

Voicexer

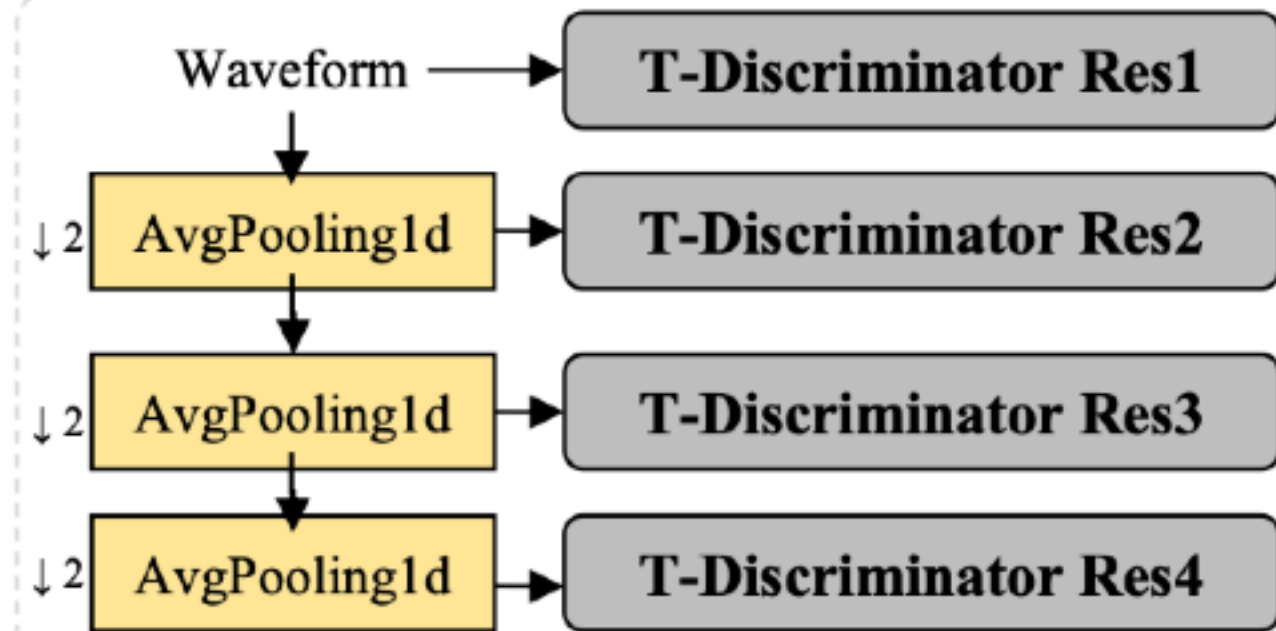
SynthesisModule-TFGANVideo-Traing

- The loss functions for generator include time domain losses L^T , frequency domain losses L^T , and discriminator losses L^D .

- Loss Function: $L_{syn} = L^T + L^F + \lambda_1 L^D$

Frequency Discriminator

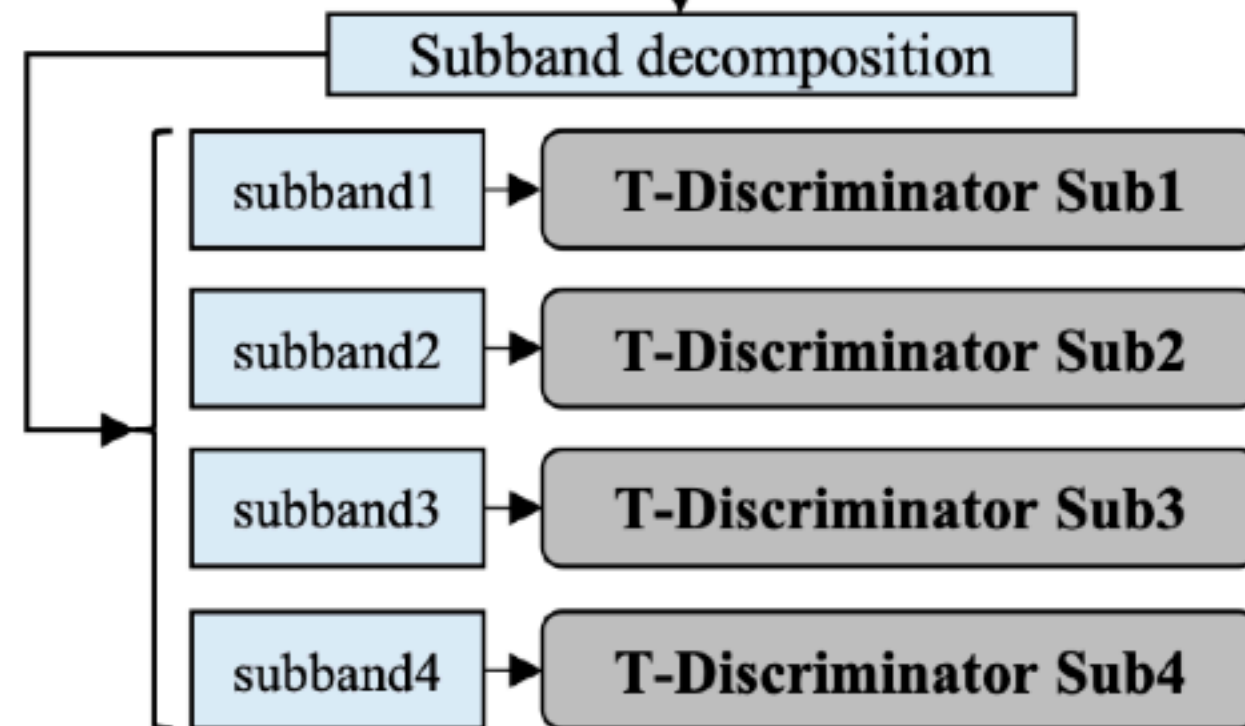
F-Discriminator



Multi-resolution discriminator

Loss functions and discriminators

Waveform



Subband discriminator

Energy Loss

Segment Loss

Phase Loss

Time
domain

Loss functions

Mel Loss

**Multi
Resolution
STFT Loss**

**Spectral
Convergence
Loss**

Frequency
domain

, and discriminator losses



osseas

The loss function for generating the main loss

Loss Function:



frequency domain

LT

LT

LD

$$\boldsymbol{L}_{syn} = \boldsymbol{L}^T + \boldsymbol{L}^F + \lambda_1 \boldsymbol{L}^D$$

TFGAN Vocoder - Training - Time Domain Losses

k	1	2	3	4
frame-length	1	240	480	960
hop-length	1	120	240	480

Table.3 Window parameter for each k