Annotated References

[1] Blankertz, B. (n.d.). *The Constant Q Transform*. Retrieved from http://doc.ml.tu-berlin.de/bbci/material/publications/Bla_constQ.pdf

This was really only used in the context of Constant Q Transformations, for which it was vital. I adapted the MATLAB code presented here into Python in order to build and understand my CQT method. It taught me about the importance of CQT and why it is chosen over its alternatives.

[2] Costantini, G., Perfetti, R., & Todisco, M. (2009). Event based transcription system for polyphonic piano music. *Signal Processing*, *89*(9), 1798-1811. https://doi.org/10.1016/j.sigpro.2009.03.024

This was good for giving me an overall understanding of polyphonic music transcription, including harmonics, spectrograms, and other techniques.

[3] Enthought. (2015, July 18). *Basic Sound Processing in Python | SciPy 2015 | Allen Downey* [Video file]. Retrieved from https://www.youtube.com/watch?v=0ALKGR0I5MA&t=414s

This video, and the subsequent github repository, helped me with sound processing and analyzing frequency time-series data.

[4] EuroPython Conference. (2016, August 4). *Anna Wszeborowska - Music transcription with Python* [Video file]. Retrieved from https://doi.org/10.5446/21107#t=00:09,01:10

This video from a conference was helpful to me because it showed me the monophonic transcription automation. Without understanding monophonic music transcription, I cannot begin attempting polyphonic transcription. Monophonic transcription does not use CNNs, but I still learnt about important concepts like

aliasing, the conversion between midi-notes and frequencies, and saving/loading numpy arrays.

[5] Fournier-S'niehotta, R., Rigaux, P., & Travers, N. (2017). Modeling Music as Synchronized Time Series: Application to Music Score Collections. *Information Systems*, *73*, 35-49. https://doi.org/10.1016/j.is.2017.12.003

This source has important information regarding time series data, and how music should be modeled in a time series fashion. It was useful for the initial part of my project, since it discusses wav and mp3 file formats.

[6] Fullstack Academy. (2017, January 18). *Digital Signal Processing (DSP) Tutorial - DSP with the Fast Fourier Transform Algorithm* [Video file]. Retrieved from https://www.youtube.com/watch?v=HJ_-5mqUZ70

This video is helpful to me for Digital Signal Processing.

[7] Jonathan Sleep, wav2mid: Polyphonic Piano Music Transcription with Deep Neural Networks, (2017), GitHub repository, https://github.com/jsleep/wav2mid

Useful for seeing spectrograms and transformations.

[8] *Music Transcription with Convolutional Neural Networks*. (n.d.). Retrieved from https://www.lunaverus.com/cnn

This source is great for explaining harmonics, the input and output of CNNs, as well as post-processing data. Images from the website were good for me to compare against my own spectrograms.

[9] Schörkhuber, C., & Klapuri, A. (2010). *Constant-Q transform toolbox for music processing*. Retrieved from https://iem.kug.ac.at/fileadmin/media/iem/projects/2010/smc10_schoerkhuber.pdf

Comparison of convolutional neural networks and other types of networks.

[10] Sleep, J. (2017, October). *Automatic Music Transcription with Convolutional Neural Networks Using Intuitive Filter Shapes*. Retrieved from https://pdfs.semanticscholar.org/3c64/386df17589af1c60007bd9877f8796216c02.pdf

[11] Ullrich, K., & Van der Wel, E. (2017, December). *Music transcription with convolutional sequence-to-sequence models*. Retrieved from https://karenullrich.info/pdfs/2017_ismir_2.pdf

Good source for description of sequence-to-sequence models and the purpose of CNN in music transcription.

[12] V. Emiya, R. Badeau and B. David, *Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle*, IEEE Transactions on Audio, Speech and Language Processing, (to be published);

The main dataset and documentation, which needs to be cited to give credit to the creators and aggregators of this data. Obtained via the MAPS website through direct message request.

[13] Verma, D. (2017, December). *Music Transcription using a Convolutional Neural Network*. Retrieved from https://medium.com/@dhruvverma/music-transcription-using-a-convolutional-neural-network-b115968829f4

This was a cool source and one of the first ones I encountered. It gives sample input, output, and talks about a method similar (use of CNNs) but different than what I plan to do. I'm not the biggest fan of the results of this project, so I want to sort of use it as a baseline. It talks about some data cleaning and subtleties in the processing (like hyperparameters, etc) which is useful.

[14] Wang, Q., Zhou, R., & Yang, Y. (2017). A Two-Stage Approach to Note-Level Transcription of a Specific Piano. *Applied Sciences*, *7*(9), 901-920. https://doi.org/10.3390/app7090901

This paper is useful to me because it outlines how to effectively use Convolutional Neural Networks with Non-Negative Matrix Factorization. The diagrams and illustrations guided me in understanding the order of the networks and how the network should be constructed such that the input and output are what I want them to be.