Adaptasi Positional Encoding pada Arsitektur Transformer untuk Sintesis Notasi Gamelan yang Koheren dan Terkendali

Arif Akbarul Huda

November 5, 2025

Rumusan Masalah

Model LSTM pembangkit notasi gamelan yang dikembangkan oleh Fanani, A.Z. dkk. (2025) gagal menangkap relasi tersirat antarnotasi. Mekanisme pengacakan Algoritma Genetika (GA) digunakan Fanani, A.Z. dkk. (2025) untuk menangani kelemahan LSTM tersebut, tetapi GA justru berpotensi merusak struktur koherensi seluruh notasi. Kondisi tersebut mengakibatkan peran notasi terhadap keseluruhan struktur terabaikan sehingga ciri khas notasi dan identitas musikal gamelan memudar seiring bertambah panjangnya sekuens notasi. Apabila hal ini tidak diatasi, koherensi tematik dalam komposisi notasi gamelan tidak terwujud. Oleh karena itu diperlukan pendekatan baru, model pembangkit notasi gamelan yang dapat mempertimbangkan peran tersirat setiap notasi dalam struktur lagu melalui mekanisme perhatian (attention mechanism).

Original Result LSTM

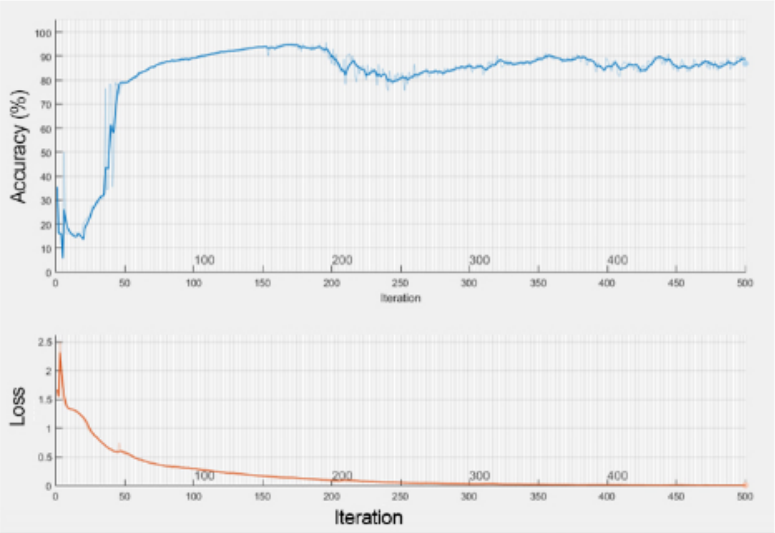

Figure: Original Training Result Gamelan with LSTM

Reverse Engineering LSTM

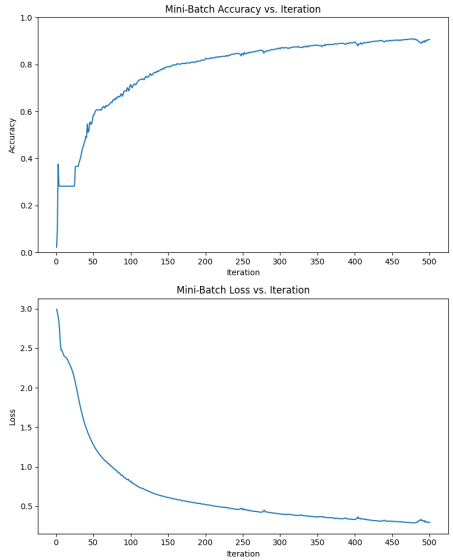

Figure: Training Result Gamelan with LSTM

Original Training Perform. Report LSTM

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:05	35.49%	1.6725	0.0100
100	100	00:06:56	89.39%	0.3051	0.0100
200	200	00:13:57	86.85%	0.0992	0.0100
300	300	00:21:01	86.98%	0.0374	0.0100
400	400	00:28:07	88.26%	0.0198	0.0100
500	500	00:35:17	86.91%	0.0143	0.0100

Figure: Original Report LSTM

Rebuild Training Perform. Report LSTM

```
PyTorch version: 2.8.0+cu126
input size: 20
hidden units: 40
num classes: 20
Using device: cuda

--- Starting Training (replicates 'trainNetwork') ---
Epoch [50/500], Iteration: 50, Loss: 1.2815, Accuracy: 58.38%
Epoch [100/500], Iteration: 100, Loss: 0.8172, Accuracy: 71.44%
Epoch [150/500], Iteration: 150, Loss: 0.6095, Accuracy: 79.09%
Epoch [200/500], Iteration: 200, Loss: 0.5202, Accuracy: 82.47%
Epoch [250/500], Iteration: 250, Loss: 0.4588, Accuracy: 84.70%
Epoch [300/500], Iteration: 300, Loss: 0.4030, Accuracy: 86.62%
Epoch [350/500], Iteration: 350, Loss: 0.3653, Accuracy: 88.10%
Epoch [400/500], Iteration: 400, Loss: 0.3291, Accuracy: 89.53%
Epoch [450/500], Iteration: 450, Loss: 0.3073, Accuracy: 90.26%
Epoch [500/500], Iteration: 500, Loss: 0.2938, Accuracy: 90.60%
--- Training Complete ---
Total Training Time: 7.57 seconds
```

Figure: Rebuild Report LSTM

Progress

- Koherensi Musikal
- Sequence Modeling
- User in the loop System

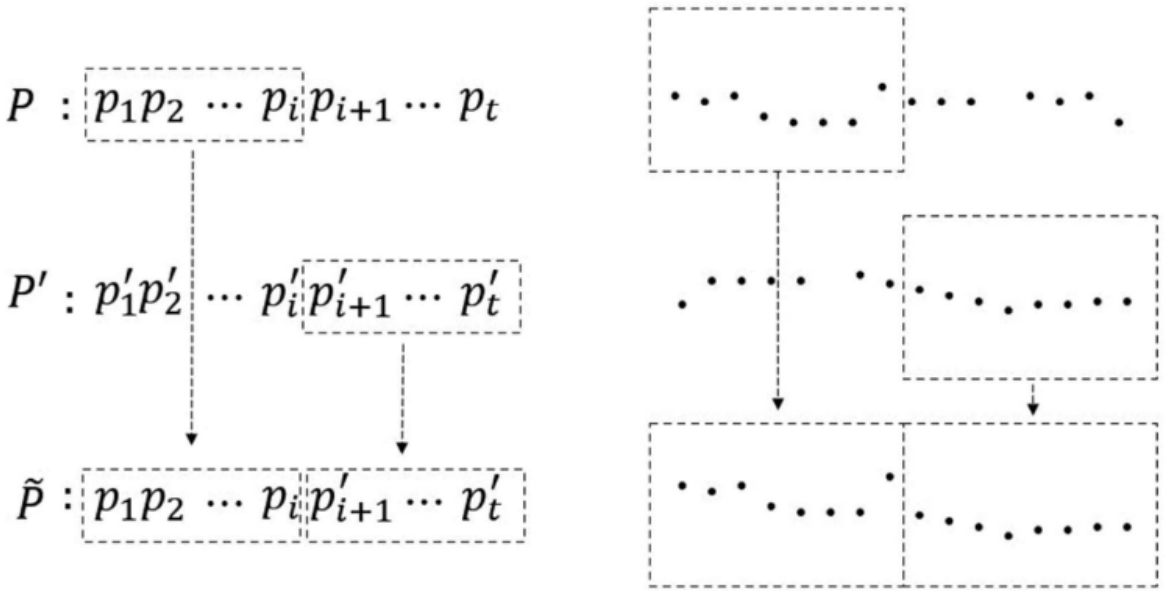

Figure: Contoh pitch mismatch method

Fu X, Deng H, Yuan X, Hu J. Generating High Coherence Monophonic Music Using Monte-Carlo Tree Search. IEEE Trans Multimedia. 2023;25:3763–72.

P.O.C

- Bagana membuktikan incoherence?
- Reverese Paper : LSTM, BiLSTM, G.A. Small prev dataset Gamelan
- Objective : Function untuk scoring tingkat coherence pada gamelan

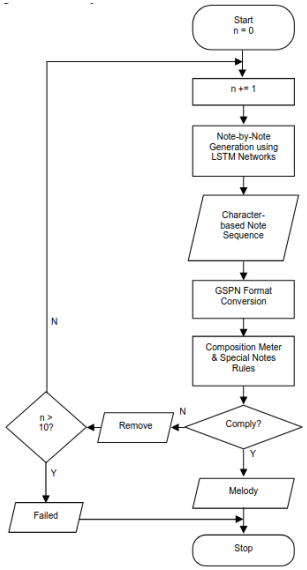

Figure: Rebuild LSTM

(Input layer 19 , Hidden layer 200 units) Attempt more than 10x. Syarif AM, Azhari A, Suprapto S, Hastuti K. Gamelan Melody Generation Using LSTM Networks Controlled by Composition Meter Rules and Special Notes. JAIT [Internet]. 2023 [cited 2025 June 9]; Available from: http://www.jait.us/show-224-1287-1.html

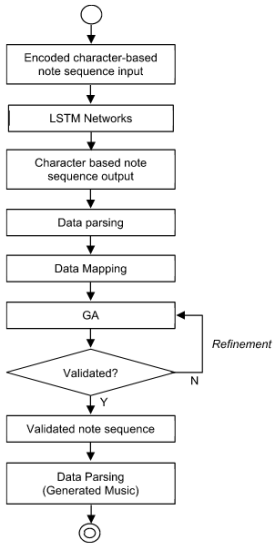

Figure: Rebuild LSTM with GA

(Input layer 17 , Hidden layer 200 units) Fanani AZ, Maulana Syarif A, Novita Dewi I, Karim A. Enhancing Creativity and Validation in Explanatory Deep Learning-Based Symbolic Music Generation: A Hybrid Approach With LSTM and Genetic Algorithms. IEEE Access. 2025;13:105280–301.