

基于深度学习的在线字临摹分析系统设计*

张承强¹, 张永爱¹, 顾兴权²

(1. 福州大学 物理与信息工程学院, 福建 福州 350116;

2. 西安电子科技大学 通信工程学院, 陕西 西安 710000)

摘要: 为方便和快速地进行字体临摹分析, 该系统将纸面手写字与名人真迹字进行相似度比较, 使用残差网络 ResNet50 模型和新的字相似度算法对手写字进行高精度识别并与名人真迹字快速地进行相似度计算。将自制的名人书法字数据集和普通中文数据集合在一起训练 ResNet50 模型, 最后结合 Web 网站和 Android 开发了一个实时在线手写字与各名人书法字进行相似度比较的系统。Android 端主要用来上传纸面手写字照片和展示处理的结果, 搭建的 Web 网站用来对图片进行识别和相似度的计算与分析。

关键词: ResNet50 模型; 字相似度算法; 中文数据集; 在线比较系统

中图分类号: TP391

文献标识码: A

DOI: 10.19358/j.issn.2096-5133.2020.02.008

引用格式: 张承强, 张永爱, 顾兴权. 基于深度学习的在线字临摹分析系统设计[J]. 信息技术与网络安全, 2020, 39(2): 40-44, 56.

Design of online word copying analysis system based on deep learning

Zhang Chengqiang¹, Zhang Yongai¹, Gu Xingquan²

(1. College of Physics and Information Engineering, Fuzhou University, Fuzhou 350116, China;

2. College of Communications Engineering, Xidian University, Xi'an 710000, China)

Abstract: In order to expediently and quickly analyse the font copying, the system compares the similarity between the hand-written characters on paper and the original characters of famous people. Using the residuals network ResNet50 model and a new word similarity algorithm to recognize the handwritten words with high accuracy and calculate the similarity quickly with the famous handwritten words. ResNet50 model is trained by combining self-made famous calligraphy character data set and common Chinese data set. At last, a real-time online handwritten word similarity comparison system is developed with the combination of web site and Android. Android mainly uploads photos of handwritten words on paper and displays the processing results, and the built web site is used to identify images and calculate and analyze the similarity of pictures.

Key words: ResNet50 model; word similarity algorithm; Chinese dataset; online comparison system

0 引言

在我国的各种古代艺术门类中, 最具标志性, 最能集中地、典型地、鲜明地、持续而广泛地表现民族精神和时代精神的艺术, 非书法莫属^[1]。数字化的书法字临摹成为了很多大众的爱好的, 数字化的手写字识别往往和深度学习相结合。深度学习是新兴的机器学习研究领域, 旨在研究怎样从数据中自动地更好地提取多层特征表示, 其核心思想是通过数据驱动的方

式, 采用一系列的非线性(如激活函数)变换, 从原始数据中提取由低层到高层、由具体到抽象、由一般到特定语义的特征^[2]。ResNet (Residual Neural Network) 是由微软研究院的 He Kaiming 等四名华人提出的强大深度学习网络^[3-4], 在 ILSVRC2015 比赛中取得冠军, ResNet 的结构可以极快地加速神经网络^[5-6]的训练, 模型的准确率也有比较大的提升, 本文的深度学习网络采用的就是 ResNet 网络。

目前, 在线的名人书法临摹基本上都是照着某名人的书法临摹再比较相似度, 且都是触屏书写^[7-8]。触屏书写存在灵敏度强弱、屏幕大小和数

* 基金项目: 国家自然科学基金(61775038); 福建省自然科学基金(2017J01758)

字量化等因素而导致的手写字细节丢失问题,而本文的系统平台是在现实的纸等实物上进行书写后再拍照进行上传,有很多的细节能够保留且很方便,只要制作的名人书法字越精细,对细节的比较就越充分。此外,现今的在线名人书法临摹也限制了用户书写行为,只能对着名人的字进行临摹,本文的系统平台除此之外,还可以把日常的手写字与多位名人字进行比较,再展示与其最相似的名人真迹字并分析其结果。本文的平台可以面向大众娱乐,方便快捷地将自己的手写字与名人字比较。

1 深度学习

1.1 ResNet 原理

ResNet 借鉴了以往网络结构的优点,使用残差学习方法解决了在信息传递时出现的信息丢失、损耗等问题,整个网络只需要学习输入与输出不同的那一部分,使得网络学习的目标和难度得到简化。与传统的卷积神经网络相比,ResNet 提出了两种映射,恒等映射(identity mapping)和残差映射(residual mapping)。图 1 是一个简单的 Residual Block,恒等映射就是图 1 中的曲线(Shortcut Connection)跳过 2 层权重(层数是可变的,也可以是 3 层)后把 X 直接送到 2 层后的 ReLU 层的映射,之所以称为恒等,是因为 X 跳过了权重层,没有经过任何计算,即 $G(X) = X$;残差映射指普通网络原来的部分。

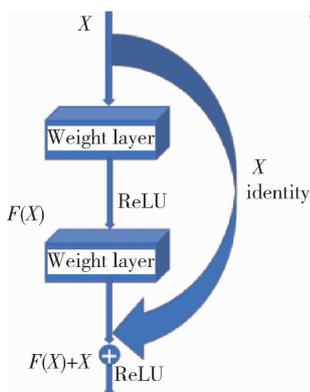


图 1 Residual Block

曲线部分使得每一个残差块很容易地学习到恒等映射结果,并且在反向传播时可以使梯度直接传播到更浅的层。在传统卷积神经网络上多次使用 Residual Block,就形成了残差网络(Residual Network, ResNet)。残差网络解决了传统极深网络的退化问题,让训练极深的网络成为可能^[9]。残差网络最常见的层数有 50 层和 101 层等。

1.2 ResNet50 模型

ResNet50 即共 50 层的残差网络,如图 2 所示,包括 conv1, conv2. x, conv3. x, conv4. x 和 conv5. x 卷积层,以及输出层。每个卷积层包含有 7×7 , 3×3 和 1×1 大小的数量不等的卷积核,以及池化(pool)层和激活(ReLU)层。

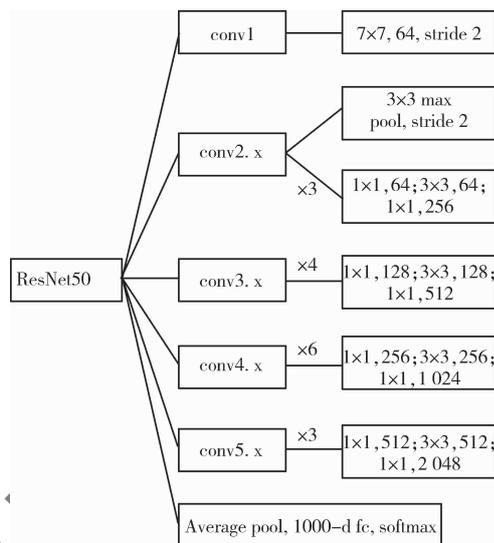


图 2 ResNet50 网络配置

2 ResNet50 模型训练与测试

本文的 ResNet50 网络使用目前流行的深度学习库 TensorFlow 搭建,首先需要对数据进行预处理,本文自制了名人书法字数据集并将其与普通中文字数据集合在一起,以增加模型训练的数据量;为了增强模型的泛化能力以提高准确率,本文将数据集进行数据增强,接着对图片进行大小尺寸的调整,本实验将图片尺寸调整为 224×224 的大小;TensorFlow 官方推荐的训练数据标准格式是 tfrecord 格式,此格式能够将图片数据和标签一起存储成二进制文件,能够在 TensorFlow 中实现快速地复制、移动、读取和存储操作,由于本文使用 TensorFlow 库进行模型搭建,可将上述的数据集制作成 tfrecord 格式^[10]。

数据预处理完成后,就开始搭建 ResNet50 网络,导入 TensorFlow 库和加载预训练模型,设置 batch_size、学习率和网络输入的大小,指定模型保存路径,损失函数选择 sigmoid 交叉熵,优化器选择 Adam,并设置其他相关参数^[11]。

设置完所有的参数后,对模型进行训练,如图 3 和图 4 所示,分别是模型训练过程中的精度曲线和损失值(loss)曲线。

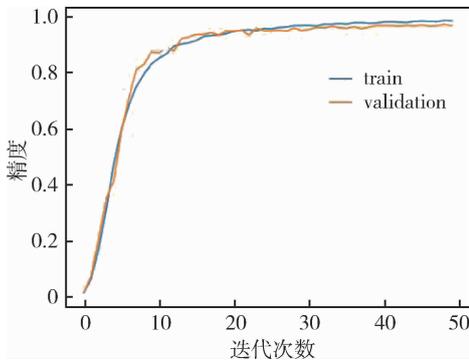


图3 迭代次数与准确率的关系图

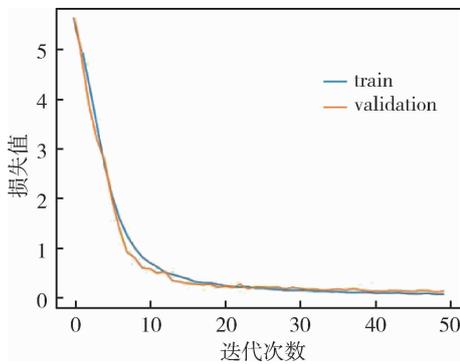


图4 迭代次数与损失值的关系图

3 字相似度算法

本文的字相似度算法比现有的传统字相似度算法简单,且符合直观感受,有易理解和计算量小等特点。

该算法分为以下几步,为便于论述,将手写字图片和名人书法字图片分别设为图片 A 和图片 B。

(1)将图片 A 和图片 B 灰度化。最初的图片是三通道的彩色图,需要将其灰度化为单通道图片以便下面只在单通道上二值化。

(2)自定义二值化。将图片 A 和图片 B 中的每个像素值大于设定的阈值(比如 100)的像素点全部设为 255,否则设为 0。对于白纸黑字的手写字,经过实验得出阈值取 100 能有很好的二值图,即小斑点和划痕相对较少。

(3)处理图片 A 和 B 中的细小斑点、小划痕、孤立点和噪音等像素,将其全部设为 255(此步骤是为了下一步截字精确)。

(4)分别将图片 A 和 B 中的字恰好截取出来,得到图片 A1 和 B1。精确截字是为了在计算字之间的相似度时减少背景的影响。

(5)将图片 A1 的尺寸设置为与图片 B1 相同,

得到图片 A2 和 B1。同尺寸后才能用步骤(7)中算法进行相似度的计算。名人书法字的尺寸不作改变是为了要保持其字的形状不变,所以将手写字设为与其同尺寸。

(6)分别将图片 A2 和 B1 颜色取反,得到图片 A3 和 B2。取反之后的图片变成了白字黑背景的图片,因为像素值为 255 表示白色,0 表示黑色,所以白字黑背景进行相似度计算时才是对字进行相似度比较,不然黑字白背景其实是对背景进行相似度计算。

(7)将图片 A3 和 B2 采用修改的均方误差算法^[12](即 PMSE)进行相似度计算,原理如式(1),得到的计算结果就是图片 A 和 B 的相似度值。

$$PMSE = 1 - \frac{1}{255^2 mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \quad (1)$$

其中, m 为图片的宽, n 为图片的高, $I(i,j)$ 和 $K(i,j)$ 分别表示两张不同图片在 (i,j) 点的像素值。

改进的均方误差算法的 Python 代码实现如下:

```
def Pmse(imageA, imageB):
    err = np.sum((imageA.astype("float") - imageB.astype("float")) ** 2)
    err /= float(imageA.shape[0] * imageA.shape[1] * 255 * 255)
    return 1-err
```

相似度算法的实际实现过程如图 5 所示。不过一般名人书法字最终的反色二值图会预先制作好存放在数据库中,这样做一方面是为了减少系统平台处理时间,另一方面名人书法字由于岁月的侵蚀可能会出现一些非理性的脏点,需要人工处理好再存于数据库中。

字相似度算法在不同手写字与不同名人真迹字之间的实际比较效果如图 6 所示,其中的名人真迹字和手写字是经过上述步骤(1)~(6)处理后的,组别 1 和组别 2 中的名人真迹字是相同的真迹字,手写字是不同的,从相似值的大小来看组别 1 更相似,也更符合人的主观感受;组别 2 和组别 3 的名人真迹字是不同的,手写字是相同的,从相似值的结果看组别 3 更相似,也更符合人的主观感受;组别 4 中的名人真迹字与手写字主观感受上看起来很像,但由于名人真迹字边缘过于粗糙而使相似度的值没有超过 0.9。通过实验认为相似度的值超过 0.88(即 88%)为手写字的结构基本准确(即相似)。

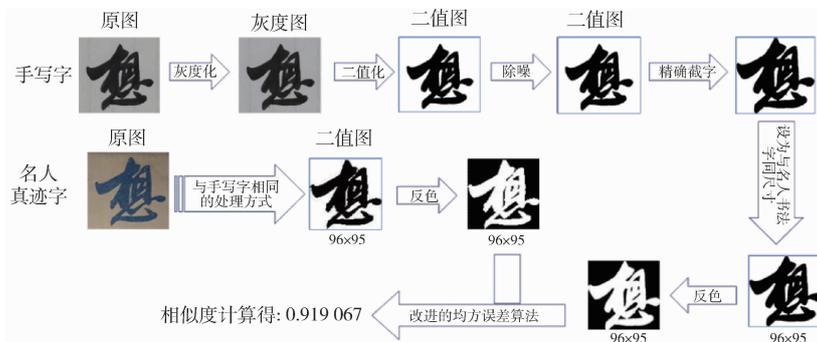


图5 字相似度计算实例流程图

组别	1	2	3	4
名人真迹字				
手写字				
相似值	0.910 621	0.762 924	0.900 055	0.886 246

图6 字相似度算法效果

本文的字相似度算法比书法字骨架提取、骨架相似度计算和点化形态相似度计算等综合性算法显然更简单和计算量更小,具有代码易实现、实时性明显和通用性(其他文字亦可)等特点,效果直观上无区别,优势明显。

4 在线 Android 平台开发

本文开发的基于深度学习的在线手写字与名

人书法字相似度比较 Android 平台,将 Android、Web 服务器、深度学习和字相似度算法结合在一起。整个平台有四大模块,分别为 Android 应用的图片上传模块^[13]、识别模块、字相似度计算模块和 Android 应用显示结果模块。Android 应用的图片上传模块:用户可以进行本地相册的图片上传或对手写字拍照后再上传;识别模块:上传的图片被 Web 网站接收并预处理后,通过已训练好的文字深度学习模型(即 ResNet50 模型)识别用户提交的图片,成功识别为某字;字相似度计算模块:在成功识别后,Web 网站^[14]会将上传的图片与数据库中各名人书法字的该字图片进行相似度的计算,并返回结果;Android 应用显示结果模块:显示内容包括用户提交的图片、最相似的名人真迹图片、相似值和分析结果等信息。在线 Android 平台结构如图 7 所示。

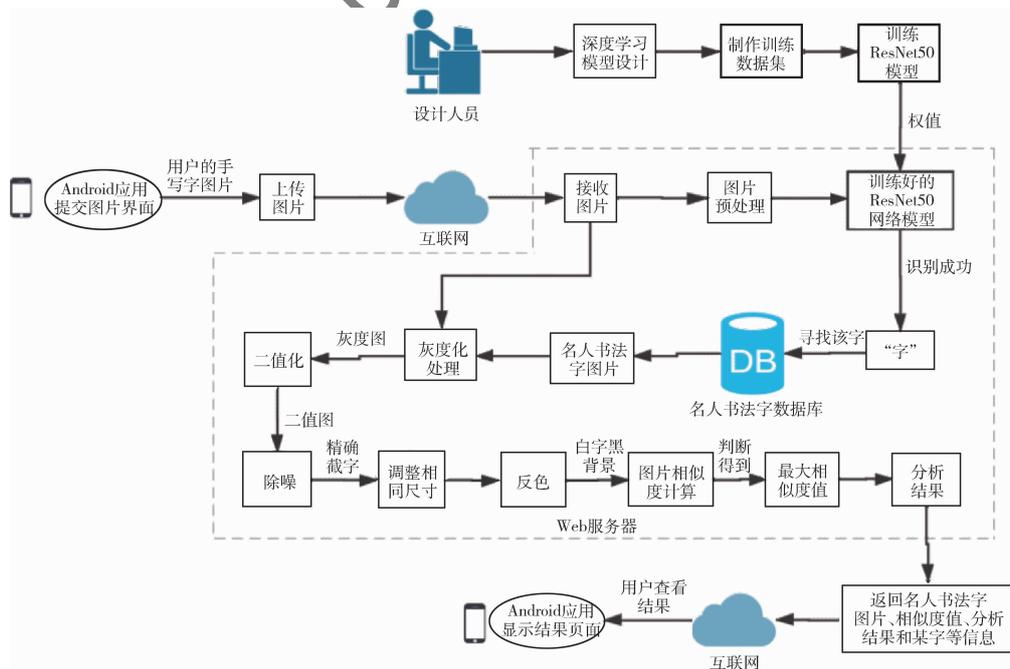


图7 在线 Android 系统平台

Android 应用的图片上传模块的显示界面如图 8 所示,这是一个用 HTML + CSS + Javascript 技术简单设计的网页页面^[15],Android 应用是一个定制的浏览器,Webview 方法已经设置了指定的 Web 服务器网址,在打开 Android 应用时会访问指定的 Web 网站,Web 网站将会响应设计好的网页页面代码,通过 HTTP 协议将代码发送到 Android 端,定制的浏览器会解析页面代码进行页面渲染。Android 应用的显示界面也可以做成微信小程序等界面^[16]。



图 8 Android 应用已选定图片的界面

Android 应用显示结果模块的显示界面如图 9 所示,图 8 上传手写字到 Web 服务器之后,服务器处理完成后会将处理结果返回给 Android 端,Android 界面会展示返回的处理结果,该界面展示的信息有识别到的手写字、相似度计算的结果、手写字图片和最相似的名人字图片以及简单评价等。



图 9 后台处理结果展示

5 结束语

本文将自制的名人真迹字数据集和普通中文数据集合在一起训练了 ResNet50 模型,训练好的 ResNet50 网络对手写汉字的识别精度非常高;改进的字相似度算法对手写字与名人真迹字的相似度计算的结果完全符合人的主观感受且计算量不大,适合做 Android 端与服务器端的在线相似度识别计算任务;最后将训练好的 ResNet50 模型和字相似度算法放在搭建好的 Web 网站上,再与 Android 应用

相结合,实现了一个实时在线手写字与名人真迹字的相似度比较系统。此外本平台若增加钢笔字和水笔等字的数据集,也可对钢笔和水笔的名人真迹字进行相似度的计算。

参考文献

- [1] 汤大民. 中国书法简史 [M]. 南京: 南京师范大学出版社, 2012.
- [2] 张军阳, 王慧丽, 郭阳, 等. 深度学习相关研究综述 [J]. 计算机应用研究, 2018, 35 (7): 1921-1928.
- [3] He Kaiming, Zhang Xiangyu, Ren Shangqing, et al. Deep residual learning for image recognition [C]. Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [4] XIE S N, GIRSHICK R, DOLLAR P, et al. Aggregated residual transformations for deep neural network [C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2017: 5987-5995.
- [5] ALBAWI S, MOHAMMED T A. Understanding of a convolutional neural network [C]. The International Conference on Engineering and Technology, 2017: 1-6.
- [6] NGUYEN K, FOOKES C, SRIDHARAN S. Improving deep convolutional neural networks with unsupervised feature learning [C]. Proceedings of IEEE International Conference on Image Processing, 2015: 2270-2274.
- [7] 常涛涛. 一种可变贴临摹练字装置: 中国, 201821088134. 3 [P]. 2019-10-29.
- [8] 广东欧珀移动通信有限公司. 在智能终端实现触屏笔画构造的方法及智能终端: 中国, 201310070640. 5 [P]. 2016-03-09.
- [9] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions [C]. Proceedings of International Conference on Computer Vision and Pattern Recognition, 2015: 1-9.
- [10] SEETALA K, BIRDSONG W, REDDY Y B. Image classification using tensorflow [M]. Springer International Publishing, 2019.
- [11] 李河伟. 一种移动式 TensorFlow 平台的卷积神经网络设计方法 [J]. 电脑知识与技术, 2017, 13 (22): 179-182.
- [12] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 105-114.

(下转第 56 页)

APP 中的不做改动。另外,在 IAP 和 APP 初始化过程中,文中所述获取中断处理函数地址的方法,APP 中还可以通过直接拷贝对应中断向量表中的内容来得到,但是在 IAP 中必须按照上面的方法,如果拷贝中断向量表的内容,得到的就是启动代码中 XXX_Handler PROC 的地址,这将导致程序进入死循环。

经过上面的设置后,分别将 IAP 和 APP 烧入对应的 Flash 空间,就可以实现两个空间的中断。如果用户有多个运行空间,也可以按照此方法来进行处理。

5 运行与测试

按照本文所述,在某国产 SWM240D8U7 单片机上分别建立了基于 CAN 通信的 IAP 和 APP 程序。程序上电后首先按照硬件默认设置从 0x0 处运行代码,该处代码完成 MCU 的启动,中断向量的定位,APP 应用标志的检测(检查 APP 代码段标志是否正常),如果一切正常则跳转到应用程序段执行,否则等待接收从 CAN 通信接口传送过来的升级数据包。经测试文章所述方法完全满足要求,在 IAP 和 APP 之间相互跳转的效果和基于硬件设置的单片机效果一致,达到了设计初衷。

参考文献

- [1] 意法半导体公司. STM32 in-application programming (IAP) using the USART[Z]. 2010.
- [2] 意法半导体公司. STM8L IAP 使用说明[Z]. 2007.
- [3] 意法半导体公司. 在 STM32L011 上通过 I²C 接口实现 IAP[Z]. 2012.
- [4] 意法半导体公司. STM32Cube 以太网 IAP 示例[Z]. 2015.
- [5] Synwit Technology Co., Ltd. SWM240 系列 MCU 数据手册[Z]. 2017.
- [6] Arm Co. Ltd. ARM ® Architecture Reference Manual [Z]. 2010.
- [7] 意法半导体公司. STM32F0x1 Reference Manual[Z]. 2012.
- [8] 意法半导体公司. 关于 AN4065 中 STM32F0 IAP 升级后的外部中断不响应问题[Z]. 2011.

(收稿日期:2019-12-24)

作者简介:

韩雨泓(1983-),男,硕士研究生,工程师,主要研究方向:信息融合与高速信号处理。

陈良勇(1977-),男,硕士研究生,工程师,主要研究方向:嵌入式系统应用。

(上接第 44 页)

- [13] Wang Jingna, Hua Guowei. Design of Android warehouse management software based on web service [J]. IOP Conference Series: Earth and Environmental Science, 2019, 252(4):1-7.
- [14] Yue Huanjing, Sun Xiaoyan, Yang Jingyu, et al. Landmark image super-resolution by retrieving web images [J]. IEEE Transactions on Image Processing, 2013, 22(12): 4865-4878.
- [15] 王成,李少元,郑黎晓,等. Web 前端性能优化方案与实践[J]. 计算机应用与软件, 2014, 31(12):89-95,147.
- [16] HAO L, WAN F C, MA N, et al. Analysis of the develop-

ment of WeChat mini program [J]. Journal of Physics: Conference Series, 2018, 1087(6):062040.

(收稿日期:2019-11-05)

作者简介:

张承强(1993-),男,硕士研究生,主要研究方向:深度学习、计算机视觉。

张永爱(1977-),通信作者,男,博士,研究员,博士生导师,主要研究方向:信息显示技术、集成电路。E-mail: yong- gaizhang@fzu.edu.cn.

顾兴权(1995-),男,硕士研究生,主要研究方向:深度学习、语音识别。

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《信息技术与网络安全》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）、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！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所