

# 基于主成分分析-BP神经网络的文物识别

王明娟, 刘风云, 林富强

上海理工大学, 上海

收稿日期: 2022年9月28日; 录用日期: 2022年10月28日; 发布日期: 2022年11月8日

## 摘要

当下中国文物面临的重大灾难, 不是所谓的文物外流, 而是文物的卖假。从金缕玉衣到汉代玉凳, 一个个的赝品将文物鉴定权威击垮。文物造假为中国带来了文物危机, 扭曲了博大精深的传统艺术, 故建立合适的文物样品识别模型尤为重要。基于此, 本文选取玻璃文物样品作为研究对象, 对其化学成分进行数学分析。首先利用加权平均预测法对受损文物的化学成分数据进行修正, 并借助斯皮尔曼相关系数对文物样品的化学成分进行相关性分析; 其次依据主成分分析数学原理提取出了文物样品的前5个主成分; 最后结合BP神经网络, 对样本空间进行重构, 构建了主成分分析-BP神经网络玻璃文物识别模型, 达到了对输入空间降维的目的。经检验, 文物模型识别的效果极好。

## 关键词

主成分分析, 斯皮尔曼相关系数, BP神经网络, 加权平均预测法

# Recognition of Cultural Relics Based on Principal Component Analysis and BP Neural Network

Mingjuan Wang, Fengyun Liu, Fuqiang Lin

University of Shanghai for Science and Technology, Shanghai

Received: Sep. 28<sup>th</sup>, 2022; accepted: Oct. 28<sup>th</sup>, 2022; published: Nov. 8<sup>th</sup>, 2022

## Abstract

At present, the major disaster facing China's cultural relics is not the so-called outflow of cultural relics, but the counterfeiting of cultural relics. From golden strands of jade clothes to Han Dynasty jade stools, fakes one by one knocked down the cultural relics identification authority. The falsification of cultural relics has brought a cultural relic crisis to China and distorted extensive and

profound traditional art, so it is particularly important to establish a suitable cultural relic sample identification model. Based on this, this article selects a batch of glass samples. Firstly, the chemical composition data of damaged cultural relics are modified by weighted average prediction method, and the correlation of chemical composition of cultural relic samples is analysed with the help of Spearman correlation coefficient. Secondly, according to the mathematical principle of principal component analysis to extract the sample of first five principal components of cultural relics. Finally, combined with the BP neural network, a principal component analysis (PCA)-BP neural network model for cultural relics recognition is constructed, by reconstructing the sample space to achieve the purpose of dimension reduction in the input space, so as to further simplify the network structure. After testing, the effect of model recognition is excellent, which greatly improves the accuracy of the cultural relic recognition model.

## Keywords

Principal Component Analysis, Spearman Correlation Coefficient, BP Neural Network, Weighted Average Prediction Method

Copyright © 2022 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

近年来, 由于生产工艺的快速发展, 文物交易市场充斥着大规模数量的仿造品, 仿造品的出现不仅严重破坏了文物交易市场的秩序, 也在很大程度上破坏了文物的传承与保留[1]。文物鉴定作为一门专业性极强的学科, 目前学者对文物鉴定具体方法的研究大多停留在理论方法阶段, 这种主观判断方式使得一些非法鉴定机构有机可乘, 故建立科学的文物识别模型对文物收藏及考古研究都具有重要意义。唐迎菲(2015)提出分类鉴定法与综合考察法来对文物样品进行鉴定, 并探讨了现代科学技术在文物鉴定中的应用[2]。王迪(2019)提出了利用微信等新媒体在书画文物鉴定中的应用, 使文物鉴定活动不再局限于某一特定场景, 以期提高书画文物鉴定的水平[3]。李擎(2019)指出三维成像技术和激光与光电透视技术这类现代技术对于文物鉴定效率的提高具有极大意义[4]。段继霞(2020)认为可以将图像识别技术与 AR 技术相结合, 利用边缘检测法来提取文物主题图像, 为用户展示更生动的文物信息[5]。郭美玉(2021)提出文物鉴定方法应紧跟时代步伐, 同时应重视新技术科技在该领域的应用[6]。

纵观以上文献, 不难发现文物识别方式逐渐由理论识别到科学技术识别转变。本文在此基础上将理论知识、科学技术与数学分析相结合, 以玻璃为例, 选取玻璃文物制品进行化学成分分析, 构建了基于主成分分析-BP 神经网络的玻璃文物识别模型, 为玻璃文物样品的鉴定提供了科学的数学依据。玻璃作为古代文化交流的物证, 其象征着中西方文化交流史上的一颗明珠, 对未知玻璃化学成分体系的确定以及类型识别是中国玻璃研究中极为重要的一个领域[7]。研究玻璃文物类型的鉴定不仅可以对中国玻璃文化有更深入的了解, 同时也可以应用数学理论来探索文物背后的经济价值与文化底蕴, 以期为文物交易市场的规范化与标准化发展提供参考。

## 2. 文物识别模型准备

### 2.1. 数据选取

本文选取了一批我国古代玻璃制品的相关数据, 共计 68 个样本, 其类型主要分为高钾玻璃和铅钡玻

璃两种，其中部分样本由于表面风化而受损，后续将对这部分数据进行修复。各个样本的主要成分与对应所占比例已知，共计 14 种化学成分，将依据这些数据进行数学分析。

### 2.2. 数据预处理

查阅文献得知，玻璃文物样品检测的各成分累加之和应介于 85%~105%之间，不在该范围内的样品不符合研究条件，故需要对样本数据进行筛选。最终剔除两个样本数据，编号为 15 和 17，剔除不符合要求的数据后，共计 66 条有效样本数据。

### 2.3. 数据修正

#### 加权平均预测模型

考虑到受埋藏环境的影响，部分玻璃文物表面受损(出现风化)，导致其化学成分比例发生变化，造成文物样品中部分受损玻璃的化学成分比例出现误差，影响了对玻璃类别的准确判断，故需要对受损玻璃文物化学成分比例数据进行修正。本文采取加权平均对受损玻璃受损前的化学成分进行预测，以得出更为准确的判断结果。

首先利用 SPSSPRO 软件对高钾类和铅钡类文物数据进行整合。接着，对高钾玻璃文化与铅钡玻璃文化风化前后的化学成分进行频率分析，使用 Matlab 求解，得到的结果如图 1 所示：

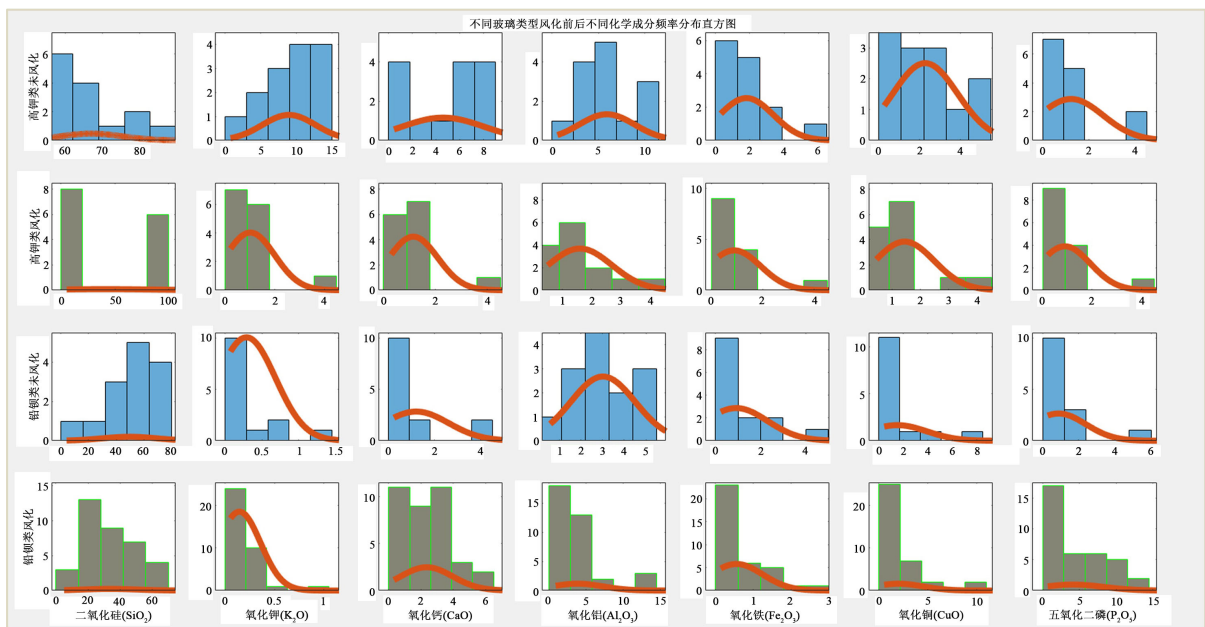


Figure 1. Histogram of frequency distribution of chemical composition of different types of glass before and after damage

图 1. 不同玻璃类型分化前后化学成分分析的频率分布直方图

在上图频率直方图的基础上进一步分析，发现部分样品化学成分未被检测到，即含量为“0”值，这些数据对我们进行预测会造成一定阻碍，因此在这里对数据进行加权平均值处理，并使用标准正态分布函数对其进行权重的分配[8]，标准正态分布曲线函数公式如下：

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (1)$$

其中  $x$  为连续型随机变量。具体计算过程如下：

按顺序定义风化之前二氧化硅( $\text{SiO}_2$ )、(氧化钠)( $\text{Na}_2\text{O}$ )、氧化钾( $\text{K}_2\text{O}$ )……二氧化硫( $\text{SO}_2$ )等共 14 个变量分别为  $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_{14}$ ，定义其权重为  $\omega_1, \omega_2, \omega_3, \dots, \omega_{14}$ ，同理将风化后化学成分定义为  $\alpha_i^*$ ，权重为  $\omega_i^*$ 。接着将化学成分的数据与标准正态分布函数相乘，计算其加权平均值如下：

$$\begin{cases} \bar{a} = \frac{\alpha_1\omega_1 + \alpha_2\omega_2 + \alpha_3\omega_3 + \dots + \alpha_{14}\omega_{14}}{\omega_1 + \omega_2 + \omega_3 + \dots + \omega_{14}} \\ \bar{a}^* = \frac{\alpha_1^*\omega_1^* + \alpha_2^*\omega_2^* + \alpha_3^*\omega_3^* + \dots + \alpha_{14}^*\omega_{14}^*}{\omega_1^* + \omega_2^* + \omega_3^* + \dots + \omega_{14}^*} \end{cases} \quad (2)$$

再定义其风化前后每个化学成分的加权占比指标为  $\mathcal{G}_i$ ，则：

$$\mathcal{G}_i = \frac{\bar{a}}{\bar{a}^*} \quad (3)$$

利用该方法对受损后文物样品受损前的化学成分进行分析，以此得到更为准确的文物化学成分含量，提高后续文物识别模型的精度。

## 2.4. 化学成分相关数据分析

### 斯皮尔曼相关系数

对于相关系数的选择，由于本文所研究的化学成分变量不满足正态分布与线性关系，故采用斯皮尔曼相关系数进行相关性分析[9]，具体定义如下：对任意两个化学成分变量进行分析，对应列向量记为  $Y$ 、 $X$ ；将各个列向量对应的元素  $Y_j$  与  $X_j$  转换为对应列向量的排序，记为  $T(Y_j)$  和  $T(X_{ij})$ ；列向量的长度记为  $N$ 。利用公式分别计算列向量  $Y$  列向量  $X_i$  对应元素  $T(Y_j)$  和  $T(X_{ij})$  之间的差异并分别求和，得到化学成分等级之间的差值  $d_i$ ：

$$d_i = \sum_{i=1}^N |T(Y_j) - T(X_{ij})|^2 \quad (4)$$

依据下面公式计算各化学成分之间的相关性  $R_i$ ：

$$R_i = 1 - \frac{6 \times d_i}{N \times (N^2 - 1)} \quad (5)$$

## 3. 主成分-BP 神经网络识别模型的构建

### 3.1. 主成分分析数学模型

当前对玻璃类别的划分依据主要依据文物样品的化学成分，本文采用主成分分析法分别提取不同类别文物样品化学成分中的主成分。用  $x_1, x_2, \dots, x_{14}$  分别表示已知文物样品中的各类化学成分，用  $i=1, 2, \dots, 58$  分别表示文物编号，第  $i$  种文物  $x_1, x_2, \dots, x_{14}$  的取值记作  $[a_{i1}, a_{i2}, \dots, a_{i14}]$ ，构造矩阵  $A = (a_{ij})_{17 \times 5}$ ，具体步骤如下：

**Step 1:** 对玻璃文物的数据进行标准化处理，将各指标值  $a_{ij}$  换成标准化指标  $a_{ij}^*$ ；

$$a_{ij}^* = \frac{a_{ij} - \mu_j}{s_j} \quad (6)$$

其中  $i=1, 2, \dots, 58$ ；  $j=1, 2, \dots, 14$ ，  $s_j, u_j$  为第  $j$  个指标的样本标准差和样本均值。

$$u_j = \frac{1}{58} \sum_{i=1}^{17} a_{ij} \quad (7)$$

$$s_i = \sqrt{\frac{1}{58-1} \sum_{i=1}^{58} (a_{ij} - \mu_j)^2} \quad (8)$$

相应的称  $x_j^*$  为标准化指标变量。

$$x_j^* = \frac{x_j - \mu_j}{s_j} \quad (9)$$

**Step 2:** 对相关性矩阵进行计算。相关性矩阵  $W = (r_{ij})_{14 \times 14}$ ，其中： $r_{ii} = 1$ ， $r_{ij} = r_{ji}$ ， $r_{ij}$  为第  $i$  个指标与第  $j$  个指标的相关系数。

$$r_{ij} = \frac{\sum_{k=1}^{58} a_{ki}^* \times a_{kj}^*}{58-1} \quad (10)$$

**Step 3:** 对特征值和特征向量进行计算。求出相关性矩阵  $W$  的特征值  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{14} \geq 0$  及对应的标准化特征向量  $\mu_1, \mu_2, \dots, \mu_{14}$ ，其中  $\mu_j = [\mu_{1j}, \mu_{2j}, \dots, \mu_{5j}]$ ，由特征向量组成 14 个新的指标变量，其中  $y_i$  代表第  $i$  个主成分。

$$y_j = u_{1j}x_1^* + u_{2j}x_2^* + \dots + u_{14j}x_{14}^* \quad (11)$$

**Step 4:** 选择前 5 个主成分来计算评价的综合值。计算特征值  $\lambda_j (j=1, 2, \dots, 5)$  的信息贡献率和累积贡献率[10]，主成分  $y_j$  的信息贡献率为：

$$b_j = \frac{\lambda_j}{\sum_{k=1}^5 \lambda_k} \quad (12)$$

$y_1, y_2, \dots, y_p$  的累积贡献率为：

$$\alpha_p = \frac{\sum_{k=1}^p \lambda_k}{\sum_{k=1}^5 \lambda_k} \quad (13)$$

当  $\alpha_p$  接近 1 时，则选择前  $p$  个指标变量  $y_1, y_2, \dots, y_p$  作为  $p$  个主成分，代替原来的 14 个指标变量，从而可对  $p$  个主成分进行综合分析。并计算综合得分：

$$Z = \sum_{j=1}^p b_j y_j \quad (14)$$

式中： $b_j$  为第  $j$  个主成分的信息贡献率，根据综合得分值就可进行评价。利用 matlab 软件求得相关性矩阵前 5 个特征值与贡献率，结果如表 1 所示：

**Table 1.** Principal component analysis results

**表 1.** 主成分分析结果

主成分序号	特征值	贡献率(%)	累积贡献率(%)
1	4.052	28.943	28.943
2	2.425	17.322	46.264
3	1.709	12.209	58.473
4	1.107	7.904	66.378
5	1.096	7.828	74.206
6	1.769	5.493	79.698

## Continued

7	1.733	5.239	84.937
8	1.552	3.943	88.880
9	1.463	3.306	92.186
10	1.420	2.999	95.185
11	1.334	2.385	97.570
12	1.216	1.541	99.111
13	1.113	1.806	99.917
14	1.012	1.083	100.00

由表 1 可知, 前 5 个特征根的累积贡献率在 74% 左右, 可初步认为这些主成分能充分表达样本数据所包含的信息。后续将选取这 5 个主成分进入 BP 神经网络进行学习。同时得出了主成分系数得分矩阵, 便于后续权重分析, 具体数值如表 2 所示:

Table 2. Component score coefficient matrix

表 2. 成分得分系数矩阵

化学成分	主成分 1	主成分 2	主成分 3	主成分 4	主成分 5
1	0.164	0.1	0.262	0.058	0.064
2	0.07	0.284	0.213	-0.16	0.091
3	0.095	0.26	-0.171	0.216	0.138
4	0.137	0.176	-0.017	0.238	0.252
5	0.11	0.243	0.08	-0.178	-0.056
6	-0.076	0.01	0.43	-0.007	0.05
7	-0.191	0.11	-0.246	-0.141	0.072
8	-0.191	-0.032	0.248	0.255	0.076
9	-0.107	0.27	-0.062	-0.075	-0.168
10	-0.172	0.133	-0.114	0.115	0.226
11	0.049	0.029	-0.141	0.725	-0.275
12	-0.114	0.019	0.302	0.304	-0.126
13	0.198	-0.197	0.002	-0.023	-0.167
14	0.034	-0.144	-0.014	0.126	0.776

由此可得 5 个主成分分别为:

$$\begin{cases} y_1 = 0.164x_1^* + 0.070x_2^* + \cdots + 0.034x_{14}^* \\ y_2 = 0.100x_1^* + 0.284x_2^* + \cdots - 0.144x_{14}^* \\ y_3 = 0.262x_1^* + 0.213x_2^* + \cdots - 0.014x_{14}^* \\ y_4 = 0.058x_1^* - 0.160x_2^* + \cdots + 0.126x_{14}^* \\ y_5 = 0.064x_1^* + 0.091x_2^* + \cdots + 0.776x_{14}^* \end{cases} \quad (15)$$

### 3.2. BP 神经网络基本原理

BP 神经网络是依据误差反向传播算法的多层前馈神经网络，其运作的基本原理是信号前向传递，误差反向传播[11]。大体上可以分为 3 层，其中输入层各神经元的输入输出关系大多为线性函数，隐层的大多为非线性函数。其运作的具体步骤是输入  $i$  个学习样本  $y_1, y_2, \dots, y_n$ ，在对应输出层样本类别已知的前提下，利用 BP 学习算法，通过梯度下降法不断调整输入层与隐含层，隐含层与输出层之间的连接权值，使其沿着目标函数的负梯度方向转变，以达到输出层误差平方和最小的目的。

### 3.3. BP 神经网络算法训练

依据上述原理，结合上述主成分分析结果，向 BP 神经网络输入主成分  $[y_1, y_2, \dots, y_5]$  设置 5 个输入层结点，从输入层经由隐层可得输出数据  $z = [z_1, z_2]$ 。BP 神经网络有两个关键函数，激活函数与误差函数，其中激活函数一般使用 S 型函数，即

$$f(x) = \frac{1}{1 + e^{-x}} \quad (16)$$

误差函数的公式为：

$$e = \frac{1}{2} \sum_{o=1}^q (d_o(k) - z_o(k))^2 \quad (17)$$

具体字母定义如表 3：

Table 3. Letter definition table

表 3. 字母定义表

符号	定义
$w_{ih}$	输入层和中间层的连接权值
$w_{ho}$	输出层和隐含层的连接权值
$B_h$	隐含层各个神经元的阈值
$b_o$	输出层各个神经元的阈值
$f$	4 激活函数
$e$	4 误差函数
$x$	4 输入向量
$g_i$	隐含层输入向量
$g_o$	隐含层输出向量
$z_i$	输出层输入向量
$z_o$	输出层输出向量
$d_o$	期望输出向量

**Step 1:** 初始化。对每个连接权值赋值，赋值区间为  $(-1, 1)$  内任意随机数，误差函数设定为  $e$ ，最大学习次数为  $M$ ，计算精度值为  $\varepsilon$ ；

**Step 2:** 任意输入  $m$  个样本，得到相应的期望输出；

$$x(m) = (x_1(m), x_2(m), \dots, x_n(m)) \quad (18)$$



$$d_o(m) = (d_0(m), d_1(m), \dots, d_n(m)) \quad (19)$$

**Step 3:** 对给每个神经元输入值和输出值进行计算:

$$\begin{cases} gi_h(m) = \sum_{i=1}^n w_{ih} x_i(m) - b_h \\ go_h(m) = f(gi_h(m)) \end{cases} \quad (20)$$

$$\begin{cases} zi_o(m) = \sum_{h=1}^p w_{ho} go_h(m) - b_o \\ zo_o(m) = f(zi_o(m)) \end{cases} \quad (21)$$

其中  $h=1, 2, \dots, p$ ;  $o=1, 2, \dots, q$ 。

**Step 4:** 分析实际值与期望值的偏差, 得到输出层各神经元误差函数的偏导数  $\delta_o(m)$ :

$$\frac{\partial e}{\partial w_{ho}} = \frac{\partial e}{\partial zi_o} \frac{\partial zi_o}{\partial w_{ho}} \quad (22)$$

$$\frac{\partial zi_o(m)}{\partial w_{ho}} = \frac{\partial (\sum_h^p w_{ho} go_h(m) - b_o)}{\partial w_{ho}} = go_h(m) \quad (23)$$

$$\begin{aligned} \frac{\partial e}{\partial zi_o} &= \frac{\partial \left( \frac{1}{2} \sum_{o=1}^q (d_o(m) - zo_o(m))^2 \right)}{\partial zi_o} \\ &= -(d_o(m) - zo_o(m)) zo_o'(m) \\ &= -(d_o(m) - zo_o(m)) f'(zi_o(m)) - \delta_o(m) \end{aligned} \quad (24)$$

**Step 5:** 借助隐含层与输出层的连接权值与输出层的  $\delta_o(m)$  得出隐含层各神经元误差函数的偏导数  $\delta_h(m)$ :

$$\frac{\partial e}{\partial w_{ho}} = \frac{\partial e}{\partial zi_o} \frac{\partial zi_o}{\partial w_{ho}} = -\delta_o(m) go_h(m) \quad (25)$$

$$\frac{\partial e}{\partial w_{ih}} = \frac{\partial e}{\partial hi_h(m)} \frac{\partial hi_h(m)}{\partial w_{ih}} \quad (26)$$

$$\frac{\partial hi_h(m)}{\partial w_{ih}} = \frac{\partial (\sum_{i=1}^n w_{ih} x_i(m) - b_h)}{\partial w_{ih}} = x_i(m) \quad (27)$$

$$\begin{aligned} \frac{\partial e}{\partial hi_h(m)} &= \frac{\partial \left( \frac{1}{2} \sum_{o=1}^q (d_o(m) - zo_o(m))^2 \right)}{\partial ho_o(m)} \frac{\partial ho_o(m)}{\partial hi_h(m)} \\ &= \frac{\partial \left( \frac{1}{2} \sum_{o=1}^q (d_o(m) - f(zi_o(m)))^2 \right)}{\partial ho_h(m)} \frac{\partial ho_h(m)}{\partial hi_h(m)} \\ &= \frac{\partial \left( \frac{1}{2} \sum_{o=1}^q \left( d_o(m) - f \left( \sum_{h=1}^p w_{ho} go_h(m) - b_o \right) \right)^2 \right)}{\partial ho_h(m)} \frac{\partial ho_h(m)}{\partial hi_h(m)} \\ &= -\sum_{o=1}^q (d_o(m) - zo_o(m)) f'(zi_o(m)) w_{ho} \frac{\partial ho_h(m)}{\partial hi_h(m)} \\ &= -\left( \sum_{o=1}^q \delta_o(m) w_{ho} \right) f'(hi_h(m)) - \delta_h(m) \end{aligned} \quad (28)$$



**Step 6:** 反向传播, 借助输出层误差函数修正连接权值  $w_{ho}(m)$ ;

$$\Delta w_{ho}(m) = -\mu \frac{\partial_k}{\partial w_{ho}} = \mu \delta_o(m) h o_h(m) \quad (29)$$

$$w_{ho}^{N+1} = w_{ho}^N + \eta \delta_o(m) h o_h(m) \quad (30)$$

**Step 7:** 借助隐含层的误差函数修正连接权值;

$$\Delta w_{ih}(m) = -\mu \frac{\partial_k}{\partial w_{ih}} = \delta_h(m) x_i(m) \quad (31)$$

$$w_{ih}^{N+1} = w_{ih}^N + \eta \delta_o(m) x_i(m) \quad (32)$$

**Step 8:** 对全局误差进行计算;

$$E = \frac{1}{2m} \sum_{k=1}^m \sum_{o=1}^q (d_o(k) - z_o(k))^2 \quad (33)$$

**Step 9:** 直至计算误差小于实际误差后停止训练。

## 4. 模型的检验

### 4.1. 模型检验的步骤

利用上述已经建立好的 PCA-BP 神经网络识别模型, 采用 PCA 原理对 BP 神经网络的输入层进行主成分分析, 在主成分累计贡献率达 74.206% 的情况下对玻璃文物化学成分样本空间进行重构, 提高了模型识别的精度。对于未知类别的玻璃文物数据, 可将主成分值以及归一化原始样本作为检验样本输入已经训练好的 PCA-BP 神经网络文物识别模型, 具体步骤如下:

**Step 1:** 首先利用主成分分析法提取出前 5 个主成分的得分 S1;

**Step 2:** 以 S 作为学习样本, 进入 BP 神经网络进行学习, 构建 PCA-BP 识别模型;

**Step 3:** 取表单 3 样本数据, 按照步骤 1 的方法计算出主成分的得分 S2;

**Step 4:** 在已构建好的 PCA-BP 识别模型输入主成分 S2, 即可输出文物样品的类别; 将输出文物的类别与文物真实类别进行对比, 结果一致。

### 4.2. 模型结果分析

由下图可以分析出, 训练数据的 R 方值接近于 99.5%, 说明利用该模型所得的拟合优度非常高, 模型具有较高的合理性, 具体分类规律见图 2。

## 5. 结语

本文利用主成分分析与 BP 神经网络相结合, 建立了玻璃文物样品类别识别模型。该方法可用于对未知玻璃文物的类别的样品进行有效鉴别。主成分分析作为一种常见的多元统计分析方法, 其通过对指标降维, 达到化繁为简的目的。BP 神经网络是一种能解决复杂的非线性关系的数据分析预测和性能优化问题的模型, 在主成分分析与 BP 神经网络相结合条件下对玻璃文物进行鉴定, 通过 BP 算法来对误差进行调整,  $\frac{\partial_e}{\partial w_{ho}} > 0$ , 此时  $\Delta w_{ho} < 0$ , 即当误差对权值的偏导数小于零时, 权值调整量为正, 使得实际输出与期望输出的差减少, 反之亦然, 由此提升鉴定的准确性。但由于 BP 算法是以局部改善进行权值调整的, 故容易陷入局部最小化问题, 同时对神经元与隐藏层个数的选择不同可能会对模型识别结果产生影响, 未来可通过改进学习率参数的调节方式来优化模型, 以期达到更好的识别效果。

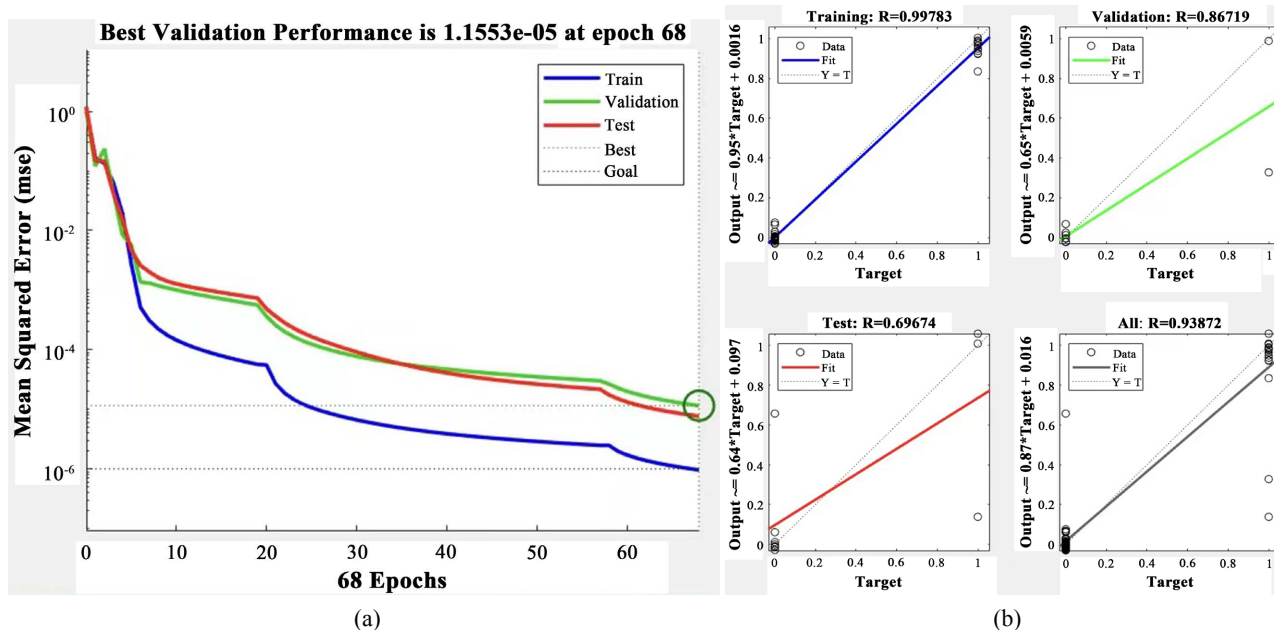


Figure 2. Fitting results after neural network training  
图 2. 神经网络训练后的拟合效果

另外本文所建立的主成分分析-BP 网络识别模型可以推广到很多实际问题中，如对煤岩表面化学成分的识别，未来也可拓展到军事上、考古文学上、股价预测上等，具有很高的应用价值。

### 参考文献

- [1] 李洪飞. 浅谈文物的鉴定与鉴赏[J]. 文物鉴定与鉴赏, 2019(15): 90-91.
- [2] 唐迎菲. 现代文物鉴定方法与技术刍议[J]. 文物鉴定与鉴赏, 2015(11): 88-89.
- [3] 王迪. 新媒体在书画文物鉴定中的应用[J]. 当代旅游, 2019(5): 277.
- [4] 李擎. 文物鉴定理论和方法的若干思考[J]. 文物鉴定与鉴赏, 2019(5): 78-79.
- [5] 段继霞. 基于边缘检测算法的文物识别研究与应用[D]: [硕士学位论文]. 太原: 中北大学, 2020. <https://doi.org/10.27470/d.cnki.ghbgc.2020.000414>
- [6] 郭美玉. 新技术条件下文物鉴定与鉴赏方法探析[J]. 中国民族博览, 2021(16): 205-207.
- [7] 薛吕. 玻璃文物保护与修复[C]//中国文物保护技术协会第七次学术年会论文集. 2012: 89-92.
- [8] 周概容. 应用统计方法辞典[M]. 北京: 中国统计出版社, 1993: 213-462.
- [9] 徐维超. 相关系数研究综述[J]. 广东工业大学学报, 2012, 29(3): 12-17.
- [10] 范杰. 主成分分析法的数值实现算法[J]. 河北工程大学学报(自然科学版), 2007(4): 103-105.
- [11] 张红玉, 丁宁, 徐江荣. BP 神经网络激励函数改进研究[J]. 杭州电子科技大学学报(自然科学版), 2017, 37(6): 62-66+90. <https://doi.org/10.13954/j.cnki.hdu.2017.06.012>