

基于一维卷积神经网络的车载语音识别研究

朱锡祥, 刘凤山, 张超, 吕钊, 吴小培

(安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039)

摘要: 针对卷积神经网络(convolution neural networks,CNNs)是二维结构,不能很好地反映出语音信号的一维特性,因此,提出使用一维模型进行语音识别研究.其通过卷积核在时间轴上的移动,在保留频带相关性的同时可以更好的满足语音信号的时变性,进而提高识别性能.最后进行了车载语音识别对比实验,结果表明一维卷积神经网络的识别率较二维卷积神经网络提高了约 10%~20%,在噪声环境下的泛化性能也明显优于后者.

关键词: 卷积神经网络; 语音识别; 网络维度; 卷积核; 泛化性

Research on in-car Speech Recognition Based on One

Dimensional Convolutional Neural Networks

ZHU Xi-xiang, LIU Feng-shan, ZHANG Chao, LV Zhao, WU Xiao-pei

(The Key Laboratory of Intelligent Computing & Signal Processing, Anhui University, Hefei 230039, China)

Abstract: Convolution neural networks(CNNs) has been the architecture of traditional convolution neural networks is two-dimensional(2D), which can not reflect the one-dimensional characteristic of speech signal. Therefore, a one-dimensional(1D)architecture for speech recognition was proposed, which can better satisfy the temporal variation while retaining band correlation by convolution along the time axis. Experiments of in-car speech recognition demonstrate that 1D CNNs can significantly outperform the 2D CNNs, with recognition rate improvement of 10% to 20%, and the generalization performance in noisy environment of 1D CNNs is also significantly better than the latter.

Key words: convolution neural networks; speech recognition; dimension of network; convolution kernel; generalization

作者简介:

朱锡祥 男, (1991-), 硕士研究生.研究方向为语音识别、神经网络等.E-mail:hixxZhu@126.com.

刘凤山 男, (1993-), 硕士研究生.研究方向为语音增强、信号处理等.

张超 男, (1983-), 博士, 讲师.研究方向为智能信息处理、视频/图像处理等.

吕钊 男, (1979-), 博士, 副教授、硕士生导师.研究方向为智能信息处理与人机交互技术等.

吴小培 男, (1966-), 博士, 教授, 博士生导师.研究方向为盲信号处理、医学信号处理等.