

Deep Composer Classification Using Symbolic Representation

Sunghyeon Kim*, Hyeyoon Lee*, Sunjong Park*, Jinho Lee, Keunwoo Choi
 {hahala25, hylee817, ryan0507, leejinho}@yonsei.ac.kr, keunwoo.choi@bytedance.com

Problem Definition



Composer Classification

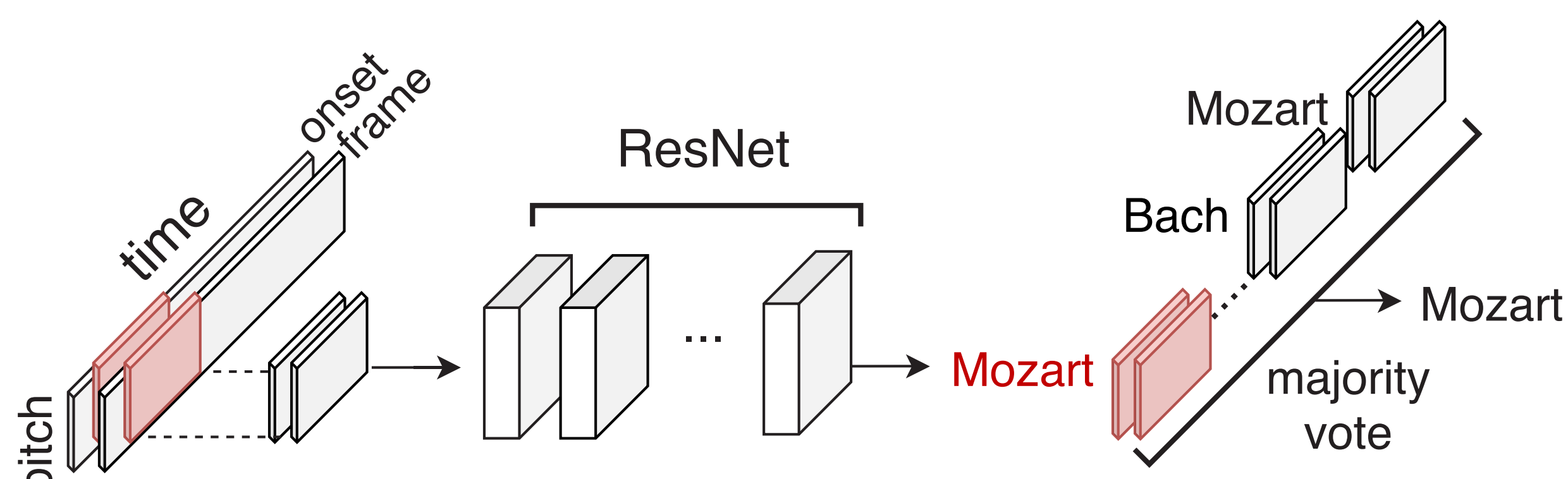


Using music on **symbolic** domain!

Why Symbolic Representation?

- Independent to timbre & acoustic recording environment.
- Focus on note-related aspects such as pitch and duration of notes.

Proposed System



MIDI in symbolic level

Input: 2 channels (onset, frame) of 2D array(time, pitch)

ResNet for learning spatial features such as pitch interval tendency (e.g. chord and voicing)

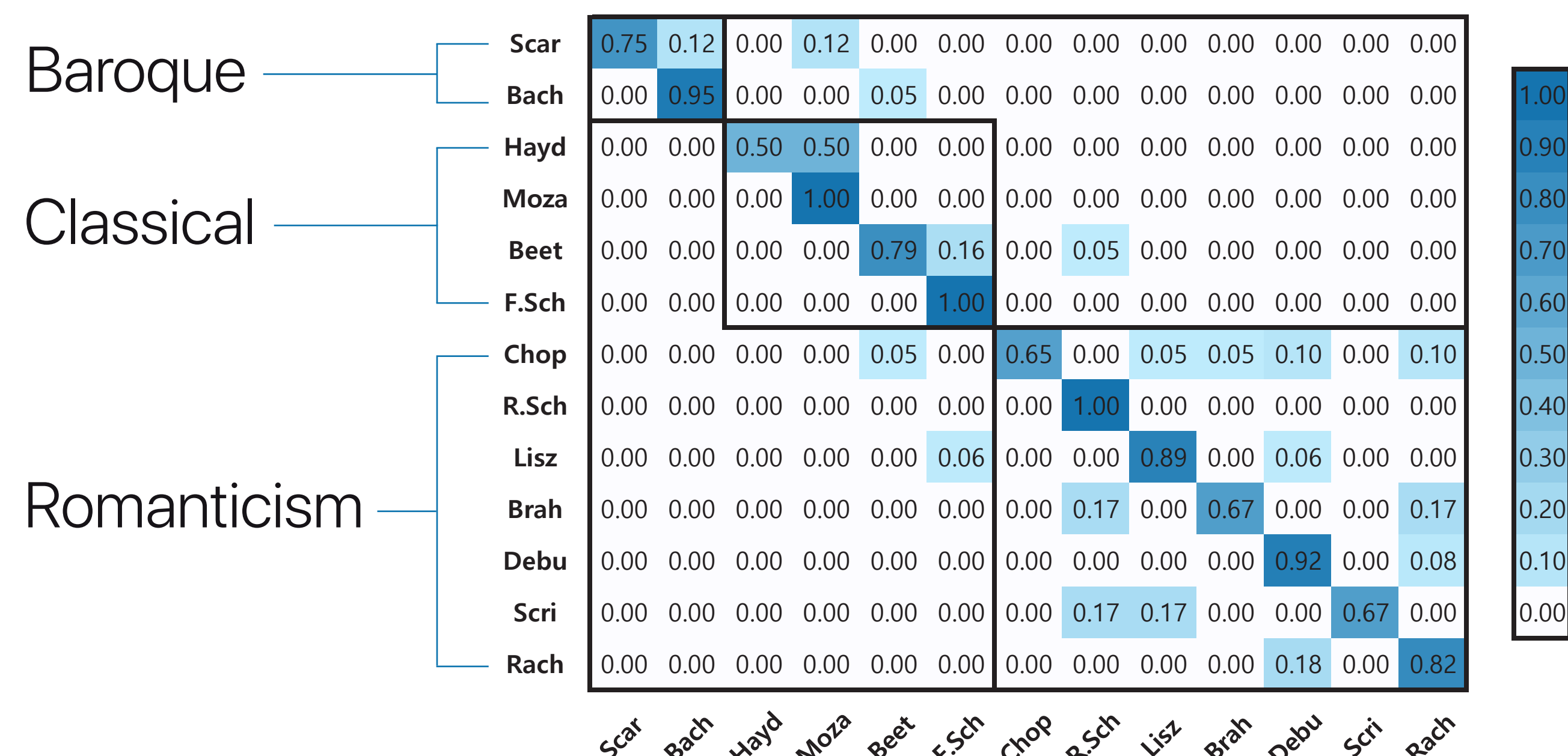
MAESTRO Dataset v2.0.0

classical MIDI performances
stratified sampling (347:158)

Composer (abb.)	Pieces	Composer (abb.)	Pieces
F. Chopin (Chop)	64	W. A. Mozart (Moza)	29
J. S. Bach (Bach)	62	D. Scarlatti (Scar)	25
L. V. Beethoven (Beet)	62	J. Haydn (Hayd)	20
F. Liszt (Lisz)	60	A. Scriabin (Scri)	19
F. Schubert (F.Sch)	58	R. Schumann (R.Sch)	18
C. Debussy (Debu)	37	J. Brahms (Brah)	17
S. Rachmaninoff (Rach)	34		

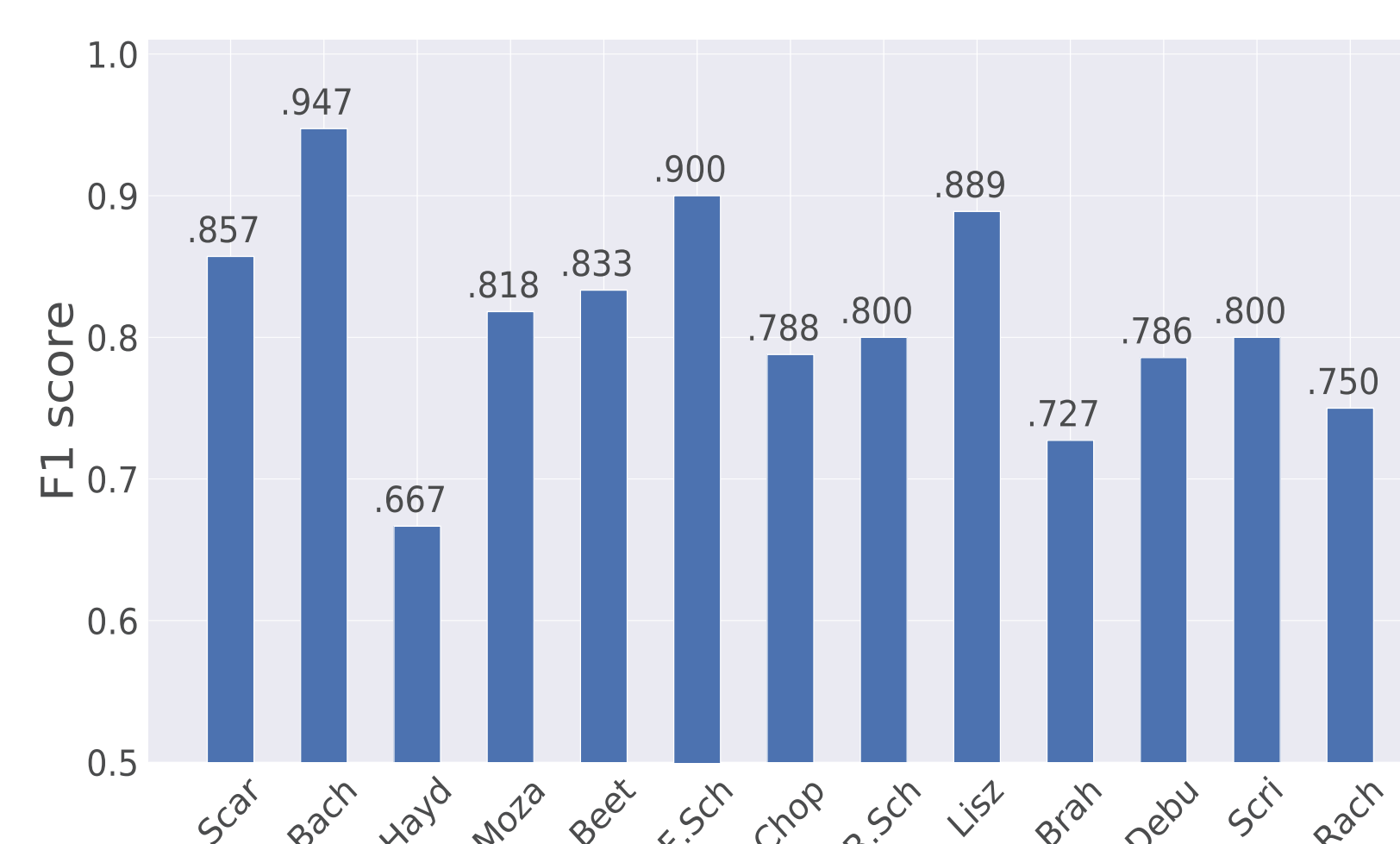
► 505 pieces by 13 composers

Results



Only 5/19 misclassifications are from different eras!

► Probably because similar musical patterns exist within the same era.



- Sorted by birth year
- The Spearman rank correlation coefficient: **-0.45**

Model performed better for relatively old classical composers!

► Probably because it's easier

No. of Segments	Onset Channel	Frame Channel
5	Used .8333	Continuous .8333
10	Omitted .7858	Binarized .8525
20		
30		
60		
90		

Number of segments per track

► Performance converges over 30 segments

Onset Channel Usage

► Having Onset information helps!

Frame Binarization

► Improved acc. By 0.0192 to 0.8525 ► velocity info didn't help

Where can I find it?

Github

<https://github.com/KimSSung/Deep-Composer-Classification>

