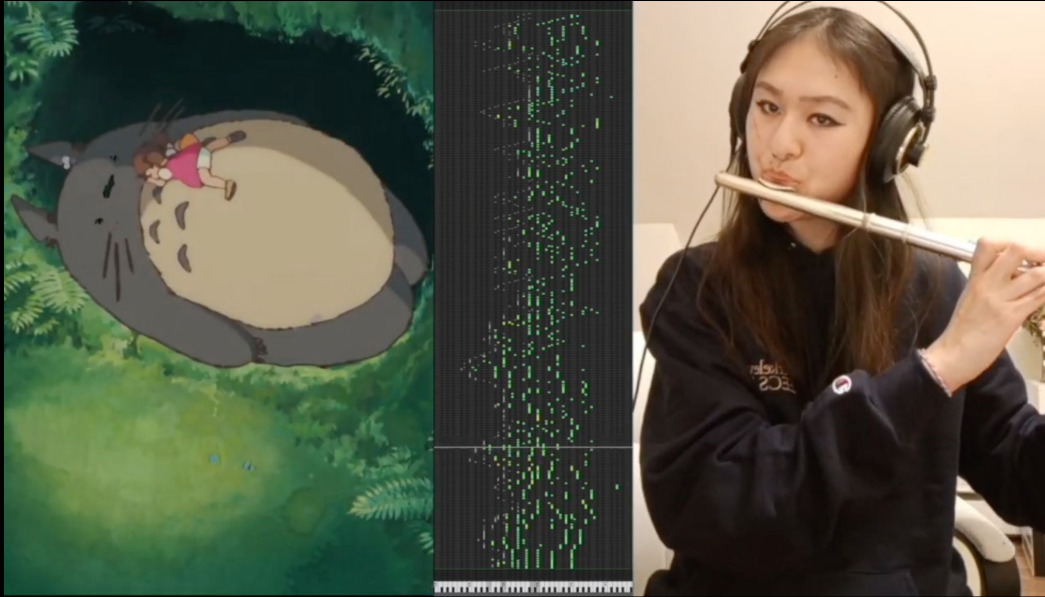


# lofi ghibli - music to train models/relax to



Group 4

Kevin Chen

Maya Shen

Kayo Yin

Kenneth Zheng

# Melody Generation



- Goal: Generate Ghibli-inspired melody
- Using the Super Piano 3 Colab notebook (Music Transformer)
- Trained on ~300 Ghibli MIDI melodies from Ghibli songs for 100 epochs
- Tried out different primer seed sequence to be continued by the model, including snippets from Ghibli songs and sequences written by us
- Final Ghibli-inspired melody:

Primer Seed



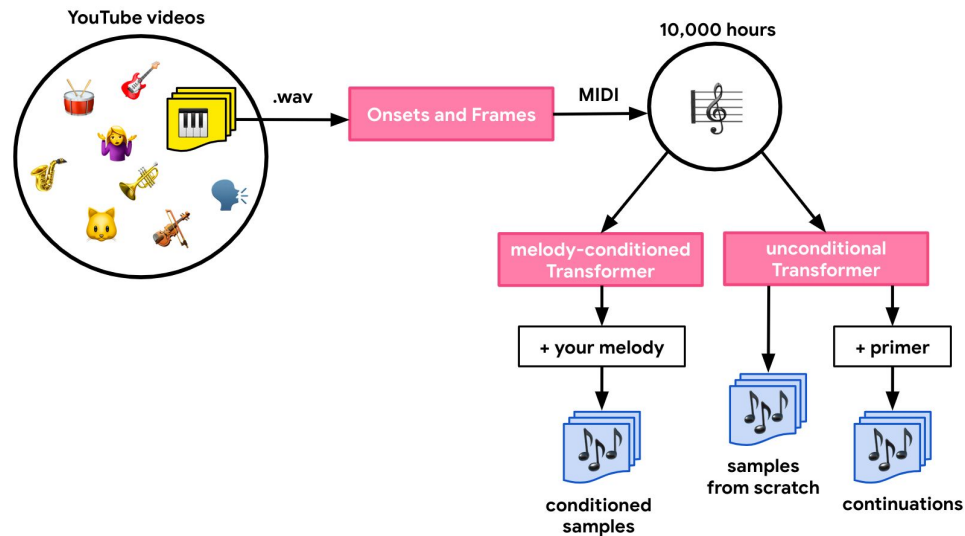
Primer Seed + Generated Output



# Accompaniment Generation



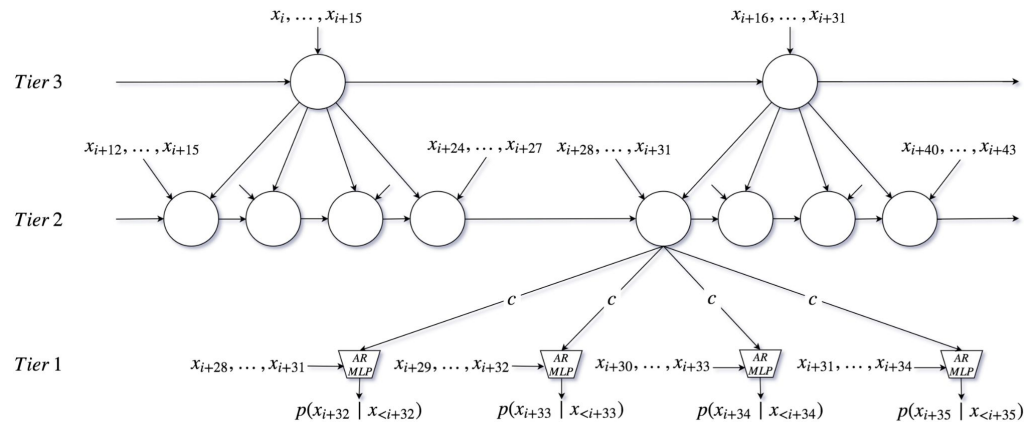
- Goal: generate the accompaniment for the melody composed by our AI model
- Use seq2seq setup of Music Transformer to generate an accompaniment conditioned on the melody



# Background Texture Generation

## SampleRNN

- Models audio at different timescales with stacked RNN layers
- Hyperparameters:
  - 3 RNN layers
  - frame sizes 16/16/4
  - trained for 170 epochs (~24hrs)
- Training data:
  - 1.5hrs of Ghibli piano music from YouTube
- Results:
  - pleasing piano-like textures, but collapsed to 1 main note/chord with little variation



# Other Explorations

## Audio-style-transfer

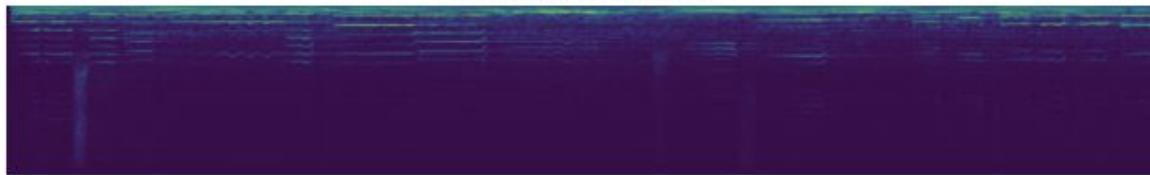
1. RandomCNN
2. Neural-style-transfer

Although we are able to generate the audio that contains the feature of the content and style audio, this music doesn't sound really great. So we end up didn't use it in our final result.

Here is result of combining Accompaniment as content and background texture from SampleRNN as style.



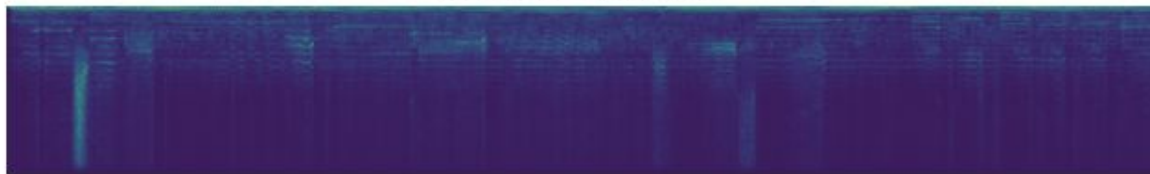
Content Spectrogram



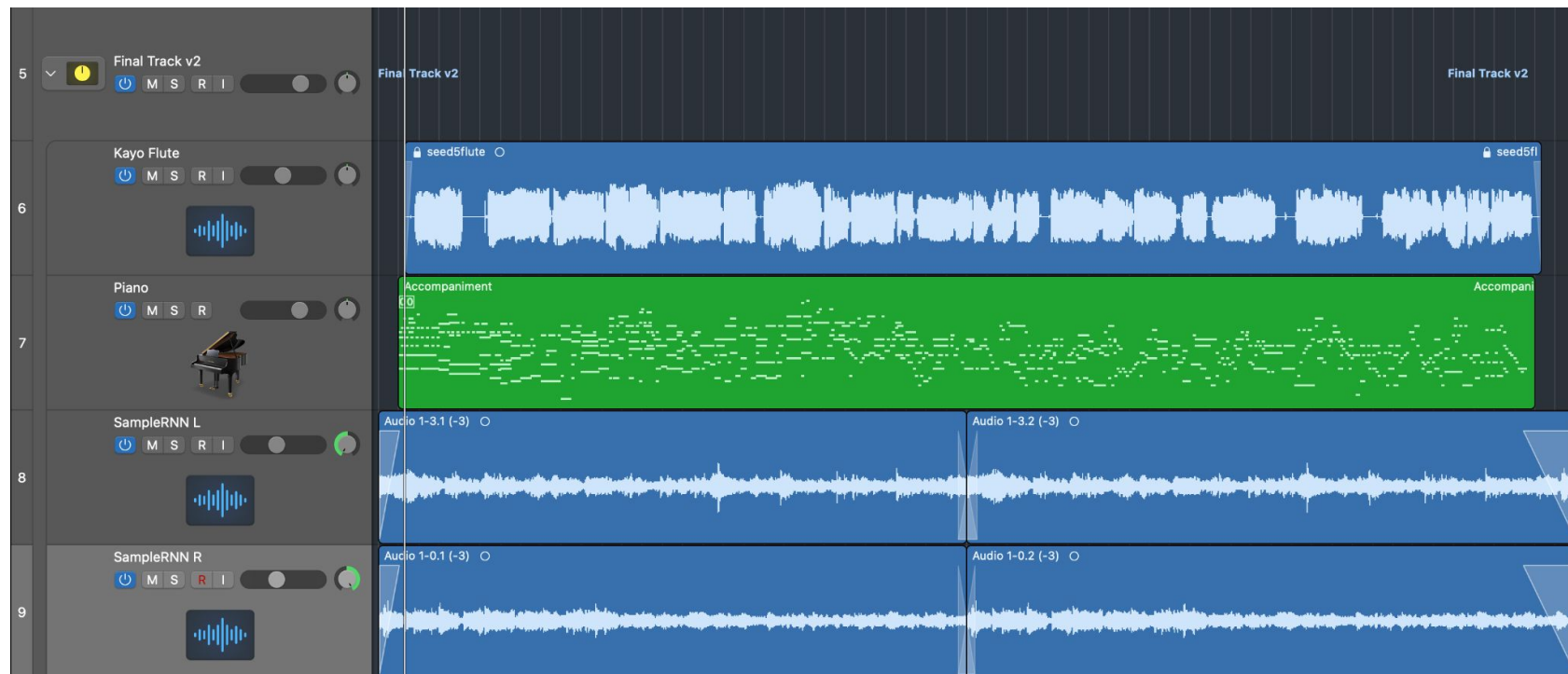
Style Spectrogram



Generated Spectrogram



# Mixing



# Result



# References

- [1] RandomCNN-voice-transfer <https://github.com/mazzzystar/randomCNN-voice-transfer>
- [2] Mehri, Soroush, et al. "SampleRNN: An unconditional end-to-end neural audio generation model." arXiv preprint arXiv:1612.07837 (2016).
- [3] Oord, Aaron van den, et al. "Wavenet: A generative model for raw audio." arXiv preprint arXiv:1609.03499 (2016).
- [4] SampleRNN implementation on GitHub, <https://github.com/wekaco/samplernn-pytorch>
- [5] Karl Heiner's Blog Post, "Generating Music with SampleRNN and WaveNet", [https://karlhiner.com/music\\_generation/wavenet\\_and\\_samplernn/](https://karlhiner.com/music_generation/wavenet_and_samplernn/)
- [6] Piano Studio Ghibli Collection 1, <https://youtu.be/HGI75kurxok> (used as training data)
- [7] Audio Style transfer [https://colab.research.google.com/github/tg-bomze/collection-of-notebooks/blob/master/Audio\\_Style\\_Transfer.ipynb](https://colab.research.google.com/github/tg-bomze/collection-of-notebooks/blob/master/Audio_Style_Transfer.ipynb)
- [8] Joe Hisaishi MIDI files <http://www5a.biglobe.ne.jp/~heefoo/midi/midi.html>
- [9] Super Piano 3 Colab notebook [https://colab.research.google.com/github/asigalov61/SuperPiano/blob/master/Super\\_Piano\\_3.ipynb](https://colab.research.google.com/github/asigalov61/SuperPiano/blob/master/Super_Piano_3.ipynb)
- [10] Cheng Zhi Anna Huang, et al. "Music Transformer." ICLR (2019).
- [11] Damon Gwinn's MusicTransformer Github repo/code <https://github.com/gwinndr/MusicTransformer-Pytorch>
- [12] Vaswani, Ashish, et al. "Attention is all you need." Advances in Neural Information Processing Systems (2017).
- [13] Generating Piano Music with Transformer  
<https://colab.research.google.com/drive/1qwMQ6BRKIlUW9kDRt8wQljkX3ZitCol?usp=sharing#scrollTo=Ql5g-x4foZls>
- [14] Huang, Y. S., & Yang, Y. H. (2020, October). Pop music transformer: Beat-based modeling and generation of expressive pop piano compositions. In *Proceedings of the 28th ACM International Conference on Multimedia* (pp. 1180-1188).
- [15] The MIDI Shrine: Anime MIDI files <http://www.midishrine.com/index.php?console=anime>