Neural Loop Combiner: Neural Network Models For Assessing The Compatibility of Loops

Bo-Yu Chen[1], Jordan B. L. Smith[2], and Yi-Hsuan Yang[1] 1 Academia Sinica, Taiwan 2 TikTok, London, UK

Demo Page: https://paulyuchen.com/Neural-Loop-Combiner-Demo/

Introduction

Assessing the Compatibility of Loops

- Help music producer to navigate the loops library efficiently by music loops compatibility estimation
- Most of previous works focus on rule-based compatibility estimation
- Neural Network can capture more complicated compatible relationship

Proposed System

Data Generation Pipeline

- Utilis the loop extraction algorithms [1, 2] to retrieve individual loops and loops used to combined before
- Apply loop refinement procedure to get rid of duplicate loops (Fig. 1)

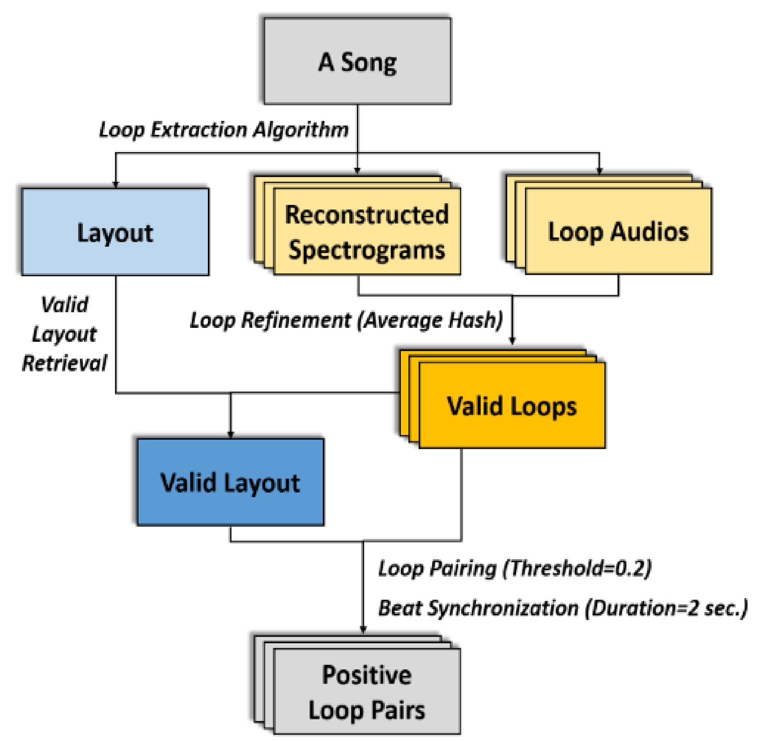

Fig. 1. The proposed data generation pipeline for positive loop pairs

Negative Sampling (Fig. 2)

- Within-song negative sampling: create negative loop pairs by shifting, rearranging, reversing one of the loops in a loops pair
- Between-song negative sampling: create negative loop pairs by choosing the loops from different songs

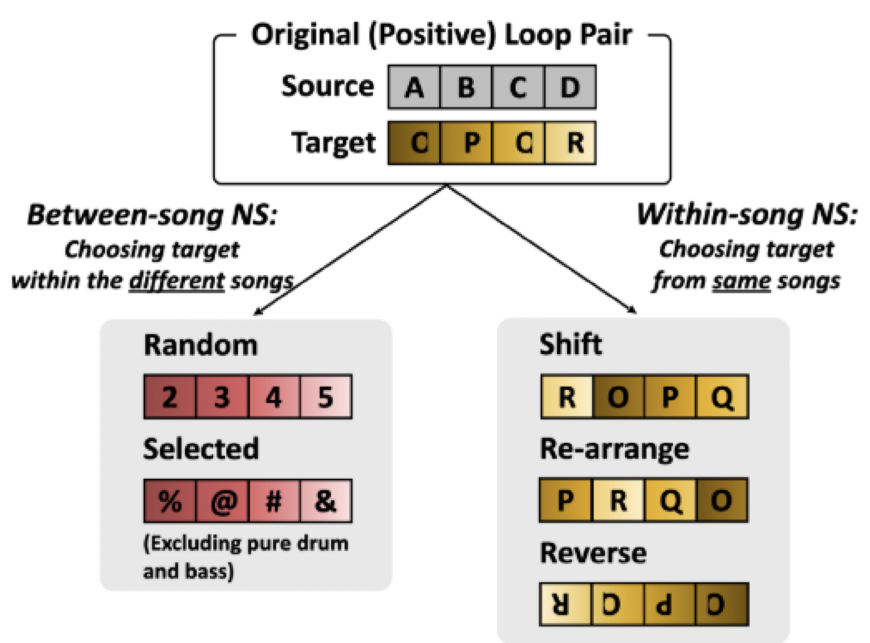

Fig. 2. Illustration of five loop-pair 'negative sampling' strategies

Models (Fig. 3)

- Train a CNN model to distinguish whether the loop combination is compatible
- Train a Siamese NN model to make the positive pair closer and push the negative pair far away in the embedding space

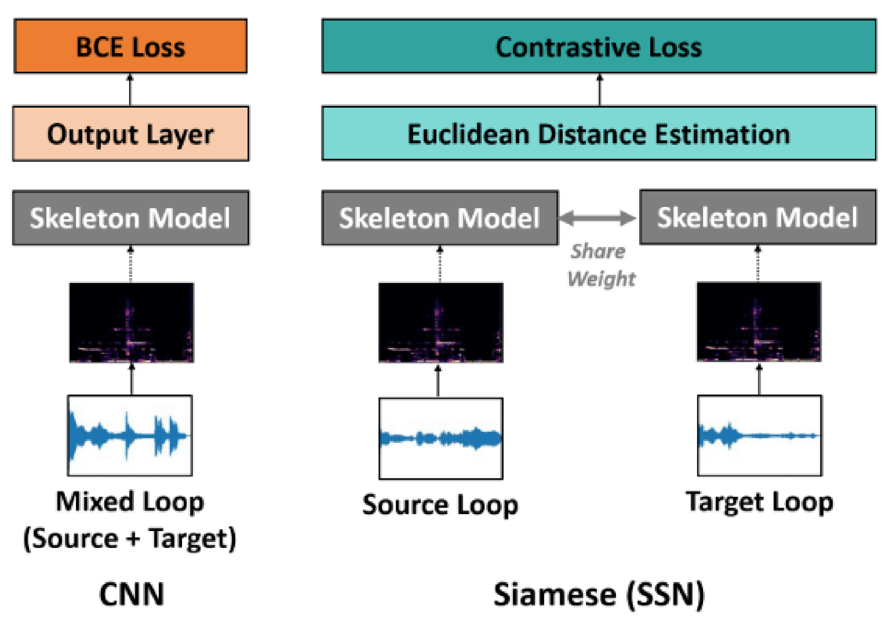

Fig. 3. Architectures of the CNN and SNN models

Datasets

Free Music Archive (FMA) datasets [1]

- Genre: Hip-Hop only

Data type	# loops	# loop pairs	# songs
Training set	9,048	12,774	2,702
Validation set	2,355	3,195	7,06
Test set	200	100	100
$\sum$	11,603	16,069	3,508

Table. 1 Statistics of the dataset

Results

Objective Evaluation (Table. 2)

- CNN works better for the classification-based metrics and Siamese NN works better for the ranking-based metrics
- CNN with reverse negative sampling performs the best in classification-based metrics
- Siamese NN with random sampling performs the best in ranking-based metrics

Model	Negative sampling	Classification-based metric: Accuracy	Classification-based metric: F1 score	Ranking-based metric: Avg. rank	Ranking-based metric: Top 10	Ranking-based metric: Top 30	Ranking-based metric: Top 50
CNN	Random	0.60	0.59	43.0	0.13	0.35	0.59
	Selected	0.59	0.59	43.1	0.13	0.29	0.62
	Reverse	**0.63**	**0.62**	41.2	0.19	0.42	0.62
	Shift	0.57	0.56	49.0	0.11	0.34	0.54
	Rearrange	0.57	0.57	47.7	0.10	0.31	0.57
Siamese NN	Random	0.51	0.47	**34.2**	**0.27**	**0.52**	**0.74**
	Selected	0.52	0.47	42.8	0.18	0.39	0.59
	Reverse	0.53	0.48	42.7	0.16	0.37	0.62
	Shift	0.53	0.52	43.0	0.16	0.41	0.65
	Rearrange	0.53	0.53	44.2	0.16	0.40	0.60

Table. 2 Objective results of different combinations of models

Subjective Evaluation (Fig. 4)

- CNN with reverse negative sampling outperforms the than other models
- Both CNN and Siamese NN outperform AutoMashUpper [4], the rule-based baseline

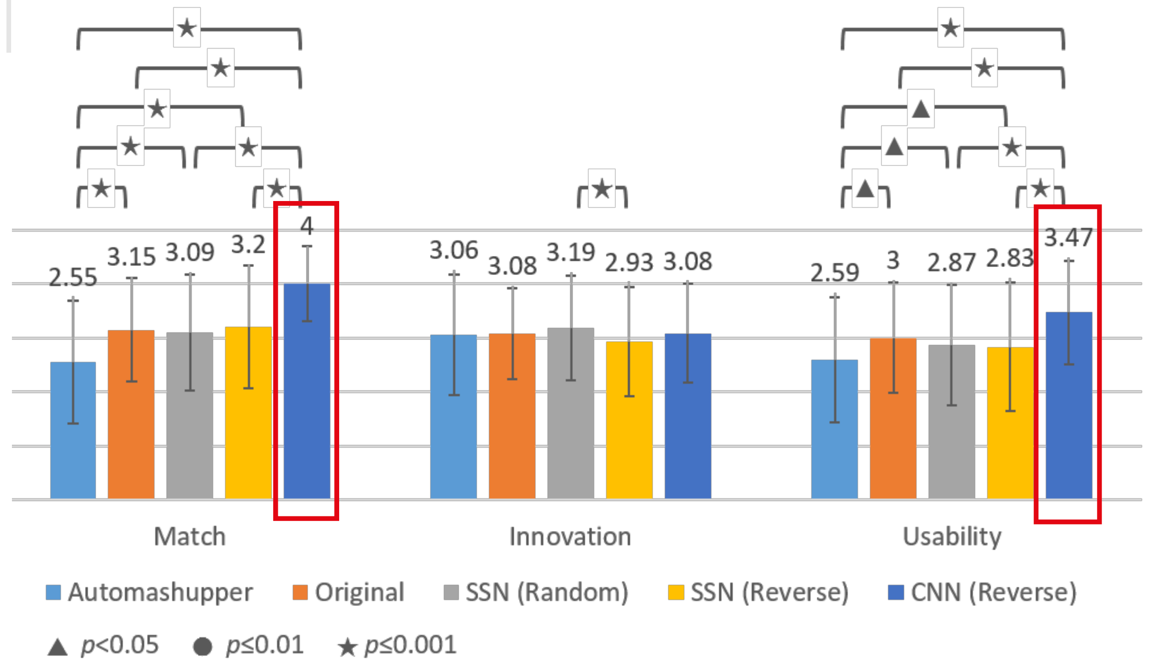

Fig. 4. The subjective evaluation results of comparing the preference among 5 models

Conclusions and Future Work

- Subjective evaluation suggests that our proposed models outperform the rule-based system AutoMashUpper [4], we therefore conclude our proposed system is effective
- We plan to investigate other objective metrics for performance evaluation and explore the relationship between loops and their arrangement by estimated layout from loop extraction algorithms [2, 3]

Reference

[1] M. Defferrard, K. Benzi, P. Vandergheynst, and X. Bresson, "FMA: A dataset for music analysis," in Proc. Int. Soc. Music Information Retrieval Conf., 2017.

[2] J. B. L. Smith and M. Goto, "Nonnegative tensor factorization for source separation of loops in audio," in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing., 2018.

[3] J. B. L. Smith, Y. Kawasaki, and M. Goto, "Unmixer: An interface for extracting and remixing loops," in Proc. Int. Soc. Music Information Retrieval Conf., 2019.

[4] M. E. P. Davies, P. Hamel, K. Yoshii, and M. Goto, "AutoMashUpper: Automatic creation of multi-song music mashups," IEEE/ACM Trans. Audio, Speech, and Language Processing, vol. 22, no. 12, p. 1726–17370, 2014.