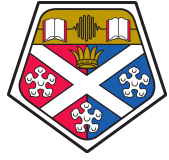
Applying Machine Learning Techniques to the Classification of Classical Orchestral Music

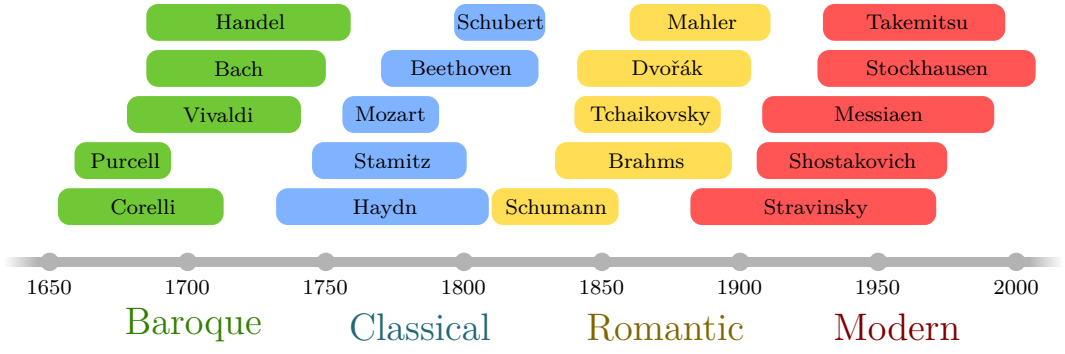
R. David Dunphy



University of Strathclyde, Glasgow

Friday 20th April, 2018

Timeline of compositional periods

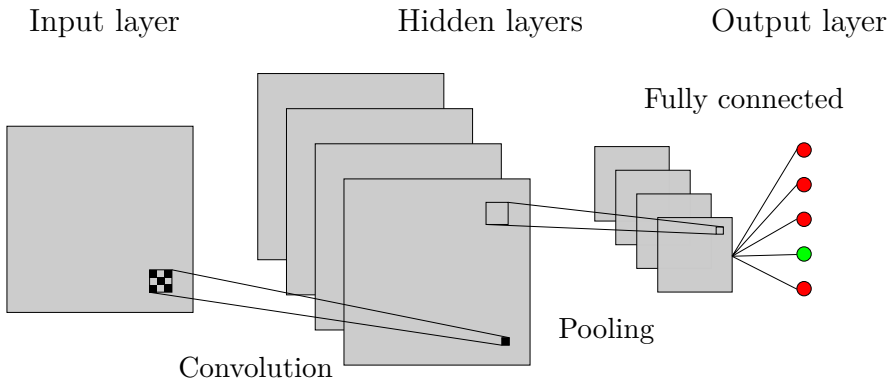


Corpus contents

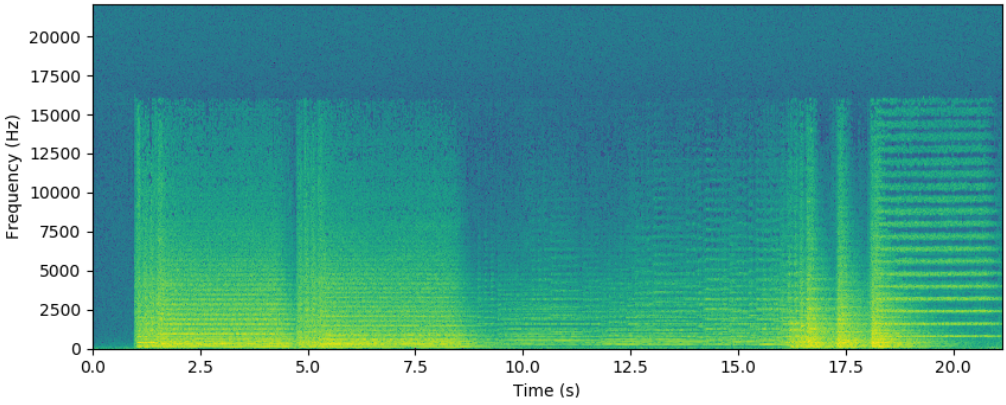
- 12 composers representing 4 periods
- 800 15-second samples per composer
- 40 hours of audio source files

Baroque	**Classical**	**Romantic**	**Modern**
Corelli	Haydn	Brahms	Stravinsky
Vivaldi	Mozart	Tchaikovsky	Shostakovich
Bach	Beethoven	Mahler	Messiaen

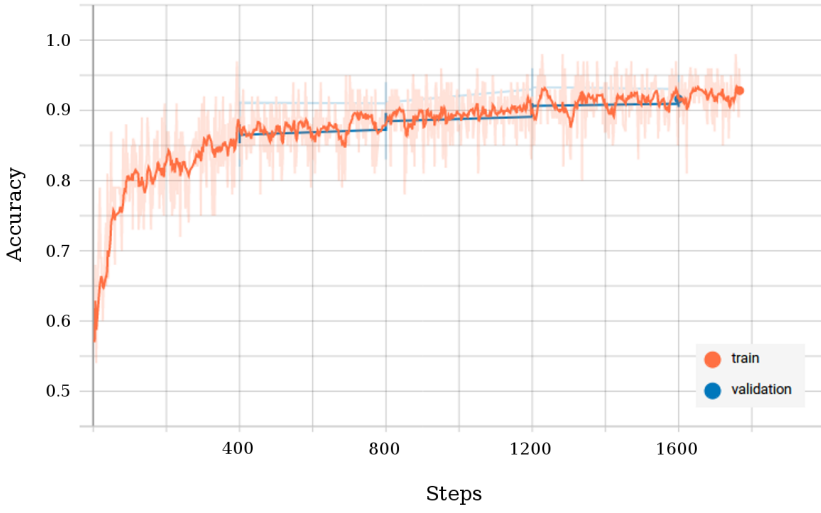
Convolutional neural network (CNN)



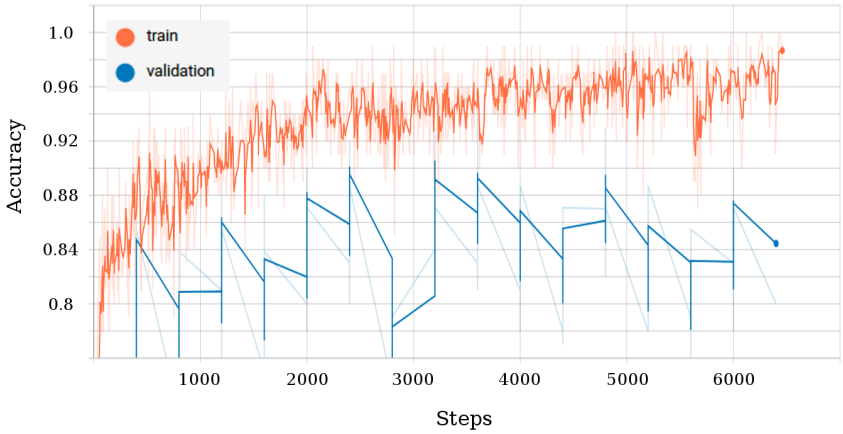
Spectrogram feature



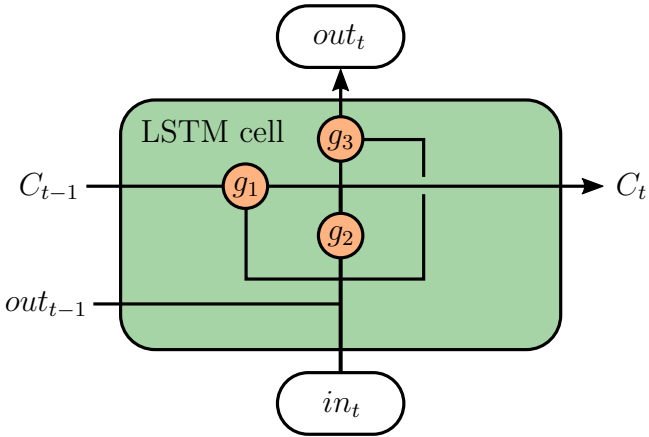
CNN accuracy on MS dataset



CNN accuracy with downsampling (16 kHz)

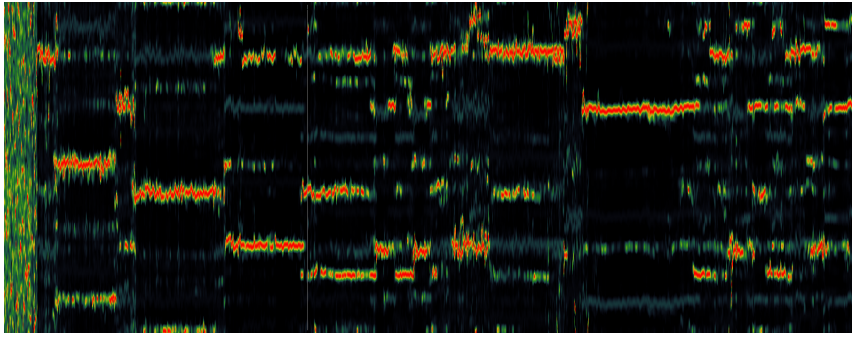


Long short-term memory network (LSTM)

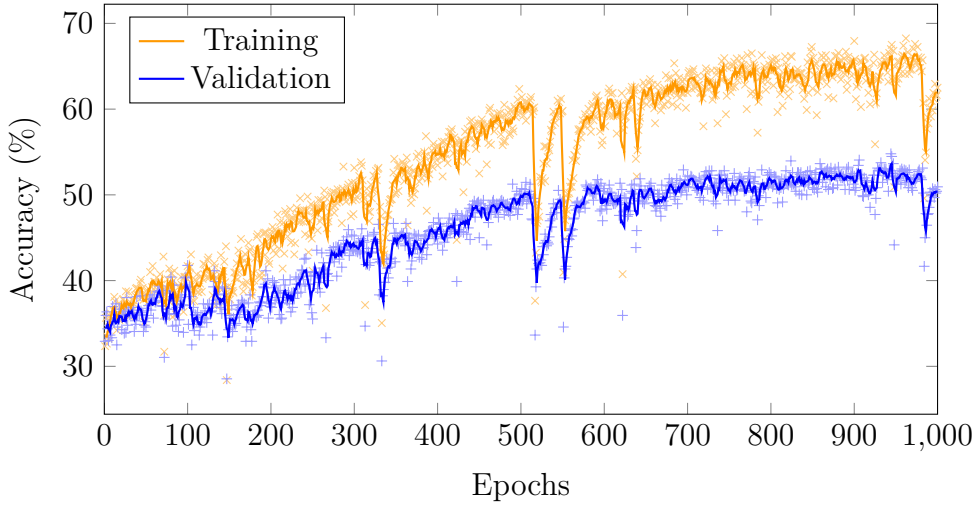


C = cell state
g_1 = forget gate
g_2 = input gate
g_3 = output gate

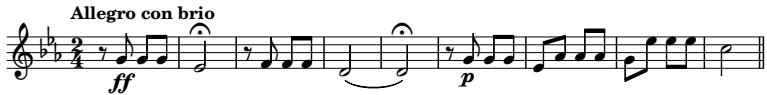
HPCP feature

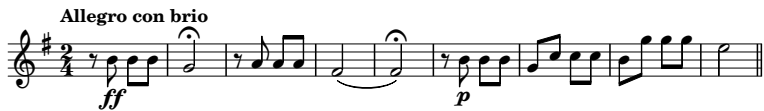


LSTM accuracy on BCRM dataset

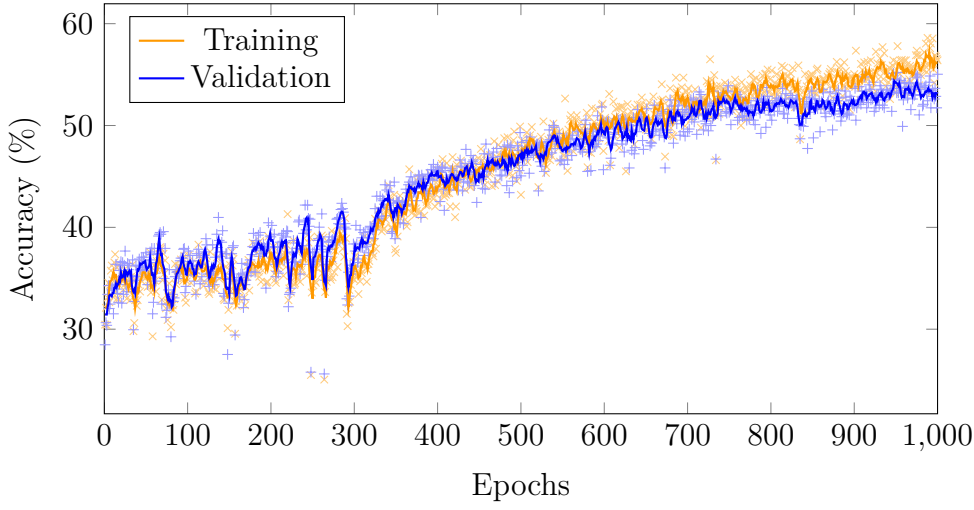


Musical transposition





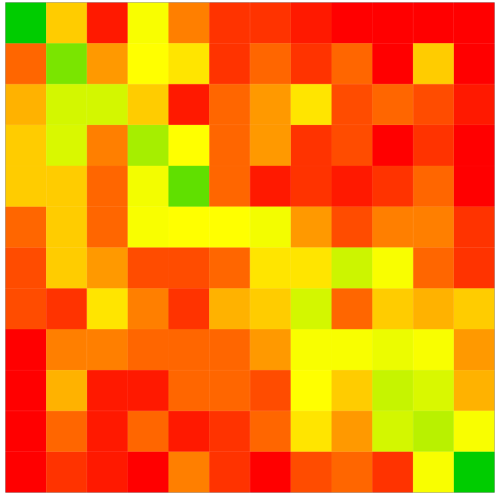
LSTM accuracy with transpositions



Confusion matrix for composer classification

		Prediction											
	($n = 960$)	Cor	Viv	Bac	Hay	Moz	Bee	Bra	Tch	Mah	Str	Sho	Mes
Actual value	Cor	**50**	8	1	11	5	2	2	1	0	0	0	0
	Viv	4	**31**	6	10	9	2	4	2	4	0	8	0
	Bac	7	17	**17**	8	1	4	6	9	3	4	3	1
	Hay	8	16	5	**24**	10	4	6	2	3	0	2	0
	Moz	8	8	4	12	**35**	4	1	1	1	2	4	0
	Bee	4	8	4	11	10	**10**	12	6	3	5	5	2
	Bra	3	8	6	3	3	4	**9**	9	18	11	4	2
	Tch	3	2	9	5	2	7	8	**17**	4	8	7	8
	Mah	0	5	5	4	4	4	6	11	**11**	13	11	6
	Str	0	7	1	1	4	4	3	10	8	**19**	16	7
	Sho	0	4	1	4	1	2	4	9	6	17	**21**	11
	Mes	0	2	1	0	5	2	0	3	4	2	11	**50**
		87	116	60	93	89	49	61	80	65	81	92	87

Confusion matrix as colour scale



Conclusions

- A large, regular corpus of samples was assembled
- CNN with spectrograms results in timbre-driven approach
- RNN with HPCP results in timbre-invariant approach
- Transposition is useful as a method to improve training
- Test accuracy of 55% for period classification
- Test accuracy of 31% for composer classification

Future work

- Resolve performance issues of CNN approach
- Improve RNN configuration and hyperparameter tuning
- Combine approaches to create CRNN