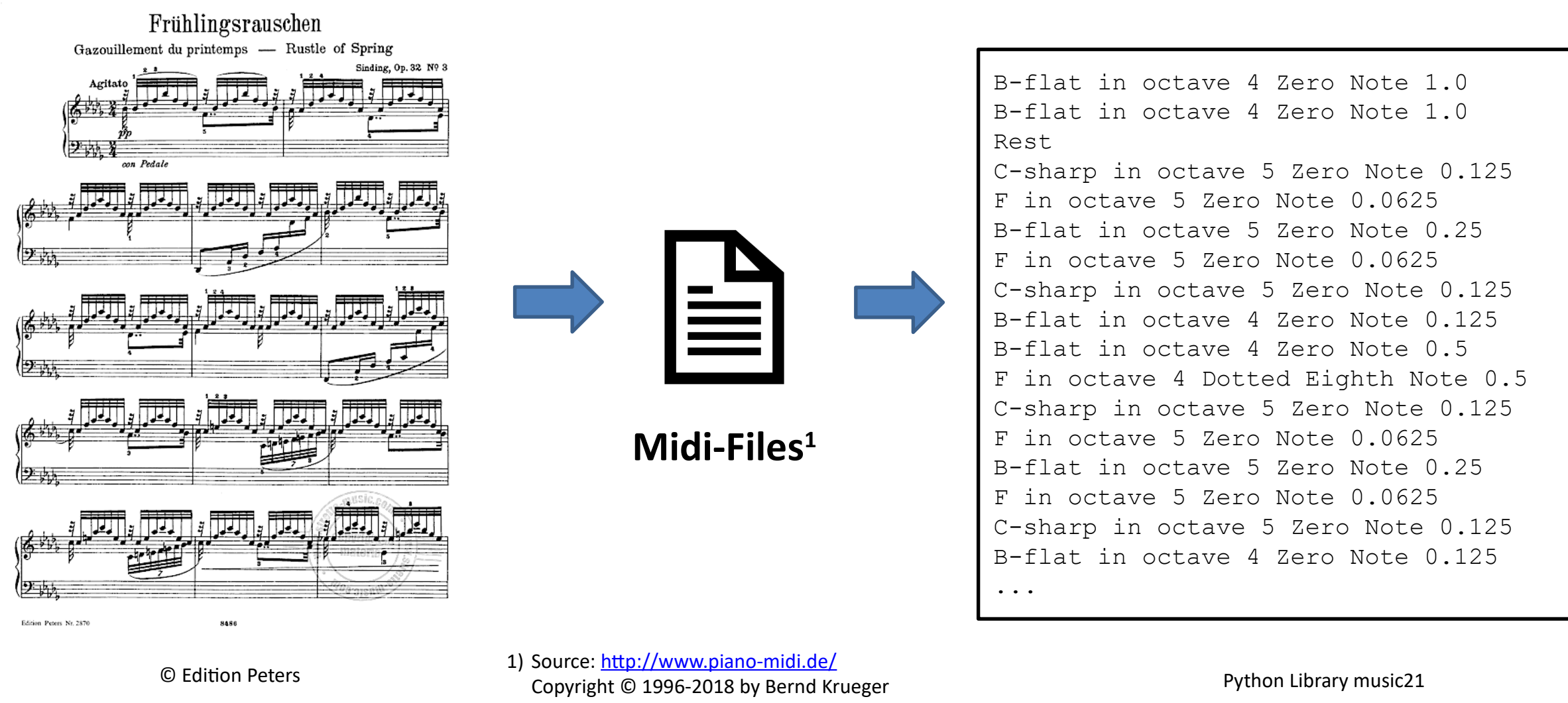


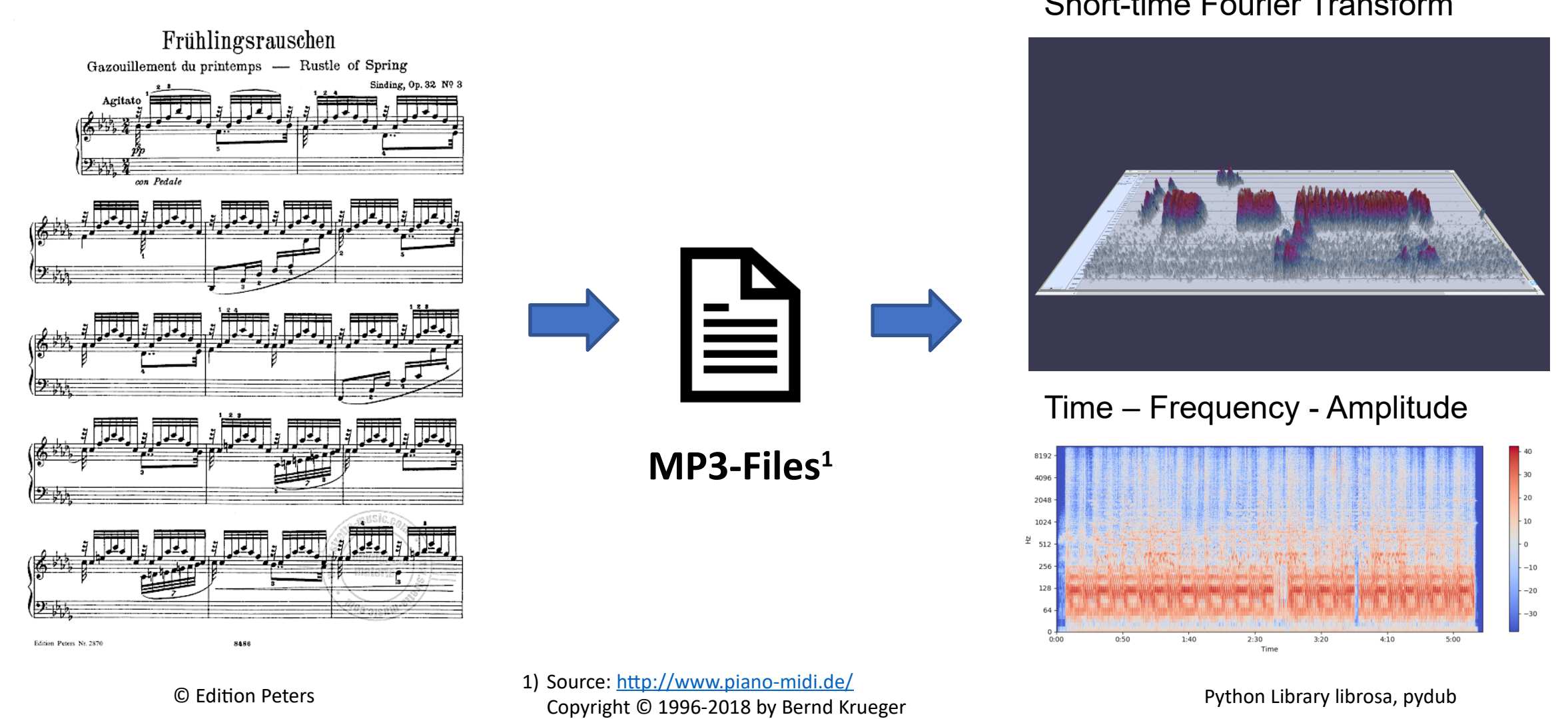
Musik Klassifikation mit Deep Learning

Project Team: Andreas Fischer, Bernd Novotny und Tobias Schieferdecker – April 2019

via Textanalyse



via Spektrogramm Bildanalyse



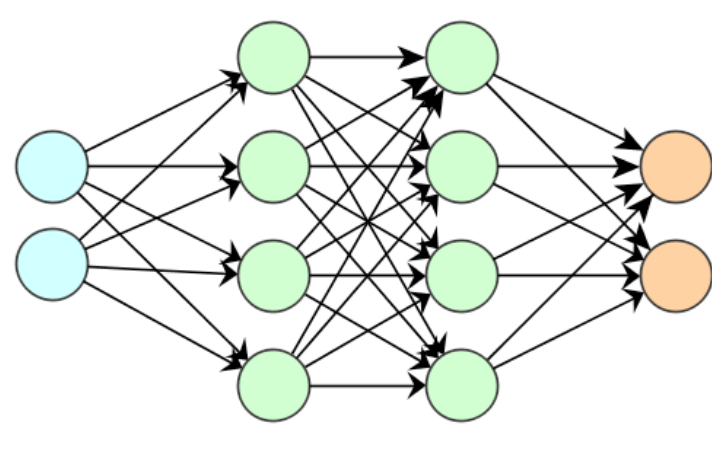
Die Daten: Facts & Figures

Input:

- 328 Werke
- Längstes: 9'917 Noten
- Kürzestes: 93 Noten

Slicing in 100 Noten Sequenzen

- 11'505 Sequenzen
- 48'460 verschiedene Noten/Akkorde/Pausen
- 100 Noten 10 – 25 Sekunden Musik



Output:

24 Komponisten (one-hot encoded)

Albeniz, Bach, Beethoven, Borodin, Brahms, Burgmueller, Chopin, Clementi, Debussy, Godowsky, Granados, Grieg, Haydn, Liszt, Mendelssohn, Moszkowski, Mozart, Mussorgsky, Rachmaninov, Ravel, Schubert, Schumann, Sinding, Tchaikovsky

Die Daten: Facts & Figures

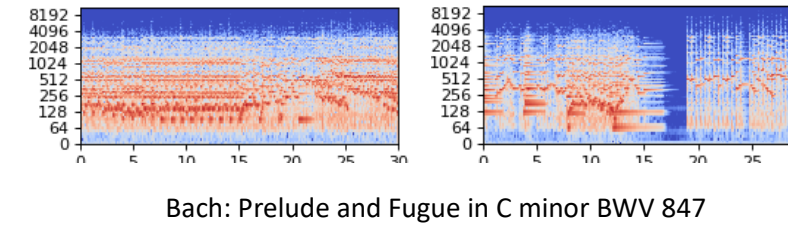
Input:

328 Werke – 1.4 GB MP3 von 25 Komponisten

Slicing in 30 Sekunden MP3's
2806 MP3 – 470 KB pro File

Generating Spectrograms

2472 Pictures 280 x 120 x 3



Set 1: 7 Komponisten

1090 Pictures:
- 697 training
- 175 validation
- 218 test

Set 2: 10 Komponisten
1879 Pictures:
- 1202 training
- 301 validation
- 376 test

Output:

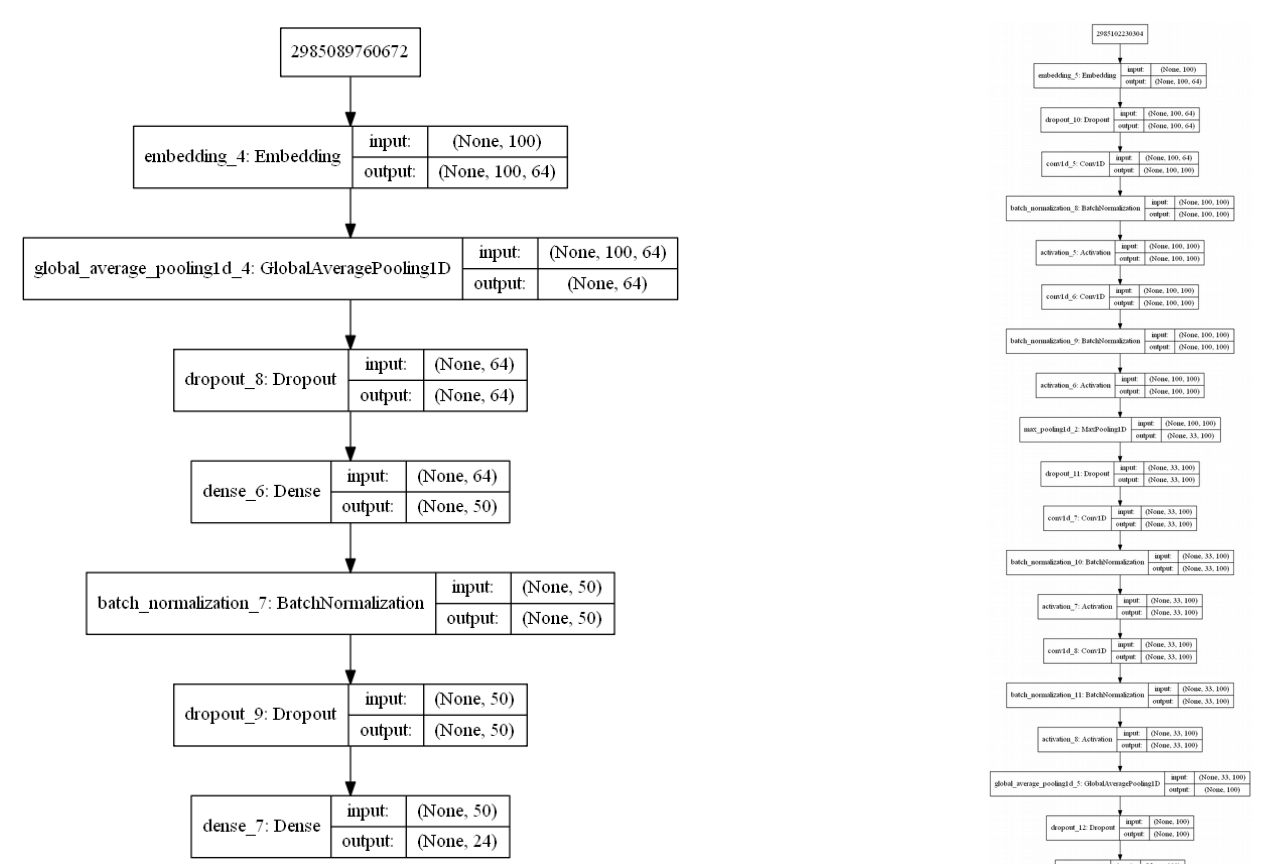
Komponisten (one-hot encoded)
7 10

- Beethoven
- Clementi
- Grieg
- Liszt
- Mendelssohn
- Mozart
- Tchaikovsky

Problematisch: Grosse Unterschiede bei Anzahl Bilder pro Komponist (80 – 421).

Auswertung

Embedding mit GlobalAveragePooling und Convolutional1D mit mehreren Layers am Besten



Test-Accuracy: 0.92133

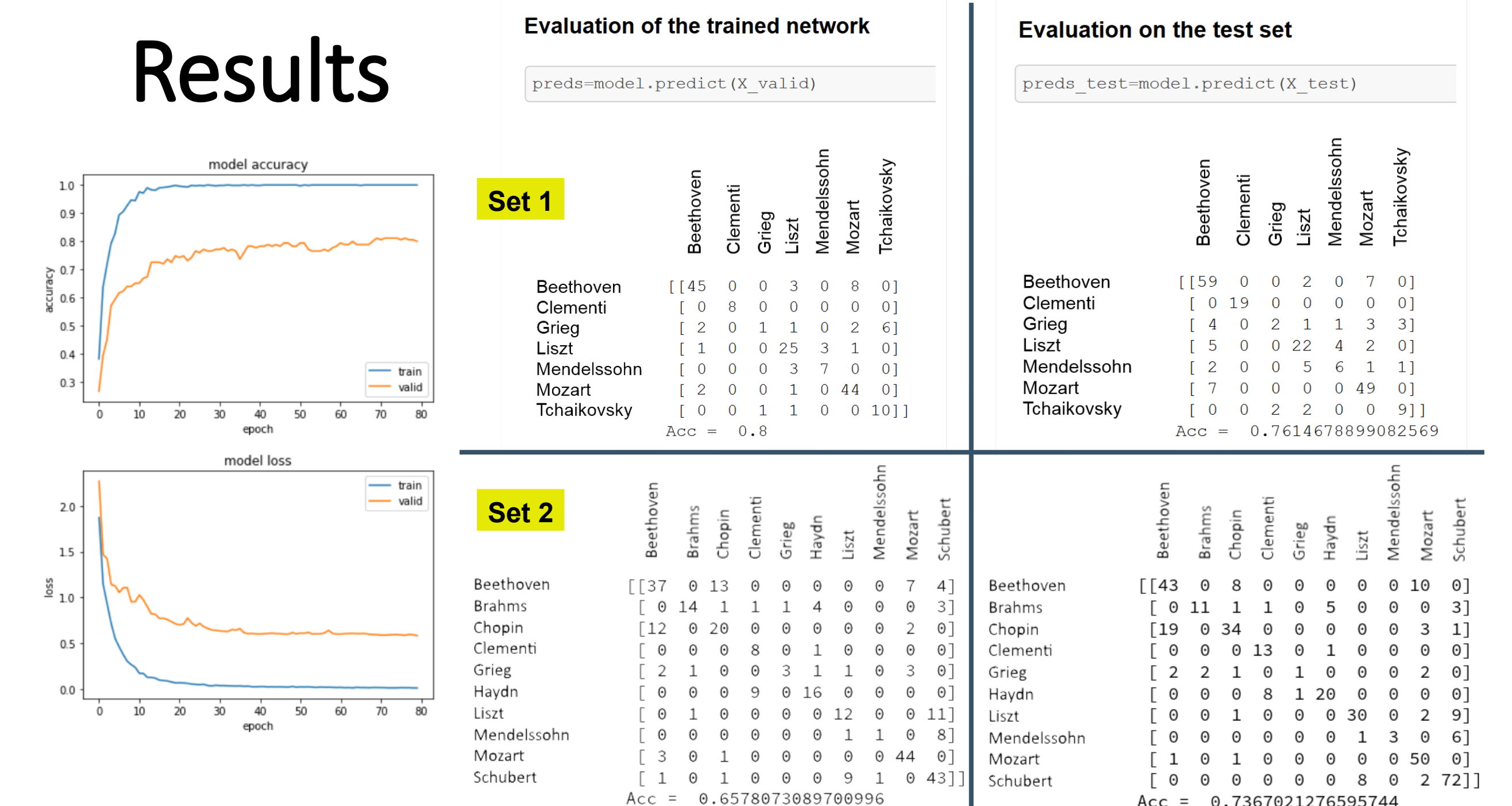
Test-Accuracy: 0.90699

Frage an die Musikexperten:
Darf man Beethoven mit Chopin verwechseln?



Quiz:
Wurde das «Frühlingsrauschen» auch falsch zugeordnet?

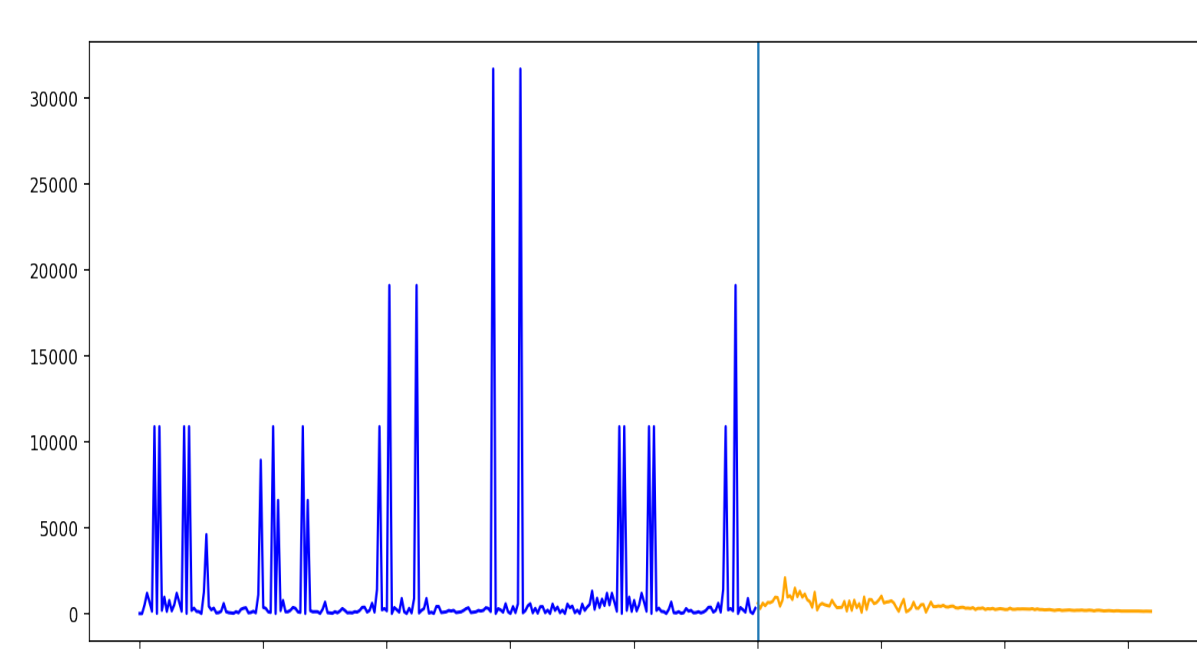
Results



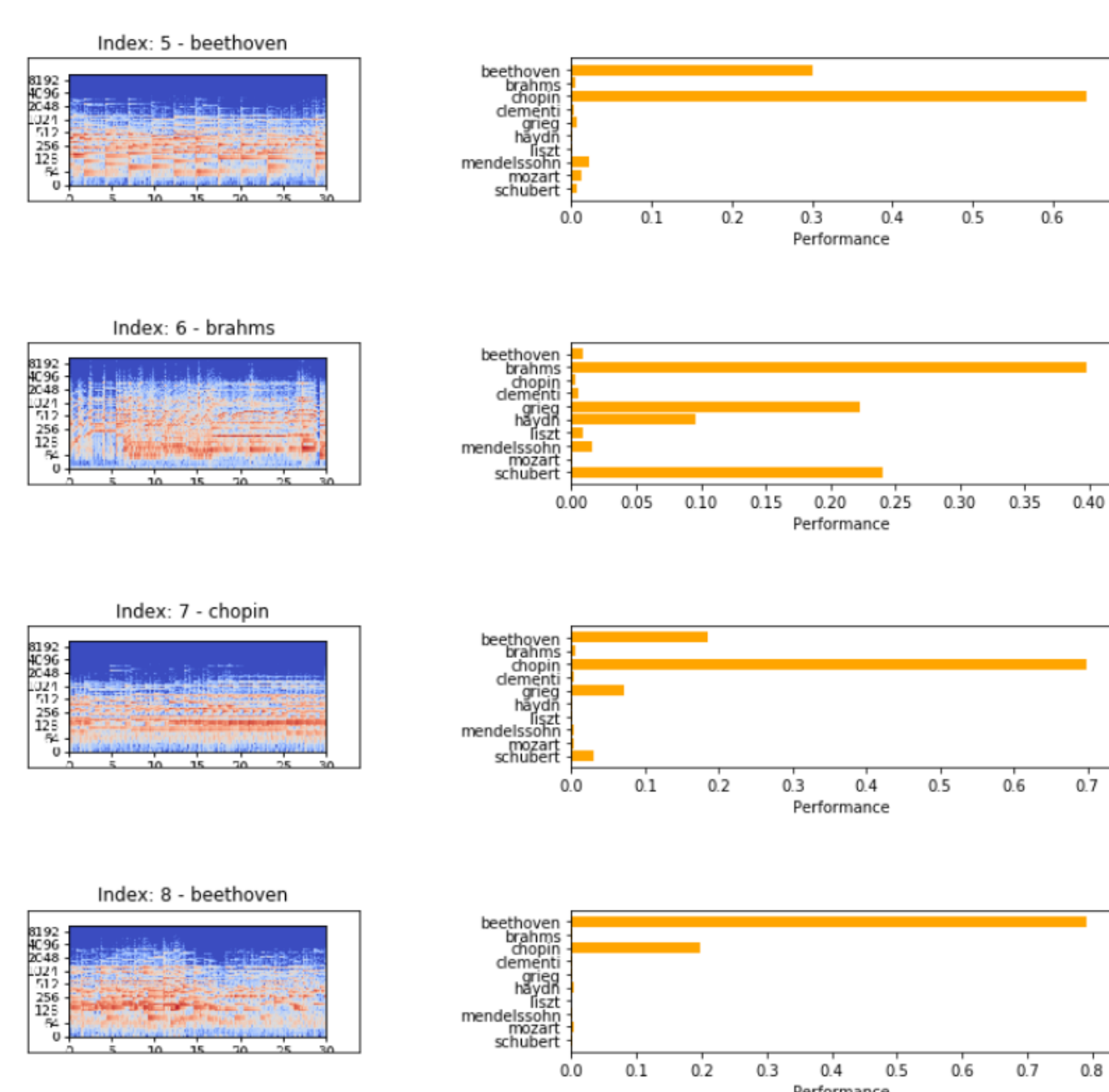
Musik-Generierung

Das geht auch. Ob es gut klingt? Immerhin scheint es abwechslungsreich und unsere Maschine mag es auf Dauer ruhiger... 😊

```
e-flat in octave 3 eighth triplet (1/3 ql) note 0.25
b-flat in octave 3 eighth note 0.125
f in octave 4 zero note 0.5
f in octave 5 zero note 0.5
d in octave 3 quarter note 1.0
g in octave 2 eighth triplet (1/3 ql) note 0.0625
f in octave 5 zero note 1.0
f-sharp in octave 3 quarter note 1.0
c in octave 5 dotted eighth note 1.0
f in octave 2 16th note 0.5
f-sharp in octave 2 zero note 0.0625
chord f in octave 4 | a in octave 4 quarter 1.0
g-sharp in octave 3 quarter note 0.25
e in octave 3 dotted quarter note 1.0
...
```



Erzeugt mit 1D Convolution with dilation rate (1/2/4/8)



Beethoven und Chopin scheinen ähnlich zu sein.

Parameter
Set 1: Total params: 4,743,277
Trainable params: 4,742,677
Non-trainable params: 600
Set 2: Total params: 4,744,180
Trainable params: 4,743,580
Non-trainable params: 600

Verbesserungen:
• Versuche mit kleineren Spektrogrammen (10 Sekunden) → mehr Bilder
• Versuche mit längeren Spektrogrammen (60 Sekunden) → deutlichere Merkmale?
• Weniger Dimensionen:
• Amplitude direkt und nicht mit RGB.
• Gleichverteilung Anzahl Bilder pro Komponisten
• CNN-Parameter und -Layers variieren