

# 人脸属性识别系统的设计与实现

王高升

(华北计算机系统工程研究所,北京 100083)

**摘要:** 目前,精确地进行人脸识别在现实场景中得到了大量的应用,如监控、门禁等,而只针对人脸属性进行的识别在娱乐、安防和社交媒体领域同样有着广泛的应用。因此,基于深度卷积神经网络 MTCNN 进行人脸检测,基于 Resnet50 网络进行人脸属性的识别,同时使用 Python 语言设计开发实现人脸属性识别应用界面的系统。通过数据实验,该系统使用基于深度卷积神经网络的人脸属性识别模型在性别、年龄以及种族属性的识别上准确率分别达到了 97.32%、71.64% 及 92.13%,同时人脸属性识别应用系统能够稳定地运行,具有一定的应用价值。

**关键词:** 深度学习;卷积神经网络;人脸检测;人脸属性识别;应用系统开发

中图分类号: TN06; TP3

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.200354

中文引用格式: 王高升. 人脸属性识别系统的设计与实现[J]. 电子技术应用, 2021, 47(4): 112-115, 125.

英文引用格式: Wang Gaosheng. Design and implementation of face attribute recognition system[J]. Application of Electronic Technique, 2021, 47(4): 112-115, 125.

## Design and implementation of face attribute recognition system

Wang Gaosheng

(National Computer System Engineering Research Institute of China, Beijing 100083, China)

**Abstract:** At present, accurate face recognition has been widely used in real scenes, such as monitoring, access control, etc., and recognition only for face attributes is also widely used in entertainment, security and social media. Therefore, the paper is based on deep convolutional neural network MTCNN for face detection and Resnet50 network for face attribute recognition. At the same time, it uses Python language to design and develop a system that implements face attribute recognition application interface. Through data experiments, the system uses a face attribute recognition model based on deep convolutional neural networks to achieve 97.32%, 71.64% and 92.13% accuracy in identifying gender, age and ethnic attributes, respectively. At the same time, the face attribute recognition application system can run stably and has certain application value.

**Key words:** deep learning; convolutional neural network; face detection; face attributes recognition; application system development

## 0 引言

人脸的属性包含了人的面部相关重要的信息,如人的年龄、性别、种族等属性信息。人脸属性的识别就是借助提取出的人脸面部的属性信息然后再进行识别的过程。最近的几年来,随着计算机视觉技术、深度学习及卷积神经网络的飞速发展与应用,出现了许多人脸检测与识别相关的应用,其中常见的一些应用场景包括:道路上的行人监控系统(如检测道路上的行人是否佩戴墨镜或者口罩等)、人脸识别的门禁系统<sup>[1]</sup>以及采用人脸识别的打卡签到系统等。虽然目前在人脸检测与人脸属性识别的方面得到了非常大的发展,不过很多的之前的研究仅仅局限于预测单个的人脸的属性(例如性别或者年龄)或者为每个的人脸属性信息都通过学习得到一个单独的用于进行识别的模型。

相比于基于 HOG-多尺度 LBP 特征的人脸性别识别<sup>[2]</sup>的 93.0% 准确率,本文采用的人脸属性识别网络在

性别属性识别的准确率上有了 4.32% 的提升。文献[2]提出了一种方向梯度直方图和多尺度局部二值模式多特征融合的人脸性别识别算法。首先,对输入图像进行裁剪和缩放得到多个分辨率的人脸图像,再分别提取 LBP 统计直方图并合成一个特征向量;然后提取目标图像头肩模型的 HOG 特征得到 HOG 特征向量;最后,将 LBP 特征向量与 HOG 特征向量合成一个新的特征向量,应用支持向量机进行训练。而相比于本文提出的采用深度卷积神经网络进行人脸检测及人脸属性的识别,本文的网络结构在性别属性识别上有着更好的效果。

同时本文采用的网络结构相比于 DEPGHAN A<sup>[3]</sup>提出的基于 MTL 的 DCNN 网络去识别人脸的属性,其网络是基于不同的任务采用不同的数据集去训练该 DCNN 网络,本文通过采用 Resnet50 深度残差卷积神经网络同时进行人脸多属性的识别,使用一个网络结构实现多个人脸属性的识别而非单一人脸属性识别,同时本文人脸

多属性识别的准确率相比于基于 MTL 的 DCNN 网络在年龄属性识别及性别属性识别上分别有 7.64% 和 6.32% 的提升, 提升显著。

## 1 基于深度学习的人脸属性识别方法

人脸属性的识别整体过程主要能够分成 2 个阶段: 人脸的检测阶段和人脸属性信息的识别阶段。由于深度学习当前在图像处理中具有较强的特征提取能力, 在学习高层的语义特征方面具有着先天的优越性能。因此, 本文主要是应用深度卷积神经网络进行人脸的检测及人脸属性信息的分类识别。

### 1.1 人脸检测

本文在进行人脸检测操作的阶段采用的是经过修改后的 MTCNN(Multi-task Cascaded Convolutional Networks) 网络<sup>[4]</sup>进行人脸检测, 其由 3 个级联的轻量级 CNN 完成: Proposal Network(P-Net)、Refine Network(R-Net) 和 Output Network(O-Net)。

MTCNN 为了兼顾性能和准确率, 避免滑动窗口加分类器等传统思路带来的巨大的性能消耗, 先使用小模型生成有一定可能性的目标区域候选框, 然后再使用更复杂的模型进行细分类和更高精度的区域框回归, 并且让这一步递归执行, 以此思想构成 3 层网络, 分别为 P-Net、R-Net、O-Net, 实现快速高效的人脸检测。在输入层使用图像金字塔进行初始图像的尺度变换, 并使用 P-Net 生成大量的候选目标区域框, 之后使用 R-Net 对这些目标区域框进行第一次精选和边框回归, 排除大部分的负例, 然后再用更复杂的、精度更高的网络 O-Net 对剩余的目标区域框进行判别和区域边框回归。图像数据先后经过以上 3 个网络的处理, 最终得到人脸检测结果, 也即是得到了人脸区域框的左上角像素以及右下角像素在输入图像中的坐标信息。

通过使用 WIDERFace 人脸数据集<sup>[5]</sup>对经过改进后的 MTCNN 网络进行有监督的训练, 并且在网络训练过程中仅进行边框回归操作, 不对关键点信息检测进行训练, 因为人脸属性识别过程不需要关键点的位置信息, 可以进一步加快网络训练与推理的速度。通过训练获得边框预测准确度比较好的深度卷积神经网络模型, 最后采用这个深度卷积神经网络模型进行人脸的检测操作过程<sup>[6]</sup>。

人脸检测阶段中改进后的 MTCNN 网络的推理过程如图 1 所示。

### 1.2 人脸属性识别

人脸属性信息的识别阶段主要是利用人脸检测阶段检测网络获得的人脸区域的图像, 对这个获得图像通过深度卷积神经网络进行信息提取并输出性别、年龄及种族人脸属性信息。

在 Resnet 网络之前, 卷积神经网络主要采用将卷积进行堆叠的方式增加网络结构的深度, 以使得网络模型的表现性能具有更好的效果。但是由于在网络深度的增

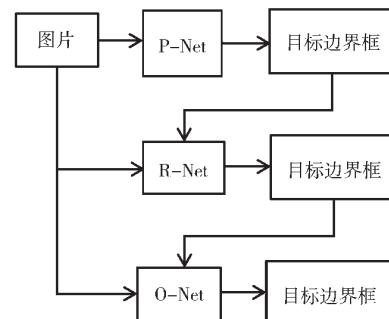


图 1 改进后的 MTCNN 网络的推理过程

加的同时, 就会出现网络模型的训练过程中梯度进行反向传播比较困难的现象, 较容易出现梯度消失或者梯度爆炸的问题, 由此, Resnet 网络引入了残差网络结构。这种网络结构极大地简化了网络的学习目标和学习的难度, 使得网络具有更强的恒等映射的能力, 从而在拓展了网络深度的同时也提升了网络模型的性能。残差结构使用跳跃连接的方式实现, 通过引入残差结构, 可以对更深的网络进行训练, 同时不会出现梯度消失或者梯度爆炸的问题。

由于深度卷积神经网络能够通过使用一系列的卷积层操作、池化操作层操作等自动地提取出输入的图像的特征, 更深的网络能够得到更具有表达能力的特征信息, 因此当前应用于特征提取的并且取得较好效果的是一些深度卷积神经网络。因此本文在人脸属性识别阶段采用改进的 Resnet50 网络<sup>[7]</sup>, 由于 Resnet50 具有较深的卷积结构, 能够很好地提取图像特征信息, 因此, 改进的 Resnet50 网络的全连接层修改成 108 个神经元, 也即是 108 维的张量输出, 其中 0~100 代表年龄, 101~102 代表性别(101: Male, 102: Female), 103~107 代表种族(对应的依次分别为白人、黑人、亚洲人、印第安人、其他类型(如拉丁人、西班牙人等))。

通过使用 UTKFace 人脸数据集<sup>[8]</sup>(其中包含性别、年龄及种族信息)及 Resnet50 的预训练模型, 采用交叉熵损失函数<sup>[9]</sup>, 对修改后的网络模型进行训练, 对数据集中的年龄、性别及种族都采用 one-hot 编码<sup>[10]</sup>, 3 种属性在训练时都是采用分类的方法。

使用交叉熵损失函数能够更好地度量出来两个概率分布之间存在的差异性, 交叉熵的公式如式(1)所示:

$$H(p, q) = - \sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (1)$$

其中,  $p$ 、 $q$  分别表示一种概率分布,  $x$  为随机变量。

交叉熵能够度量出来相同的一个随机变量中的两个存在差异的概率分布之间的差别的程度, 交叉熵的值越小, 两个不同的概率分布之间的差别越小。在进行网络训练时, 输入网络的数据与标签信息已经完全确定, 而且目标的标签信息采用了独热编码的形式, 一次可以把输入网络的标签信息看作为真实的概率的分布  $P(x)$ ,

而  $Q(x)$  即是网络预测的概率的分布, 而网络模型的训练过程其实就是不断地把网络预测的概率与真实分布的概率之间的差异减小的过程, 所以采用交叉熵损失函数来计算网络的损失, 通过不断地更新网络参数, 来减小网络预测损失的大小, 提高网络模型预测的精度。

在使用训练好的网络模型时, 性别、种族和年龄属性均采用分类的方法得到结果, 如年龄属性多分类使用的方法主要是对全连接网络层的 0~100 的输出张量首先使用 Softmax 对网络输出的数据进行归一化的操作, 得到对应人脸年龄的概率大小, 根据概率的大小排序, 得到人脸的年龄属性结果信息。

人脸属性识别阶段的 Resnet50 网络推理过程如图 2 所示。人脸检测网络训练过程的损失变化如图 3 所示。进行人脸属性分类识别的 Resnet50 网络的训练过程中损失变化如图 4 所示。

通过人脸检测阶段及人脸属性识别阶段, 可以得到人脸的性别、年龄及种族属性信息, 整个人脸属性识别

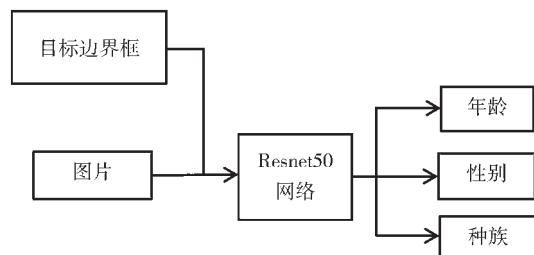


图 2 改进后的 Resnet50 网络模型的推理过程

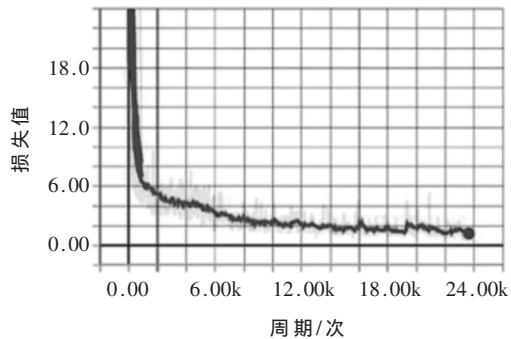


图 3 人脸检测网络训练过程损失变化

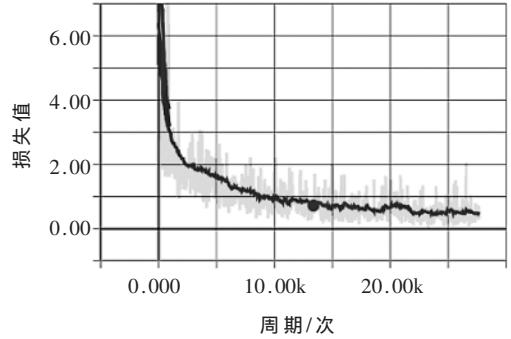


图 4 Resnet50 网络训练过程中损失变化

过程采用端到端的方式, 进一步使得网络的精度得到提升, 并且能够预测多种人脸的属性信息<sup>[11]</sup>。

## 2 实验结果与分析

实验中采用的数据集由 WIDERFace 和 UTKFace 人脸数据集组成, 其中测试集的占比约为 5%。

人脸检测阶段采用的是 WIDERFace 人脸数据集进行网络的训练, 由于人脸属性识别过程不需要人脸的关键点位置信息, 因此人脸属性识别系统中人脸检测阶段改进后的 MTCNN 网络只使用了数据集中人脸的边框位置信息进行网络模型的训练, 训练后的人脸检测网络模型推理结果得出的只有人脸的边框位置信息<sup>[12]</sup>。

人脸属性识别阶段采用的是 UTKFace 人脸数据集对改进后的 Resnet50 网络进行网络的训练, 改进后的 Resnet50 网络采用了 UTKFace 数据集中的 age、gender、race 信息, 并对这些数据进行了数据清洗、one-hot 编码等数据预处理过程, 然后加载 Resnet50 网络的预训练模型进行迁移学习训练, 采用预训练模型可以加速网络收敛, 大幅度减少网络训练所需要耗费的时间。训练后的网络模型在推理过程中输出的是 108 维的特征向量, 这个结果分别代表年龄、性别及种族信息。人脸属性识别中, 性别、种族及年龄均采用分类方法, 性别是二分类, 种族是多分类(识别的种族包含 5 种), 年龄属性同样是多分类任务。

人脸属性识别网络经过端到端的训练之后, 各属性识别的准确度如表 1 所示。

从表 1 中可以看出, 网络模型在性别属性的识别预测的精度最高, 达到 97.32%; 种族识别预测的准确度为 92.13%; 年龄识别预测的准确度则达到了 71.64%。改进后的网络模

表 1 验证集各属性识别的准确度

属性名	准确度/%
性别	97.32
年龄	71.64
种族	92.13

型进一步提高了人脸各属性信息识别的准确率, 得到了更好的识别效果。

在实际的场景中, 种族及年龄的识别预测受到光线及角度的影响较为严重, 光照对种族识别预测的影响比较严重, 主要原因是在采用的种族识别预测的数据集中, 种族的识别预测更多地是根据肤色来进行, 因此光照会造成结果有很大的偏差<sup>[13]</sup>。同时人脸的不同角度对于人脸属性的识别也有一定的影响, 所以通过加入人脸矫正过程可能对人脸属性识别的准确率有一定的提升。光照及角度对识别的影响需要进一步通过预先进行图像处理等方式来降低对识别网络进行预测产生的影响<sup>[14]</sup>。

## 3 人脸属性识别系统的实现

人脸属性识别系统主要是采用 Python 及其 Tkinter 库进行系统的开发与实现, 并把人脸属性识别网络及训练好的网络模型加入到系统中, 最后通过 Pyinstaller 把整个系统程序打包成在 Windows 操作系统下能够直接

运行的可执行程序(人脸属性识别系统.exe)<sup>[15]</sup>。

人脸属性识别应用系统的主界面如图 5 所示, 界面功能主要有摄像头的选取功能, 通过下拉列表可选取系统运行时采用电脑自带的摄像头或者使用外接的摄像头设备, 在开启摄像头时该设置不可进行修改, 可在关闭摄像头之后再次进行选取。通过开启摄像头或者关闭摄像头功能按钮, 可以对指定的摄像头设备进行数据读取或者关闭摄像头设备。右侧为对识别到的人脸采用网络模型进行运算后输出的人脸属性信息进行详细的显示以及检测到的人脸图片显示。系统程序采用多线程方式进行人脸属性识别与界面信息显示。系统在进行人脸属性识别时未检测到人脸时的系统界面如图 6 所示, 识别成功时的系统界面如图 7 所示(注: 人像图片来自公开人脸数据集 LFW), 当视频中出现多个人脸时只对其中面积最大的人脸进行属性识别。系统在进行人脸属性识别时网络模型的运算速度能够达到约 20 f/s, 系统整体运行速度较快, 操作简单, 运行状态稳定。

#### 4 结论

本文详细论述了人脸属性识别系统设计与实现以



图 5 人脸属性识别系统主界面



图 6 未检测到人脸时系统界面



图 7 识别成功时的系统界面

及网络改进方法, 相比于传统的识别方法, 文中所提出的基于深度学习的人脸属性识别方法采用了端到端的网络训练方法, 进一步地提高了人脸属性识别的准确度; 本文的网络采用了较深的卷积神经网络结构, 通过深度卷积神经网络能够更好地提取出人脸图像中的人脸特征信息, 采用交叉熵损失函数训练网络模型, 提高了人脸属性的识别准确率<sup>[16]</sup>, 使得人脸属性识别系统能够很好地进行人脸属性的识别。但是该识别方法会受到一定程度上的光照及角度的影响, 因此, 后续研究可在该方法的基础上加入图像预处理方法来减少光照及人脸角度问题产生的影响。同时通过更多的数据集的收集, 后续可以加入更多的人脸属性信息进行识别。而对于人脸属性识别系统, 后续可以对系统界面进行美化以及进一步增加各种相关功能, 使得人脸属性识别系统能够对更多的属性进行识别<sup>[17]</sup>。

本文的网络结构及系统设计方面还有许多可以改进和优化的地方, 随着深度学习及计算机视觉的快速发展, 将会出现更多计算机视觉相关成果的产品。

#### 参考文献

- [1] 宿静宜, 刘久付, 杨明海. 采用人脸身份识别的智慧门禁系统[J]. 软件导刊, 2019, 18(4): 32–35.
- [2] 闫敬文, 江志东, 刘蕾. 基于 HOG-多尺度 LBP 特征的人脸性别识别[J]. 扬州大学学报(自然科学版), 2016, 9(3): 58–62.
- [3] DEHGHAN A, ORTIZ E G, SHU G, et al. Dager: deep age, gender and emotion recognition using convolutional neural network[J]. arXiv: 1702.04280, 2017.
- [4] Zhang Kaipeng, Zhang Zhanpeng, Li Zhifeng. Joint face detection and alignment using multi-task cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499–1503.
- [5] YANG S, LUO P, LOY C C, et al. WIDER FACE: a face detection benchmark[C]. IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2016.

(下转第 125 页)

- Research and Applications , 2020 , 28(7) : 629–638.

[2] 陈敏曦.激活光伏“平价时代”[J].中国电力企业管理 , 2018(16) : 26–29.

[3] 江华 , 金艳梅 , 叶幸 , 等.中国光伏产业 2019 年回顾与 2020 年展望[J].太阳能 , 2020(3) : 14–23.

[4] RAKESH N , MALAVYA U. Maximizing the power output of partially shaded solar PV array using novel interconnection method[C].International Conference on Innovative Mechanisms for Industry Applications.IEEE , 2017.

[5] SAGAR G , PATHAK D , GAUR P , et al. A Su Do Ku puzzle based shade dispersion for maximum power enhancement of partially shaded hybrid bridge-link total-cross-tied PV array[J].Solar Energy , 2020 , 204 : 161–180.

[6] FEMIA N , GRANOZIO D , PETRONE G , et al. Predictive & adaptive MPPT perturb and observe method[J].IEEE Transactions on Aerospace and Electronic Systems , 2007 , 43(3) : 934–950.

[7] 张永革 , 石季英 , 张文 , 等.复杂遮阴条件下光伏系统 MPPT 控制改进 PSO 算法仿真研究[J].中国电机工程学报 , 2014 , 34(S1) : 39–46.

[8] 刘幸幸 , 王鼎.自动追光型便携式太阳能供电器[J].黑龙江科技信息 , 2015 , 19(10) : 51–52.

[9] 范英 , 华权冰 , 冯德思 , 等.太阳能便携式野外供电器的设计[J].无锡职业技术学院学报 , 2016 , 15(1) : 46–48.

[10] 张臻 , 沈辉 , 李达.局部阴影遮挡的太阳电池组件输出特性实验研究[J].太阳能学报 , 2012 , 33(1) : 5–12.

[11] 戚军 , 张晓峰 , 张有兵 , 等.考虑阴影影响的光伏阵列仿真算法研究[J].中国电机工程学报 , 2012 , 32(32) : 131–138.

[12] 孟安波 , 袁志坚 , 徐海波 , 等.复杂阴影条件下光伏阵列的建模与仿真分析[J].电气应用 , 2014 , 33(13) : 28–32.

[13] DÍAZ-DORADO E , CIDRÁS J , CARRILLO C. Discretized model for partially shaded PV arrays composed of PV panels with overlapping bypass diodes[J].Solar Energy , 2017 , 157 : 103–115.

[14] 张明锐 , 蒋利明 , 欧阳丽.局部阴影条件下光伏阵列旁路二极管和阻塞二极管的影响和作用[J].电气技术 , 2017(8) : 1–5.

[15] 吴亚盼 , 刘丁璞 , 李沛泽 , 等.新型光伏组件旁路二极管功能实验方法分析[J].电源技术 , 2020 , 44(5) : 714–717.

[16] MICHÁ , DANIEL N , RICARDO T , et al. The influence of solar spectrum and concentration factor on the material choice and the efficiency of multijunction solar cells[J]. Scientific Reports , 2019(9) : 1–13.

(收稿日期:2020-09-11)

#### 作者简介：

杨炀(1985-),男,博士,主要研究方向:电力电子技术、航天器电源系统仿真。

刘青(1985-),女,博士,副研究员,主要研究方向:航天器电源、特种电源。

王子才(1932-),男,博士,教授,主要研究方向:伺服系统、复合驱动控制系统及变阻尼大摩擦系统的控制技术等。

(上接第 115 页)

- [6] 吴纪芸,陈时钦.一种改进的MTCNN人脸检测算法[J].软件导刊,2019,18(12):78–81.
  - [7] HE K,ZHANG X,REN S,et al.Deep residual learning for image recognition[C].IEEE Conference on Computer Vision and Pattern Recognition(CVPR),2016.
  - [8] ZHANG Z,SONG Y,QI H.Age progression/regression by conditional adversarial autoencoder[C].IEEE Conference on Computer Vision and Pattern Recognition(CVPR),2017.
  - [9] 谢娟英,刘然.基于深度学习的目标检测算法研究进展[J].陕西师范大学学报(自然科学版),2019,47(5):1–9.
  - [10] 徐培超,陈雯柏,陈祥凤.多任务及Resnet网络在人脸多属性识别中的应用[J].小型微型计算机系统,2018,39(12):162–166.
  - [11] 庄妮.基于深度神经网络的人脸属性识别方法研究[D].厦门:厦门大学,2018.

- [12] 肖冰,王映辉.人脸识别研究综述[J].计算机应用研究,2005,22(8):1-5.
  - [13] 杜波.人脸识别中光照预处理方法研究[D].北京:中国科学院计算技术研究所,2005.
  - [14] 刘新业.基于人脸识别的人脸图像预处理研究[J].电子测试,2008(2):83-85.
  - [15] CHAUDHARY B.Tkinter GUI application development blueprints(Second Edition)[M].Packt,2018.
  - [16] 任进军,王宁.人工神经网络中损失函数的研究[J].甘肃高师学报,2018,23(2):65-67.
  - [17] 杨俊钦,张雨楠,林实锋,等.基于深度学习的人脸多属性识别系统[J].现代计算机(专业版),2019(5):54-57,61.

---

作者简介

王高升(1994-),男,硕士研究生,主要研究方向:计算机视觉、软件工程

## 版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所