

改动了三处：在第一个卷积之前加入DTCWT小波变换，在Dilated-Dense Block（四层dilation convolution）加入残差，之后在第二个卷积加入前馈。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Domain | PESQ | STOI | SSNR | CSIG | CBAK | COVL | Para(maillion |
| TSTNN（改进前的模型） | T | *2.96* | *95* | 9.70 | 4.33 | 3.53 | 3.67 | 0.92 |
| Ours（小波+前馈+残差） |  | 2.96 | 94 | 5.47 | 4.33 | 3.27 | 3.67 | 1..45 |
| 去掉小波 |  | 2.92 | 94 | 5.37 | 4.27 | 3.27 | 3.61 | 1.45 |
| 去掉前馈 |  | 3.02 | 95 | 9.53 | 4.33 | 3.55 | 3.68 | 0.92 |
| 去掉残差 |  | 2.89 | 93 |  | 4.24 | 3.18 | 3.58 | 1.45 |
| 只有小波 |  |  |  |  |  |  |  |  |
| 只有残差 |  |  |  |  |  |  |  |  |

跑了五个消融实验，最好的效果是小波加上残差，由于只加上小波和只有加上残差的实验是在太仓服务器那边跑的，还没有跑测试集，但是就训练集的效果可以明显得到残差加上小波（也就是只去掉前馈网络）的模型效果最好，并且其中PESQ、CBAK、COVL的实验结果要优于改进前的模型。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Domain | PESQ | STOI | SSNR | CSIG | CBAK | COVL | Para(maillion |
| SEGAN | T | 1.97 | 91 | 1.73 | 3.48 | 2.94 | 2.80 | 43.2 |
| MMSEGAN |  | 2.53 | 93 |  | 3.8 | 3.12 | 3.14 |  |
| SERGAN |  | 2.16 | 93 |  | 3.48 | 2.94 | 2.80 | 43.2 |
| Wavenet | T |  |  |  |  |  |  |  |
| MetricGAN |  | 2.86 |  |  | 3.99 | 3.18 | 3.42 | 1.86 |
| CRGAN |  | 2.92 | 94 |  | 4.16 | 3.24 | 3.54 |  |
| DCCRN |  | 2.68 | 94 |  | 3.88 | 3.18 | 3.27 | 3.70 |
| PHASEN |  | 2.99 |  |  | 4.21 | 3.55 | 3.62 |  |
| MHSA-SPK |  | 2.99 |  |  | 4.21 | 3.15 | 3.42 |  |
| T-GSA |  | 3.06 | 94 |  | 4.18 | **3.59** | 3.62 |  |
| Wave U-Net | T | 2.40 |  | 9.97 |  |  |  | 10 |
| CTS-Net |  | 2.92 |  |  | 4.25 | 3.46 | 3.59 | 4.35 |
| GaGNet |  | 2.94 | 95 |  | 4.26 | 3.45 | 3.59 | 5.94 |
| DEMUCS (small) | T | 2.93 | 95 |  |  |  |  | 18.9 |
| DEMUCS (large) | T | **3.07** | 95 |  |  |  |  | 33.5 |
| TSTNN（改进前的模型） | T | *2.96* | *95* |  | 4.33 | 3.53 | 3.67 | 0.92 |
| 去掉前馈 |  | 3.02 | **95** |  | **4.33** | 3.55 | **3.68** | **0.92** |

我还找了一些在相同数据集上跑的模型进行对比，上表中的模型是从2017-2021年的一些在VoiceBank + DEMAND 数据集上的表现，如上表黑色字体模型所示，可以看到我们的模型要优于一些模型。