

模糊会计账目与票据的数字辨识研究

陈 清

曲阜市老年大学, 山东 曲阜

收稿日期: 2022年10月18日; 录用日期: 2023年3月7日; 发布日期: 2023年3月15日

摘 要

会计账目及票据是财务事务的根本依据, 因此它的准确性和清晰度至关重要, 甚至涉及法律问题。手写体造成的难以辨认识别的票据、污损票据等情况时常出现。于是, 如何准确识别模糊会计账目、票据数字成了财务部门一个重要的难题, 同时也是审计部门与执法部门的一个课题。需要识别故意污损会计账目和票据也是审计部门与执法部门的基本功。本文给出了基于Hopfield神经网络理论的数字识别原理、步骤及算法, 并给出了详细的运算程序, 可以为企事业单位和相关政府部门在理论和实践上对财务数据的准确性提供支持及预警。

关键词

财务票据, 数据识别, 神经网络

Research on the Digital Identification of Fuzzy Accounting Accounts and Bills

Qing Chen

Qufu University for the Aged, Qufu Shandong

Received: Oct. 18th, 2022; accepted: Mar. 7th, 2023; published: Mar. 15th, 2023

Abstract

Accounting accounts and bills are the fundamental basis of financial affairs, so its accuracy and clarity is very important, and even involves legal issues. Handwriting caused by difficult to recognize the bill, defaced bills and other situations often occur. Therefore, how to accurately identify the fuzzy accounting accounts, bills in the number of an enterprise can not be ignored, it is also an audit and law enforcement departments of a subject. The need to identify intentional defacement of accounting accounts and notes is also a fundamental skill of auditing and law enforcement departments. In this paper, the principle, procedure and algorithm of number recognition based on

Hopfield neural network theory are given, and the detailed operation program is also given, it can provide support and early warning for enterprises and relevant government departments to the accuracy of financial data in theory and practice.

Keywords

Financial Bills, Data Identification, Neural Network

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

票据是重要的商业凭证，是市场经济的产物，它便利了市场交易和财富的转移，极大地促进了市场经济的发展。折叠的票据、沾水的票据，或者手写体造成的难以辨认识别的票据，或者故意部分污损的票据，这些在现实社会中会经常碰到。这些财政票据是事业单位、国家权利机关等从事财务活动或根据规定征收非税收入而向特定对象开具的凭证，这些凭证与税务发票一起构成了会计事物的重要凭证，同时也是审计、财政等监督部门有效开展监督检查活动的依据。

会计账目是财务部门用来按照不同类别不间断记录经济财务情况，反映会计条目动态变化及其结果的凭证与证据。该类账目，如资产类或负债类等，出现污损，沾染油污；日记账中的现金帐或银行帐目因为折叠造成破损或部分断裂时，就涉及该类问题中的模糊数字识别问题。

模糊数字识别问题需要科学的方法。而 Hopfield 神经网络因为它部分具有类人脑记忆的特点，所以已被大量应用于图片、数字及面部识别等各类识别问题中[1] [2]。Hopfield 神经网络有离散型和连续型之分，前者应用更为广泛和成熟，是一种单层的递归神经网络，网络中所有神经元都有联系，已知和结果也有反馈关系，又称为全信息网络[3]。其结果值总会作为网络的已知值，在激励函数的作用下，再产生新的输出，如此循环往复，于是网络状态会发生不断的变化。如果它是稳定的，则这个迭代的运算过程所导致的已知与结果间的差值会越来越小，当它们相等时，迭代运算结束。Hopfield 本人一开始给出的是一个二值神经网络，其选择的激励函数为阶段函数，神经元的已知、结果是离散型的，如取{0, 1}或者{-1, 1}。因此，所得结果的离散值 1 和 0 或者 1 和-1，它们分别表示神经元处于工作情形和休息情形。每个神经元都可取值 1 或-1，当某个神经元所受的外部影响大于发生变化的临界点时，神经元就取值 1，否则神经元为-1。

在日常生活中，经常遇到带噪声(干扰)的字符识别问题，如汽车车牌受污染等造成的字体模糊不清，录像难以辨认。手写体造成的难以辨认识别的票据、污损票据等情况也时常出现。于是，如何准确识别模糊会计账目、票据数字就成了财务部门一个需要重视的问题[4] [5]，同时也是审计部门与执法部门的一个课题。需要识别故意污损会计账目和票据也是审计部门与执法部门的基本功。本文基于文献[6] [7]给出的神经网络及其稳定与控制理论，得到了基于 Hopfield 神经网络理论[8] [9] [10]的数字识别原理、步骤及算法，并给出了详细的运算程序，可以为企业和相关部门在理论和实践上对财务数据的准确性提供支持及预警。

2. 神经网络数学模型

离散 Hopfield 神经网络模型分为两类：

1) 串联模式：处于某个时刻 t 时，只有神经元 j 动了，而其他神经元不动，即：

$$y_j(t+1) = f\left(\sum_i w_{ij} y_i(t) + x_j - \theta_j\right);$$

$$y_j(t+1) = y_i(t), i \neq j$$

2) 并联模式：在任一时刻 t ，所有的神经元的状态都动了

$$y_j(t+1) = f\left(\sum_i w_{ij} y_i(t) + x_j - \theta_j\right);$$

$$j = 1, 2, \dots, n$$

其中 θ_i 为第 i 个神经元的阈值。下面给出一个具有部分人脑记忆特点的离散型 Hopfield 神经网络模型，其功能与特性是能准确的辨识阿拉伯数字，并且当这些阿拉伯数字被外界噪声深度影响和干扰后，也会有良好的辨识效果。

我们选取矩阵直观地模拟这 10 个阿拉伯数字，即把每个数字都写成矩阵的形式，有数字的黑色部分为 1，无数字的空白部分为-1。当带有噪声干扰的数字矩阵作为网络的输入时，网络的输出是最接近准确数字的目标矩阵，从而完成数字的识别任务。例如准确的数字 1，3 的矩阵表示分别为：

在程序中数字 1 和 3 的代码分别是

array_one=

```
[-1 -1 -1 1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1;
-1 -1 -1 1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1;
-1 -1 -1 1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1];
```

array_three=

```
[-1 1 1 1 1 1 1 1 -1; -1 1 1 1 1 1 1 1 -1; -1 -1 -1 -1 -1 -1 1 1 -1;
-1 -1 -1 -1 -1 -1 1 1 -1; -1 1 1 1 1 1 1 1 -1; -1 -1 -1 -1 -1 -1 1 1 -1;
-1 -1 -1 -1 -1 -1 1 1 -1; -1 1 1 1 1 1 1 1 -1; -1 1 1 1 1 1 1 1 -1];
```

其他数字的矩阵与代码可以类似给出。

带噪声干扰的数字矩阵是指以上代码矩阵的某些值发生了变化，模拟这种变化的方法很多，这里仅仅选择最常用的确定噪声法和随机噪声法。将受到噪声干扰的矩阵输入到创建好的 Hopfield 神经网络中。下面从 Matlab 神经网络工具箱中选取 Newhop() 函数用于创建该网络，其命令为：

```
net=newhop(B)
```

以上字母 B 是一个矩阵，net 为生成的神经网络。其激活函数选取 satlins()，而仿真 Sim() 函数用于图像模拟仿真，其命令为

```
[Y, Af, E, perf]=sim(net, P, Ai, B)
```

```
[Y, Af, E, perf]=sim(net, {R BN}, Ai, B)
```

以上字母 P, R 为已知向量，Y 为结果。下面通过例子详细说明它们的用法。

本文接下来选择阿拉伯数字 1 和 3 (其他的阿拉伯数字类似)，首先将它们写成矩阵演练样本 B：

```
%% 训练样本，
```

```
B=[array_one, array_three];
```

```
%% 创建离散型 Hopfield 神经网络
```

```
net=newhop(B)
```

这里先用人为篡改的办法篡改某些值，达到模拟带干扰噪声的目的。例如，经过人工修改后的数字 1 和数字 3 的带噪声点阵变为，

```
%% 确定噪声
noisy_array_one=
[-1 -1 -1 1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1; -1 1 -1 1
1 1 -1 -1 -1; -1 -1 -1 -1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 -1 -1 -1;
-1 -1 -1 -1 1 1 -1 -1 -1; -1 -1 -1 1 1 1 1 -1 -1];
noisy_array_three=
[-1 1 1 1 1 1 1 1 -1; -1 1 1 1 1 1 1 1 -1; -1 -1 -1 -1 -1 1 1 1 -1; -1
-1 -1 -1 -1 1 1 -1; -1 1 1 1 1 1 1 1 -1; -1 -1 -1 -1 1 -1 1 1 -1; -1 1 -1 1
-1 -1 1 1 -1; -1 1 1 1 1 1 1 1 -1; -1 1 1 1 1 1 1 1 -1];
```

以上方法的缺点是需要人工做多次的修改。而随机噪声产生法可以比较方便达到模拟各种带噪声的矩阵。这种办法通过随机的方法来找出需要修改的地方，而后再对该地方实行修改。因为点阵中的数字只有 1 和 -1，所以修改它们只能互换来完成。带干扰噪声的阿拉伯数字 1 和 3 的图像随机过程如下

```
%% 随机噪声
noisy_array_one=array_one;
noisy_array_three=array_three;
for i=1:100
a=rand;
If a<0.1
noisy_array_one(i)=-array_one;
noisy_array_three(i)=-array_three;
end
End
```

3. 仿真图形与结论

利用以上准备，便可以对带干扰噪声的数字矩阵进行辨识，其程序如下

```
%% 仿真测试
noisy_one={(noisy_array_one)'};
identify_one=sim(net,{9,9},{},noisy_one);
identify_one{9}'
noisy_three={(noisy_array_three)'};
identify_three=sim(net,{9,9},{},noisy_three);
identify_three{9}'
```

确定噪声的结果如图 1 所示。

图 2 所示是干扰度为 0.1 (即 10% 的数字图像动了，其他没动) 时的辨识结果，图 2 第 3 行显示辨识结果较好。继续研究发现，随着干扰度的提高，数字图像的辨识效果逐渐模糊。其结果表明，当干扰越强时，这里的 Hopfield 神经网络方法已经很难奏效。

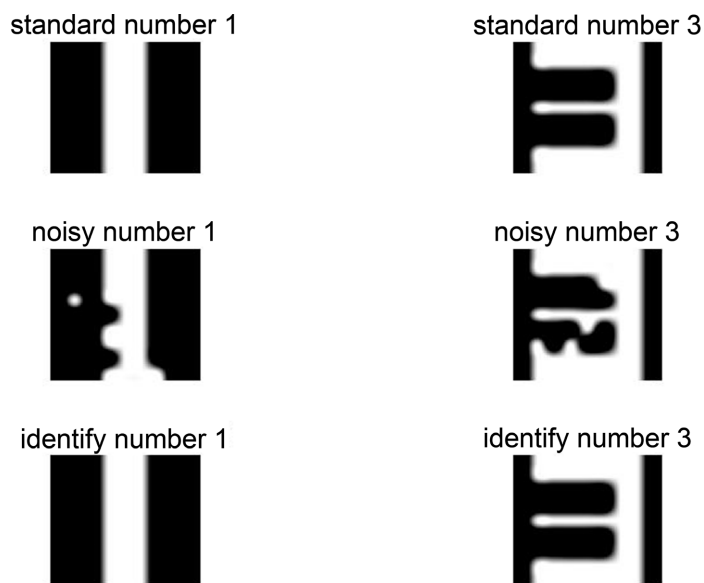


Figure 1. Determine noise figure

图 1. 确定噪声图

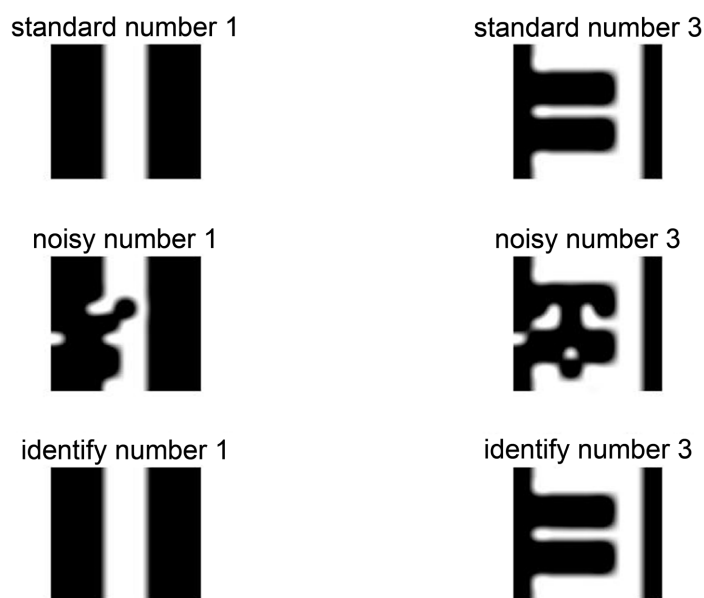


Figure 2. Random noise figure

图 2. 随机噪声图

近年来越来越多的学者和应用技术人员把 Hopfield 神经网络应用到各个领域[11] [12], 因此可以在一定程度上有效弥补传统方法的缺陷。但是通过上面也可以看出, 当干扰达到一定的程度时, 会出现收敛到错误的局部极小值, 而非全局极小值问题。于是, 寻求不同方法的结合, 如何扬长避短成了研究的热点。将一些优化算法与本文的方法相结合, 会使其联想记忆特点更显现, 应用实践性更强。例如, 遗传算法具有的全局观可以较好地帮助神经网络跳出伪稳定点的问题。

基金项目

山东省自然科学基金项目(ZR2021MA043)。

参考文献

- [1] 董长虹. Matlab 神经网络与应用[M]. 第 2 版. 北京: 国防工业出版社, 2007.
- [2] 崔永华, 左其亨. 基于 Hopfield 网络的水质综合评价及其 Matlab 实现[J]. 水资源保护, 2007, 23(3): 14-16.
- [3] 崔泽乾, 白云杰, 韩阳, 陈丽芳. 基于 Hopfield 网络的手写体汉字识别系统[J]. 华北理工大学学报(自然科学版), 2021(43): 123-131.
- [4] 谢志华, 谢昊宇. 对现金流量表若干理论问题的反思[J]. 财会月刊, 2022(15): 15-121.
- [5] 冯春龙. 企业现金流量管理改进策略探究[J]. 中国集体经济, 2022(19): 151-153.
- [6] 焦李成. 神经网络系统理论[M]. 西安: 西安电子科技大学出版社, 1990.
- [7] 郭英新. 非线性随机时滞神经网络的稳定性分析与脉冲镇定[M]. 北京: 科学出版社, 2017.
- [8] Guo, Y. and Cao, J. (2019) Stability and Stabilization for Stochastic Cohen-Grossberg Neuralnetworks with Impulse Control and Noise-Induced Control. *International Journal of Robust and Nonlinear Control*, **29**, 153-165.
<https://doi.org/10.1002/rnc.4379>
- [9] Guo, Y. (2018) Globally Robust Stability Analysis for Stochastic Cohen-Grossberg Neural Networks with Impulse Control and Time-Varying Delays. *Ukrainian Mathematical Journal*, **69**, 1049-1060.
<https://doi.org/10.1007/s11253-017-1426-3>
- [10] Guo, Y. (2017) Exponential Stability Analysis of Traveling Waves Solutions for Nonlinear Delayed Cellular Neural Networks. *Dynamical Systems—An International Journal*, **32**, 490-503.
<https://doi.org/10.1080/14689367.2017.1280447>
- [11] 代小红. 模糊模式的手写数字识别技术研究与实现[J]. 重庆大学学报, 2011, 34(6): 117-122.
- [12] 符哲夫. 基于注意力机制的模糊数字识别研究[J]. 计算机应用及软件, 2022, 39(12): 208-212, 259.