|  |  |  |  |
| --- | --- | --- | --- |
| 卷积块 | 单声道模型 | 双声道模型 | |
| 零填充（1\*1） | 音频输入 | 音频输入 | |
| 批标准化 | 卷积块（32） | 声道1 | 声道2 |
| ReLU | 最大池化（3\*3） | 单声道  模型 | 单声道  模型 |
| 卷积核（3\*3） | 卷积块（64） |
| 零填充（1\*1） | 最大池化（3\*3） | 连接 | |
| 批标准化 | 卷积块（128） | 全连接层（ReLU,1024） | |
| ReLU | 最大池化（3\*3） | 全连接层（softmax,15） | |
| 卷积核（3\*3） | 卷积块（256） |  | |
|  | 整体平均池化 |  | |
|  | 全连接层（ReLU,1024） |  | |
|  | 全连接层（softmax,15） |  | |