

基于深度可分离卷积神经网络的关键词识别系统

王 帅, 彭意兵, 何顶新

(华中科技大学 自动化学院, 湖北 武汉 430070)

摘 要: 关键词识别系统是智能语音交互系统的重要组成部分. 本文使用 Google 语音命令数据集, 探索使用传统卷积神经网络和深度可分离卷积神经网络在关键词识别任务中的应用, 对两种卷积神经网络模型从识别率、计算量、内存消耗进行对比, 并提出适用于受限设备的低资源、较高识别率的网络模型. 实验结果显示无论传统卷积神经网络还是深度可分离卷积神经网络在关键词识别任务中的表现都优于传统的隐马尔科夫模型和全连接深度学习模型, 而深度可分离卷积神经网络进一步优于传统卷积神经网络.

关键词: 关键词识别; 卷积神经网络; 深度可分离卷积神经网络; 受限设备

Keywords spotting system based on deepwise separable

convolutional neural network

WANG Shuai, PENG Yi-bing, Newest He

(College of Automation, Huazhong University of Science and Technology, Wuhan 430070, China)

Abstract: The keyword spotting system is an important part of the intelligent voice interaction system. We explore the application of convolution neural networks and depthwise separable convolution neural networks to the keyword spotting task, using the Google Speech Commands Dataset as our benchmark. We will make comparison of recognition rate, calculation amount, and storage consumption for two convolutional neural network models and propose a network model with low resource and high recognition rate for restricted devices. The experimental results show that both the traditional convolutional neural networks and the deep separable convolutional neural networks perform better than the traditional Hidden Markov model and deep learning model based on fully connected neural networks in the keyword spotting task, while the depthwise separable convolutional neural networks is more superior to the convolutional neural networks.

Key words: keywords spotting; convolution neural network; depthwise separable convolution neural network; restricted device

作者简介:

王 帅 男, (1993-), 硕士研究生. 研究方向为语音识别、深度学习. E-mail: 13659829769@163.com.

彭意兵 男, (1994-), 硕士研究生. 研究方向为深度学习、网络安全.

何顶新 男, (1966-), 硕士, 教授. 研究方向为系统仿真、控制工程.