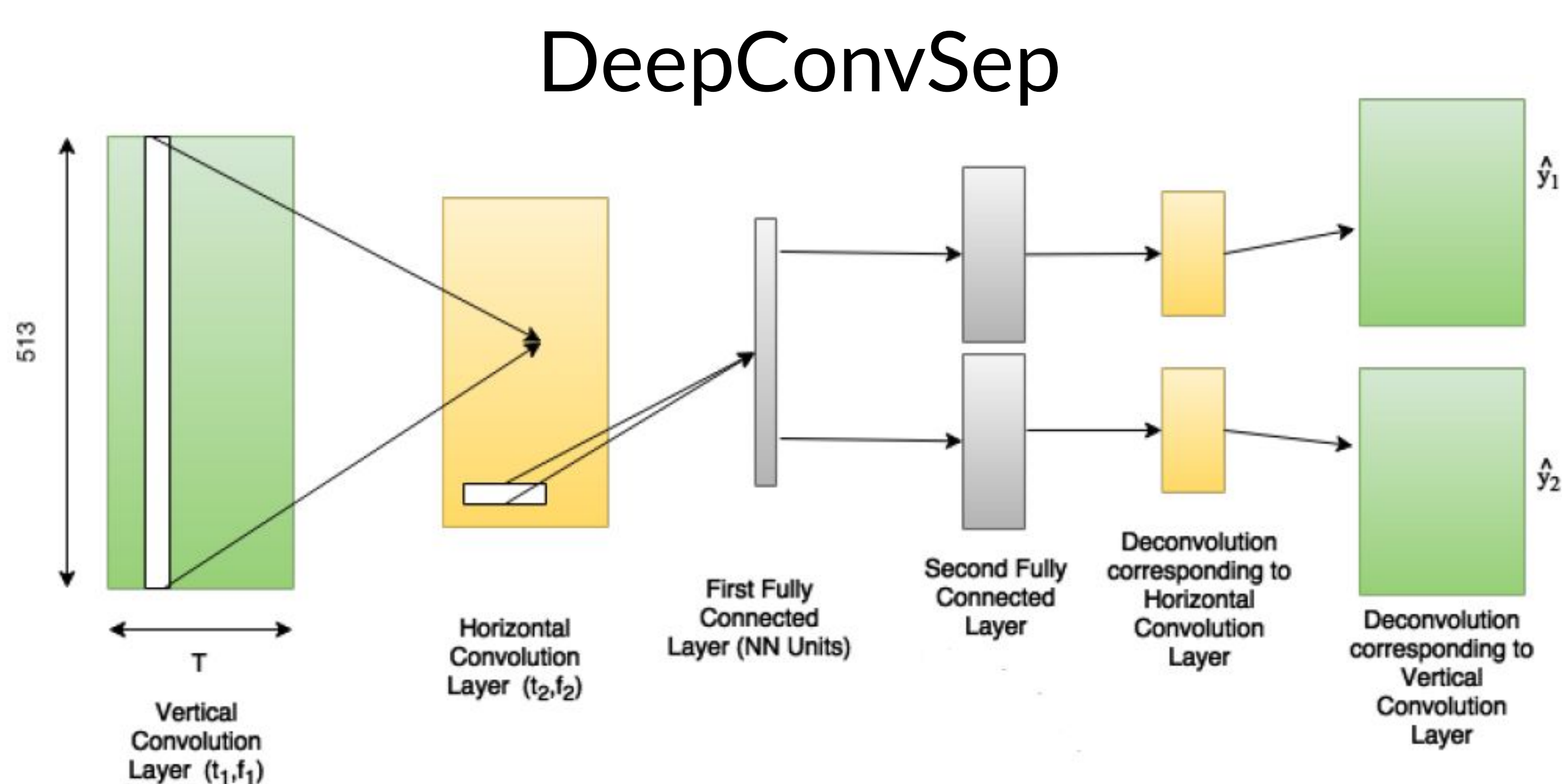
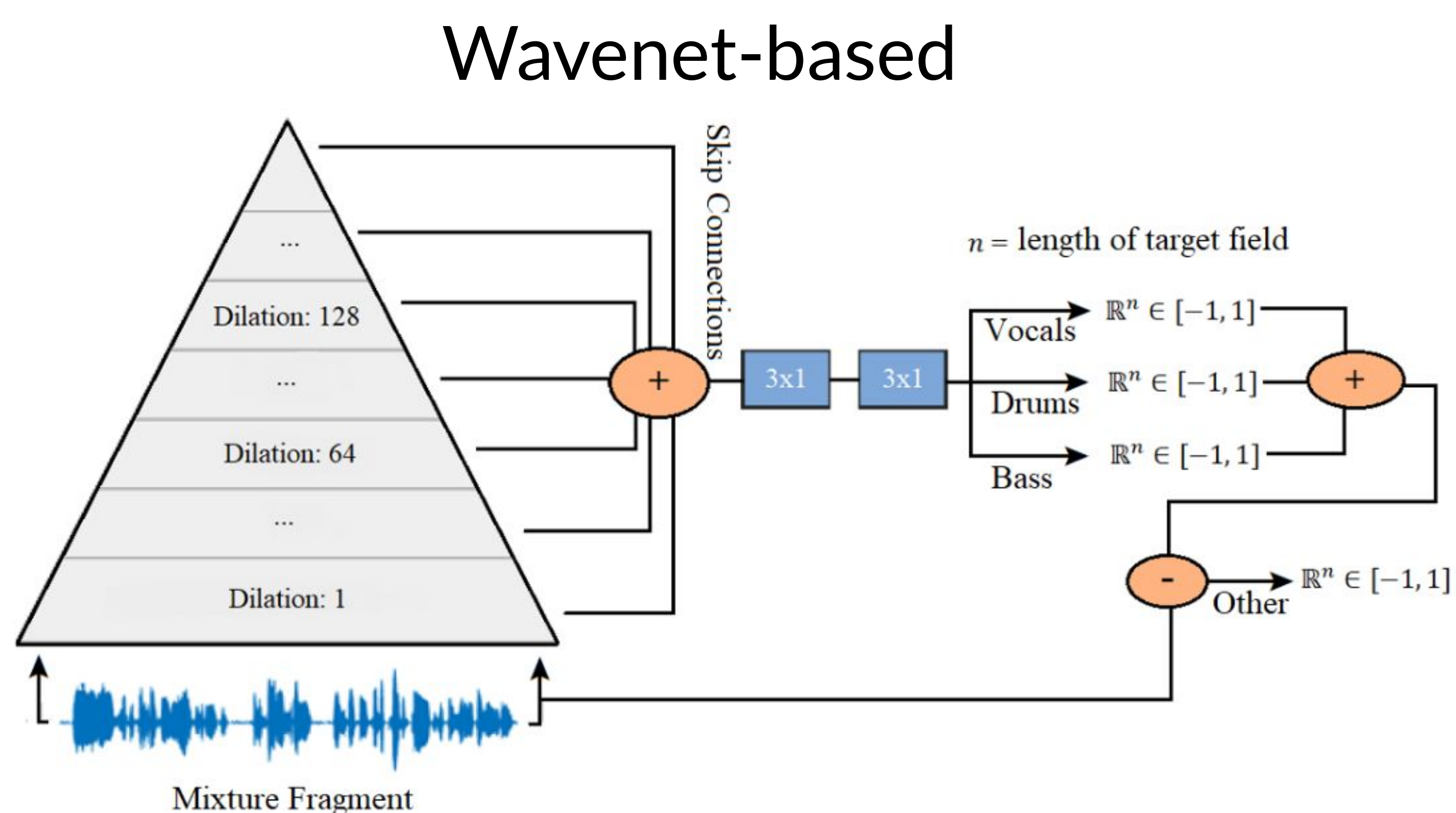


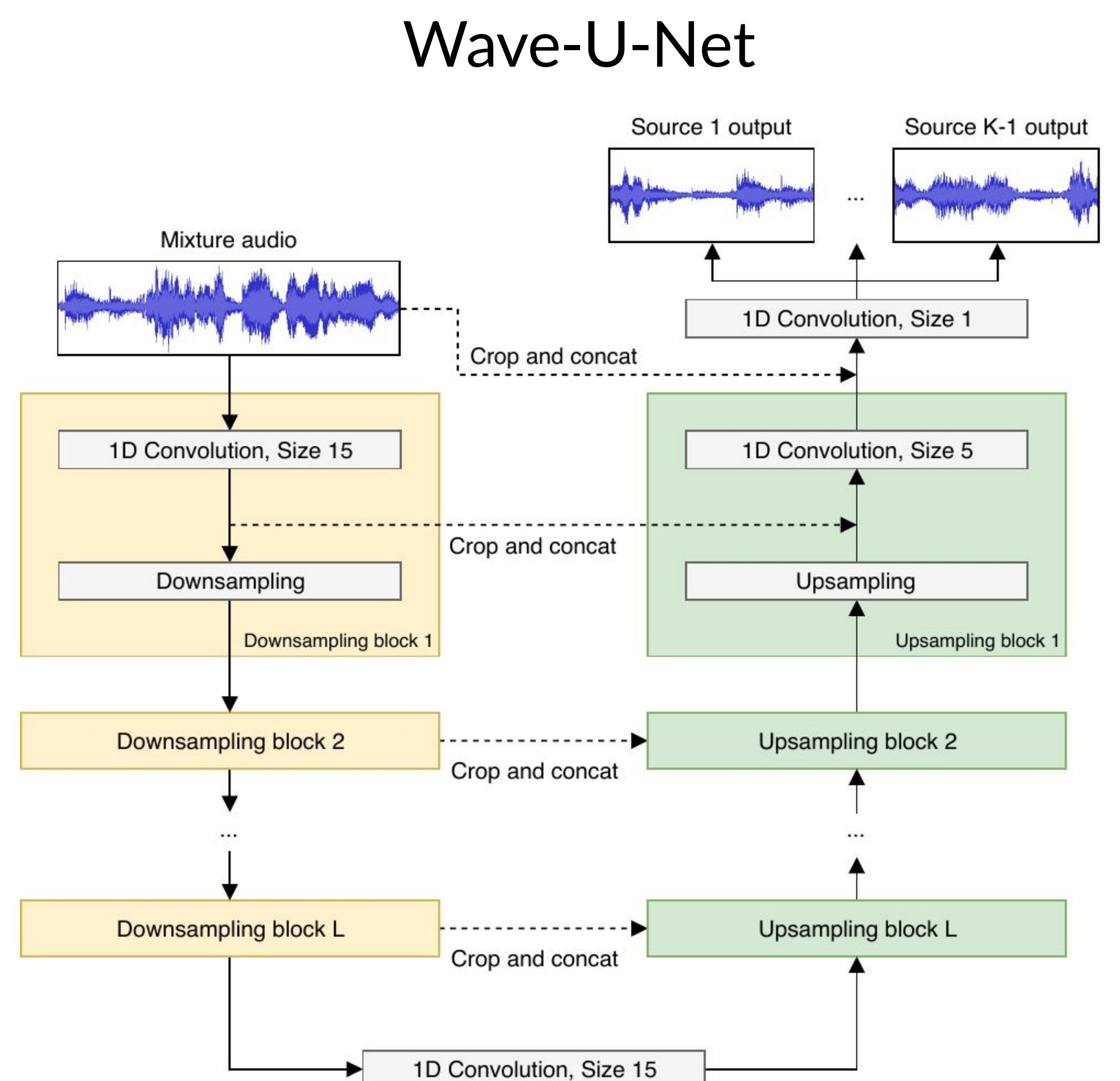
END-TO-END MUSIC SOURCE SEPARATION: IS IT POSSIBLE IN THE WAVEFORM DOMAIN?

YES!

Francesc Lluís, Jordi Pons,
and Xavier Serra



(Chandna et al., 2017)



(Stoller et al., 2018)

Multi-Instrument Source Separation

<i>BSS_eval</i>	Vocals			Drums			Bass		
	SDR	SIR	SAR	SDR	SIR	SAR	SDR	SIR	SAR
Wavenet	3.35	11.25	5.24	4.13	13.23	5.00	2.49	6.53	5.77
DeepConv	2.38	4.45	8.39	3.19	6.69	6.58	0.27	1.92	7.46

<i>MOS</i>	Vocals	Drums	Bass
Wavenet	2.4±0.9	2.9±1.1	2.4±1.0
DeepConv	2.3±0.9	2.5±0.7	1.8±0.8
p-value	0.423	0.018	<0.001

Singing-Voice Source Separation

<i>BSS_eval</i>	Vocals			Accompaniment		
	SDR	SIR	SAR	SDR	SIR	SAR
Wavenet	3.67	12.14	5.24	10.64	14.43	13.22
Wavenet+50%	4.49	13.52	6.17	11.39	16.37	13.49
Wave-U-Net	4.60	14.30	5.54	11.87	16.08	14.20

<i>MOS</i>	Vocals
Wavenet+50%	3.0±1.0
Wave-U-Net	3.3±0.85
p-value	0.049

*50%: percentage of forced fragments containing singing voice during training



Listen to **examples!**
Link to the **paper & code**



Universitat
Pompeu Fabra
Barcelona

MTG
Music Technology
Group

