



Thoughts about WaveNet

(Oct 19, 2016)

YANG Jiancheng



Outline

- **I. WaveNet**
- **II. Thoughts**



- **I. WaveNet**

- **Amazing Effective**

- a) **Multi-speaker Speech Generation**
- b) **Text-To-Speech (TTS)**
- c) **Music**
- d) **Speech Recognition**

<https://deepmind.com/blog/wavenet-generative-model-raw-audio/>



• I. WaveNet

- Raw Audio



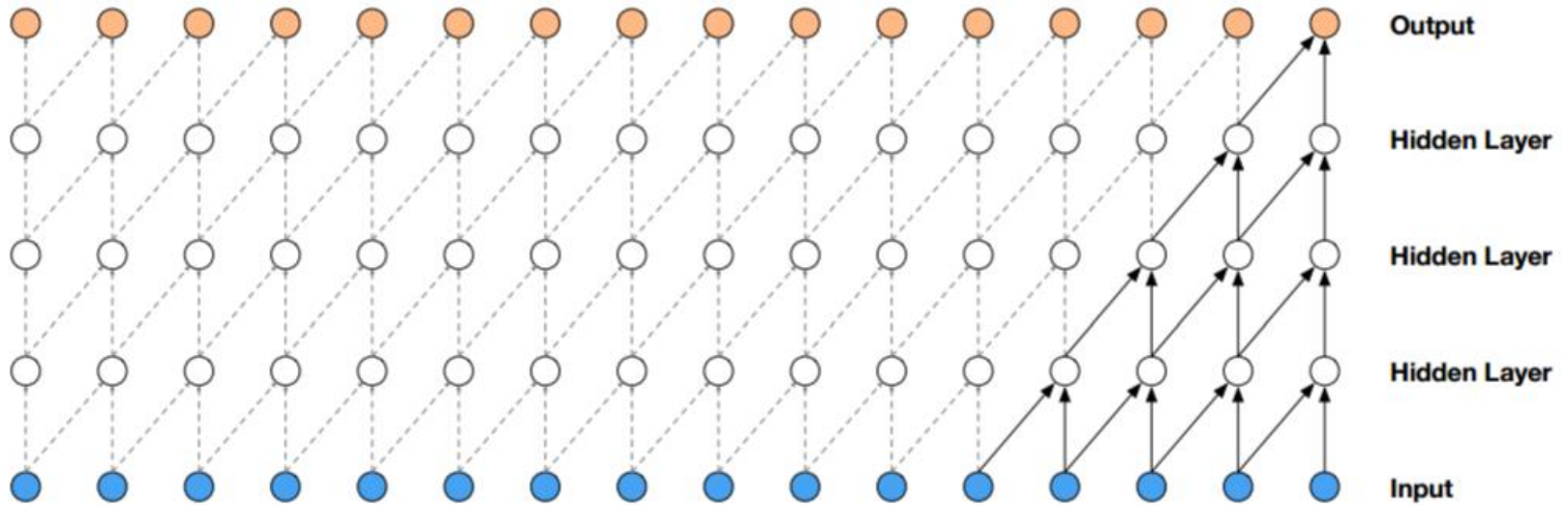
1 Second





• I. WaveNet

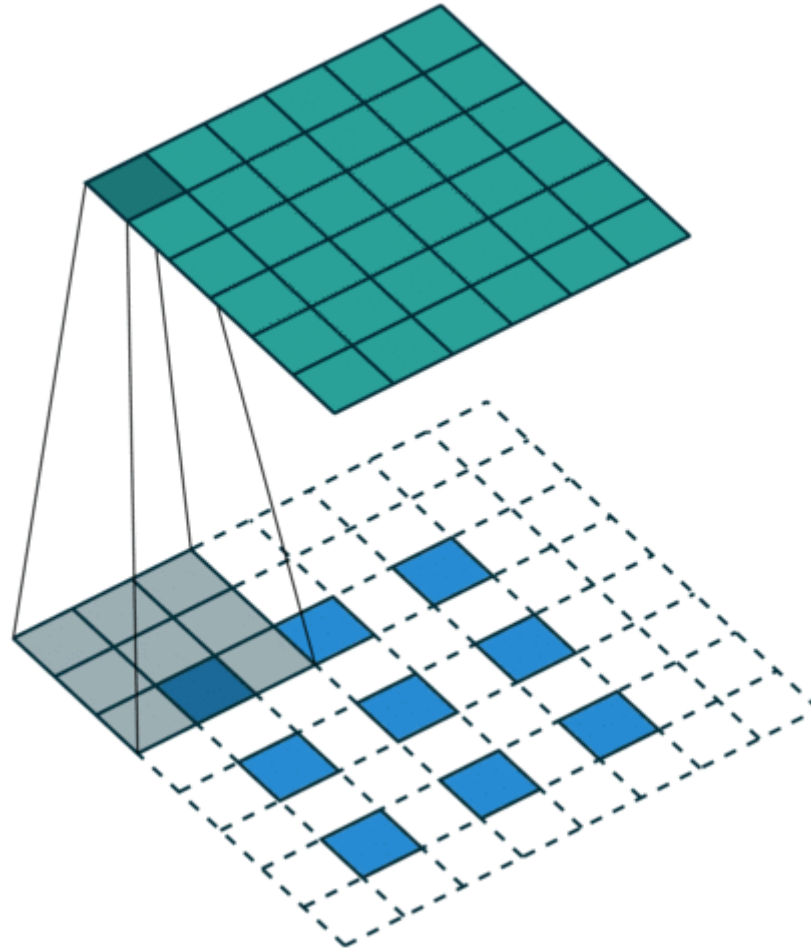
- Naïve Version





• I. WaveNet

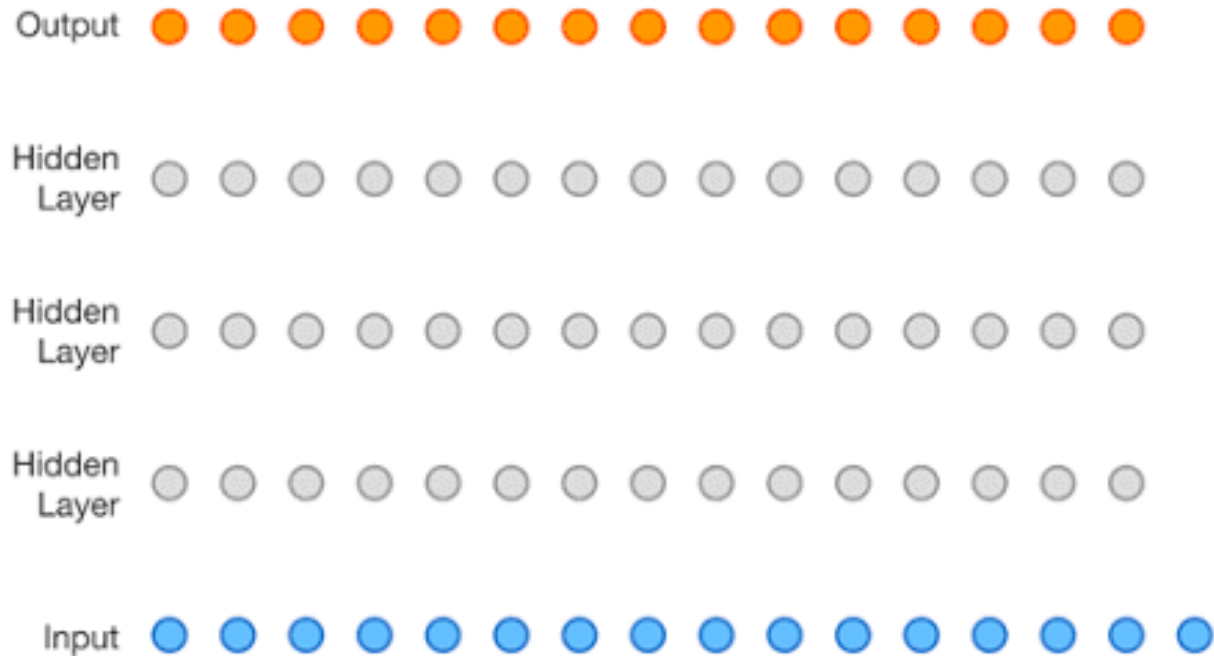
- Atrous (Dilated) Convolution





• I. WaveNet

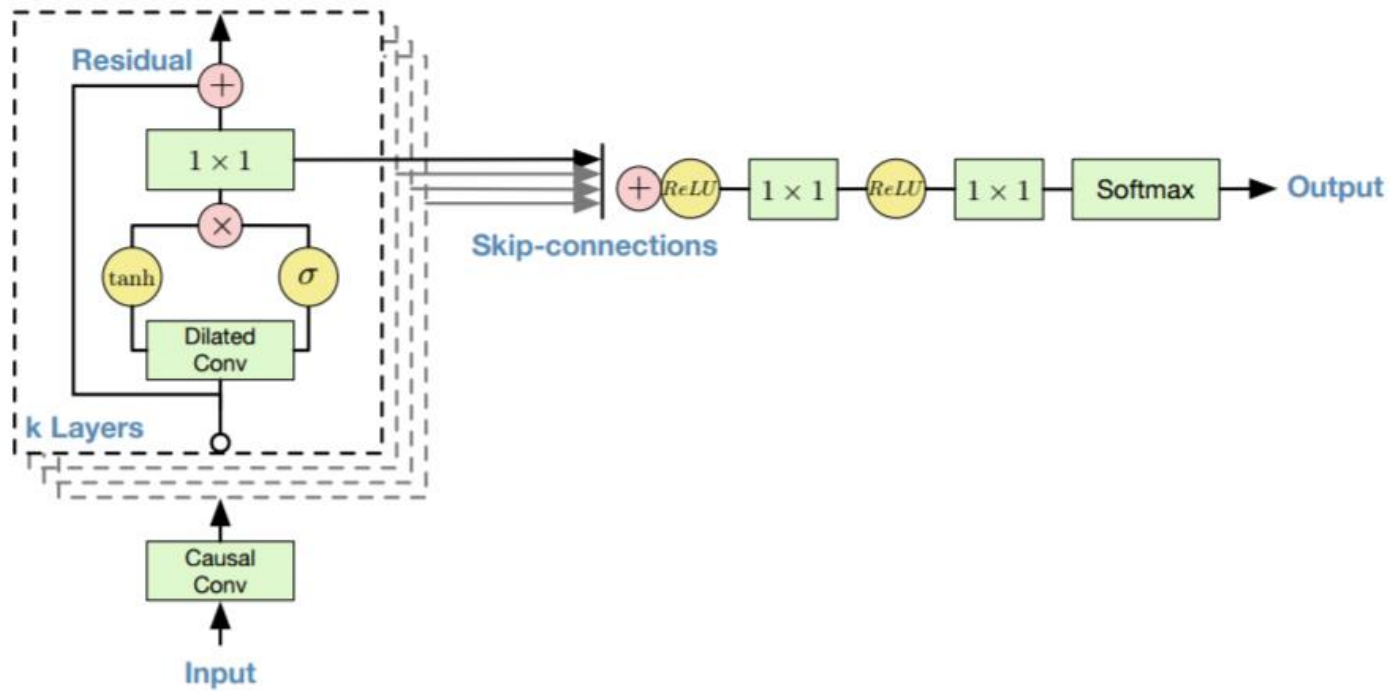
- Dilated Causal Convolutions





• I. WaveNet

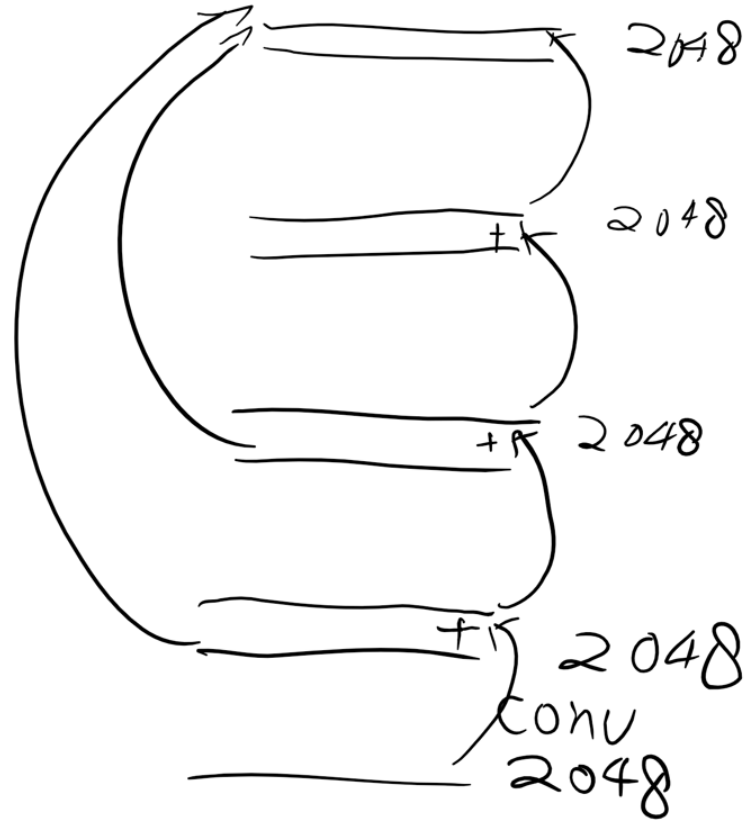
- Structure and tricks





• I. WaveNet

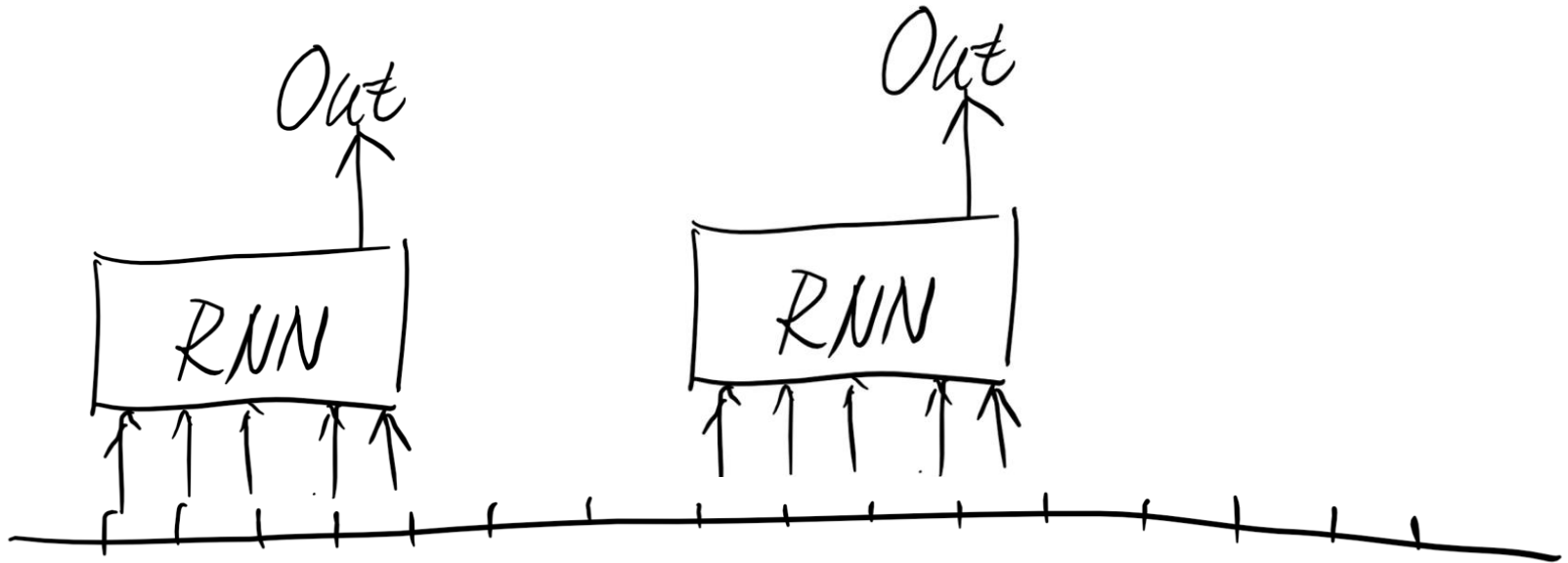
- Structure and tricks





• II. Thoughts

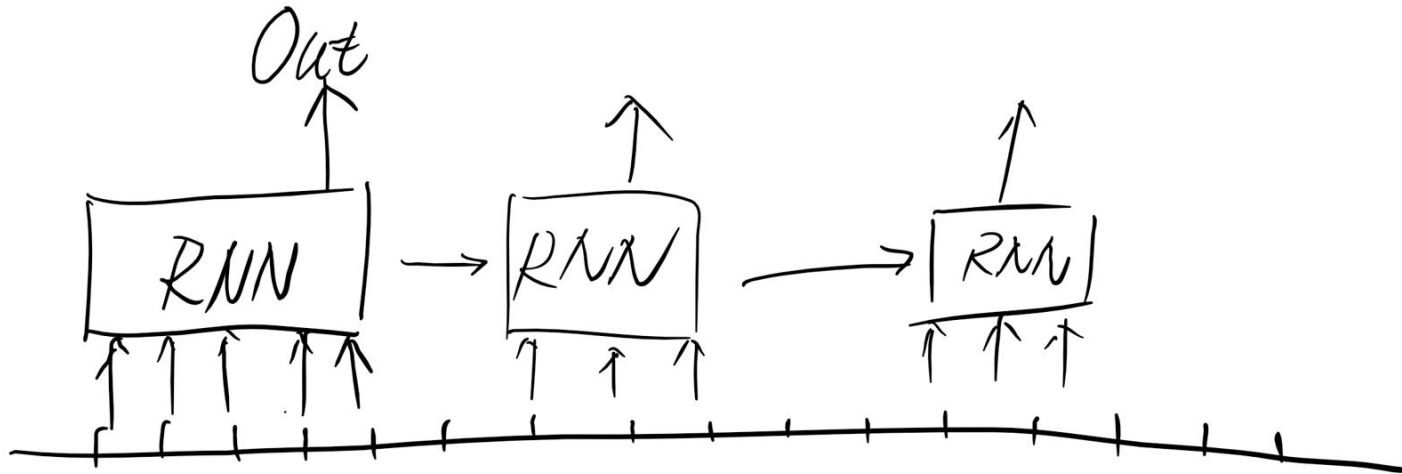
- RNN Conv Kernel (Multiple Scale)





• II. Thoughts

- RNN in RNN (Multiple Scale)





Bibliography

- [WaveNet: A Generative Model for Raw Audio](#)



Thanks for listening!

