



Phonetic-based CNN for language identification

汇报人 齐诏娣

学校 江苏师范大学

导师 顾明亮教授

C 目录 ontents



1 语种识别简介

2 Phonetic特征

3 CNN模型

4 总结



1 语种识别简介

- 1.1 语种识别框架
- 1.2 常用特征
- 1.3 深度瓶颈特征提取
- 1.4 研究内容
-

1.1 语种识别系统框图

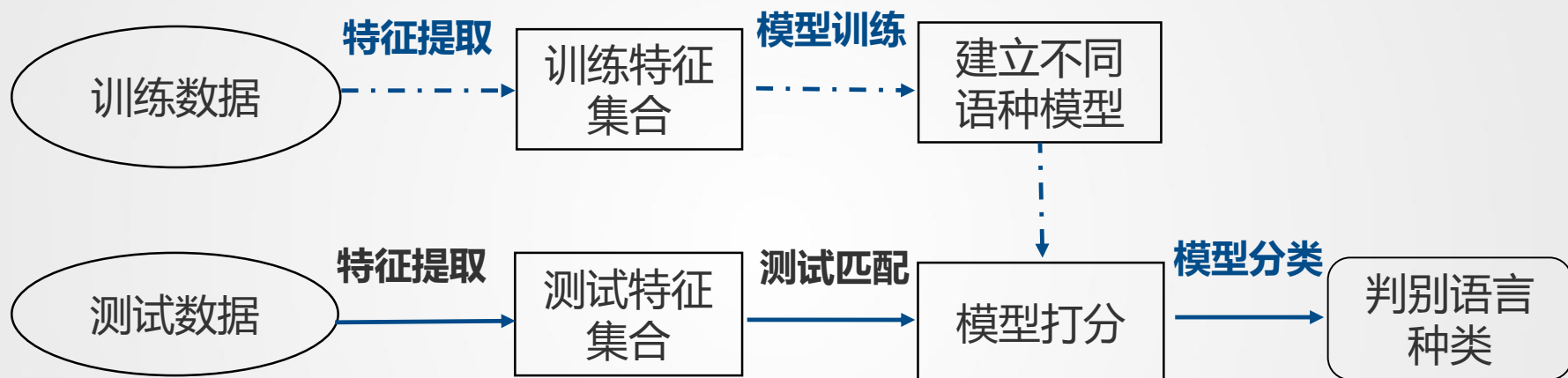


图1.1 语种识别系统框图



1.2 常用特征

特征提取层次(低 → 高)



2002,姜洪臣
2007,许慧燕

声学特征
(SDC,MFCC)

1997,蔡伟和
2007,Campbell

音素搭
配特征

2007,钱盛友
2008,李丹

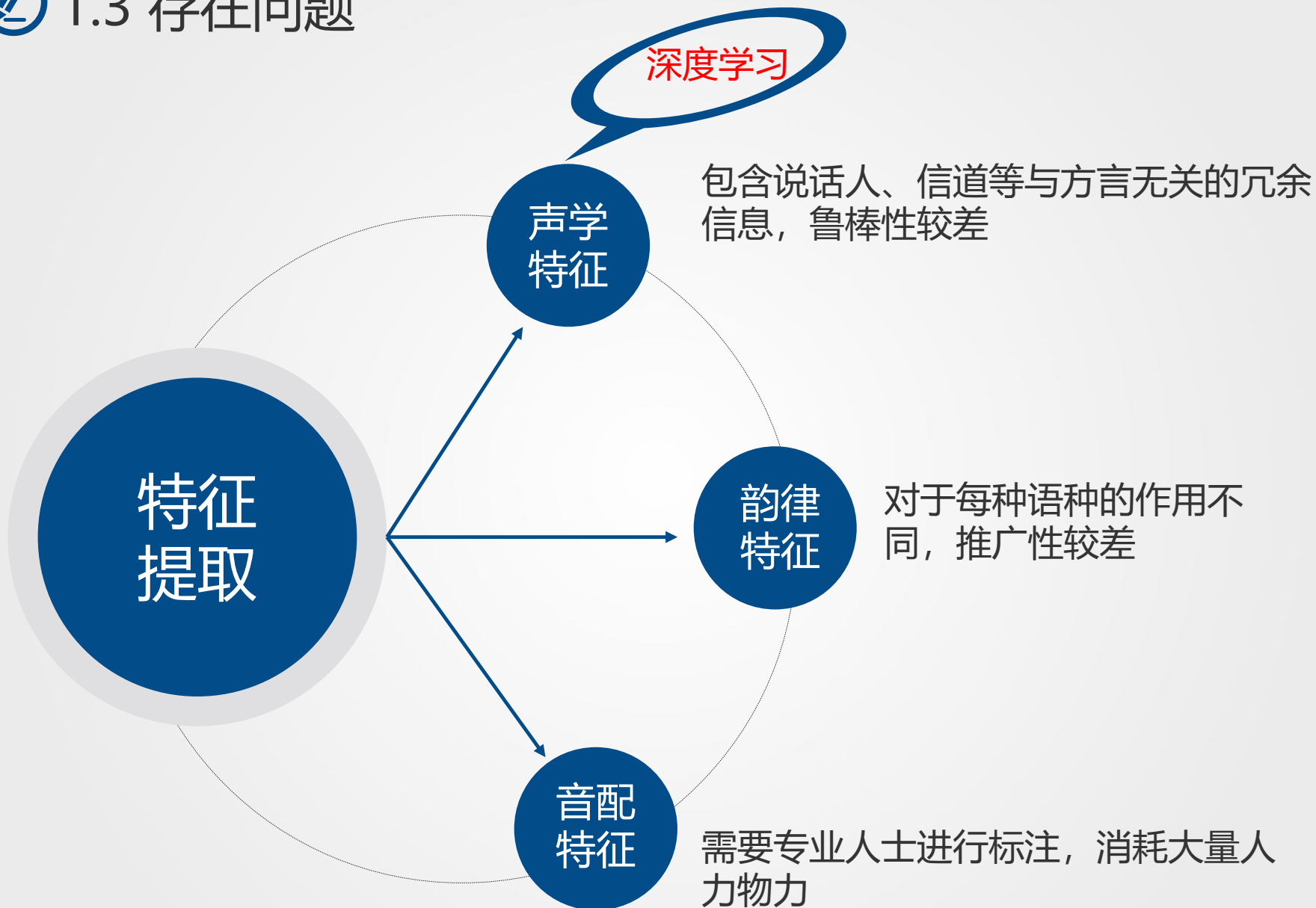
韵律特征

2005,马斌
2006,顾明亮

融合特征

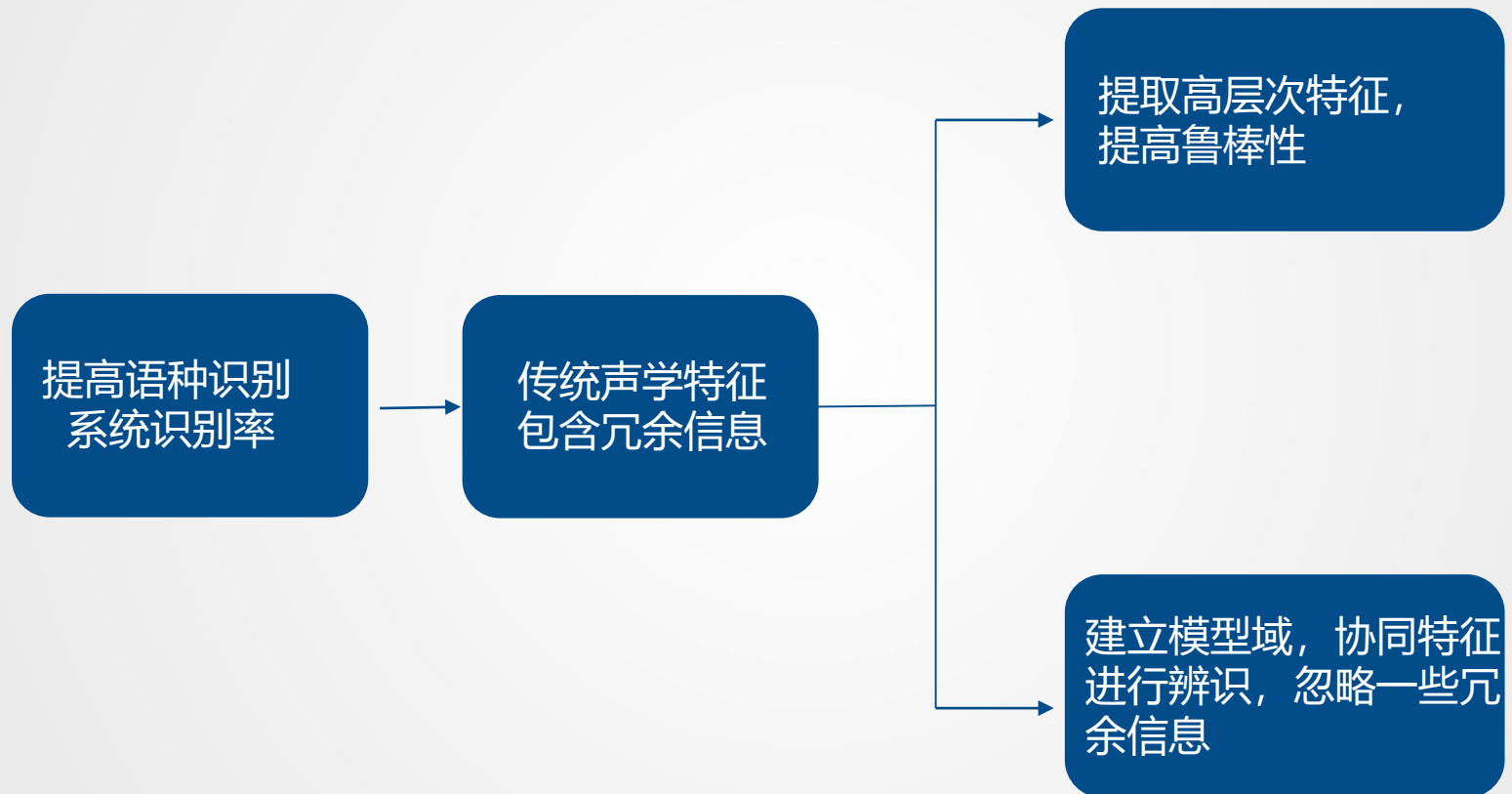


1.3 存在问题





1.4 研究内容



2 Phonetic特征



2.1 Phonetic特征

2.2 Phonetic for LID

2.1 Phonetic 特征

定义:

语音特征，是声学特征的高层特征，从底层声学特征中分析得来。

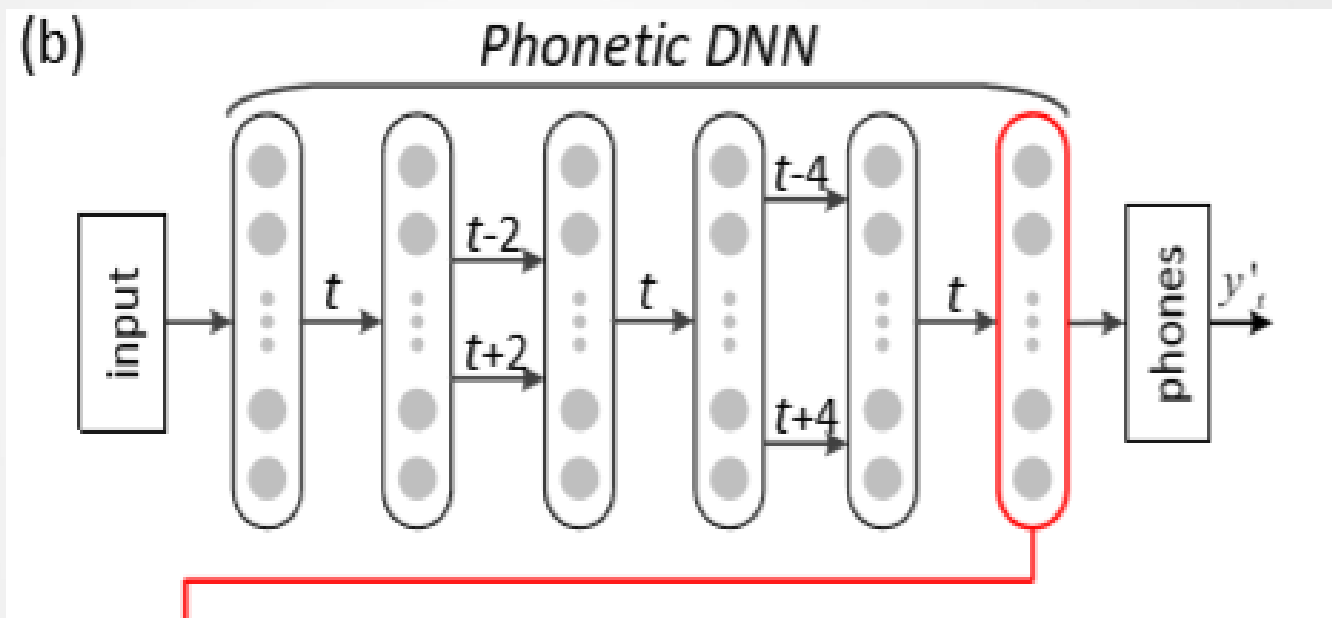


图2.1 Phonetic的特征提取模型, 图来自
《Phonetic temporal neural model for
language identification 》



2.2 Phonetic for LID

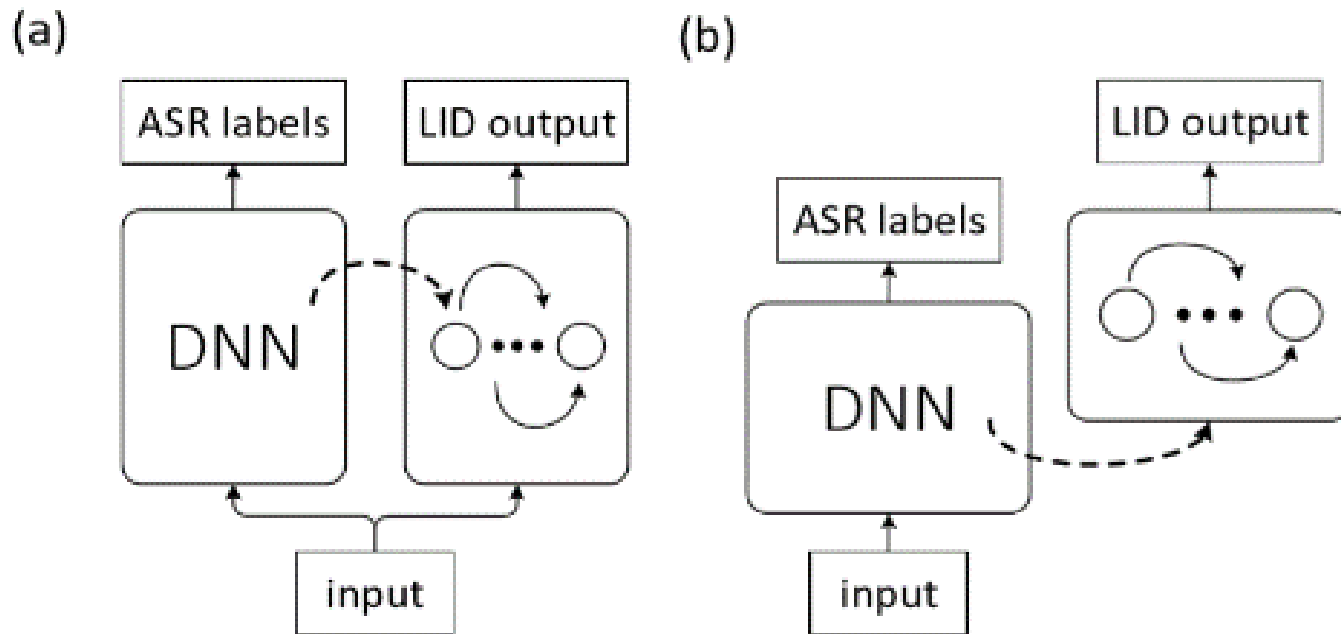


Fig. 1: LID models employing phonetic information: (a) the phonetically aware model; (b) the PTN model. Both models consist of a phonetic DNN (left) to produce phonetic features and an LID RNN (right) to make LID decisions.

图2.2 Phonetic for LID, 图来自《Phonetic temporal neural model for language identification》



3 CNN模型

- 3.1 CNN简介
- 3.2 CNN模型
- 3.3 实验结果分析

3.1 CNN简介

定义:

卷积神经网络的基本结构由输入层、卷积层、池化层（也称为取样层）、全连接层及输出层构成。

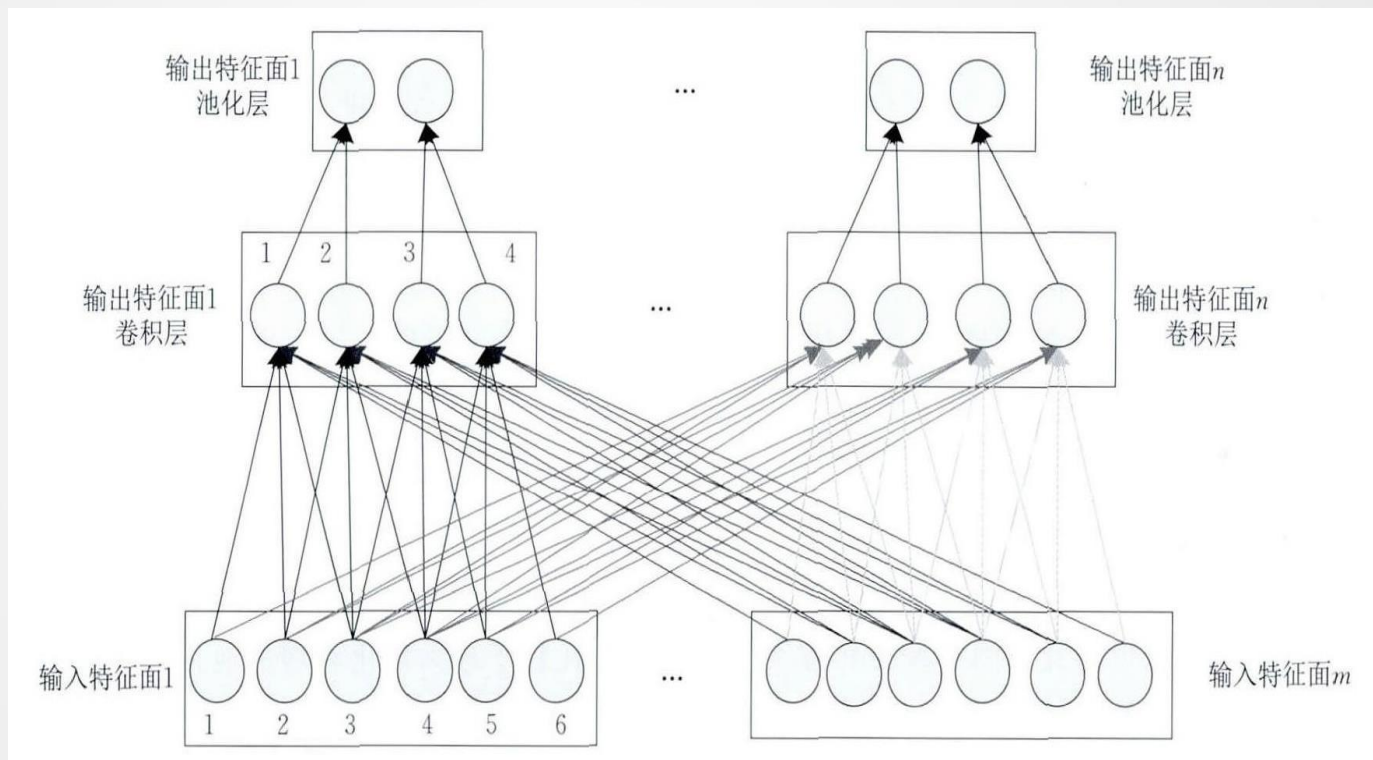


图3.1 CNN结构图



CNN流程图

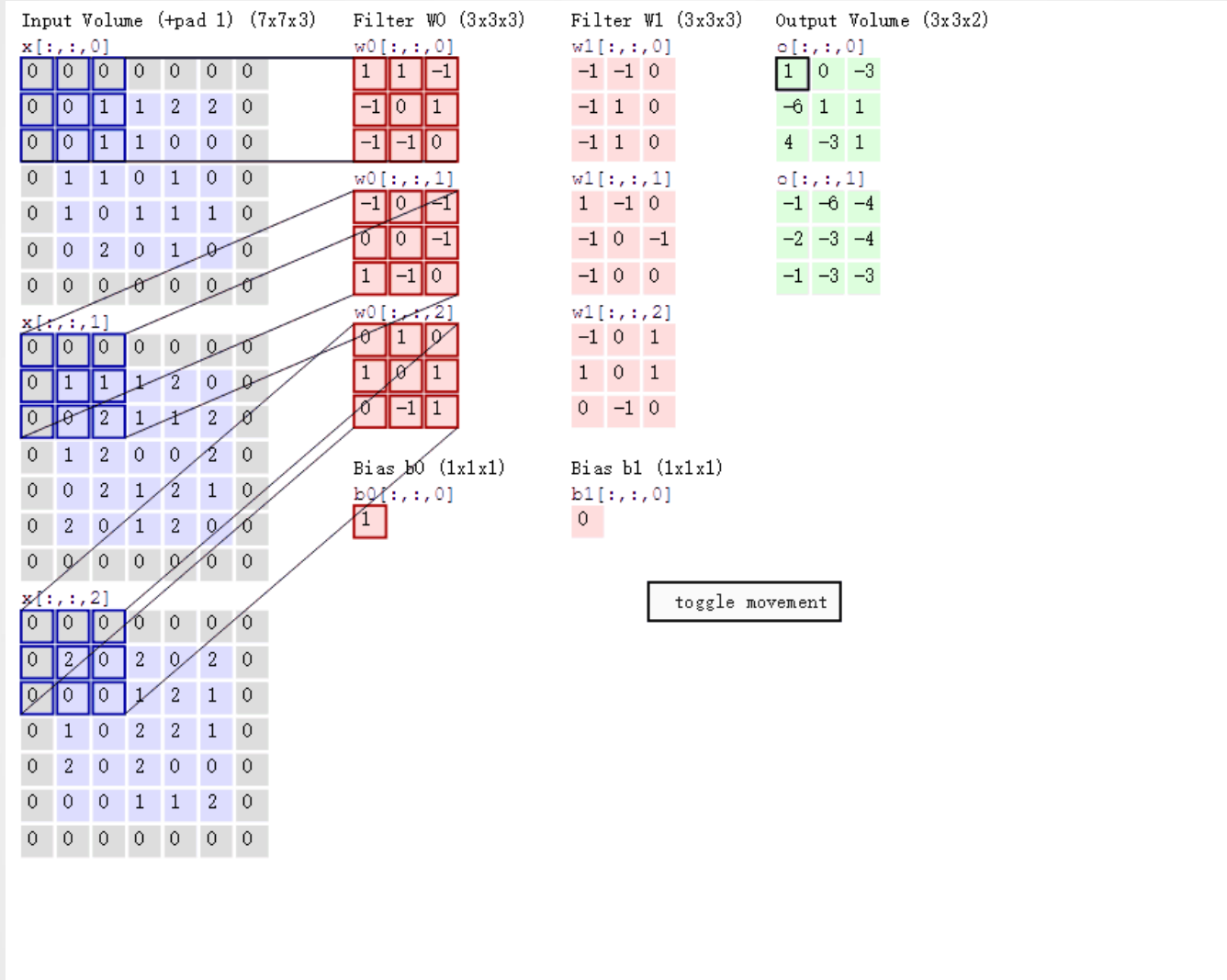


图3.2 CNN流程图

3.2 CNN模型

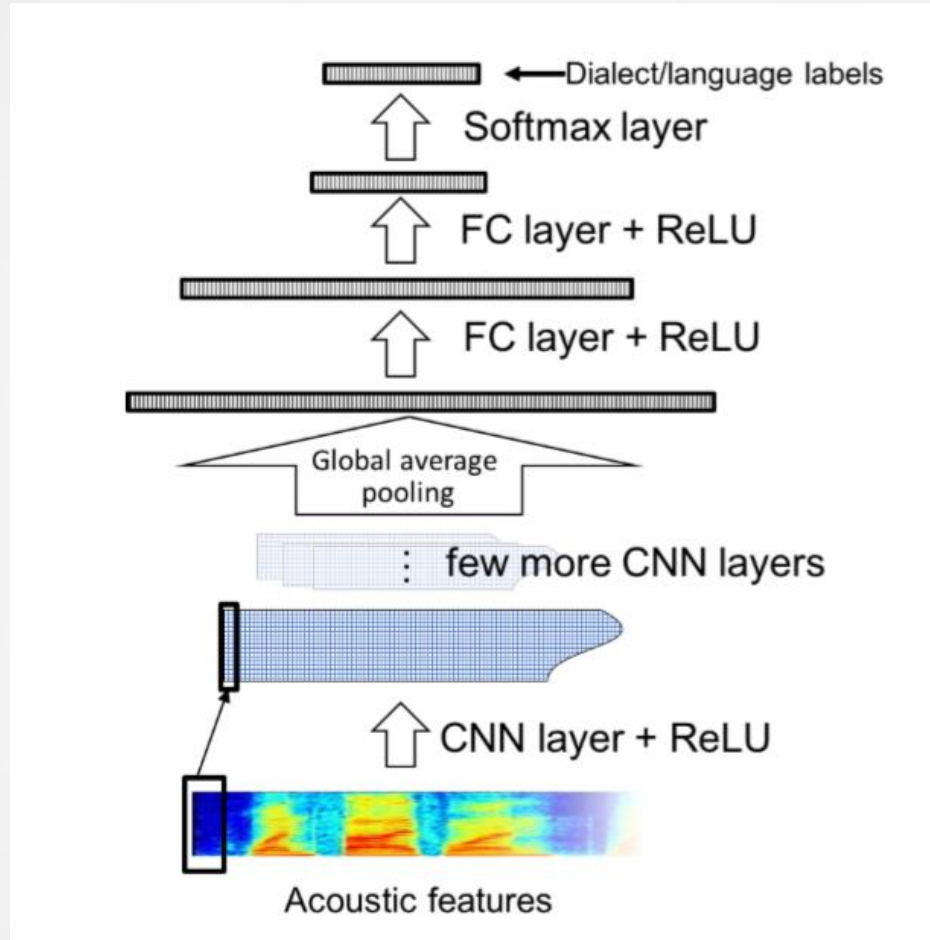


图3.3 CNN for LID

3.3 实验结果分析



实验结果：ACC：53%

分析原因：参数在调试中.....

4 总结

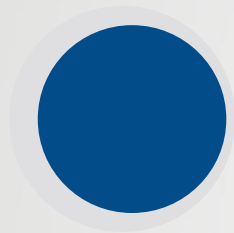


● 4.1 研究创新

● 4.2 总结



4.1 研究创新



改进End-to-End CNN for LID 的特征输入，由于底层声学特征具有极大的冗余信息，语种区分信息不明显，所以基于PTN网络的基础上，提出Phonetic-based CNN模型。

4.2 总结



采用高层次语音信息，减少语种识别的冗余性。以及利用CNN共享卷积核，对高维数据处理无压力的优点，共同组建LID系统。



感谢各位老师批评指正!