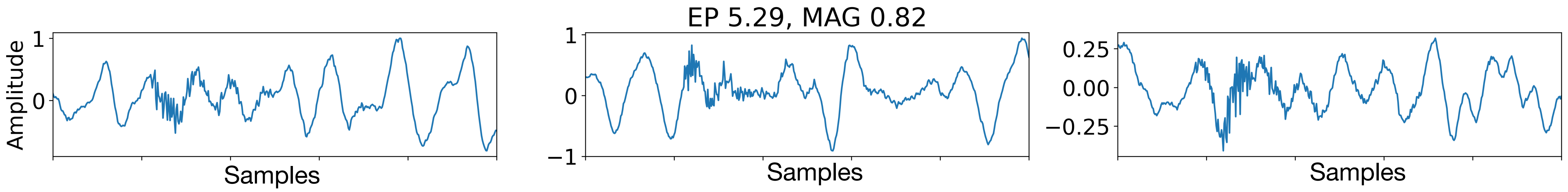


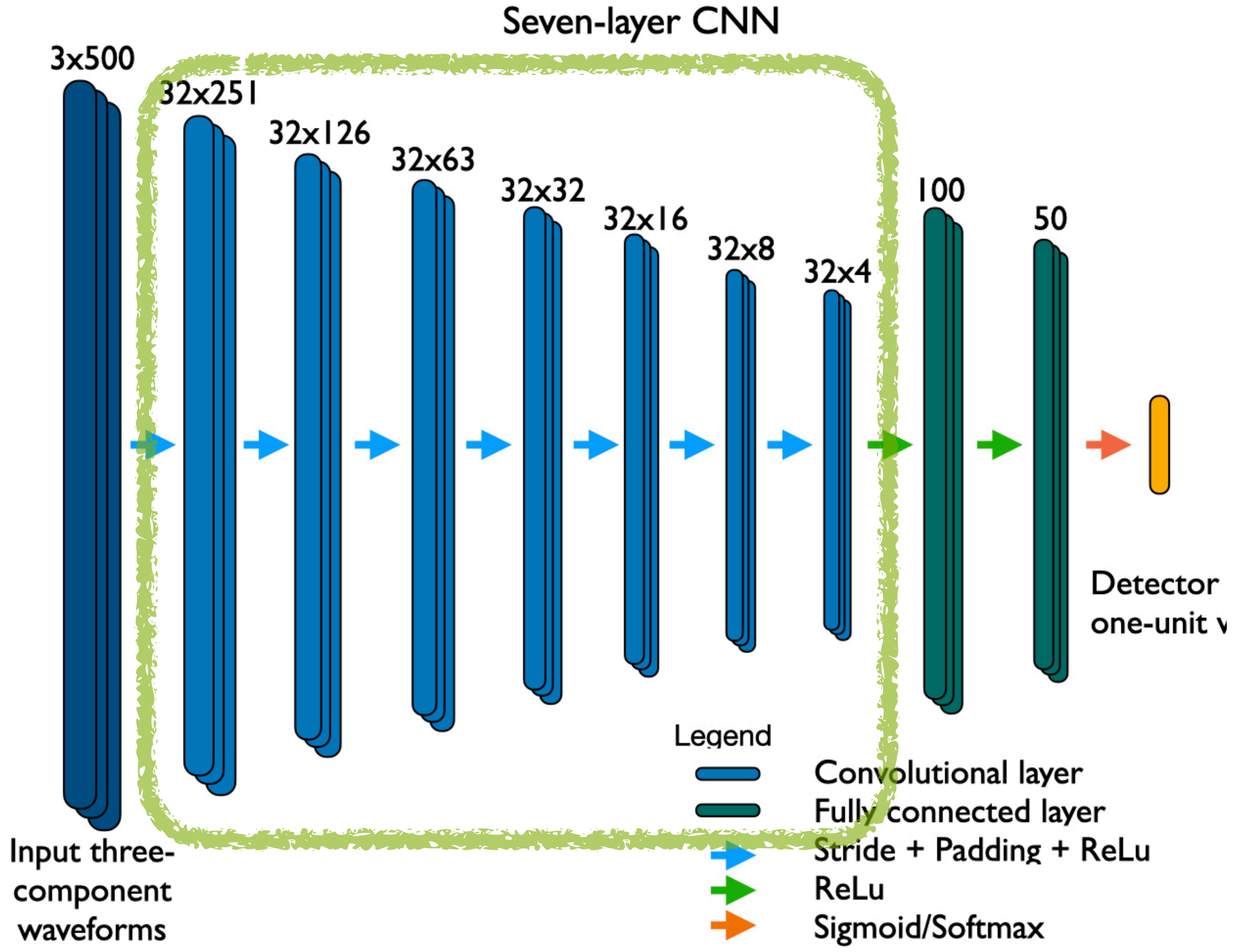


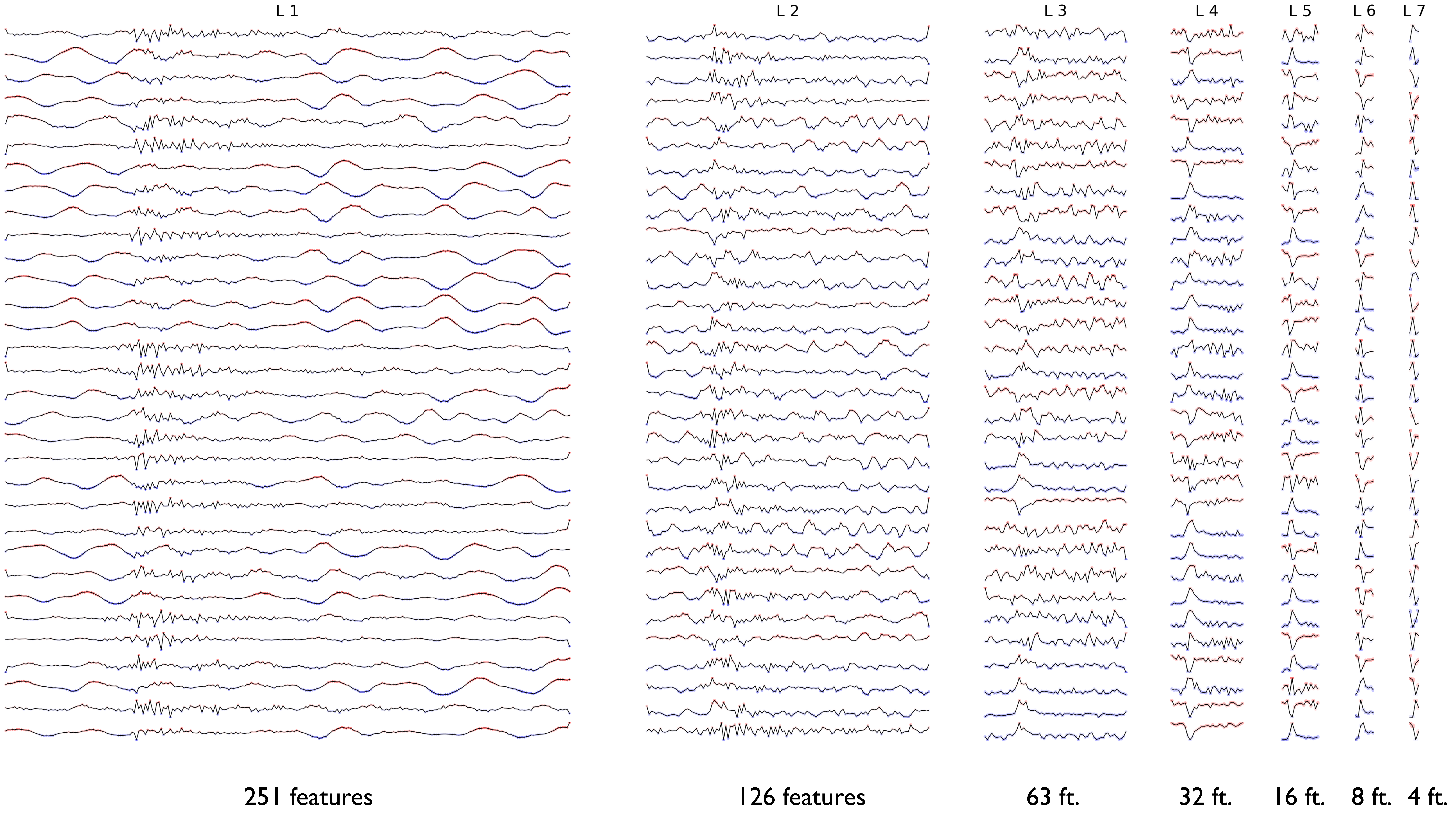


Kernel visualisation

Feature map visualisation



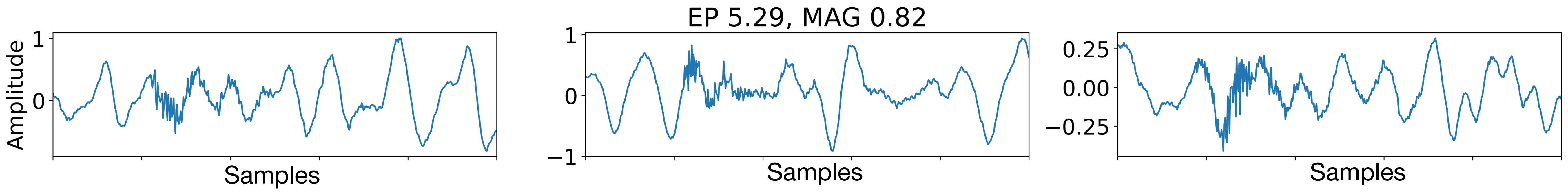


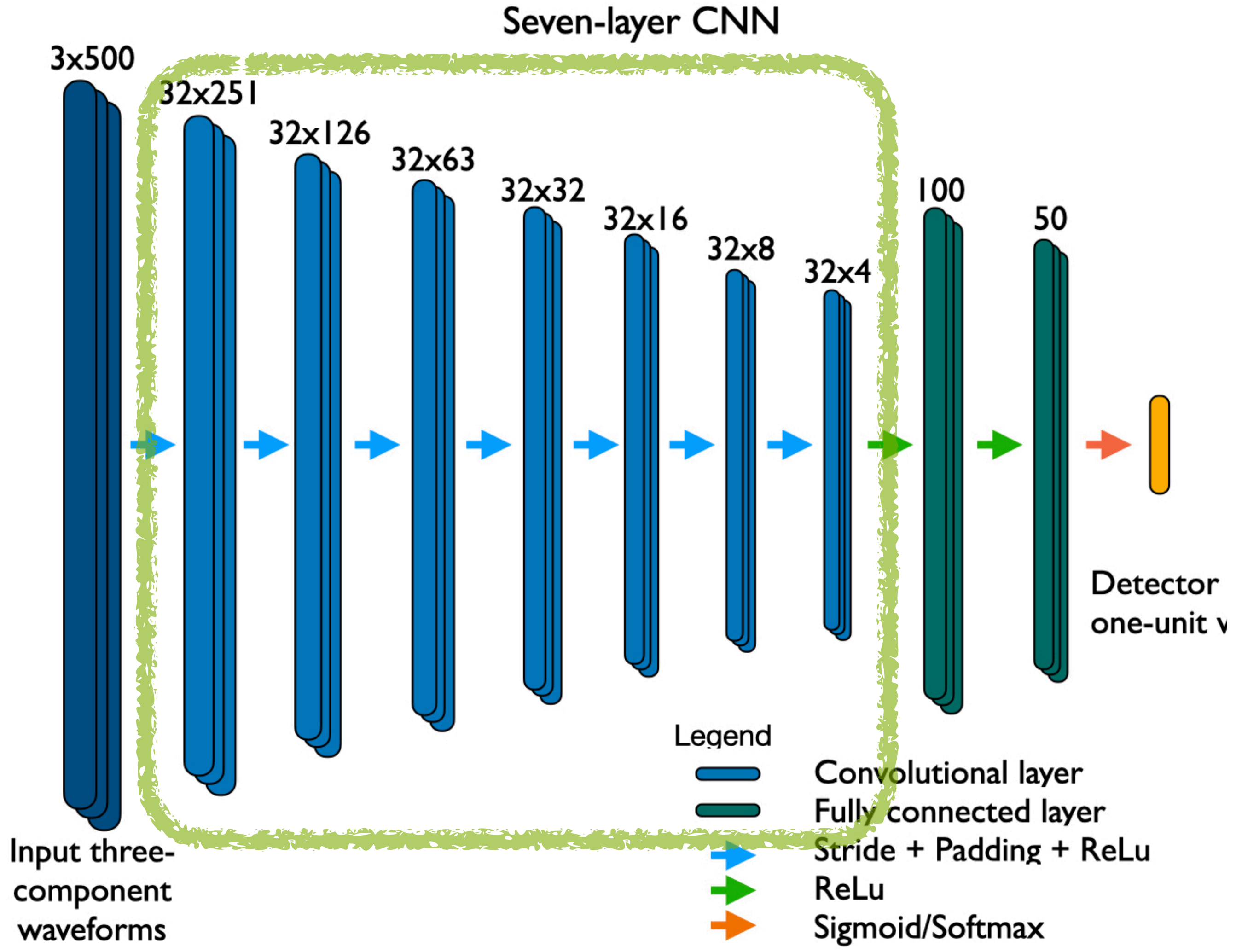


Kernel visualisation

Feature map visualisation

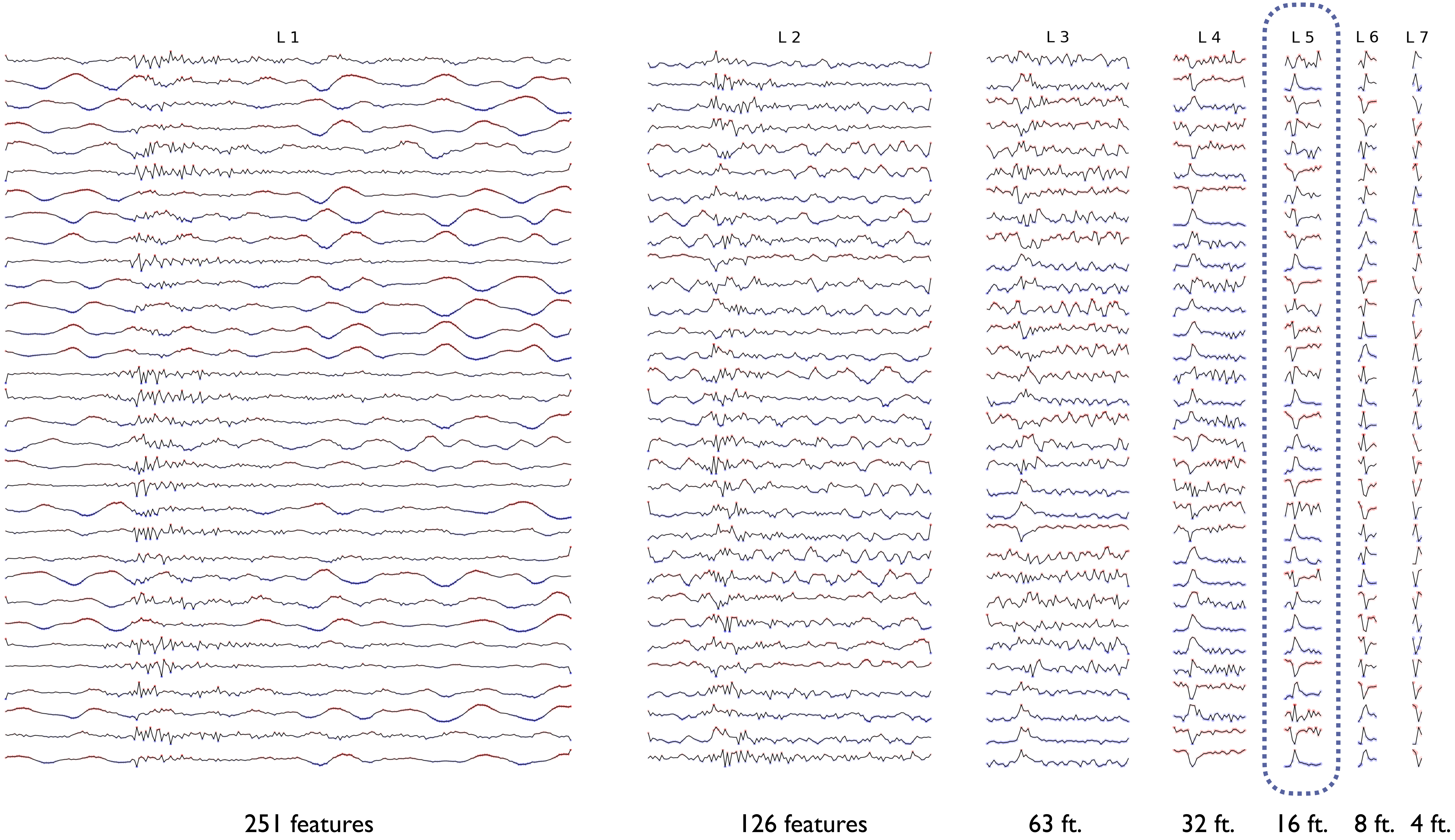
Outline | Introduction | CNN detector | Interpretation | Feature visualisation | Backward optimisation | Layer-wise relevance method





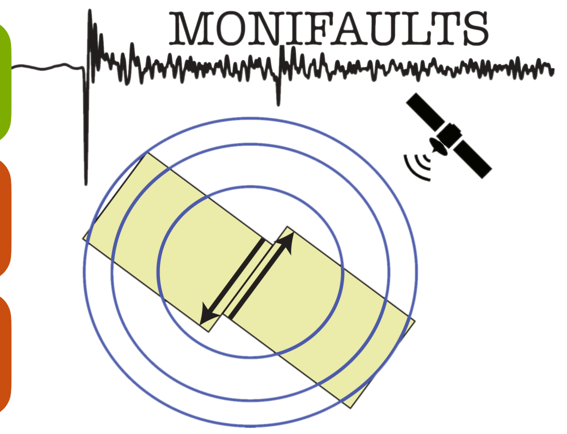
Kernel visualisation

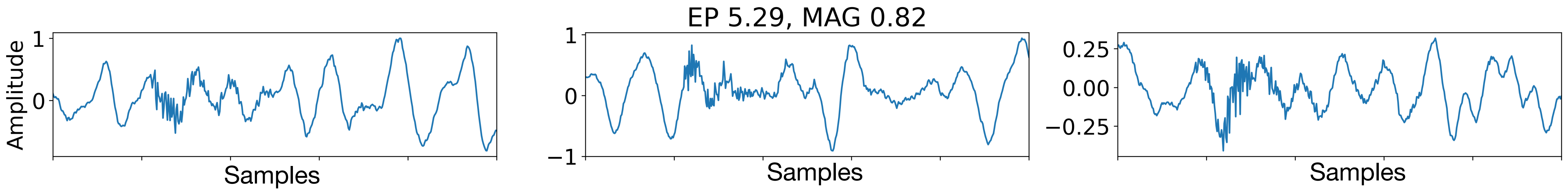
Feature map visualisation

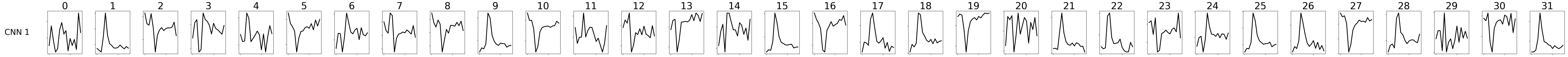


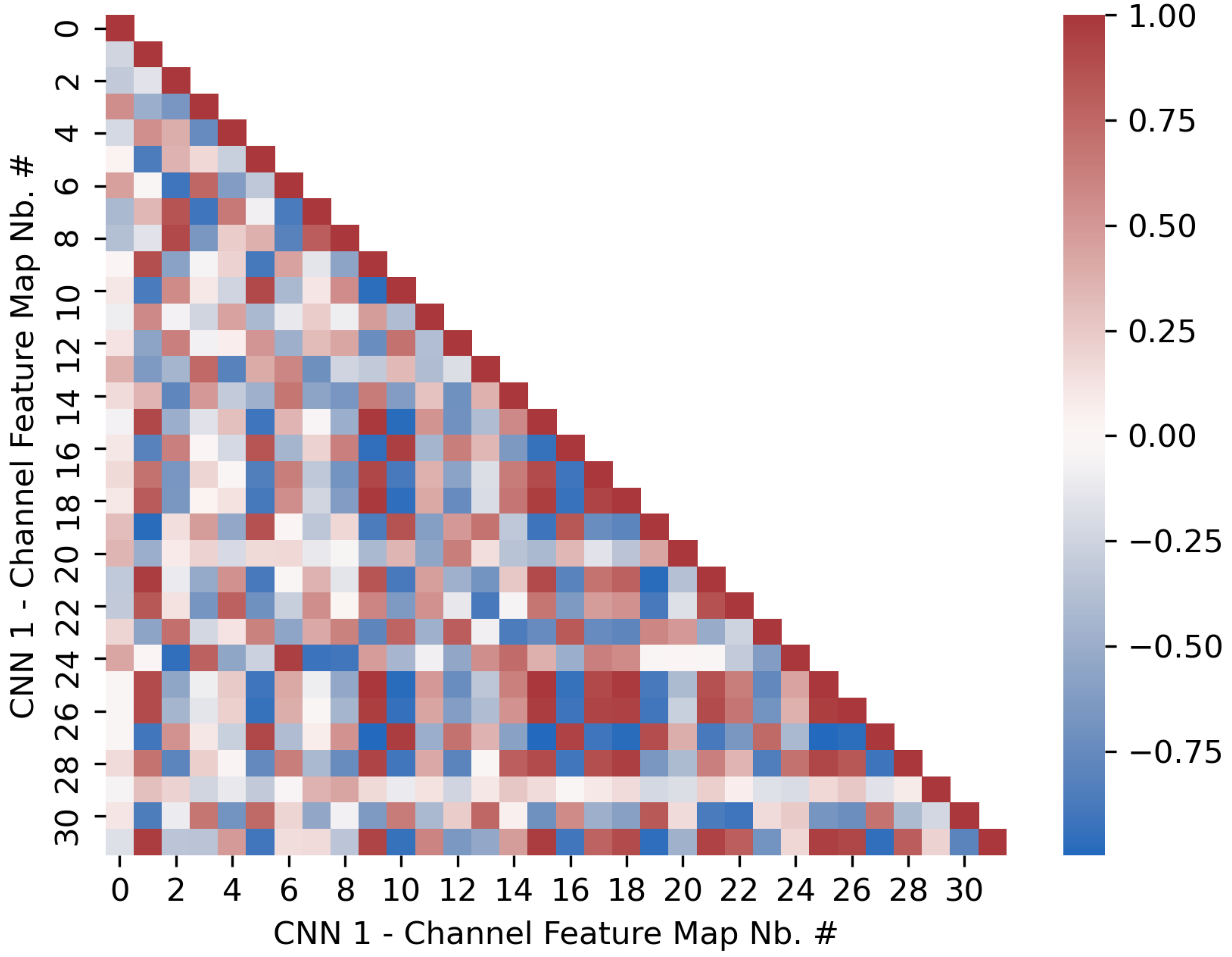
Z component

Y axis is the same per column/channel


MONIFAULTS

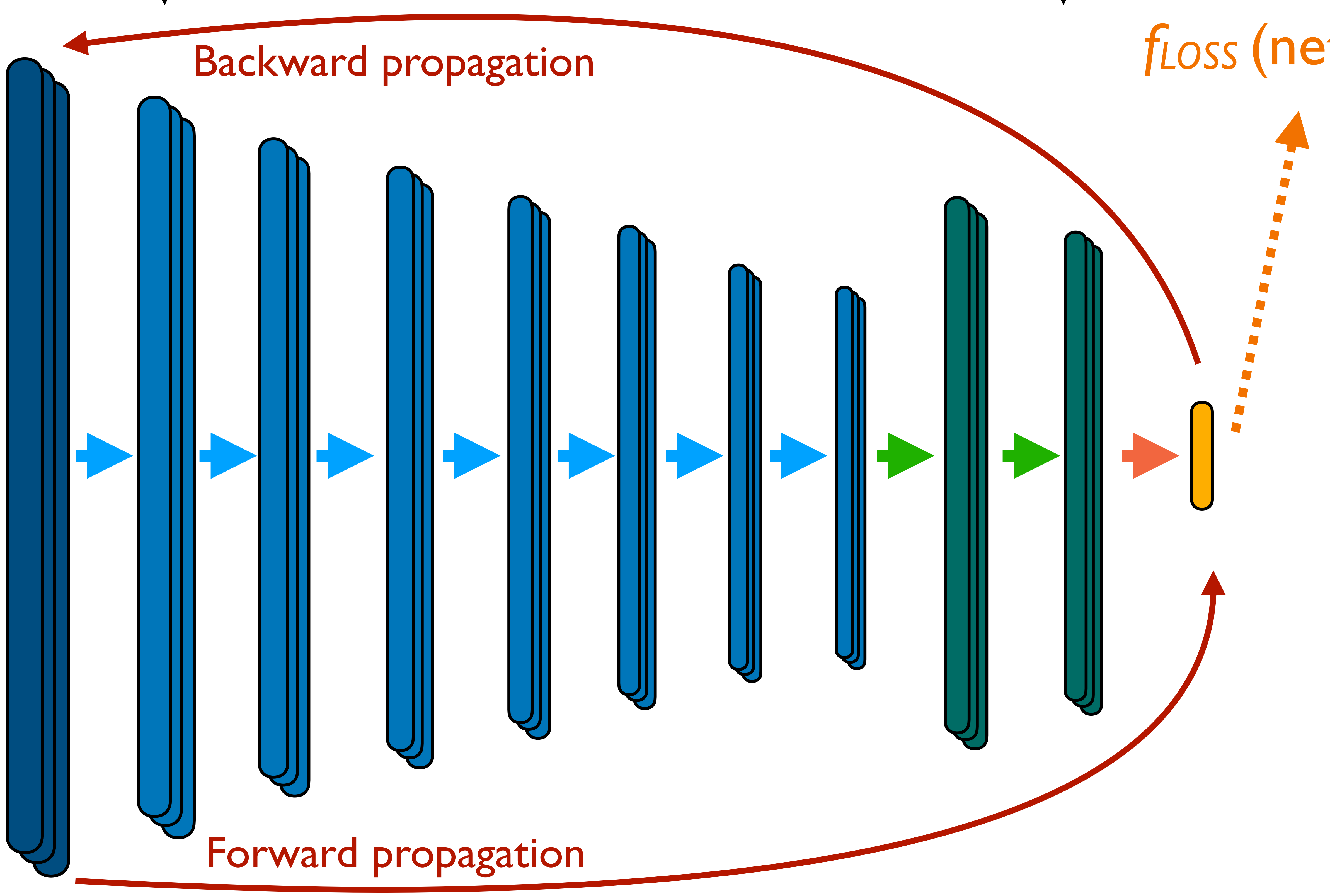

Amplitude
1
0
Samples
EP 5.29, MAG 0.82
1
0
−1
Samples
0.25
0.00
−0.25
Samples


CNN 1
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31


CNN 1 - Channel Feature Map Nb. #
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
CNN 1 - Channel Feature Map Nb. #
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75

1. We freeze weights and biases of hidden layers.

2. We either use
a) zero,
b) random,
c) real input
and forward propagate it to the output layer.

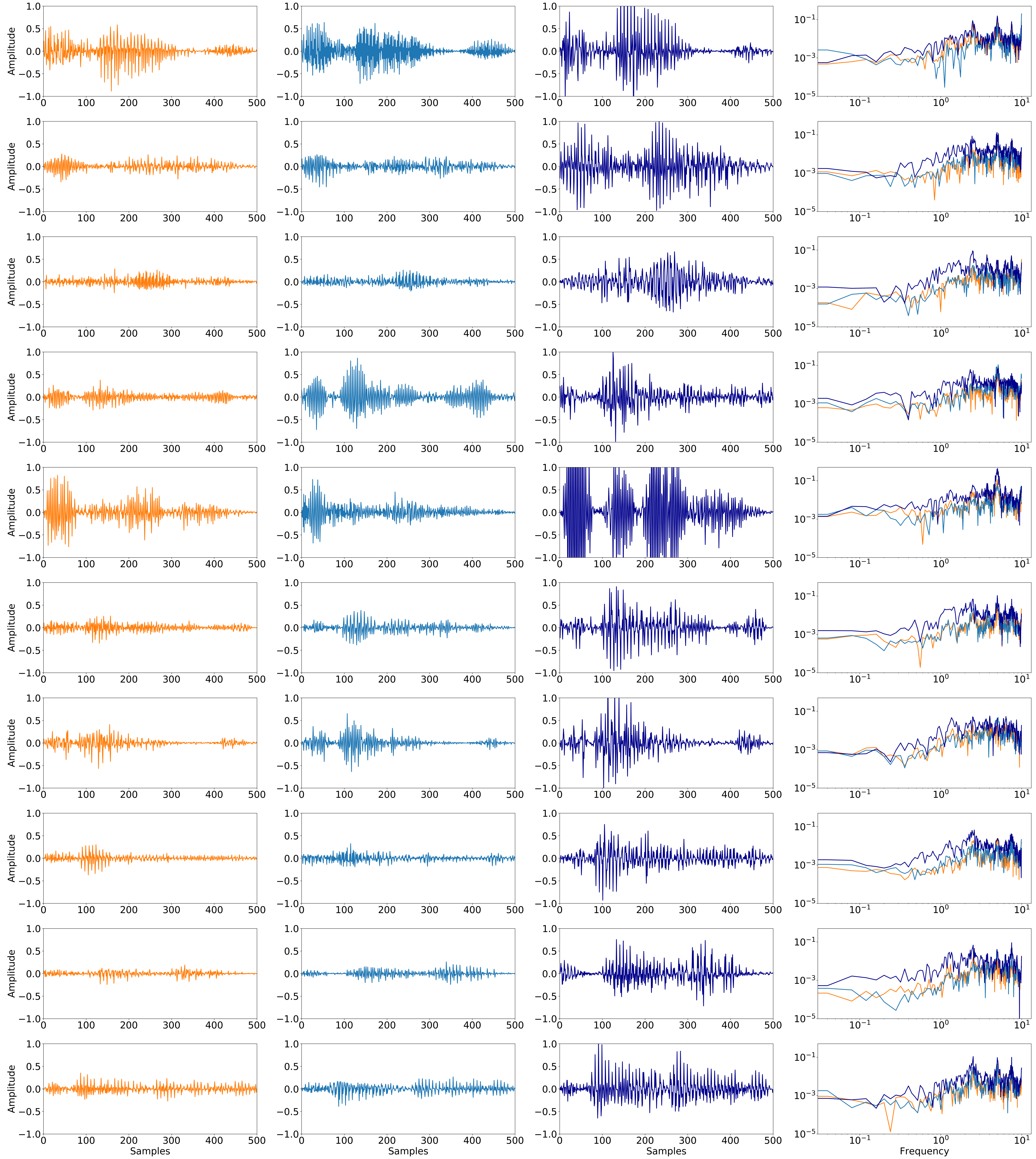


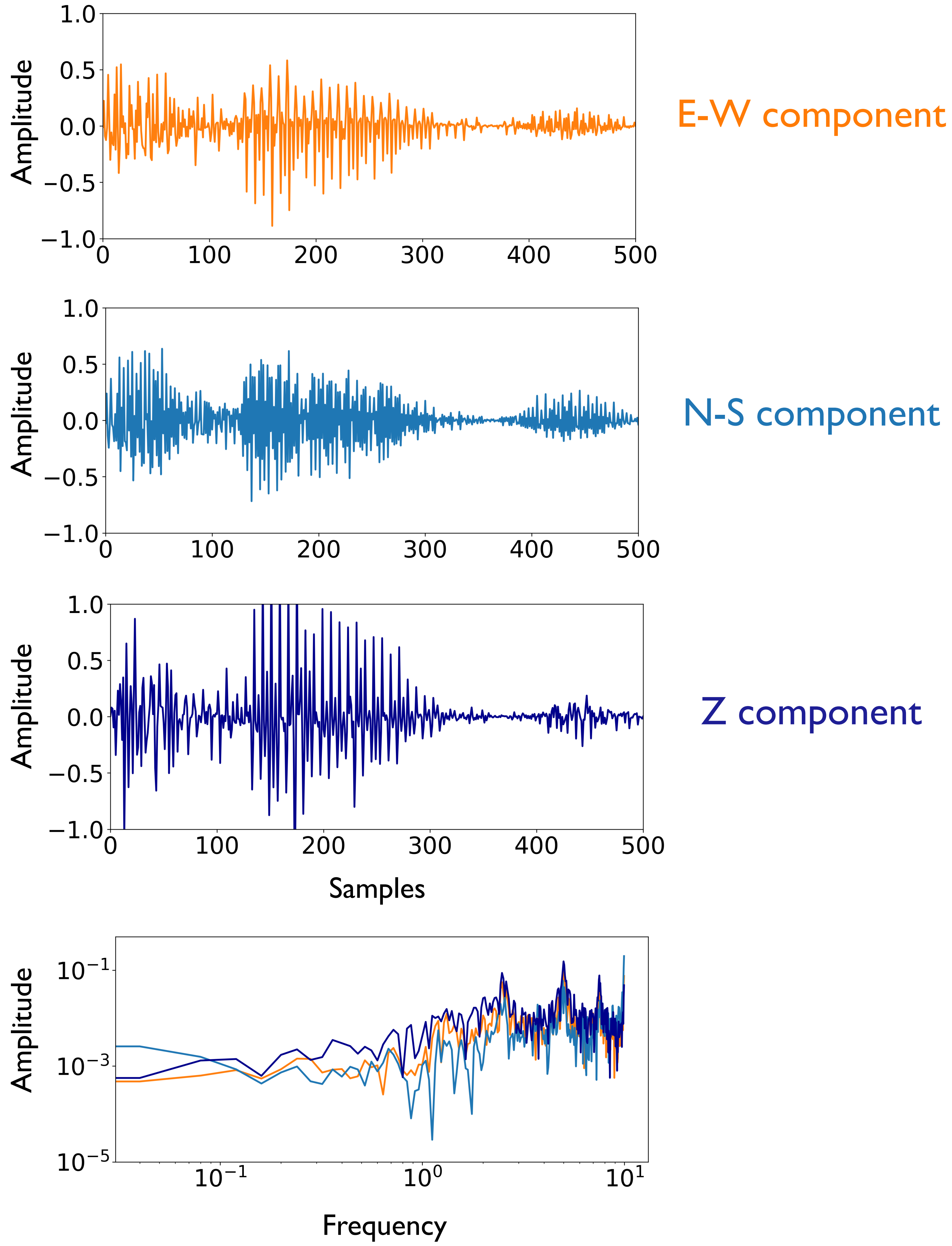
The loss function
f_{LOSS} (network output, desired output)

3. We calculate f_{LOSS} by setting desired output to 1 for investigating the earthquake optimal input or 0 for the noise one.

4. The value f_{LOSS} is backpropagated to update the input.

5. The forward and backward iterations are continued until f_{LOSS} converges.


Amplitude
Samples
Frequency

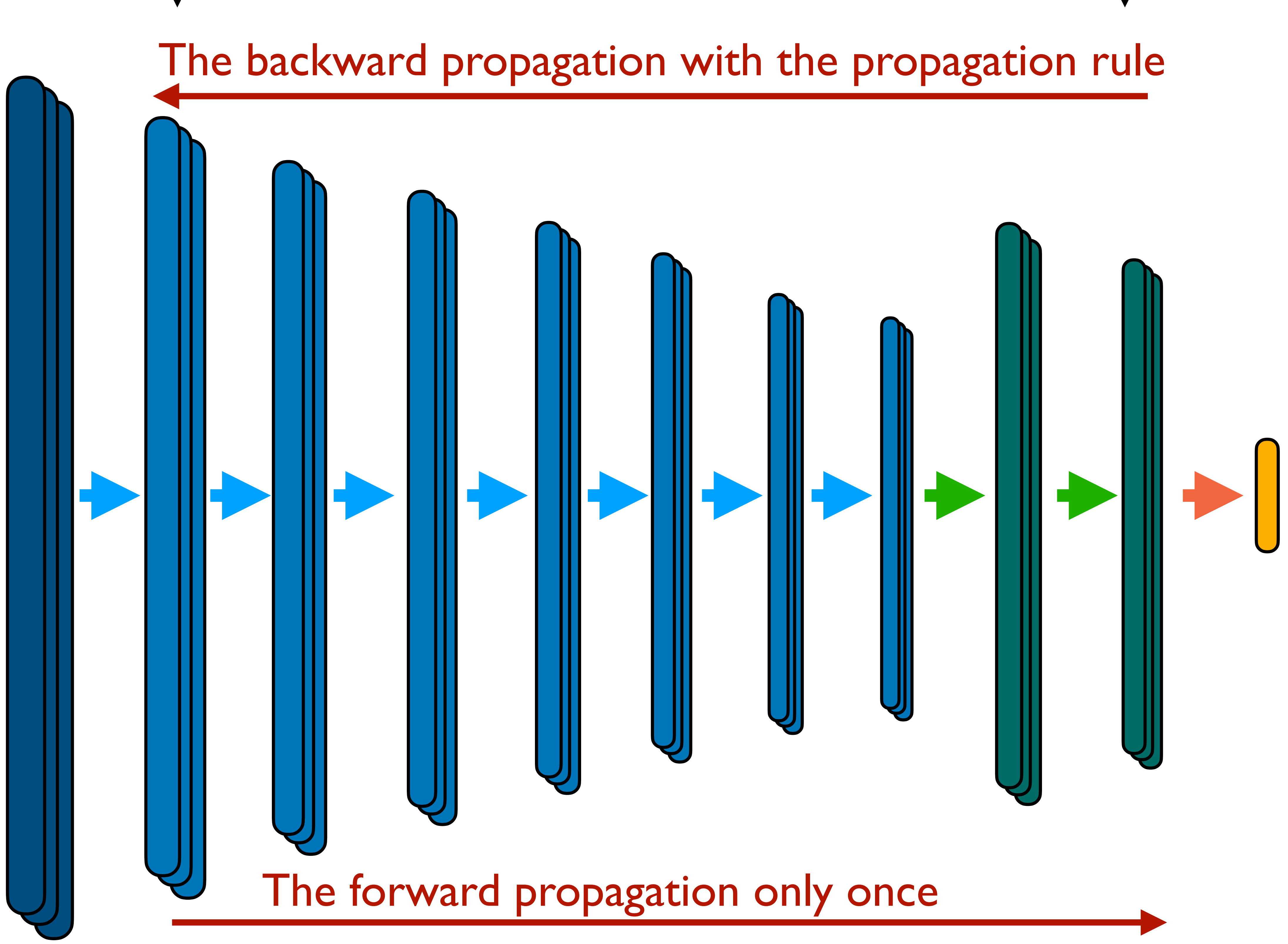

E-W component
N-S component
Z component
Amplitude
Samples
Frequency

2. We input real earthquake sample.

3. Using the frozen weight and biases we obtain the network output by the forward propagation.

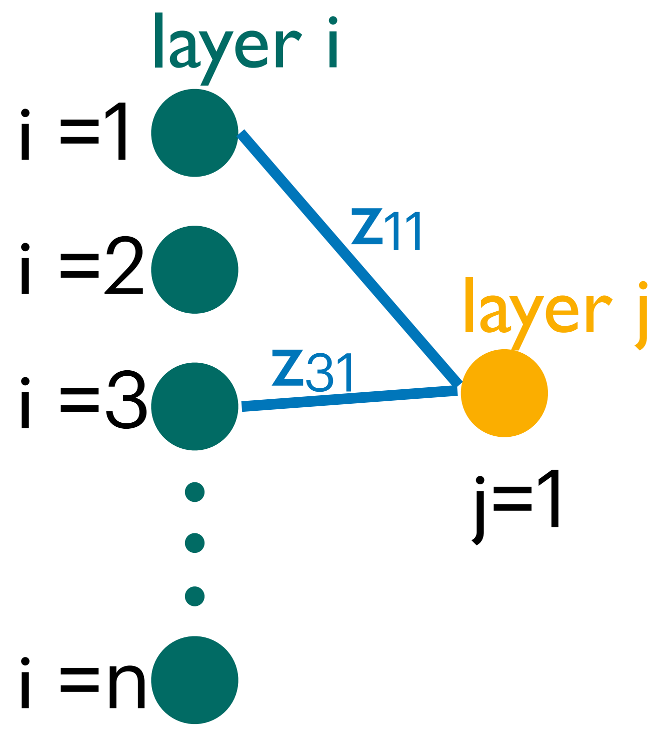
4. We calculate the relevance *R* from the previous layer using the relevance propagation rule called the LRP-β.

1. We freeze weights and biases of hidden layers.



The propagation rule LRP-β

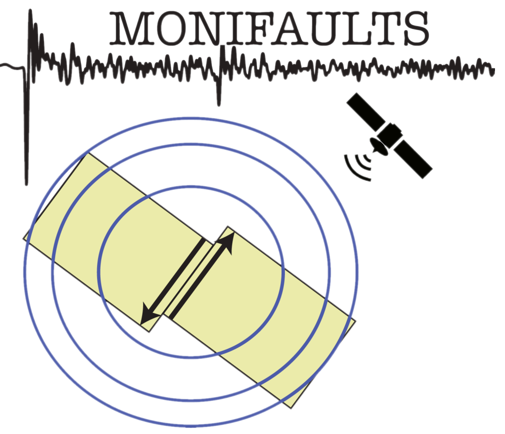
$$R_{l,l+1}^{i\to j} = \left((1+\beta)\frac{z_{ij}^{+}}{z_j^{+}} - \beta\frac{z_{ij}^{-}}{z_j^{-}} \right) R_{l+1}^{j}$$



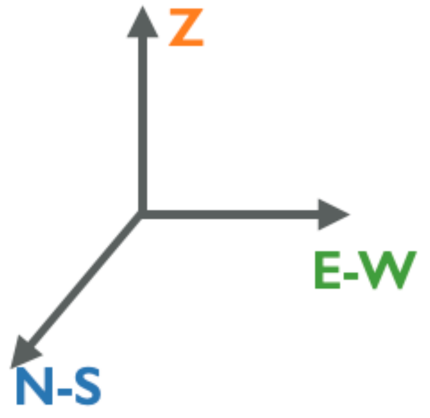
β=0 - we propagate only positive relevance

β=1 - we propagate only negative relevance

5. We propagate the relevance from the hidden layers up to the input layer.



The propagation rule LRP-β-0

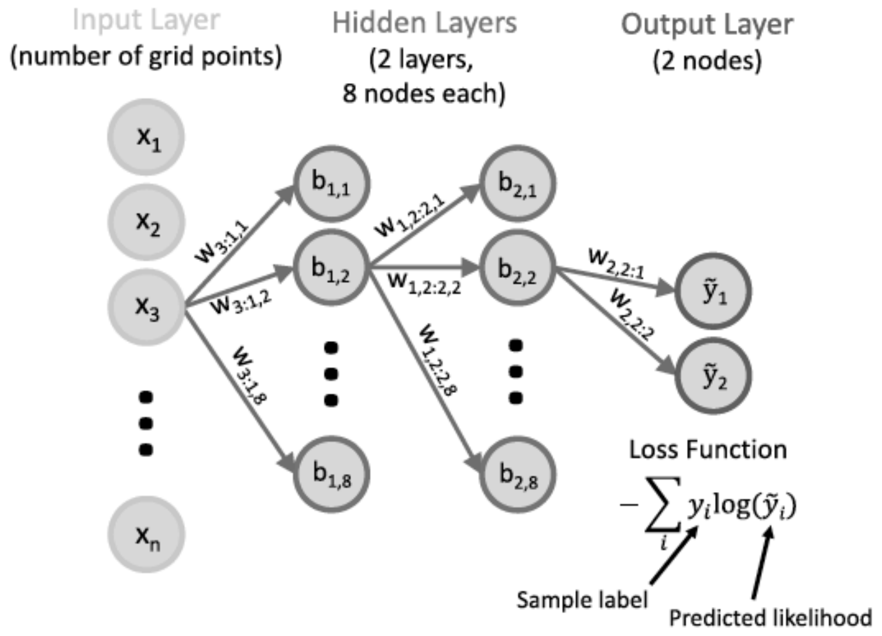


Basics of neural networks, deep learning models

Developing CNN detector

What does interpretation stand for in DL?

Majstorović et. al., 2021, JGR

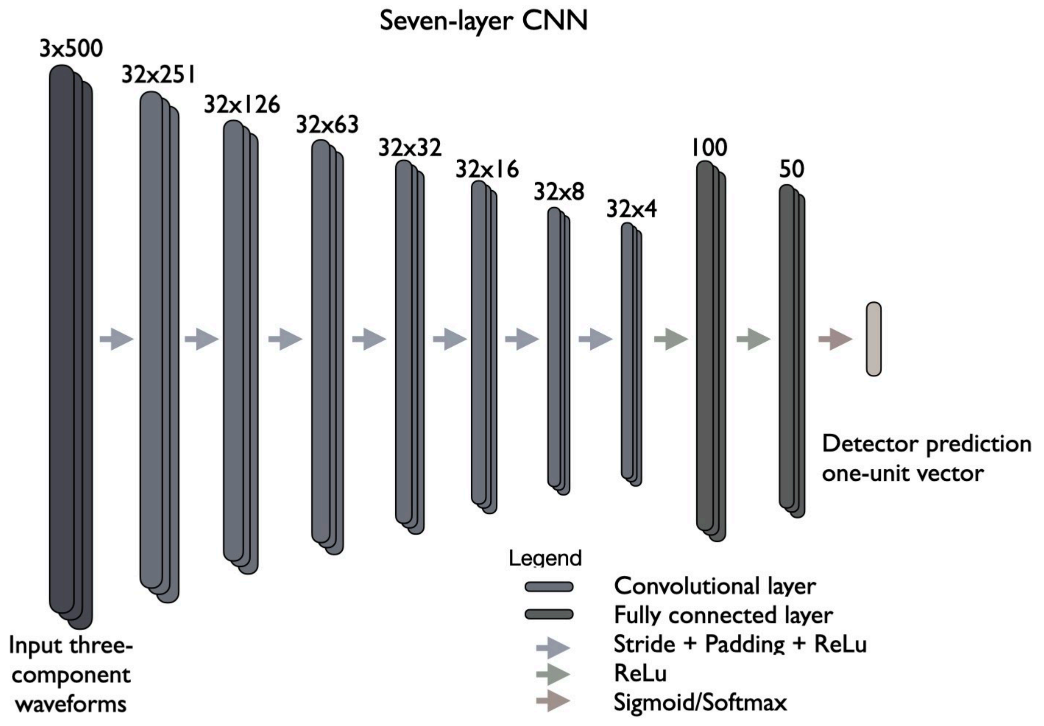


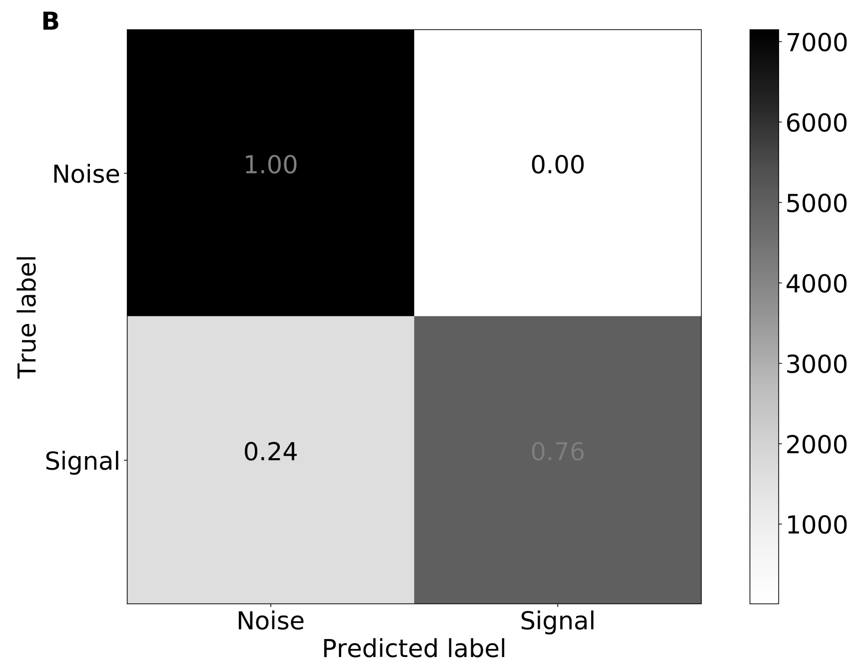
1 Training data

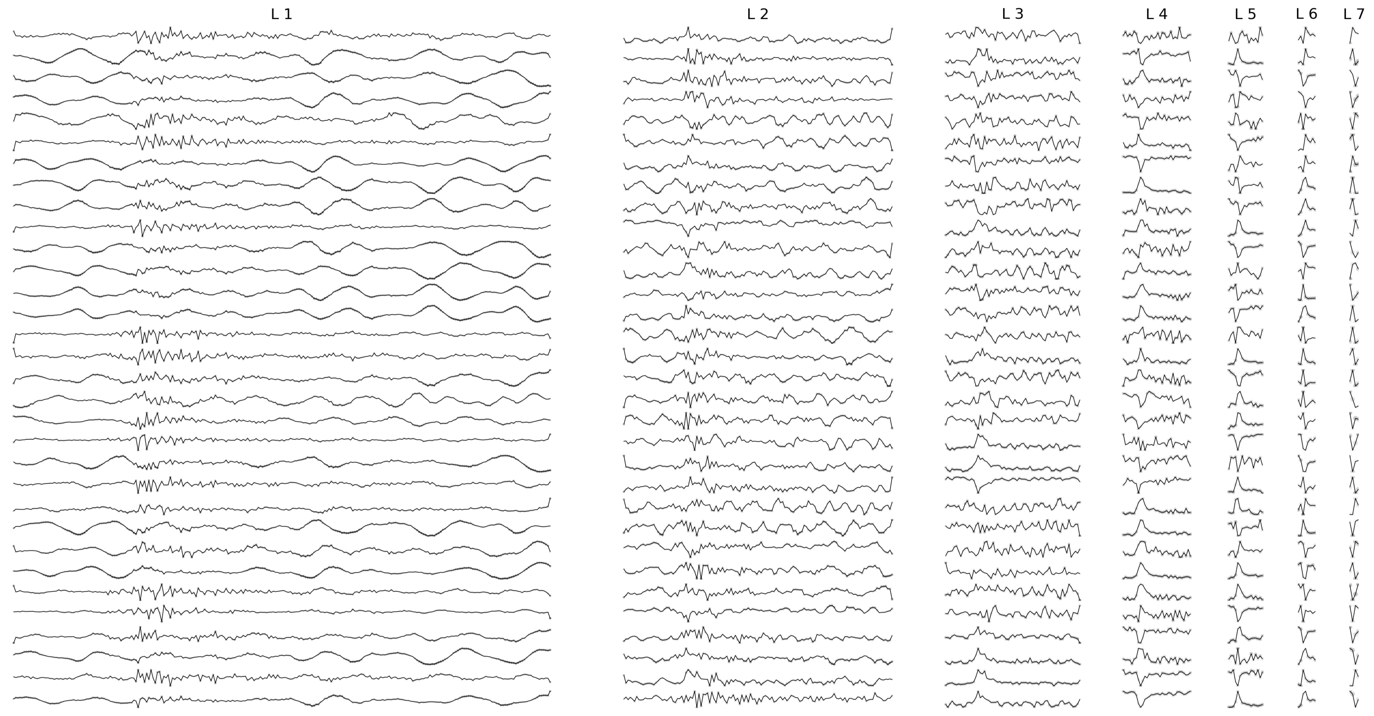
2 Training process

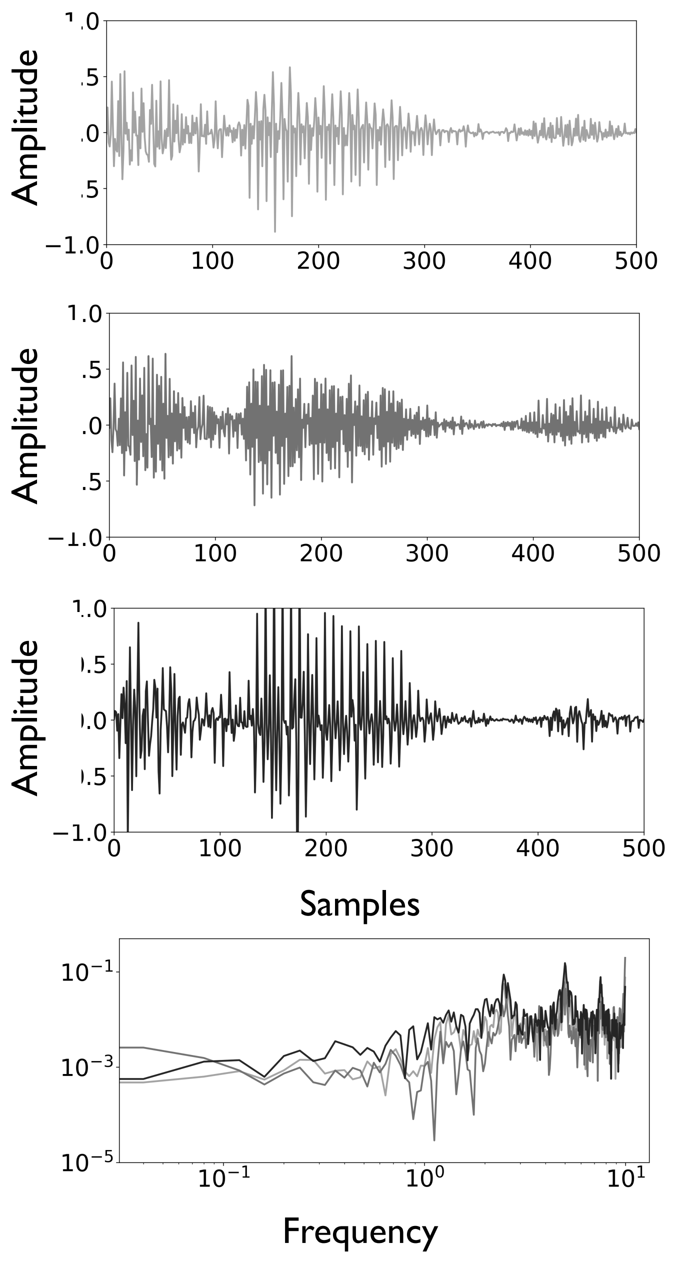
2 Modelling approach

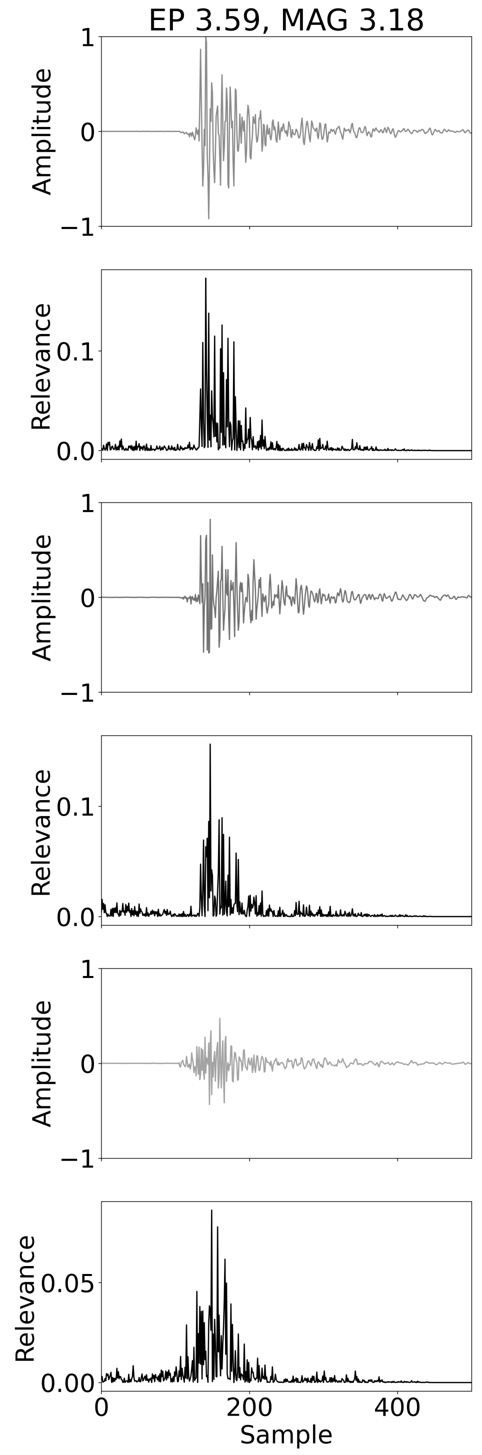
3 NN architecture











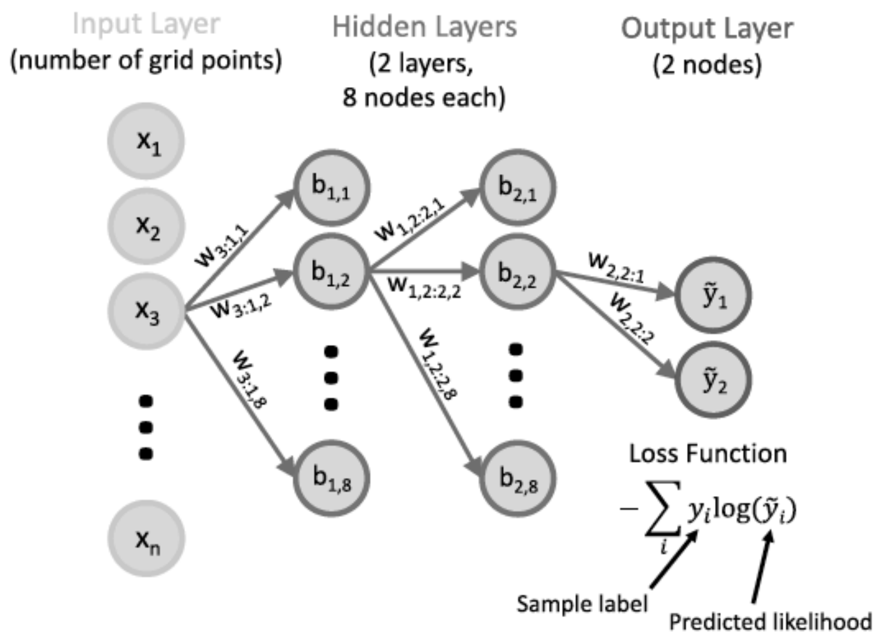
Take away message: there are endless options to explore data with the interpretation methods.

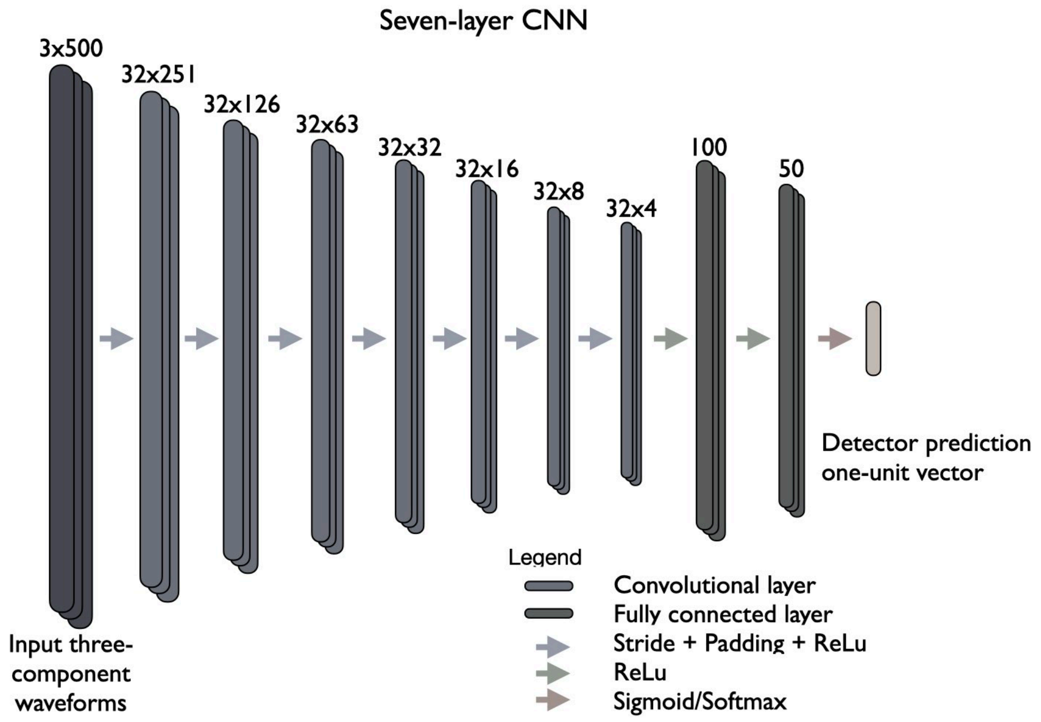
Introduction: Basics of neural networks, deep learning models

CNN detector: Developing CNN detector

Interpretation: What does interpretation stand for in DL?

Majstorović et. al., 2021, JGR



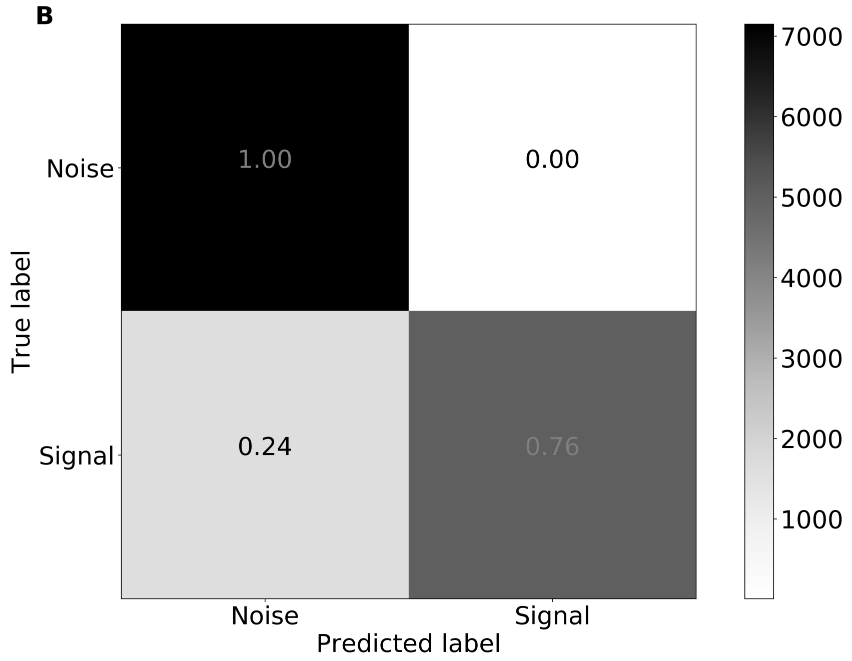


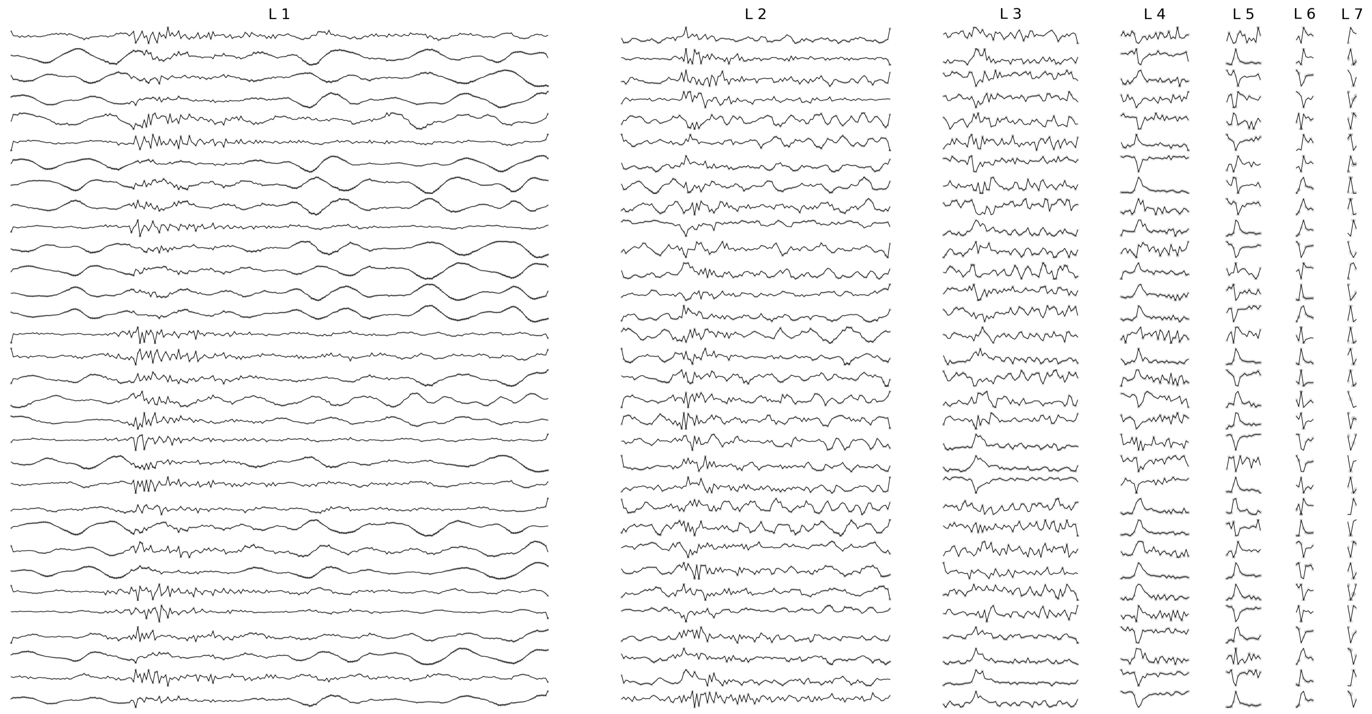
1 Training data

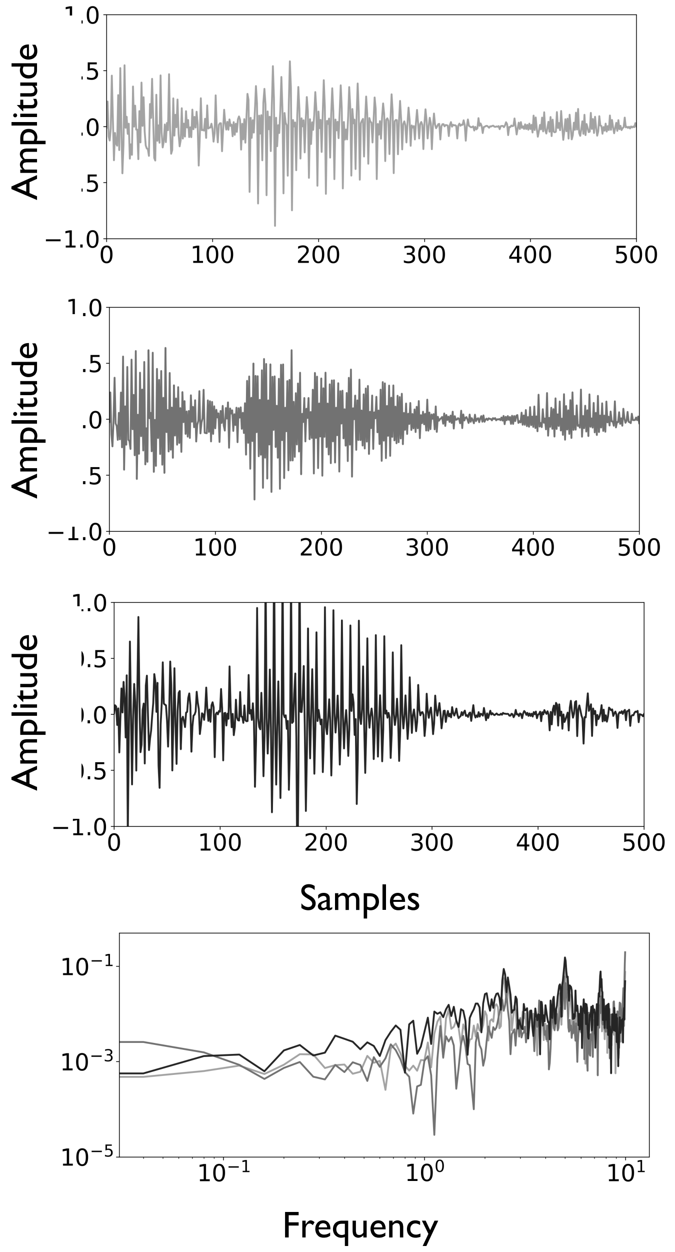
2 Training process

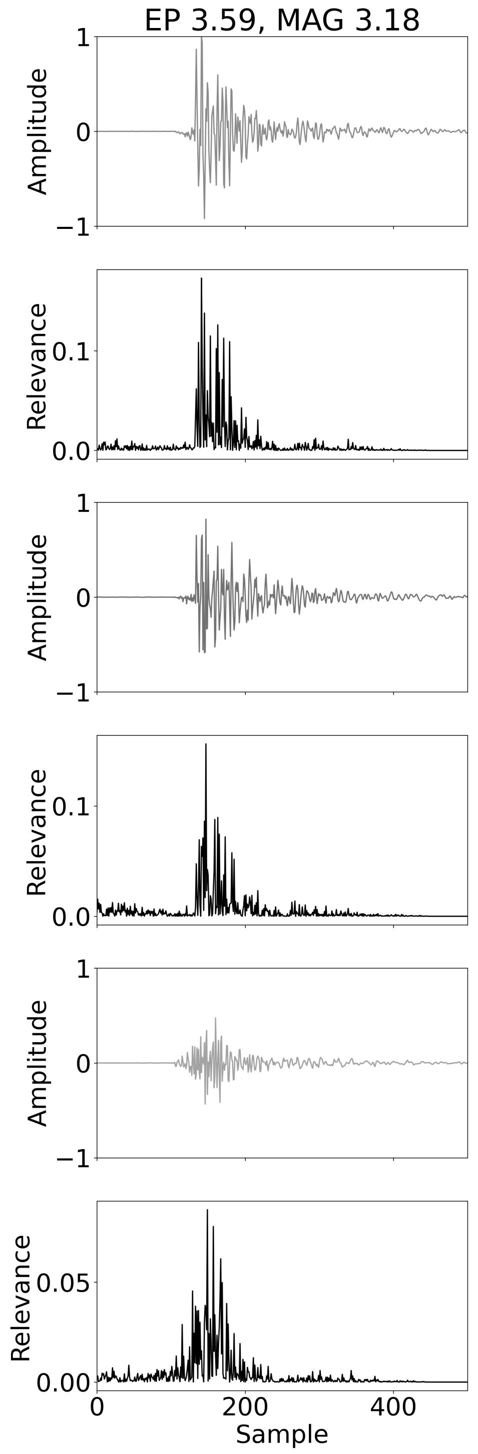
2 Modelling approach

3 NN architecture









Take away message: there are endless options to explore data with the interpretation methods.

Thank you for your attention!