

Music Generation with Deep Learning Techniques

Student: Lee Yu Sheng Daniel Supervisor: Dr Alexei Sourin

Project Objective:

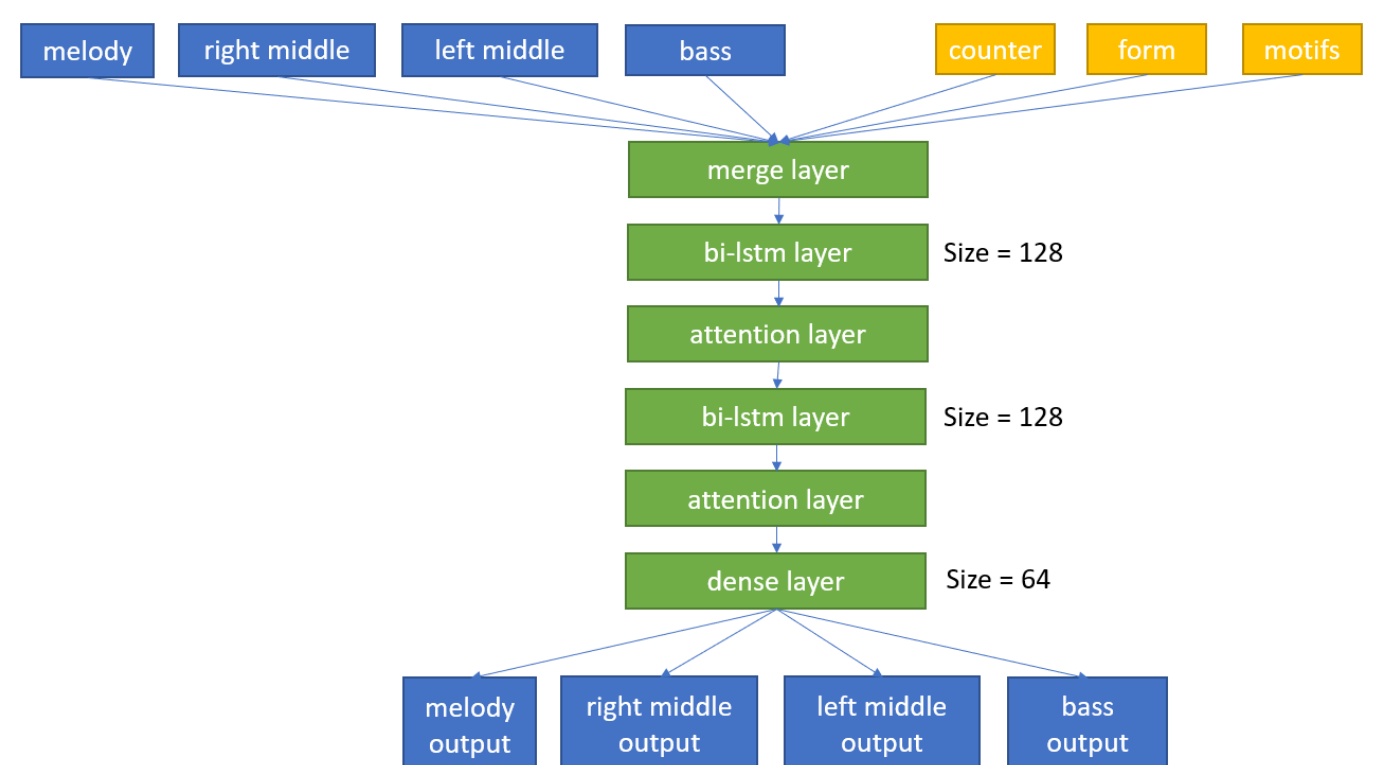
Existing research was able to improve short-term structure of generated music (note to note transitions) but lacked long-term structure (flow and progression of a piece). Conditioning has shown to potentially improve long-term structure but has not been fully explored. This project aims to improve the overall structure of generated music by using relevant conditioning inputs, accompanied by an appropriate model architecture.

Approach:

Bar counter, repeated motifs, and form were chosen as the 3 conditioning inputs to be tested. This was complemented with a model of Bi-LSTM and attention layers. Anime piano MIDI files were used as the dataset and trained for 50 epochs.



<- Scan QR code to listen to generated samples!



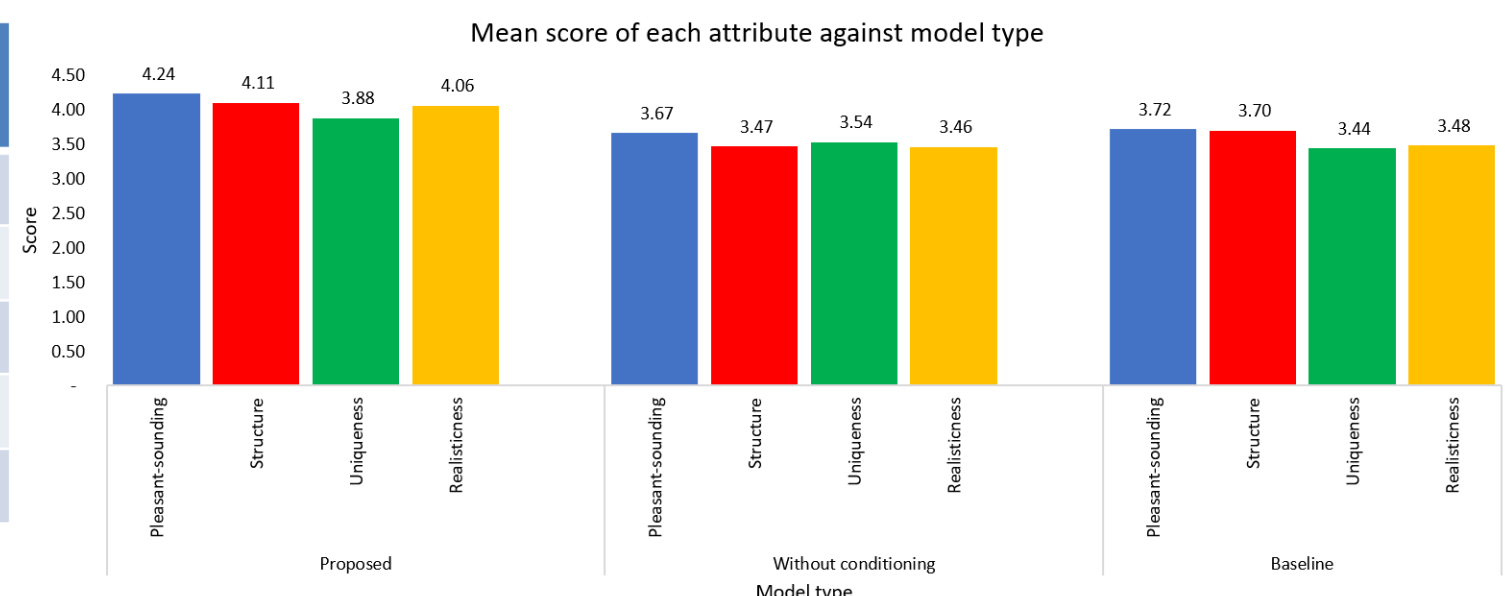
Overview of model

Results:

The proposed model was compared against 2 other models. Quantitative and qualitative analysis showed that the proposed model performed the best among them. These results verified that the proposed model was able to generate music with better structure and produced more coherent and melodious music.

	Proposed	Without conditioning	Baseline LSTM
Loss	0.8364	1.1248	2.2366
Melody accuracy	0.8817	0.8358	0.6955
Right middle accuracy	0.9441	0.9232	0.8609
Left middle accuracy	0.9662	0.9596	0.9202
Bass accuracy	0.9235	0.9026	0.7959

Quantitative analysis (test loss and test accuracies)



User study results