

Hyper Recurrent Neural Network: Condition Mechanisms for Black-box Audio Effect Modeling

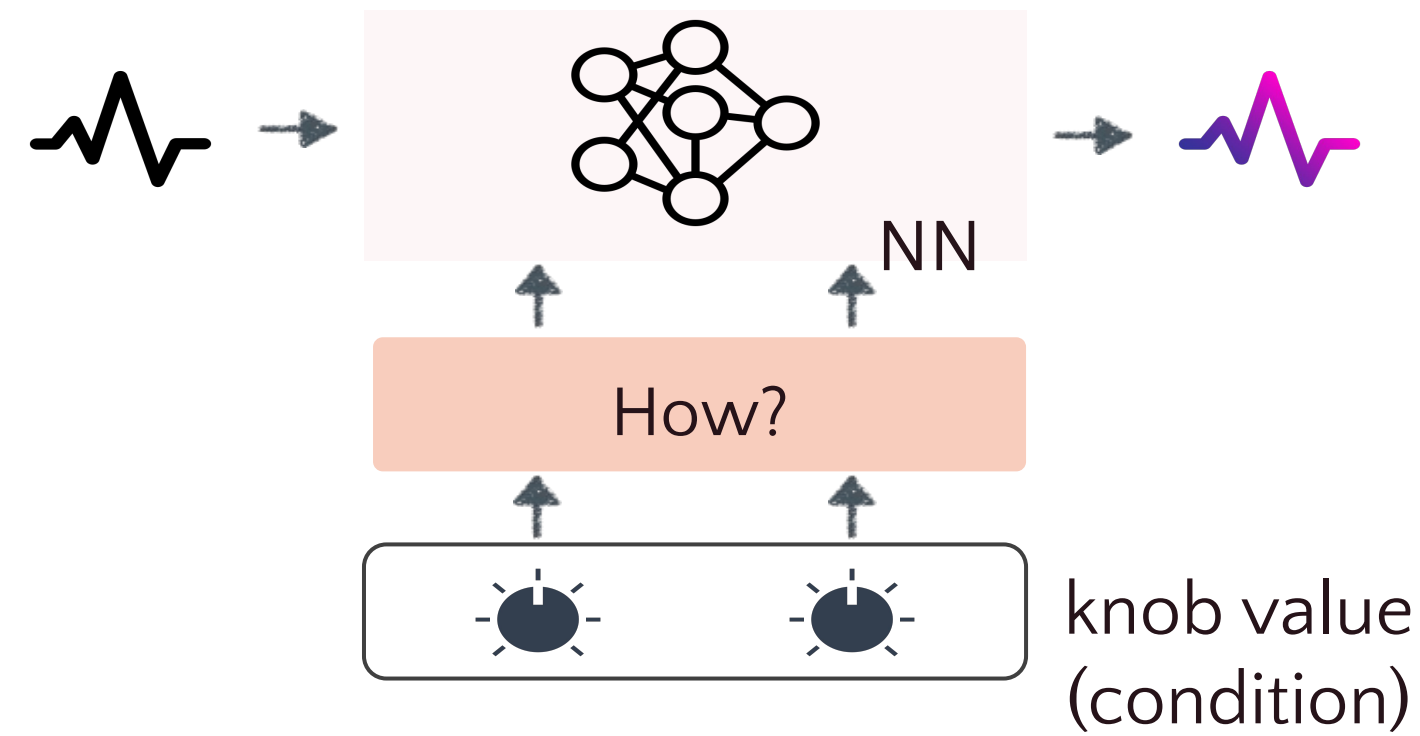
Yen-Tung Yeh¹, Wen-Yi Hsiao², Yi-Hsuan Yang¹

¹National Taiwan University, ²Independent Researcher



Overview

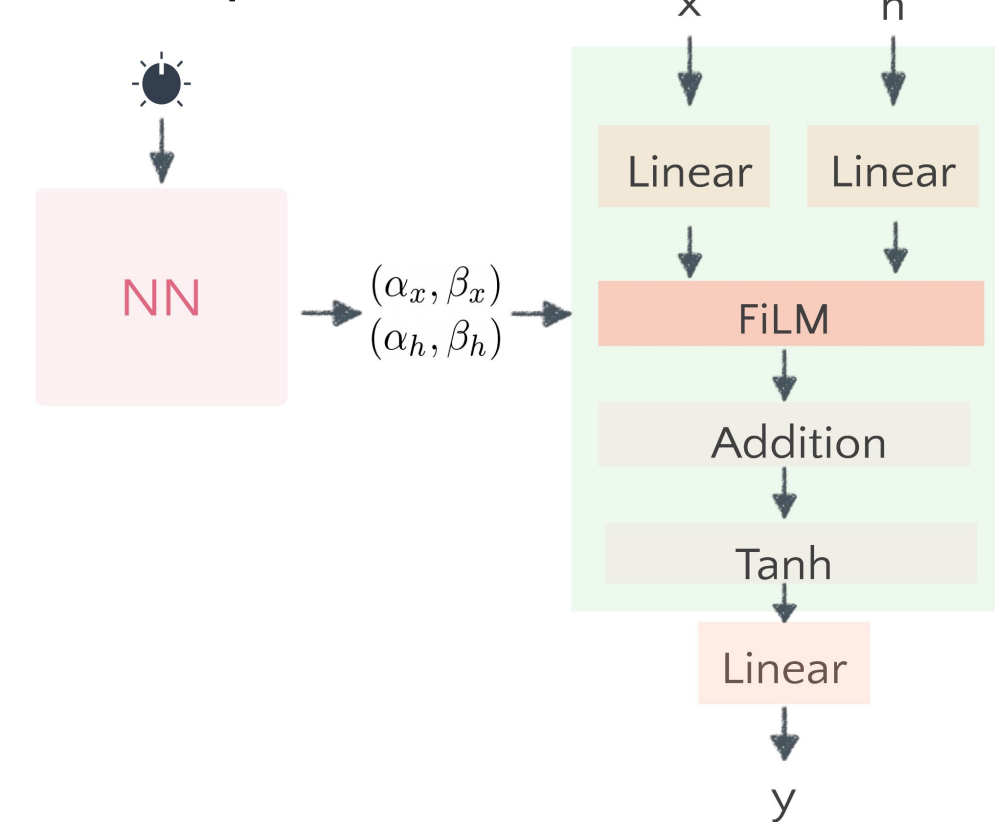
We propose novel condition mechanisms for RNN-based black-box audio effect modeling



Method

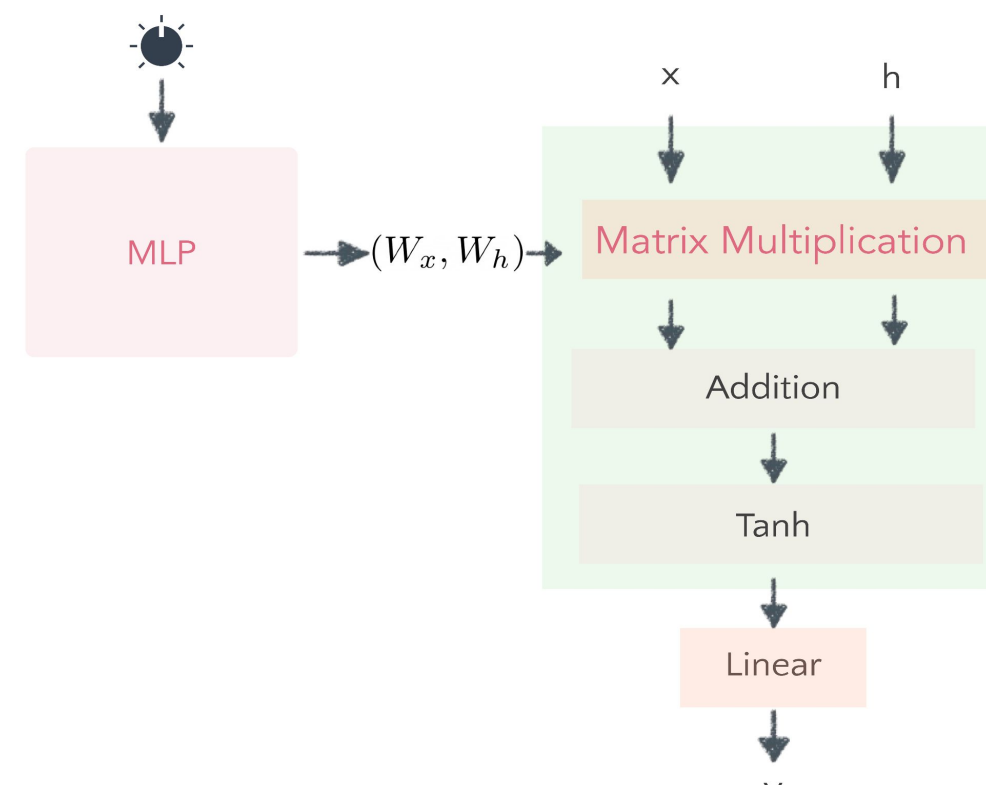
FiLM-RNN

Inject conditioning value via element-wise linear multiplication



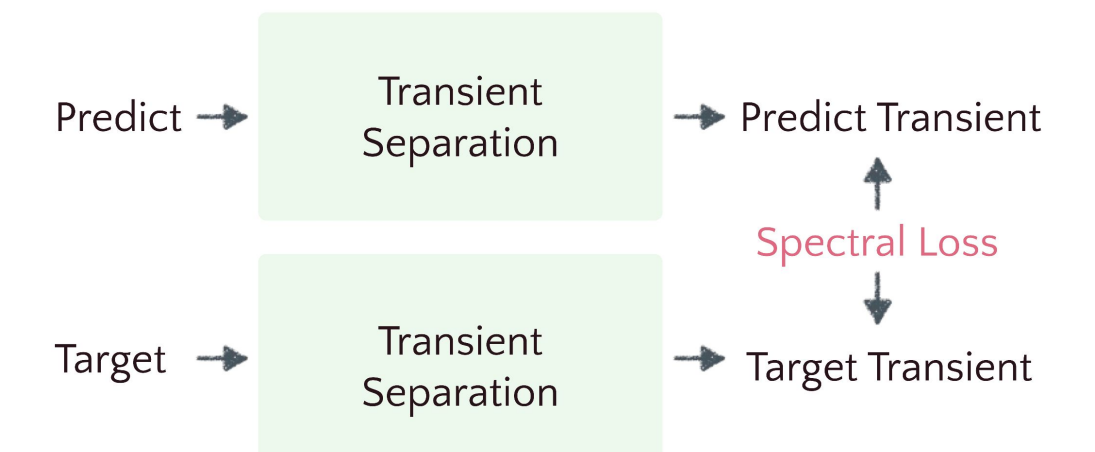
StaticHyper-RNN

Directly determine the RNN's weights based on the conditioning value



Transient Metric

We measure reconstruction of transient via TMS algorithm (Transient Modeling Synthesis)



Results

Boss OD-3 (Overdrive)

Model	Condition	OD-3 (Overdrive)						Params
		L1	STFT	LUFS	CF	RMS	Transient	
GRU	Concat	0.120	1.933	0.455	2.932	1.096	27.338	3585
	FiLM	0.011*	0.536†	0.176	0.676*	0.401	12.504	17217
	StaticHyper	0.017	0.698	0.165	1.650	0.318†	12.347†	30369
	DynamicHyper	0.150	0.428*	0.075*	0.883	0.153*	11.308*	20289

LA2A Compressor

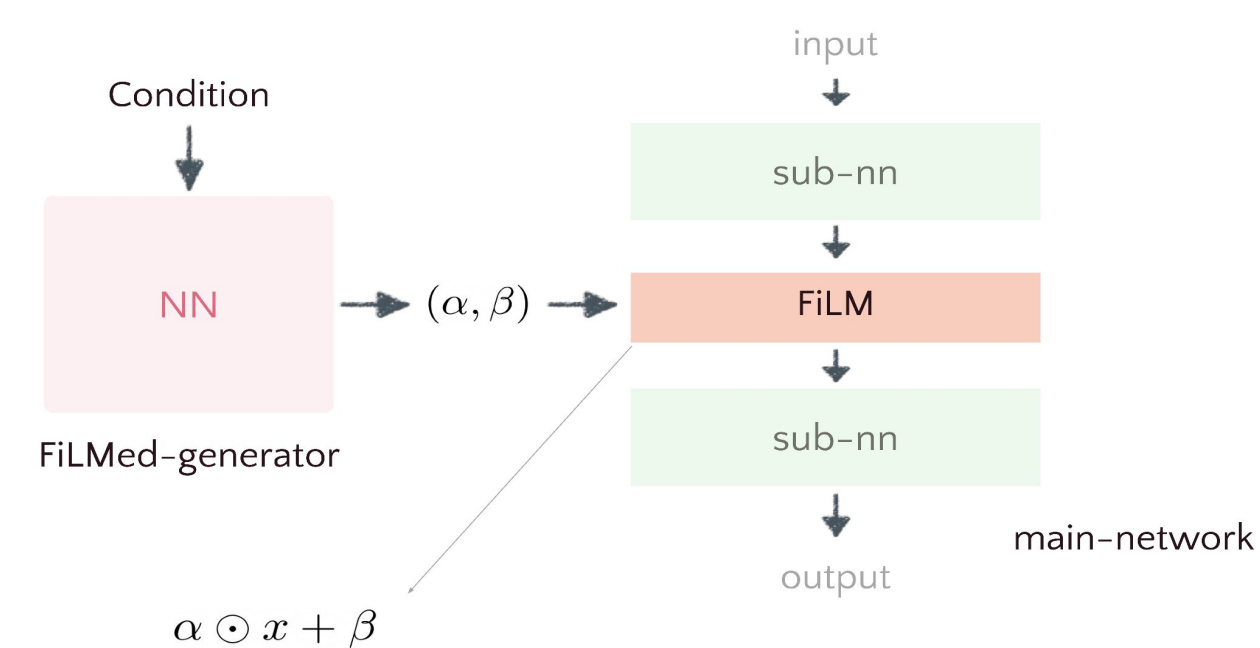
Model	Condition	LA2A (Compressor)						Params
		L1	STFT	LUFS	CF	RMS	Transient	
GRU	Concat	0.108	0.507	0.716	2.081	1.640	21.002	3489
	FiLM	0.011†	0.597	1.383	2.006†	3.081	15.825*	17185
	StaticHyper	0.008*	0.371*	0.543	2.386	1.211	20.437	30361
	DynamicHyper	0.109	0.377†	0.377†	1.919*	0.819†	19.826†	20169

Computation

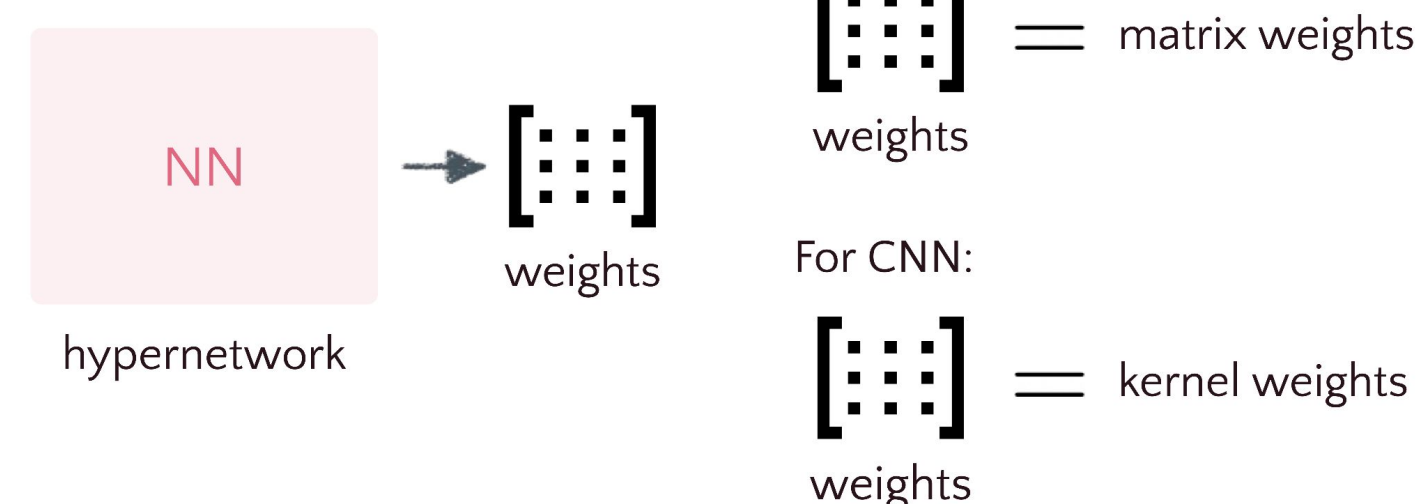
Models	GFLOPs
Concat-GRU	0.325
FiLM-GRU	0.307
StaticHyper-GRU	0.003
DynamicHyper-GRU	1.907

Background

FiLM

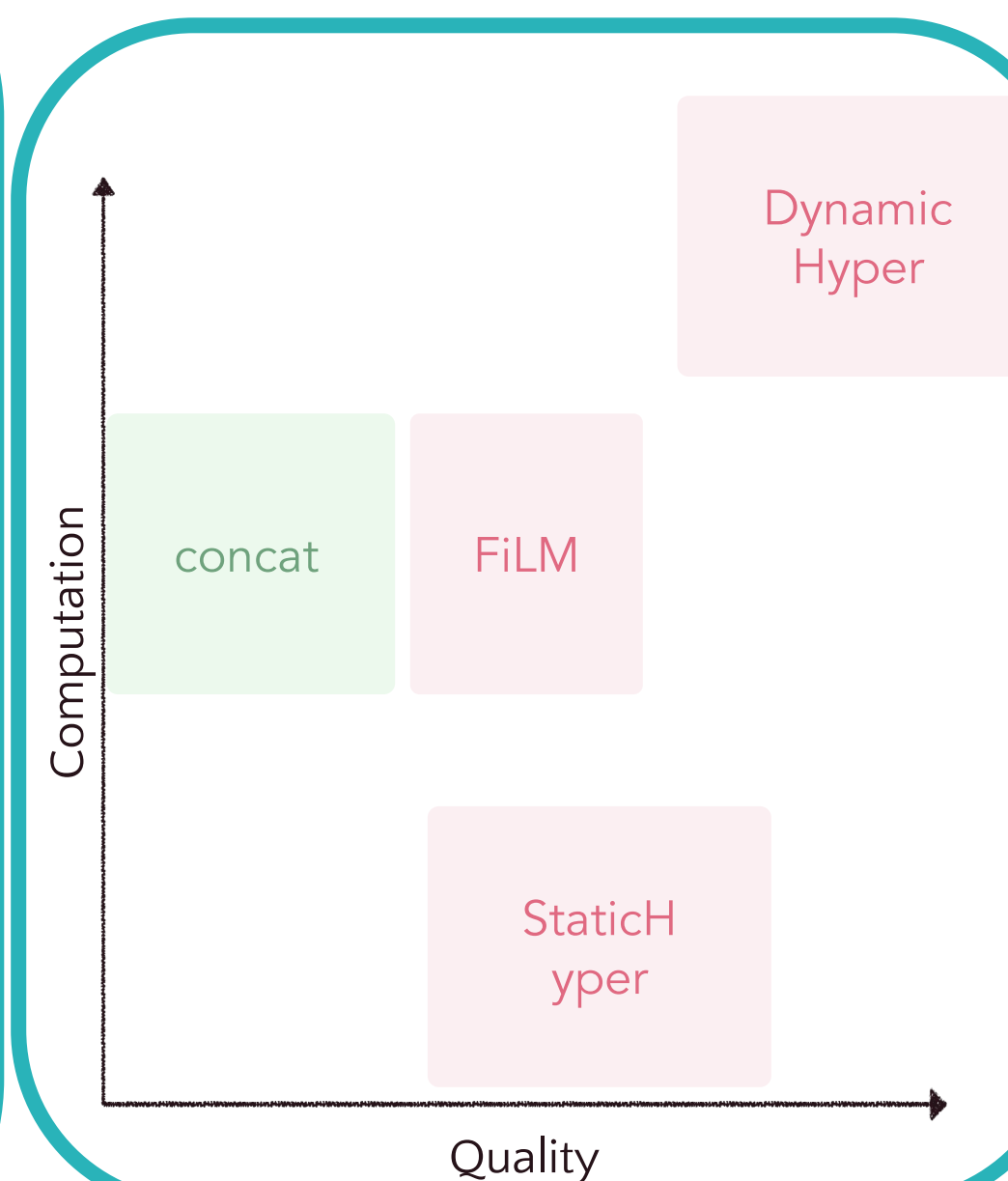
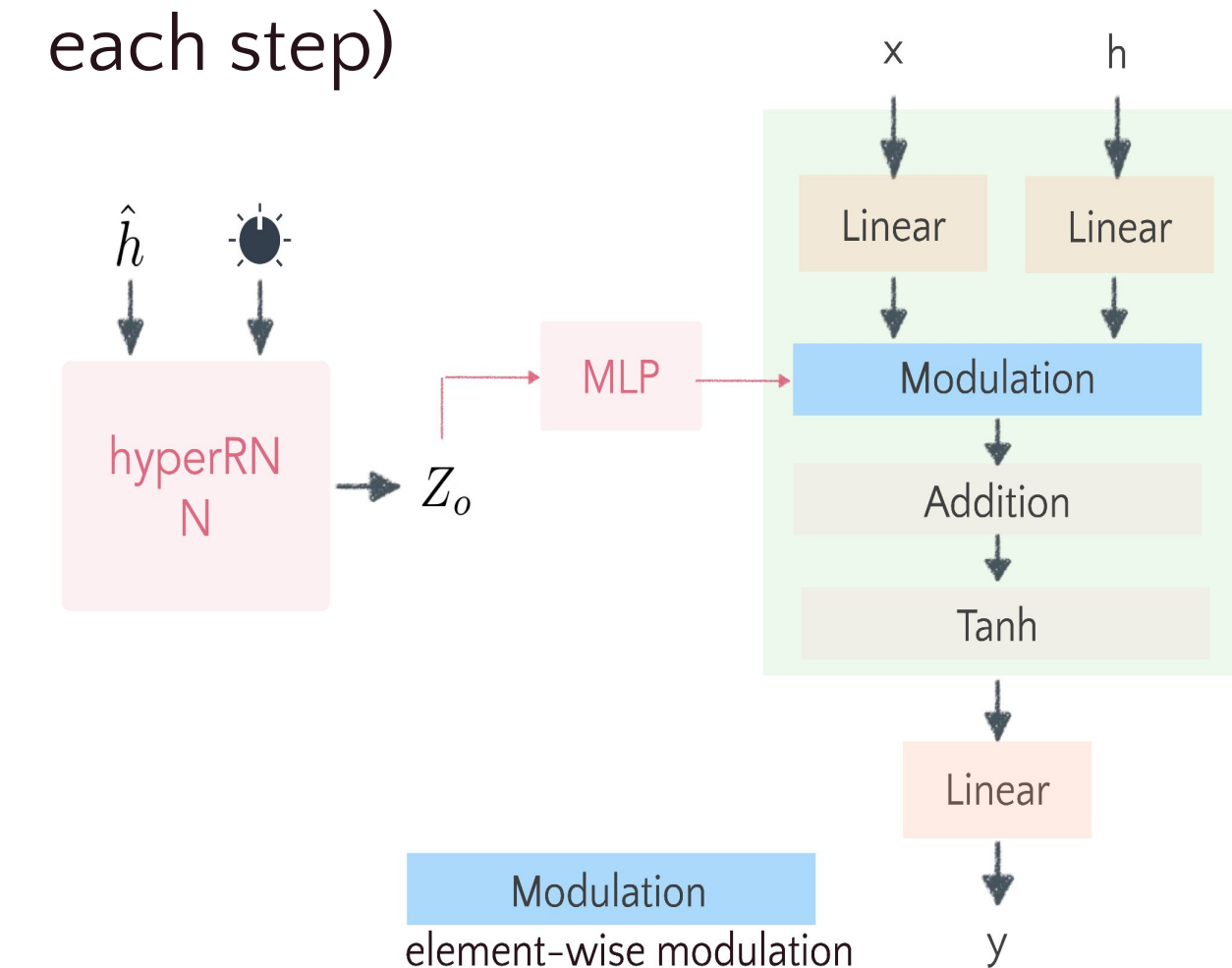


HyperNetwork



DynamicHyper-RNN

Modulate the RNN's weights for each audio sample (using different weights at each step)



Conclusion & Future work

- Advanced conditioning mechanism can improve the emulation quality
- Real-time plugin implementation is needed
- Transient metric can be improved
- Time-varying effect is not included in our experiments

code: <https://github.com/ytsrt66589/pyneuralfx>

paper: <https://arxiv.org/abs/2408.04829>

demo: <https://yytung.notion.site/HyperRNN>

