

# Face Orientation Recognition Based on Multi-hidden Layer BP Neural network

Xue Liu

Jiangsu Intellectual Property Protection Center (Jiangsu Patent Information Service Center), Nanjing, Jiangsu, 210036, China

## Abstract

Face orientation recognition using Back Propagation neural network has been a research focus, but most of them are limited to single hidden layer method, which leads to low recognition rate. In order to achieve high-precision face orientation recognition, a multi-hidden layer BP neural network model is created. The multi-hidden layer method can effectively solve the problems of low recognition rate of face orientation and result fluctuation, and has the characteristics of fast and stable, and the recognition rate of double hidden layer can reach 98.00%, the R value of regression analysis is 0.97341, and the mean square error is below the threshold of  $1.00e-06$ . The simulation results show that the multi-hidden BPNN can significantly improve the face orientation recognition rate, and the face orientation recognition model established has high reliability and strong prediction ability.

## Keywords

Feature vector extraction; Back propagation; Multiple hidden layers; Neural network; Face orientation recognition; Simulation

## 基于多隐含层反向传播神经网络的人脸朝向识别

刘雪

江苏省知识产权保护中心（江苏省专利信息服务中心），中国·江苏南京 210036

## 摘要

利用反向传播神经网络进行人脸朝向识别一直是研究热点，但大多局限于单隐含层方法，导致识别率不高。为实现高精度人脸朝向识别，创建多隐含层BP神经网络模型。多隐含层方法有效解决人脸朝向识别率低、结果波动等问题，具有快捷、稳定的特点，且双隐含层识别率达到98.00%，回归分析R值为0.97341，均方差MSE值低于阈值 $1.00e-06$ 。仿真结果表明，多隐含层反向传播神经网络能够显著提升人脸朝向识别率，所建立的人脸朝向识别模型可靠性高、预测能力强。

## 关键词

特征向量提取；反向传播；多隐含层；神经网络；人脸朝向识别

## 1 引言

人脸识别主要指通过智能系统分析人脸特征信息，实现身份鉴别，但摄像头摄取的图像有时达不到预期要求，譬如在拍摄范围内无法识别人物正脸，往往造成匹配失败。因此，人脸朝向识别研究具有较强的应用性，也是现阶段模式识别、图像处理等领域研究热点<sup>[1]</sup>。

反向传播神经网络BPNN（Back Propagation Neural Network）具有较强的非线性映射能力、容错能力、泛化能力以及自适应、自学习能力，是人脸朝向识别的一项重要技术<sup>[2]</sup>。本课题采用多隐含层BP神经网络，以Matlab作为编写工具和运行环境，建立人脸朝向识别模型，相对取得了较好的效果。

## 2 人脸朝向识别模型

### 2.1 BP神经网络

#### 2.1.1 BP神经网络原理

BP神经网络通常由输入层、（单层或多层）隐含层以及输出层构成，能够完成从输入源数据至输出结果的深度非线性映射。它包含两个环节：前向传播与误差反向传播。当一个模式自输入层进入网络后，将在BPNN中传递，借助权值线路经过隐含层中神经元，通过激活函数、隐含层中所有神经元对该模式加权求和，而后利用权值线路到达输出层神经单元，经过相关处理，再在输出端建立一个输出模式。BP神经网络在多次训练过程中能够逐步调整自身结构参数，以期输出结果与实际情况之间误差符合要求，该网络能以尽可能高的精度靠近一个非线性函数，适合解决现实中复杂问题。

【作者简介】刘雪（1978-），男，中国安徽宿州人，硕士，从事图像处理、网络安全研究。

## 2.1.2 改进的 BP 神经网络

### 1) LM 算法

传统的 BP 神经算法存在收敛速度迟缓、易于“陷入”局部极值等问题，可运用 LM 算法进行改进，它是梯度下降算法同高斯-牛顿算法的融合。LM 算法计算公式如下：

$$\Delta x = -[J^T(x)J(x) + \mu I]^{-1} J(x)e(x) \quad (1)$$

(1) 式中 I 是单位矩阵。如果比例系数  $\mu$  的值为 0，LM 算法就转换成高斯-牛顿算法；如果  $\mu$  的值较大，就与梯度下降算法近似。所以，LM 算法收敛时间最短、鲁棒性好，能够显著增强 BPNN 网络收敛效果。

### 2) 双曲正切 S 型传递函数

在 BPNN 中，传递函数属于神经元及神经网络的核心内容，双曲正切 S 型传递函数 tagsig 具有非线性放大增益特点。其计算公式如下：

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

## 2.1.3 BP 神经网络模型可靠性验证

由于 BPNN 学习与记忆呈不稳定性，利用回归分析结果 R 验证 BPNN 模型可靠性。其计算方法如下：

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

其中，回归平方和 SSR 体现自变量 x 取值对因变量 y 结果的影响，总平方和 SST 体现 y 的 n 个值与其平均值总误差。R 值位于 0~1 之间，离 1 越近，模型越可靠；R 值超过 0.8 的模型可靠性较高，但 R 值若为 1，则为过拟合。

## 2.2 人脸朝向识别模型构建

### 2.2.1 人脸特征提取

当人脸朝向变化时，眼睛在图像中位置将有相应变动。在人脸朝向识别模型中，BPNN 输入端是人脸特征量，因此需提取表达眼睛位置信息的特征向量。该提取函数  $edge(f, 'roberts')$  是利用 roberts 算子对图像边缘检测，利用 floor() 函数把图像分割成 6 行、8 列，转换成 0~1 矩阵，然后读取分割后的图像第 2 行眼睛数据。

这里使用的 roberts 算子定位精度高，而常用的 sobel 算子优势是抑制噪声，由于本课题中图像已经过降噪、滤波等预处理，因此 roberts 算子边缘检测效果高于 sobel 算子。

运用 roberts 算子检测图像  $f(x,y)$  边缘，其实就是计算  $(x,y)$  处的梯度。梯度属于二维的离散型数集，利用求函数偏导方式得到图像的偏导数，也就是  $(x,y)$  处的最高变化率，如公式 (4)、(5) 所示。

$$g_x = \frac{\partial f(x,y)}{\partial x} = f(x+1,y) - f(x,y) \quad (4)$$

$$g_y = \frac{\partial f(x,y)}{\partial y} = f(x,y+1) - f(x,y) \quad (5)$$

检测图像  $f(x,y)$  边缘的结果是目标图像  $M(x,y)$ ，是在上

述工作基础上通过公式 (6) 计算。

$$M(x,y) = mag(\nabla f) = \sqrt{g_x^2 + g_y^2} \quad (6)$$

### 2.2.2 模型建立

#### 1) 输入层和输出层、隐含层设计

输入层：人脸图像提取后特征信息是 8 个列向量，因此输入层神经元个数为 8。

输出层：人脸 5 个朝向分别用 3 个元素 (0 或 1) 的数组表示，譬如 (1, 0, 0) 表示朝右，因此输出层神经元个数为 3。

隐含层：网络隐含层的层数及神经单元数的确定是个难点，也是本文重点。下面两个方法可供参考：

方法一：研究人员 fangfaGorman 通过公式 (7) 描述隐含层节点数 s 和模式数 N 之间关系：

$$s = \log_2 N \quad (7)$$

方法二：著名的 Kolmogorov 定理显示，隐含层节点数 s 可用公式 (8) 计算，其中 n 是输入层节点数。

$$s = 2n + 1 \quad (8)$$

上述方法得到的隐含层神经单元数目可作为人脸朝向识别模型中初始值，进行试凑。

#### 2) 函数选择

BPNN 训练函数采用 Trainlm (前文所述的 LM 算法)，传递函数选用 tansig，最后一层隐含层至人脸朝向识别模型输出层，采用线性 pureline 函数。如此，可较好地保证模型的快速、稳定，精度也高。

#### 3) 学习率等因素设置

学习率过小阻碍训练速度，过大易导致网络波动，通常在 0.01-0.8 选择，这里为了保证 BPNN 模型训练质量，取较小值 0.01。另外，训练数阈值定为 1000，误差阈值定为  $1e-6$ 。

### 2.2.3 处理流程

基于 BPNN 的人脸朝向识别处理流程：

1) 提取人脸特征向量：将图像中人眼位置信息的向量取出，每个图像分为 6 行 8 列，BPNN 的输入内容就是一个 8 维的向量。

2) 建立训练集 / 测试集：在 50 幅人脸图像中随机选择 30 幅当作训练集，其余 20 幅当作测试集。

3) 建立 BPNN：应用 newff() 方法构建 BP 神经网络，其中函数通过 “{ 'tansig', 'purelin' }, 'trainlm' ” 形式设置。

4) 训练 BPNN 网络：把人脸图像训练集输入已创建的 BPNN 网络，训练过程中自动调整网络中线路权值，直到符合要求，迭代结束。

5) 人脸朝向识别检测：对随机选择的图像中人脸进行朝向检测，并计算出识别率。

## 3 实验仿真与分析

### 3.1 仿真识别过程

为探讨基于多隐含层 BPNN 模型的人脸朝向识别方法

的有效性,分别针对不同隐含层选择四个数量神经元,进行仿真实验,每项实验重复10-15次,取平均值。这里以五隐含层[61 35 21 13 7]为例,阐述人脸朝向识别实验过程。

### 3.1.1 BPNN 训练进度

从人脸朝向识别实验过程中BPNN训练可以看出,神经网络训练迭代次数Epoch阈值设为1000,实际迭代9次,训练时间为14秒;误差阈值是 $1.00e-06$ ,实际值为 $2.12e-07$ ;梯度阈值是 $1.00e-07$ ,实际值为 $0.000578$ ;阻尼因子阈值是 $1.00e+10$ ,实际值为 $1.00e-09$ ;泛化能力检查次数阈值是6,实际检查3次。这里误差是指均方差,实际值为 $2.12e-07$ ,低于阈值 $1.00e-06$ ,表示人脸朝向识别模型的泛化能力强,即模型预测能力较强。

### 3.1.2 BPNN 误差变化曲线

在BPNN训练时,误差目标Goal与训练过程目标越来越靠近;泛化能力检查MSE与其目标Best之间误差开始时逐渐变小,随着训练的进行,训练误差继续变小,迭代6次后(此时模型训练效果最佳),泛化能力检查MSE反而渐渐上升,连续上升3次后,模型训练停止(此时已迭代9次),这是为了避免过拟合。

### 3.1.3 BPNN 训练状况

BPNN训练实际迭代9次,梯度实际值为 $0.00057831$ ,阻尼因子实际值为 $1.00e-09$ ;当BPNN训练结果达到最佳后,泛化能力检查3次(这里默认阈值为6),BPNN训练结束,防止出现过拟合倾向。

### 3.1.4 BPNN 回归分析

回归分析用来检验模型中两个变量之间相关性,回归分析结果值R值越大,越接近1(但不等于1),模型可靠性就越高。训练过程回归分析R值为 $0.99925$ ,泛化检查过程R值为 $0.99056$ ,测试过程R值为 $0.99856$ ,总体预测过程R值为 $0.99759$ ,都超过0.8,且已接近1,说明模型可靠性很高;另外,这几项R值都没达到1,说明没出现训练过度情况,不存在过拟合问题。

## 3.2 多隐含层不同结构实验结果对比

### 3.2.1 单隐含层 BPNN 人脸朝向识别

当神经元数为7、6、5、4时,人脸朝向识别率最高仅为86.50%,且实验时结果波动性大;R值经常为1,表现出过拟合,表示单隐含层BPNN模型泛化能力较弱,即人脸朝向识别能力较差;对于单隐含层最优结构[5],实验中均方差MSE值有时达到0.0312,误差值较高,从另一个侧面说明单隐含层BPNN模型泛化能力较弱。

### 3.2.2 双隐含层 BPNN 人脸朝向识别

双隐含层采用[15 7]、[13 7]、[11 7]、[9 5]等结构试凑,人脸朝向识别率最高为98.00%,对应隐含层结构为[11 7],总体预测过程R值为 $0.97341$ ,说明模型可靠性满足要求;

迭代次数为8,执行时间为1秒。另外,实验过程中均方差MSE值达到 $1.00e-07$ 以下值,明显低于阈值 $1.00e-06$ ,表示该模型泛化能力较强,即人脸朝向识别能力较强。

### 3.2.3 三隐含层 BPNN 人脸朝向识别

三隐含层采用[40 18 8]、[31 15 7]、[30 11 7]、[15 9 5]等结构试凑,人脸朝向识别率最高为96.21%,对应隐含层结构为[15 9 5],总体预测过程R值为 $0.96157$ ,说明模型可靠性满足要求;迭代次数为10,执行时间为1秒。实验过程中均方差MSE值低于阈值 $1.00e-06$ ,表示该模型泛化能力较强。

### 3.2.4 四隐含层 BPNN 人脸朝向识别

四隐含层不同结构BPNN模型人脸朝向识别实验结果。四隐含层采用[53 31 15 7]、[39 21 13 7]、[35 19 13 7]、[29 19 13 7]等结构试凑,人脸朝向识别率最高为97.50%,对应隐含层结构为[53 31 15 7],预测过程R值为 $0.98782$ ,说明模型可靠性满足要求;迭代次数为9,执行时间为8秒。实验过程中均方差MSE值低于阈值,表示该模型泛化能力较强。

### 3.2.5 五隐含层 BPNN 人脸朝向识别

五隐含层采用[95 47 23 11 5]、[61 35 21 13 7]、[43 26 15 9 5]、[25 17 11 7 5]等结构试凑,人脸朝向识别率最高为93.52%,对应隐含层结构为[61 35 21 13 7],预测过程R值为 $0.98715$ (这里取多次实验平均值,与前文图4中R值有区别),说明模型可靠性满足要求;迭代次数为9,执行时间为14秒。实验过程中均方差MSE值低于阈值,表示该模型泛化能力较强。

## 3.3 不同隐含层人脸朝向识别实验对比

在多隐含层中,双隐含层BPNN人脸朝向识别效果最高,达到98.00%,对应隐含层结构为[11 7];隐含层数为三、四、五时,识别率有所降低,但仍高于93%;而且,这部分BPNN模型R值都大于0.96(接近1但不等于1),可靠性都很高。另外,多隐含层MSE值均低于阈值,表示多隐含层模型泛化能力较强,即人脸朝向识别能力较强。

## 4 结论

实验结果表明,所创建的多隐含层反向传播神经网络人脸朝向识别效果好,可靠性高,预测能力强,为当前利用反向传播神经网络进行人脸朝向识别提供了一种可参考的解决方案,具有较高的实际应用价值。

### 参考文献

- [1] 刘柯,李旭健.一种人脸眼部特征提取和定位的方法[J].计算机仿真,2020,37(7):426-431.
- [2] 许佳,田维坚,樊养余.基于深度学习的人脸关键点识别定位方法仿真[J].计算机仿真,2020,37(6):434-438.