

# A Fast and Efficient Algorithm for Recognizing Cigarette Laser Security Code

Limin Chen, Tianlan Qi, Yinfei Du, Jianwei Zhao<sup>#</sup>

Department of Mathematics, College of Sciences, China Jiliang University, Hangzhou 310018, China

<sup>#</sup>Email: zhaojw@amss.ac.cn

## Abstract

A fast algorithm for recognizing the cigarette laser security code with high accuracy has been designed based on the idea of randomness. Firstly, it extracts the features of every character in the cigarette image code as the training and testing samples, using the histogram equalization, median filter, expansion and corrosion, and image segmentation. Secondly, it applies the random weight network to train the classifier, which helps to identify the cigarette image code. Finally, some experimental results show that the proposed algorithm based on the idea of randomness has faster recognition speed and higher recognition accuracy for identifying the cigarette laser security code than those methods based on BP, RBF and SVM algorithms.

**Keywords:** Laser Security Code Recognition; Random Weight Networks; Image Processing

## 一种快速有效的卷烟激光打码防伪识别算法<sup>\*</sup>

陈丽敏, 戚天兰, 杜殷飞, 赵建伟

中国计量学院 理学院数学系, 浙江 杭州 310018

**摘要:** 设计了一种快速、识别精度高的基于随机权思想的卷烟激光打码防伪识别算法。该算法首先利用直方图均衡化、中值滤波和膨胀与腐蚀、图像分割等预处理技术提取卷烟打码图像的单个字符特征作为训练样本和测试样本, 其次利用随机权网络算法训练分类器, 然后利用学习好的分类器识别卷烟打码图像。最后实验结果显示, 本文所提的基于随机权思想的卷烟激光打码防伪识别算法比基于 BP、RBF 和 SVM 算法的识别方法具有更快的识别速度和更高的识别精度。

**关键词:** 激光打码防伪识别; 随机权前向网络; 图像处理

## 引言

烟草行业是一个高税收高利润的行业, 在巨额利益的诱惑之下, 一些不法分子铤而走险、大肆制售假冒卷烟牟取暴利。为了有效打击假冒、走私、串货卷烟等违法行为, 维护国家根本利益和保障消费者的身体健康, 国家对烟草实行专卖。合法的烟草生产企业和经销商户会在烟盒上打上包含喷码日期、供货来源地、零售户性质、供货对象等信息的激光编码(见图 1), 以便执法人员通过卷烟喷码专用识别器了解该烟草销售的范围<sup>[1-2]</sup>。

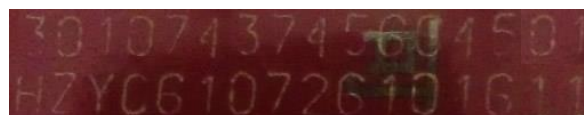


图 1 卷烟激光打码采集图

一般地, 卷烟喷码专用识别器包括三部分: 拍摄部分、信息处理部分和信息表述部分。其核心技术是信息处理部分, 即对所拍摄的图像进行识别, 最终输出信息字符串。现有的烟包喷印防伪数码识别技术的

<sup>\*</sup>基金资助: 受国家级大学生创新创业训练计划项目资助。

信息处理部分主要利用反向传播(BP)算法实现对图像特征的识别,如胡承东、速永仓等人采用 Zernike 矩阵特征和 BP 算法相结合的方法设计卷烟打码识别系统<sup>[3]</sup>。但由于 BP 算法是基于最速下降法,通过迭代算法来调节网络参数,所以 BP 算法具有收敛速度慢与容易落入局部极小等缺陷<sup>[4]</sup>。这大大增加了算法的计算成本和降低了算法的学习速度,从而影响识别效果。

文献[5-6]提出了快速、有效的前向神经网络训练算法,该算法的基本思想是随机选择隐层权值和阈值,而外权通过线性系统直接求解。由于该算法突破了传统迭代学习算法学习速度慢的问题,因此它具有很快的学习速度。本文针对卷烟激光打码识别技术存在的鉴别速度慢的问题,试图通过图像处理、智能计算中的随机权网络算法等,设计一种识别速度快、识别精度高的卷烟激光打码识别方法。

## 1 基于随机权网络的卷烟激光打码防伪识别方法

在本节中,我们将提出一种新的卷烟激光打码防伪识别方法,该方法由图像预处理、分析图像并提取图像特征和基于随机权网络分类器的识别算法组成,其步骤示意图见图 2。



图 2 本文提出的卷烟激光打码防伪识别算法的流程图

### 1.1 图像预处理

利用 1600 万像素的索尼 NEX-3N 数码相机拍摄 100 张不同光照、不同角度下的云南卷烟厂生产的红河烟的激光打码防伪图片,并用以下图像预处理步骤建立数字 0 至 9 以及字母 H、Z 的图像数据库:

- (1)将彩色激光打码防伪图片转化为灰度图,并采用灰度修正和直方图均衡化等方法进行图像增强处理<sup>[7]</sup>;
- (2)采用中值滤波和膨胀与腐蚀等方法进行图像降噪处理<sup>[7]</sup>;
- (3)利用阈值分割法进行图像二值化处理<sup>[7]</sup>;

(4)利用二值图像对 X 轴垂直投影进行图像分割提取单个字符(见图 3),各分割图像大小均为 50×50。把图片矩阵向量化,得到图片数据库,作为后续分类器的训练样本和测试样本。



图 3 预处理后得到的分割图像

### 1.2 设计基于随机权网络的图片分类器

通过图像预处理获得激光打码防伪识别图片中的各个数字特征图后,需要将它们输入到分类器中训练分类器,然后利用学习好的分类器识别激光打码图片。

给定训练样本集  $\{(x_i, t_i)\}_{i=1}^N \subseteq \mathbb{R}^{2500} \times \mathbb{R}$ , 则具有一个隐层节点的前向神经网络的输出可以表示为

$$f(x) = \sum_{i=1}^N \beta_i g(w_i \cdot x + b_i) + \beta_0, \quad x \in \mathbb{R}^{2500} \quad (1)$$

其中  $\beta_i \in \mathbb{R}$  是连接隐层第  $i$  个神经元与输出神经元的权值,  $w_i \in \mathbb{R}^{2500}$  和  $b_i \in \mathbb{R}$  是隐层权值参数,  $i=1, 2, \dots, N$ ,  $g: \mathbb{R} \rightarrow \mathbb{R}$  是网络的激活函数,  $w_i \cdot x$  表示向量  $w_i$  与  $x$  的内积。那么, (1)式中的神经网络逼近  $N$  个样本点  $\{(x_i, t_i)\}_{i=1}^N \subseteq \mathbb{R}^{2500} \times \mathbb{R}$  是指存在  $\beta_i, w_i, b_i (i=1, 2, \dots, N)$  和  $\varepsilon_j (j=1, 2, \dots, N)$ , 使得

$$\sum_{i=1}^N \beta_i g(w_i \cdot x_j + b_i) + \beta_0 = t_j + \varepsilon_j, \quad j=1, 2, \dots, N. \quad (2)$$

本文中的参数将利用随机权网络<sup>[6]</sup>中隐层权值随机选、外层权值通过求解线性系统的方法来解决。即先随机选择隐层权值  $w_i$  和  $b_i$ ，则 (2) 转化成关于外权的线性方程组。令

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \\ \beta_0 \end{bmatrix}, \quad T = \begin{bmatrix} t_1 \\ \vdots \\ t_N \end{bmatrix}, \quad E = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}, \quad H = \begin{pmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_N \cdot x_1 + b_N) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_N \cdot x_N + b_N) & 1 \end{pmatrix}_{N \times \tilde{N}},$$

则上述线性方程组可写成

$$H\beta = T + E. \quad (3)$$

当  $N = \tilde{N}$  且隐层输出矩阵  $H$  可逆时，外权  $\beta$  可以通过  $\beta = H^{-1}T$  来求解。在大多数情况下，训练样本个数远远多于隐层结点个数，此时，外权  $\beta$  可以通过

$$\hat{\beta} = H^+ T \quad (4)$$

求解，这里  $H^+$  是隐层输出矩阵  $H$  的 Moore-Penrose 伪逆<sup>[8]</sup>。目前为止，有很多方法计算矩阵的 Moore-Penrose 伪逆，包括正交化方法、正交投影方法、迭代方法和奇异值分解方法等等。当  $H^T H$  非奇异时， $H^+ = (H^T H)^{-1} H^T$ 。

本文提出的基于随机权网络的防伪识别算法的具体步骤如下，

### 1.3 卷烟激光打码防伪识别算法设计：

(1) 采集红河烟卷烟激光防伪码，利用灰度处理、图像增强、去噪、分割二值化等图像处理的方法提取图像特征，制作样本数据库（即数字 0~9 以及字母 H、Z、Y 的分割后的图像，共 13 组每组 20 张，以向量的形式存储）；

(2) 选取隐层含有  $N$  个神经元的前向神经网络，利用随机权网络的思想确定神经网络的权值，即随机选取隐层权值  $w$  和偏置值  $b$ ，并用线性方程组求解外权  $\hat{\beta} = H^+ T$ ；

(3) 利用训练好的分类器识别激光防伪码。

## 2 实验

本文进行的试验样本是利用 1600 万像素的索尼 NEX-3N 数码相机拍摄的 100 张不同光照、不同角度下的云南卷烟厂生产的红河烟的激光打码防伪图片，通过图像预处理得到 13 组每组 40 张分割后的字符图像（数字 0~9 以及字母 H、Z、Y 分割后的图像各 20 个，图像大小均为  $50 \times 50$ ）。各组随机选取 25 张图像数据作为训练样本，剩下的 15 张图像数据作为测试样本。

本文(1)中的隐层激活函数  $g$  取为

$$g(t) = \frac{1}{1 + e^{-t}}, \quad t \in \mathbb{R}. \quad (5)$$

下列表 1 是本文提出的基于随机权网络的识别算法与基于反向传播(BP)算法的识别方法在精度和速度上的比较结果。

从表 1 可见，本文算法在训练时间均比基于 BP 算法的训练时间来得少，接近于 0，且训练精度与识别精度也比 BP 算法高，均在 95% 之上，识别效果很好。当隐层神经元个数为 50 时，识别速度与精度达到最大。

表 2 是本文提出的基于随机权网络的识别算法与基于 BP、RBF<sup>[9]</sup> 和 SVM<sup>[10]</sup> 等算法的识别方法在精度和速度上的比较结果。

从表 2 可见，本文算法在识别精度和识别速度上都比现有的经典方法 BP、RBF 和 SVM 算法效果好，其次是 SVM 算法、BP 算法和基于聚类的 RBF 算法。

表 1 本文算法与 BP 算法在识别精度和识别时间上的比较

隐层神经元 个数 ( $N$ )	训练时间秒 (s)		训练精度 (%)		识别精度 (%)	
	BP算法	本文算法	BP算法	本文算法	BP算法	本文算法
10	0.532	0.022	90.4	92.9	95.7	97.7
30	0.592	0.011	91.4	95.3	94.8	96.3
40	0.643	0	87.7	98.6	92.6	95.3
50	0.663	0	91.3	98.8	95.5	99.5
60	0.695	0.001	92	97.5	94.7	99
70	0.712	0.001	90	97.3	94.9	99
80	0.878	0.001	92.7	96.4	95.4	99.1

表 2 本文算法与 BP、RBF 和 SVM 算法在识别精度和识别时间上的比较

训练时间 (s)				识别时间 (s)			
BP算法	基于聚类的 RBF算法	SVM算法	本文算法	BP算法	基于聚类的 RBF算法	SVM算法	本文算法
0.663	0.71	0.201	0	0.224	0.25	0.175	0
训练精度 (%)				识别精度 (%)			
BP算法	基于聚类的 RBF算法	SVM算法	本文算法	BP算法	基于聚类的 RBF算法	SVM算法	本文算法
91.3	78.4	93.1	98.8	95.5	91.2	97.4	99.5

### 3 结语

本文针对卷烟激光打码防伪识别问题设计了一种快速、识别精度高的算法。该算法首先利用直方图均衡化、中值滤波和膨胀与腐蚀、图像分割等预处理技术提取卷烟打码图像的单个字符特征作为训练样本和测试样本，其次利用随机权网络算法训练分类器，然后利用学习好的随机权前向网络识别卷烟打码图像。最后仿真实验结果显示，本文所提的基于随机权网络的卷烟激光打码防伪识别算法比基于 BP、RBF 和 SVM 算法的识别算法具有识别速度更快、识别精度更高的优点。

### REFERENCES

- [1] Anonymous. Coding for the cigarette handles security identification ID [J]. Chinese security reports, 2010, 18-19.
- [2] International tobacco-Oriental Tobacco Network, [www.eastobacco.com/gjyc/](http://www.eastobacco.com/gjyc/).
- [3] C.D. Hu, Y.C.Su, L.Z.Cao. Coding cigarette character recognition system [J]. Mechanical Engineering and Automation, 2010(3): 117-119.
- [4] D. E. Rumelhart, G. E. Hinton, R. J. Williams. Learning representations by backpropagating errors [J]. Nature, 1986, 323: 533-536.
- [5] Y. H. Pao, G. H. Park, D. J. Sobajic. Learning and generalization characteristics of the random vector functional-link net [J]. Neurocomputing, 1994, 6(2): 163-180.
- [6] W. F. Schmidt, M. A. Kraaijveld, R. P.W. Duin. Feed forward neural networks with random weights [J]. Proceedings 11th IAPR International Conference on Pattern Recognition. Vol. II. Conference B: Pattern Recognition Methodology and Systems, 1992, pp.1-4.
- [7] Y.Wang, X.Zhou. Face Recognition for image preprocessing [J]. Science and Technology Information, 2011, 16: 94-94.

- [8] C. R. Rao, S. K. Mitra. Generalized inverse of matrices and its applications [M]. 1971, Wiley, New York.
- [9] J. E. Meng, S. Q. Wu, J. W. Lu, H. L. Toh. Face recognition with radial basis function (RBF) neural networks [J]. IEEE Transaction on Neural Networks, 2002, 13(3): 697-710.
- [10] Y. Fu, Y. W. Wang, W. Q. Wang, W. Gao. Content-based natural image classification and retrieval using SVM [J]. Chinese Journal of Computer, 2003, 26(10): 1261-1265.

## 【作者简介】



<sup>1</sup> 陈丽敏（1991-），女，汉，本科生，  
专业：数学与应用数学，中国计量学院  
理学院。Email: 253258084@qq.com



<sup>2</sup> 戚天兰（1991-），女，汉，本科生，  
专业：数学与应用数学，中国计量学院  
理学院。Email: 1391778601@qq.com



<sup>3</sup> 杜殷飞（1991-），男，汉，本科生，  
专业：数学与应用数学，中国计量学院  
理学院。Email: 354957997@qq.com



<sup>4</sup> 赵建伟（1977-），女，汉，博士，教  
授。研究方向：智能算法、模式识别。  
学习经历：2006 年博士毕业于中科院系  
统与科学研究院，现任教于中国计量学  
院理学院数学系。  
Email: zhaojw@amss.ac.cn