

摘要: 近年来,我国儿童青少年的近视率呈上升趋势,引起近视的原因有很多,不正确的学习姿势也会增加近视发病的风险,可能还会影响儿童的成长发育。许多家长都十分重视儿童的学习坐姿问题,但却由于各种原因难以做到时刻的监督。针对这一问题,设计了一个基于人工智能的儿童坐姿监测系统。系统的核心为微型电脑,控制摄像头拍摄照片,再通过坐姿识别模型识别照片中的坐姿,在儿童保持低头等错误坐姿时,发出语音提醒其纠正。系统能利用卷积神经网络识别坐姿,及时且有针对性地提醒纠正,并且程序移植到了嵌入式平台,体积小巧,操作简便。

关键词: 不良坐姿;人工智能;图像分类;语音提醒

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1006-883X(2021)01-0019-05

收稿日期: 2020-10-28

基于人工智能的儿童坐姿监测系统

李晓彤 程进 邹小平 李光东 王晓岚 任宝凯 王一飞 刘宝宇 王骏琦 周子肖 余兴

北京信息科技大学北京市传感器重点实验室,北京 100101

0 引言

2019年4月,国家卫健委疾控局副局长张勇在新闻发布会上指出,2018年全国儿童青少年总体近视率为53.6%,近视防控任务艰巨^[1]。2020年4月21日,习近平总书记在陕西省安康市平利县老县镇中心小学考察时说“现在孩子普遍眼镜化,这是我的隐忧^[2]。”近年来,我国儿童青少年近视发病率呈上升趋势,原因之一就是不正确的读写姿势。另外,不良坐姿还会影响青少年的成长发育,造成脊柱弯曲等问题。教师和家长都十分注意监督和提醒儿童保持正确的学习姿势,但对许多家长来说难以做到时时刻刻的监督。

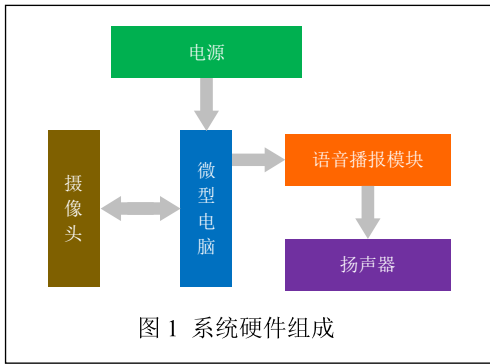
针对此类问题,目前市场上有支撑起儿童头部、强制保持眼睛与书本距离和通过座椅背带强制儿童背部沿座椅挺直的产品等,均为功能单一、强制性的产品,容易影响儿童学习时的舒适度,分散注意力。本文设计了基于人工智能的儿童坐姿监测系统,通过摄像头和神经网络识别儿童在读书学习时的错误坐姿,并及时通过语音进行提醒。

对儿童错误坐姿的识别通过卷积神经网络实现。卷积神经网络(convolutional neural network, CNN)是一种深层前馈神经网络,在20世纪60年代,由Hubel和Wiesel首先提出,Fukushima K^[3]在1980年提出第一个卷积神经网络的实现网络,目前已成功应用于图像识别,在大型图像处理方面表现出色^[4-5]。本系统移植到微型电脑(树莓派),通过摄像头拍摄照片,针对儿童学习时比较容易出现的低头、歪头、趴在桌子上、高低肩等4种错误坐姿,通过神经网络识别,在检测到儿童保持错误坐姿时,发出相应的语音,提醒儿童纠正。

1 系统设计

1.1 系统构成

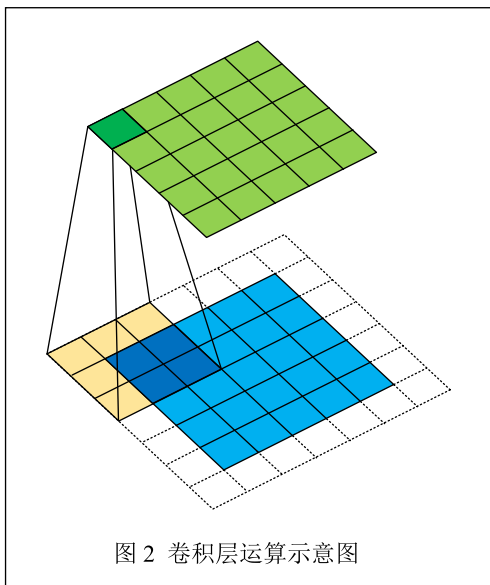
本系统由微型电脑、摄像头、语音播放器、扬声器4部分组成,如图1所示。微型电脑控制摄像头拍照并保存,利用神经网络识别照片中的坐姿,在坐姿错误时,微型电脑控制语音播放器,通过扬声器播放语音提醒。



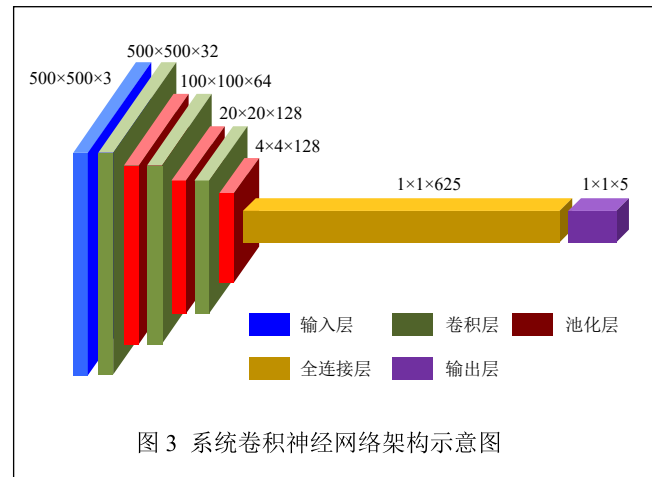
1.2 卷积神经网络模型

拍摄的坐姿照片通过卷积神经网络进行识别。卷积神经网络由卷积层、池化层、全连接层和输出层构成。与普通人工神经网络不同，卷积神经网络中包含由卷积层和池化层构成的特征抽取器。卷积层中包含一组卷积滤波器，将输入图像分解为像素矩阵，卷积滤波器作为权值矩阵，从上到下，从左到右地在输入数据上“滑动”，对滑过区域的像素矩阵进行矩阵乘法，结果作为单个输出像素值，形成特征图，如图2所示^[6-7]。通过卷积可使原信号特征增强，同时降低噪音。卷积层将特征提取后，在池化层计算一个区域特征的最大值或平均值，可以将相似特征合并，从而降维，减少数据处理量^[8]。

图3为本系统的卷积神经网络架构。首先



将输入的图像分解为 3×3 的像素矩阵块，在卷积层与卷积滤波器进行矩阵乘法运算，结果组成特征矩阵，并通过0填充，使特征矩阵与原图像保持相同的尺寸；之后通过池化层对特征矩阵 5×5 的矩阵块取最大值，形成新的矩阵，实现降维的目的，在减少数据处理量的同时保留有用的信息，并且可以避免过拟合的问题。输入的图像经过这样3个“卷积→池化”的过程，连接成一个向量，输入到全连接层，最终得到输出结果。



1.3 数据集

本系统应用于儿童学习时坐姿的识别，现存数据集并不适用，因此我们自行创建了坐姿数据集，用于对卷积神经网络的训练^[9]。首先通过树莓派摄像头录制视频，再从视频中按帧截取图片^[10]，采集5个人的4种错误坐姿和正确坐姿共5类照片，如图4所示。在拍摄视频时，移动摄像头，从不同距离和角度进行拍摄，截图得到不同拍摄距离和角度的坐姿图片，减少拍摄距离和角度对系统坐姿识别准确度的影响，将系统放置在前方任意位置，均能较为准确地识别坐姿。5种坐姿包括4种错误坐姿和正确坐姿，每种均拍摄5个人的5段30秒视频，每段视频以时间间隔1.5秒分别截取出20张图片，组成数据集，每种坐姿100张图片，共500张。

1.4 语音提醒模块

本文中采用YS-M3语音播报模块连接扬声器实现语音提醒的功能，在语音播报模块中存入4段语音，分别对应系统能够识别的4种典型错误坐姿，在系统检测到儿童保持错误坐姿时，根据识别出的坐姿类型播放相应的语音进行提醒。另外，提醒语音可以定制父母语音，增加提醒效果。



2 坐姿识别模型训练

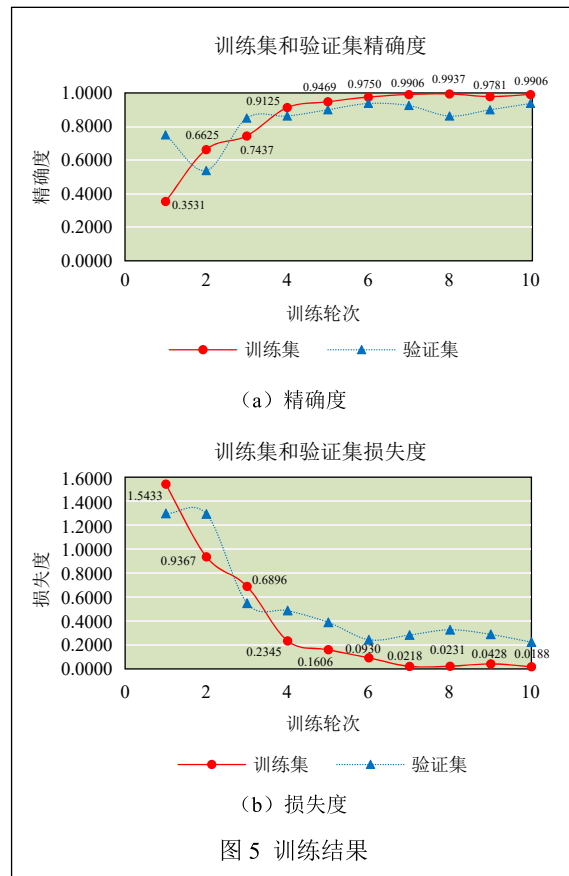
采用前文中自行创建的坐姿数据集对坐姿识别模型进行训练, 导入正确坐姿和4种不良坐姿各100张, 共500张图片, 打乱顺序后对卷积神经网络进行训练, 训练的结果如图5所示。从图中可以看到, 随着训练轮次的增加, 损失度逐渐降低, 精确度不断提高, 最终训练集的精确度达到99%, 验证集的精确度达到93%左右。

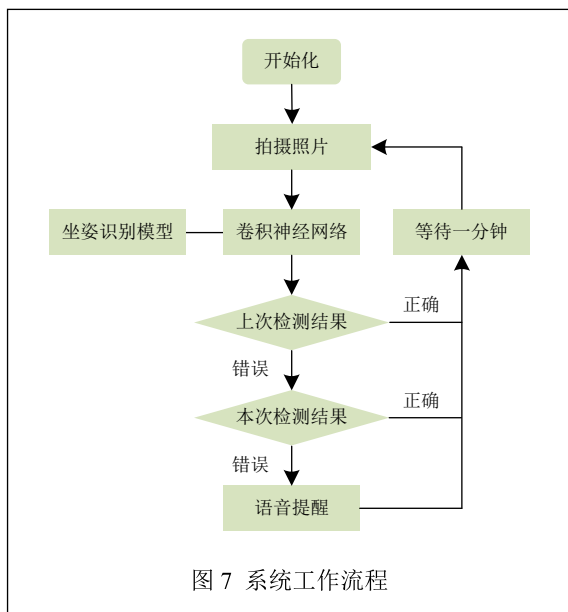
在以往进行人体姿势或坐姿识别的文献中, 在数据导入神经网络训练之前, 需要先进行图像分割、图像形态学处理和空域滤波处理^[9]等, 本文利用卷积神经网络可直接将图片导入神经网络进行训练, 大大减少了数据预处理的运算量, 系统的实现不需要具有强大运算能力的硬件, 可移植到嵌入式平台, 能够较快且较为准确地识别错误坐姿。

3 系统安装与工作流程

摄像头和语音播报模块均与微型电脑连接, 封装在外壳中, 使用时将装置放置在儿童前方, 使摄像头正对儿童, 连接微型电脑电源即可启动系统, 如图6所示。

系统的工作流程如图7所示。连接微型电脑电源, 打开开关, 系统程序开始运行, 微型电脑控制摄像头每分钟拍摄一张照片, 保存在微型电脑中, 通过卷积神经网络识别照片中的坐姿, 在系统检测到坐姿错误, 但上一次检测结果为正确时, 不会播放语音, 而是将此次结果保存, 再继续等待1分钟, 拍照、识别, 若连续两次检测到坐姿不正确, 则系统判断儿童正在保持错误坐姿, 微型电脑根据最新一次的坐姿识别结果





控制语音播报模块播放相应的语音，提醒儿童纠正，如儿童长时间保持低头姿势，系统会发出“把头抬起来，不要离书本太近哦”提醒。

4 结束语

本系统利用卷积神经网络识别坐姿，并将程序移植到嵌入式平台，系统体积小、成本低、易于安装、操作方便，能够识别儿童在学习时的坐姿，在儿童长时间保持错误姿势时通过语音提醒纠正，提醒语音与错误姿势相对应，具有针对性，并且提醒语音可定制父母语音，提升提醒效果，有助于预防因学习坐姿不正确而导致近视等问题。目前，自行创建的坐姿数据集包含数据量较小，坐姿图片场景较为单一，为进一步提高系统的实际应用性能，可增加更多场景下拍摄的更多人的坐姿图片数据，导入卷积神经网络进行训练，降低背景场景对系统坐姿识别的影响。

参考文献

[1] 白剑峰. 儿童青少年总体近视率为 53.6%, 我国将更有针对性地开展近视干预 [N]. 人民日报, 2019, 4, 30.
 [2] 王家源, 李小伟. 后疫情阶段如何加强青少年近视防控 [EB/OL]. (2020-08-03)[2020-12-25]. <http://www.nhc.gov.cn/guihuaxxs/zxxjkcj/202008/ec559c4f0a674eab8cba860e56d705a1.shtml>.

[3] FUKUSHIMA K, MIYAKE S. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position[J]. Pattern Recognition, 1982, 15(6): 455-469.
 [4] 张曰花, 王红, 马广明. 基于深度学习的图像识别研究 [J]. 现代信息技术, 2019, 3(11): 111-112, 114.
 [5] 王禹嫣. 深度学习在图像识别中的应用与发展 [J]. 中国战略新兴产业, 2017, 132(48): 127-128.
 [6] 高华照. 基于深度卷积神经网络的图像识别算法研究 [D]. 吉林大学, 2018.
 [7] 黄旭. 基于判别式深度学习的坐姿视觉识别方法研究 [D]. 湖南大学, 2018.
 [8] 王瑞. 基于卷积神经网络的图像识别 [D]. 河南大学, 2015.
 [9] 陈瑞强. 基于人体姿势分析的智能监控系统研究 [D]. 燕山大学, 2010.
 [10] 周全. 基于坐姿图像识别的教学反馈系统设计及实现 [D]. 江西财经大学, 2018.

A Monitoring System for Children's Sitting Position Based on Artificial Intelligence

LI Xiaotong, CHENG Jin, ZOU Xiaoping, LI Guangdong, WANG Xiaolan, REN Baokai, WANG Yifei, LIU Baoyu, WANG Junqi, ZHOU Zixiao, YU Xing

(Beijing Sensor Key Laboratory, Beijing Information Science & Technology University, Beijing 100101, China)

Abstract: In recent years, the myopia rate of children and adolescents in China is on the rise. There are many reasons for myopia. Incorrect learning posture will increase the risk of myopia, and may also affect the growth and development of children. Many parents attach great importance to children's learning posture, but for various reasons, it is difficult to achieve constant supervision for many of them. To solve this problem, this paper designs a monitoring system of children's sitting position based on artificial intelligence. The core of the system is a micro computer, which controls the camera to take photos, and then recognizes the sitting position in the photos through the sitting posture recognition model. When the children maintain the wrong sitting position, such as lowering their head, they will be reminded to correct it by voice. The system can use the convolutional neural network to identify the sitting position, timely and targeted to remind

correction, and the program has been transplanted to the embedded platform, small in size and easy to operate.

Key words: undesirable posture; artificial intelligence; image classification; voice remind

作者简介

李晓彤: 北京信息科技大学北京市传感器重点实验室, 硕士研究生, 研究方向为物联网技术

通讯地址: 北京市北四环中路 35 号教 2 楼 117 实验室
邮编: 100101

邮箱: 252240783@qq.com

程进: 北京信息科技大学北京市传感器重点实验室, 研究员, 主要从事光纤声学传感器方面的研究。

邹小平: 北京信息科技大学北京市传感器重点实验室, 硕士生导师, 研究员, 主要从事物联网方面的研究。

李光东: 北京信息科技大学北京市传感器重点实验室, 硕士研究生, 研究方向为物联网技术。

王晓岚: 北京信息科技大学北京市传感器重点实验室, 硕士研究生, 研究方向为物联网技术。

任宝凯: 北京信息科技大学北京市传感器重点实验室, 硕士研究生, 研究方向为物联网技术。

王一飞: 北京信息科技大学北京市传感器重点实验室, 硕士研究生, 研究方向为物联网技术。

刘宝宇: 北京信息科技大学北京市传感器重点实验室, 硕士研究生, 研究方向为物联网技术。

王骏琦: 北京信息科技大学北京市传感器重点实验室, 硕士研究生, 研究方向为物联网技术。

周子肖: 北京信息科技大学北京市传感器重点实验室, 硕士研究生, 研究方向为物联网技术。

余兴: 北京信息科技大学北京市传感器重点实验室, 硕士研究生, 研究方向为物联网技术。