



结合 Bi-2DPCA 与 CNN 的美式手语识别

杨明羽, 叶春明

(上海理工大学 管理学院, 上海 200093)

摘 要: 针对现有美式手语(ASL)识别算法准确率低和模型训练时间长的问题, 提出一种结合双向二维主成分分析(Bi-2DPCA)与卷积神经网络(CNN)并基于贝叶斯优化的识别算法。利用 Bi-2DPCA 算法对原始图像做数据降维处理, 提取行、列方向的特征图, 使用卷积神经网络对特征图进行训练分类, 同时采用贝叶斯优化算法对模型超参数进行自动调参。在 24 分类 ASL 数据集上的实验结果表明, 该算法的识别准确率达到 99.15%, 训练时间相比传统 CNN 算法减少 90.3%。

关键词: 美式手语识别; 双向二维主成分分析; 卷积神经网络; 贝叶斯优化; 自动调参

开放科学(资源服务)标志码(OSID):



中文引用格式: 杨明羽, 叶春明. 结合 Bi-2DPCA 与 CNN 的美式手语识别[J]. 计算机工程, 2021, 47(12): 278-284.

英文引用格式: YANG M Y, YE C M. American sign language recognition combining with Bi-2DPCA and CNN[J]. Computer Engineering, 2021, 47(12): 278-284.

American Sign Language Recognition Combining with Bi-2DPCA and CNN

YANG Mingyu, YE Chunming

(School of Business, University of Shanghai for Science and Technology, Shanghai 200093, China)

[Abstract] The existing algorithms for American Sign Language (ASL) recognition are limited in the recognition accuracy, and require much time for model training. To address the problem, a Bayesian Optimization (BO)-based algorithm that combines Bidirectional Two-Dimensional Principal Component Analysis (Bi-2DPCA) and Convolutional Neural Network (CNN) is used to optimize model parameters. The Bi-2DPCA algorithm is used to reduce the dimensionality of the original image data, and extract the feature maps in the row and column directions. Then the convolutional neural network is used to train and classify the feature maps. Finally, the Bayesian optimization algorithm is used to adjust the model hyperparameters automatically. The experimental results On 24 classified ASL data sets show that the algorithm achieves a recognition accuracy of 99.15%, and reduces the running time by 90.3% compared with the traditional CNN algorithms.

[Key words] American Sign Language (ASL) recognition; Bidirectional Two-Dimensional Principal Component Analysis (Bi-2DPCA); Convolutional Neural Network (CNN); Bayesian Optimization (BO); automatic tuning

DOI: 10.19678/j.issn.1000-3428.0059945

0 概述

手语通过摆出不同的手臂与手的姿势, 辅以表情及其他肢体动作传达信息, 是聋哑人重要的沟通语言, 也是帮助他们融入社会的重要工具^[1]。美式手语(American Sign Language, ASL)是适用于美国、加拿大等地区的手语语言^[2-3]。在 ASL 中, 包含静态手势和动态手势, 以 J 和 Z 表示动态手势, 其余字母表示静态手势。但由于人类的手非常灵活, 可以摆

出很多形状, 因此增加了手势识别的难度。

近年来, 许多学者在 ASL 图片的识别问题上, 使用流行的神经网络进行研究。ASHA 等^[4]提出一种通过提取静态 ASL 图片的特征直方图、统计量度等进行神经网络训练的识别算法, 最高识别率为 98.17%。QUTASHAT 等^[5]通过使用霍夫变换和神经网络开发了 ASL 语言翻译及标志系统, 识别准确率为 92.3%。ADITHYA 等^[6]通过使用卷积神经网络(Convolutional Neural Network, CNN)^[7]对 ASL 数据集进行训练预测,

基金项目: 国家自然科学基金(7184003); 上海市科委“科技创新行动计划”软科学重点项目(20692104300); 上海理工大学科技发展基金(2018KJFZ043)。

作者简介: 杨明羽(1997—), 男, 硕士研究生, 主研方向为图像识别、人工智能、智能算法; 叶春明, 教授、博士生导师。

收稿日期: 2020-11-09 **修回日期:** 2020-12-24 **E-mail:** 447661483@qq.com

准确率达到94.7%。但在使用神经网络训练时,往往需要花费大量的时间在训练模型上^[8]。特别是近年来CNN的网络层数不断加深^[9]和可训练参数骤增,更凸显了模型训练时间过长的问题。因此,一些学者提出了将主成分分析(Principal Component Analysis, PCA)与神经网络相结合的观点。吴伟^[10]提出PCA与CNN相结合识别ASL图片的方法,通过在CNN前插入PCA层进行特征降维,大幅提升了训练速度,且识别率达到94.45%。钟健等^[11]提出一种PCA与自组织神经网络SOM相结合的识别方法,使用SOM作为分类器进行图片分类,同样显著提高了识别速度,且识别率稳定在90%左右。上述方法使用的PCA均在一维数据中使用,针对图片等二维数据,学者们又发展了二维主成分分析(Two-Dimensional PCA, 2DPCA)^[12]和双向二维主成分分析(Bidirectional 2DPCA, Bi-2DPCA)^[13]方法。相比于一维的PCA,改进后的2DPCA和Bi-2DPCA对图像的降维效果更好,提升速度的效果更明显。

在基于神经网络的各种模型中,超参数的选择对模型的性能影响很大,不恰当的超参数或将直接影响模型的应用效果。近几年在ImageNet比赛中获得成功的GoogleNet、ResNet-52和DenseNet都需要对超参数进行精细地调整,这3个模型分别有78、150、376个超参数^[14],若手动调参则会浪费非常多的时间。因此,在优化超参数时可使用粒子群优化(Particle Swarm Optimization, PSO)算法^[15]、遗传算法(Genetic Algorithm, GA)^[16]、贝叶斯优化(Bayesian Optimization, BO)算法^[17-19]等自动调参算法。

为提高ASL图片识别准确性同时加快模型训练速度,本文提出基于Bi-2DPCA与CNN的美式手语识别算法。利用Bi-2DPCA对原始图片进行降维预处理,在此基础上提取特征矩阵,并将降维后的数据输入CNN网络进行模型训练。同时,使用贝叶斯优化方法,通过设计超参数的选取及其范围优化CNN的超参数。

1 相关知识

1.1 卷积神经网络

卷积神经网络(CNN)用于处理空间上有关联的数据^[20],其使用多个较小的卷积核(filter),有规律地扫描数据。CNN的网络连接是稀疏连接,卷积核使用的是共享参数,这极大地减少了网络可训练参数的数量,加快了网络训练速度。CNN的基本结构由输入层、卷积层、激活层、池化层、丢弃层(Dropout)、全连接层和输出层构成。在卷积层后加入批归一化层(Batch Normalization, BN)可以进一步增强网络的性能和泛化能力^[21-22]。

卷积层中有多个卷积核,对于前一网络层传入的图像或特征图,在其每一图层都有一个卷积核进行扫描。卷积核计算公式为:

$$y_i^k = \left(\sum_{j \in M_i} X_j^{k-1} W_{ji}^k + b_i^k \right) \quad (1)$$

其中: y_i^k 为第 k 层第 i 个卷积核的特征映射; M_i 为第 $k-1$ 层所有进行卷积计算的特征映射; X_j^{k-1} 为第 $k-1$ 层第 j 个特征映射; W_{ji}^k 为第 k 层当前计算的权重矩阵; b_i^k 为偏置值。

激活层将线性的数据非线性输出,常用的激活函数有Sigmoid、Tanh、ReLU等。对比其余2个函数,ReLU函数能够有效缓解梯度消失和梯度爆炸的问题^[23]。

BN层对一批数据进行标准化,使数据符合以0为均值、以1为标准差的分布,计算公式为:

$$H_k^{i'} = \frac{H_k^i - \mu_{\text{batch}}^i}{\sigma_{\text{batch}}^i} \quad (2)$$

其中: $H_k^{i'}$ 为第 i 个卷积核BN后的数据; H_k^i 为第 i 个卷积核BN前的数据; μ_{batch}^i 和 σ_{batch}^i 分别为第 i 个卷积核在当前batch的平均值和标准差。之后,有学者又加入了缩放因子 γ_i 和偏移因子 β_i ,对数据在BN后进行缩放和偏移,得到了新的标准化数据 X_k^i ,增加了非线性表达能力,其计算公式为:

$$X_k^i = \gamma_i H_k^{i'} + \beta_i \quad (3)$$

池化层向下采样可增加感受野,缓解过拟合。目前,常用的池化方法有最大值池化和均值池化。

Dropout层在训练时舍弃一定比例的神经元,这使得网络的训练参数变少,增加了网络的泛化能力,减少了某些神经元的依赖性。当网络中存在BN层时,Dropout的比例可适当减少。

CNN使用Adam优化器对Softmax的交叉熵函数进行优化。Adam定义了一阶动量 m_t 和二阶动量 V_t ,分别为当前 t 次迭代时梯度 g_t 的一次函数与二次函数, β_1 与 β_2 为超参数,设置为0.9与0.999。一阶和二阶动量表达式分别如式(4)和式(5)所示:

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t \quad (4)$$

$$V_t = \beta_2 \cdot V_{t-1} + (1 - \beta_2) \cdot g_t^2 \quad (5)$$

1.2 Bi-2DPCA算法

PCA算法核心思想是将样本投影到一个超平面上,并使所有样本的投影尽可能分开,即最大化投影点的方差。2DPCA针对图片等二维矩阵,不需要将图片数据打平为一维数据,极大地减少了维度^[12]。

利用2DPCA算法对图片 X_i (m 行 n 列)进行特征提取,其中投影基 W (n 行 k 列)的列宽远小于图片矩阵,由此可得到最后的投影矩阵为:

$$Y = XW \quad (6)$$

最后的投影矩阵 Y 为 m 行 k 列,实现了图片在列维度的降维。其中, W 矩阵为原数据中心化后的协方差矩阵前 k 个特征值组成的特征向量。协方差矩阵 G_i 表示为:

$$\mathbf{G}_i = \frac{1}{N} \sum_{j=1}^N (\mathbf{X}_i - \bar{\mathbf{X}})^T (\mathbf{X}_j - \bar{\mathbf{X}}) \quad (7)$$

其中: N 为样本总数; $\bar{\mathbf{X}}$ 为样本均值。计算 \mathbf{G}_i 矩阵的特征值和特征向量,取前 k 个特征值组成的特征向量构成 \mathbf{W} 矩阵,即可计算最后的投影矩阵 \mathbf{Y} 。

由于2DPCA只对列维度进行降维,整体维度仍比较大,因此出现了改进算法Bi-2DPCA,其将2DPCA处理后的数据 \mathbf{Y} (m 行 k 列)继续投影到某组基 \mathbf{C} 上,得到最终的投影矩阵 \mathbf{U} ,表示为:

$$\mathbf{U} = \mathbf{C}^T \mathbf{Y} \quad (8)$$

其中, \mathbf{C} 为由新样本构造的协方差矩阵 \mathbf{G}'_i 前 j 个特征值组成的特征向量。 \mathbf{G}'_i 表示为:

$$\mathbf{G}'_i = \frac{1}{N} \sum_{i=1}^N (\mathbf{Y}_i - \bar{\mathbf{Y}})(\mathbf{Y}_j - \bar{\mathbf{Y}})^T \quad (9)$$

最后得到行变换 \mathbf{C} 与列变换 \mathbf{W} ,将行列变换合并。对于图片 \mathbf{X}_i ,降维后的特征矩阵 \mathbf{A} 的维度为 j 行 k 列,其中 j 和 k 远小于原始图片的 m 和 n ,则行列变换矩阵公式为:

$$\mathbf{A} = \mathbf{C}^T \mathbf{X} \mathbf{W} \quad (10)$$

若对特征矩阵进行复原显示,则可得到:

$$\mathbf{X}' = \mathbf{C} \mathbf{A} \mathbf{W}^T \quad (11)$$

在Bi-2DPCA算法中也存在2个超参数,即列降维选取的 k 取值与行降维选取的 j 取值,其经验选取为特征值贡献率为0.9~0.99的特征数^[12]。

1.3 贝叶斯优化

贝叶斯优化(BO)算法基于顺序模型,其目标函数不需要具有可导、连续等数学性质,并且如果是昂贵的黑盒函数,也可以在比较少的迭代次数内计算得到最佳的取值^[17]。

在BO算法中,最重要的2个部分是概率代理模型和采集函数。概率代理模型用于替代评估代价高昂的目标函数,常用的代理模型有贝塔-伯努利(Beta-Bernoulli)模型、线性模型、高斯过程(Gaussian Processes, GP)模型等。GP模型具有较好的灵活性、可拓展性和可分析性,是BO中应用最广泛的概率代理模型。该模型由均值函数和协方差函数构成,如式(12)所示:

$$f(\mathbf{x}) \sim \text{GP}(m(\bar{\mathbf{x}}), k(\bar{\mathbf{x}}, \bar{\mathbf{x}}')) \quad (12)$$

其中: $m(\bar{\mathbf{x}})$ 通常为0; $k(\bar{\mathbf{x}}, \bar{\mathbf{x}}')$ 采用Matérn协方差函数^[18]。 $k(\bar{\mathbf{x}}, \bar{\mathbf{x}}')$ 表示为:

$$k(\bar{\mathbf{x}}, \bar{\mathbf{x}}') = \left(1 + \frac{\sqrt{3}r}{l}\right) \exp\left(-\frac{\sqrt{3}r}{l}\right) \quad (13)$$

其中: r 为 $|\bar{\mathbf{x}} - \bar{\mathbf{x}}'|$; l 为尺度参数。

采集函数从搜索域 \mathbf{X} 中选择评估点,将其映射到实数空间 $\alpha: \mathbf{X} \rightarrow \mathbb{R}$ 。该函数由已观测数据集 $\mathbf{D}_{1:t}$ 的后验分布组成,通过最大化该函数寻找下一个评估点 \mathbf{x}_{t+1} ,如式(14)所示:

$$\mathbf{x}_{t+1} = \max_{\mathbf{x} \in \mathbf{X}} \alpha_t(\mathbf{x}; \mathbf{D}_{1:t}) \quad (14)$$

该过程基于提升的策略EI(Expected Improvement),采集函数为:

$$\alpha_t(\mathbf{x}; \mathbf{D}_{1:t}) = \begin{cases} (v^* - \mu_t(\mathbf{x})) \phi\left(\frac{v^* - \mu_t(\mathbf{x})}{\sigma_t(\mathbf{x})}\right) + \\ \sigma_t(\mathbf{x}) \phi\left(\frac{v^* - \mu_t(\mathbf{x})}{\sigma_t(\mathbf{x})}\right), \sigma_t(\mathbf{x}) > 0 \\ 0, \sigma_t(\mathbf{x}) = 0 \end{cases} \quad (15)$$

其中: v^* 为当前最优函数值; $\phi(\cdot)$ 为标准正态分布密度函数; $\mu_t(\mathbf{x})$ 为均值; $\sigma_t(\mathbf{x})$ 为标准差。

2 B2DPCA-CNN算法

本文提出Bi-2DPCA-CNN算法,算法流程如图1所示。

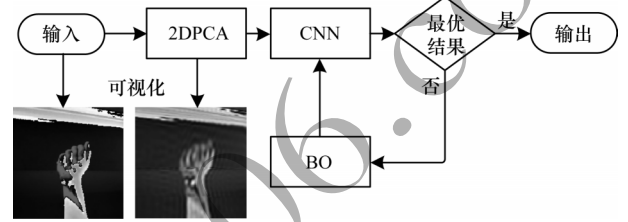


图1 Bi-2DPCA-CNN算法流程

Fig.1 Procedure of Bi-2DPCA-CNN algorithm

2.1 数据预处理

本文使用的ASL数据集来源于Kaggle,为24分类样本,对应A~Z(除去J、Z)的24个英文字母,每一类有3 000张图片,共72 000张图片。每个字母对应的照片集都在光照、姿势、位置等方面有所不同,可以较好地模拟真实识别场景。图2展示了ASL数据集部分图片及其字母表示。

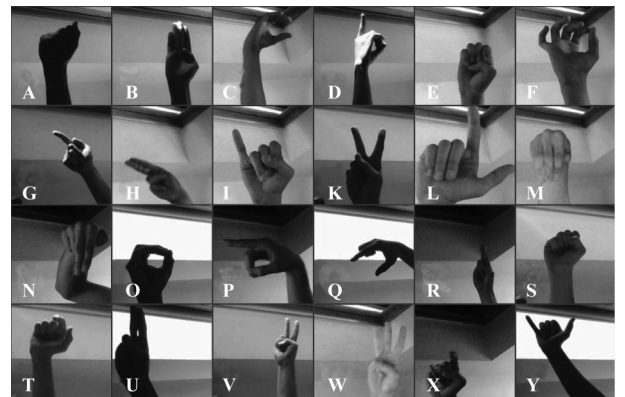


图2 ASL数据集部分图片及其字母表示

Fig.2 Some pictures of ASL dataset and their letter representations

图2中每张图片的原始尺寸大小为200像素×200像素,在进行特征估计并排序后,前30位特征柱状图如图3所示。可以看出,排名高的特征值远大于其他特征值。

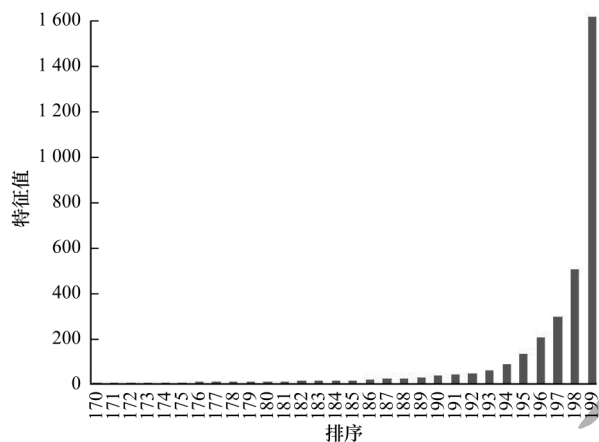


图3 前30个特征柱状图

Fig.3 Top 30 feature histograms

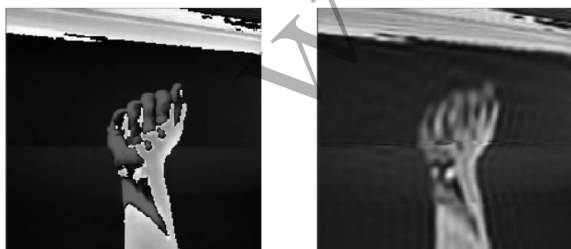
若对图像进行复原,则可以直观地看出降维后的效果,图4~图6分别为原始灰度图与降维到100、50、25维度复原图像的对比图。可以看出降维到25的复原图像仍然可以还原手势的形状,所以,Bi-2DPCA降维参数选择为25,即200像素×200像素的原图降维成25像素×25像素的尺寸。



(a)原图 (b)复原图

图4 100维度原图与复原图对比

Fig.4 Comparison of 100-dimensional original image and restored image



(a)原图 (b)复原图

图5 50维度原图与复原图对比

Fig.5 Comparison of 50-dimensional original image and restored image



(a)原图 (b)复原图

图6 25维度原图与复原图对比

Fig.6 Comparison of 25-dimensional original image and restored image

2.2 CNN网络及参数设计

如图7所示,本文设计包含多个卷积层和池化层的卷积神经网络。加入BN层进行数据批标准化,并在Dropout层进行丢弃处理,避免过拟合,最后接入全连接层和Softmax分类器进行分类。

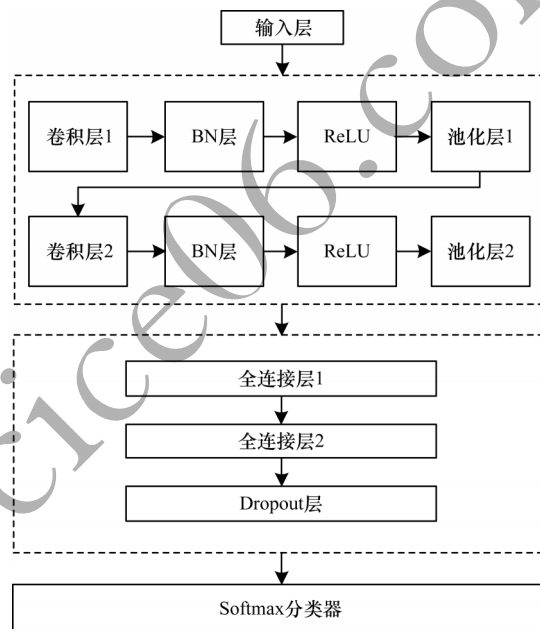


图7 包含多个卷积层和池化层的CNN结构

Fig.7 Structure of CNN with multiple convolution layers and pooling layers

卷积层、池化层和Dropout层参数项如下,参数范围如表1所示。

1)卷积层1:卷积核的大小为dim1,数量为num_conv1。

2)池化层1:池化大小为pool。

3)卷积层2:卷积核的大小为dim2,数量为num_conv2。

4)池化层2:池化大小为pool。

5)Dropout层:丢弃率为drop。

6)学习率:设置为优化参数lr。

表1 参数范围

Table 1 Range of parameters

参数项	参数类型	参数范围
dim1	整型	(3,5)
num_conv1	整型	(8,16)
dim2	整型	(3,5)
num_conv2	整型	(16,32)
pool	整型	(2,3)
drop	浮点型	(0.1,0.3)
lr	浮点型	(0.005,0.020)

3 实验与结果分析

3.1 实验设置与评价指标

本文实验使用 Windows10 系统,编程语言为 Python,实验平台使用 TensorFlow2.1 版本,采用 GPU 加速,处理器为 Inter i7-8750H 处理器,显卡为 NVIDIA GTX1060,6 GB 显存。实验使用的样本为 72 000 张图片,其中 60 000 张作为训练集,12 000 张作为测试集,模型迭代 50 次,保存模型每次迭代的准确率与损失。评价指标选取准确率、训练时间、测试时间和 F1 值。

3.2 Bi-2DPCA 优化验证

分别使用以下模型对 ASL 数据集识别并进行实验分析,验证 Bi-2DPCA 对 CNN 整体性能的提升:1)文献[10]中的 PCA-CNN 模型(下文简称 a-CNN 模型);2)文献[6]中的 CNN 模型(下文简称 b-CNN 无改进模型);3)文献[6]中加入 Bi-2DPCA 的 CNN 模型(下文简称 b-CNN-2DPCA 模型);4)本文设计的 CNN 网络无改进模型(下文简称 CNN 无改进模型);5)本文设计的 Bi-2DPCA-CNN 模型。

5 种模型的实验评价指标如表 2 所示。可以看出:a-CNN 模型虽然使用 PCA 算法,但未使用 Bi-2DPCA,仍需要花费大量时间将图片变为一维向量进行样本矩阵的降维,且计算的协方差矩阵维度较大,无法准确地进行特征向量估计,影响了准确率与 F1 值,其准确率和 F1 值分别为 95.82%、95.67%,测试时间为 158 s;b-CNN-2DPCA 模型使用 3 轮卷积池化结构,由于存在较多的可训练参数,训练时间较长,在进行 Bi-2DPCA 改进后,准确率为 98.32%,提升了 0.28%,F1 值为 98.17%,提升了 0.54%,训练时间缩减了 91.5%,测试时间从 244 s 减少到 96 s;本文提出的 Bi-2DPCA-CNN 模型相比于 CNN 无改进模型,准确率为 98.28%,降低了 0.03%,F1 值为 98.26%,提升了 0.02%,训练时间从 3 707.3 s 缩短到 358.2 s,缩减了 90.3%,测试时间从 232 s 减少到 90 s,缩减了 61.2%。图 8 为 Bi-2DPCA-CNN 模型的准确率与损失曲线。

表2 不同模型的评价指标

Table 2 Evaluation indicators of different methods

模型	准确率/%	F1 值/%	训练时间/s	测试时间/s
a-CNN 模型	95.82	95.67	782.1	158
b-CNN 模型	98.04	97.63	4 713.6	244
b-CNN-2DPCA 模型	98.32	98.17	403.2	96
CNN 无改进模型	98.31	98.24	3 707.3	232
Bi-2DPCA-CNN 模型	98.28	98.26	358.2	90

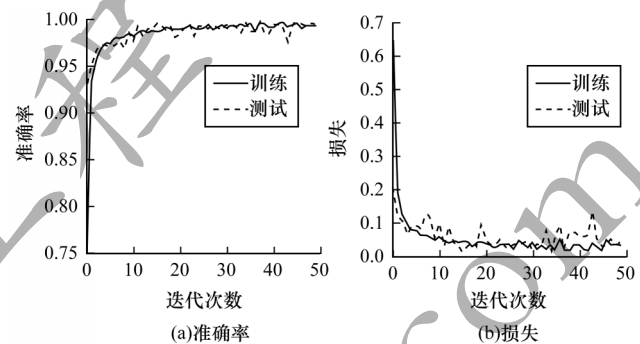


图8 Bi-2DPCA-CNN 模型的准确率与损失曲线

Fig.8 Accuracy and loss cruves of Bi-2DPCA-CNN model

3.3 贝叶斯优化验证

在 Bi-2DPCA-CNN 的基础上继续使用贝叶斯算法优化网络存在的超参数。在上文 2.2 节中,设计了 7 组变量并分别设定了范围,本文使用贝叶斯优化算法,将 7 组变量作为算法的输入,将 Bi-2DPCA-CNN 的准确率作为输出进行自动调参迭代。在进行 5 次初始化寻优和 10 次自动寻优后,得出模型最优参数。贝叶斯优化后的模型最优参数如表 3 所示。

表3 Bi-2DPCA-CNN 模型最优参数

Table 3 Optimal parameters of Bi-2DPCA-CNN model

参数项	最优参数
dim1	3
num_conv1	16
dim2	3
num_conv2	28
pool	3
drop	0.2
lr	0.01

使用贝叶斯优化后的最优模型与未使用贝叶斯优化的模型评价指标如表 4 所示。可以看出:经过贝叶斯优化后的模型性能更好,准确率为 99.15%,提升了 0.87%,F1 值为 99.02%,提升了 0.76%;对比文献[6]未经贝叶斯优化的模型,准确率提升了 1.11%,F1 值提升了 1.39%。

表4 贝叶斯优化前后Bi-2DPCA-CNN性能对比

Table 4 Performance comparison of Bi-2DPCA-CNN before and after Bayesian optimization %

方法	准确率	F1值
贝叶斯优化前	98.28	98.26
贝叶斯优化后	99.15	99.02

图9为贝叶斯优化后Bi-2DPCA-CNN模型的准确率与损失曲线。可以看出:该模型使用Bi-2DPCA算法对数据进行预处理,在保留原图特征的同时降低了尺寸大小,显著减少了运行时间。在使用贝叶斯优化后,准确率与F1值都有所提升。同时,2D2DPCA算法与其他模型结合也有较好的结果,能够有效减少训练时间。

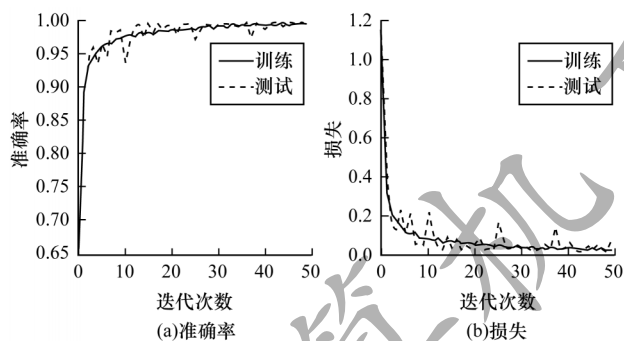


图9 贝叶斯优化后Bi-2DPCA-CNN的准确率与损失曲线

Fig.9 Accuracy and loss cruve of Bi-2DPCA-CNN after Bayesian optimization

4 结束语

针对现有算法识别ASL数据集训练模型速度慢和识别准确率低的问题,本文提出Bi-2DPCA-CNN算法。在图片预处理阶段使用Bi-2DPCA算法对原始图片特征提取降维,保留原图重要特征并减小图片尺寸。在此基础上,设计多层卷积、池化结构的卷积神经网络进行特征图的识别分类,并加入批归一化层与丢弃层防止过拟合。同时,使用贝叶斯优化方法对卷积神经网络中的超参数进行优化,以实现模型的最优化。在对24分类ASL图片进行识别时,该算法在准确率、训练时间、F1值方面较文献[6,10]方法均具有优势,验证了算法的有效性,并且2DPCA算法得到了特征值矩阵,对不同识别任务可获得基于当前数据集的特征值,在其他图像识别问题上也有可拓展性。本文工作仅针对ASL数据集英文静态字母,并未对动态字母及其他手语动作进行识别,下一步研究将聚焦于动态手势识别,包括单张图片出现多个动态手势的复杂情况,以及手势遮挡等问题。

参考文献

- [1] 郝子煜,阿里甫·库尔班,李晓红,等. 基于CapsNet的中国手指语识别[J]. 计算机应用研究, 2019, 36(10): 3157-3159.
HAO Z Y, KUERBAN A, LI X H, et al. Chinese finger language recognition using CapsNet[J]. Application Research of Computers, 2019, 36(10): 3157-3159. (in Chinese)
- [2] CLEBESON C D S, JORGE L A S, RAQUEL F V. Dynamic gesture recognition by using CNNs and star RGB: a temporal information condensation[J]. Neurocomputing, 2020, 400: 238-254.
- [3] TAO W J, MING C L, YIN Z Z. American sign language alphabet recognition using convolutional neural networks with multiview augmentation and inference fusion[J]. Engineering Applications of Artificial Intelligence, 2018, 76: 202-213.
- [4] ASHA T, DIXIT S K. COHST and wavelet features based static ASL numbers recognition[J]. Procedia Computer Science, 2016, 92: 455-460.
- [5] QUTASHAT M, MOUSSA H, BAYAN T, et al. American Sign Language(ASL) recognition based on Hough transform and neural networks[J]. Expert Systems with Applications, 2005, 32(1): 24-37.
- [6] ADITHYA V, RAJESH R. A deep convolutional neural network approach for static hand gesture recognition[J]. Procedia Computer Science, 2020, 171: 2353-2361.
- [7] 柯鹏飞,蔡茂国,吴涛. 基于改进卷积神经网络与集成学习的人脸识别算法[J]. 计算机工程, 2020, 46(2): 262-267, 273.
KE P F, CAI M G, WU T. Face recognition algorithm based on improved convolutional neural network and ensemble learning[J]. Computer Engineering, 2020, 46(2): 262-267, 273. (in Chinese)
- [8] ZHANG Y F, SHI L, WU Y, et al. Gesture recognition based on deep deformable 3D convolutional neural networks[J]. Pattern Recognition, 2020, 107: 1-5.
- [9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Image net classification with deep convolutional neural networks[C]// Proceedings of International Conference on Neural Information Processing Systems. Cambridge, USA: MIT Press, 2012: 1097-1105.
- [10] 吴伟. 基于SAE-PCA模型的ASL字母识别方法研究[D]. 厦门: 厦门大学, 2014.
WU W. Research on ASL letter recognition method based on SAE-PCA model[D]. Xiamen: Xiamen University, 2014. (in Chinese)
- [11] 钟健,何韦颖,谭汉松. 基于PCA降维结合机器学习算法的人机交互手势识别研究[J]. 机床与液压, 2020, 48(6): 181-186.
ZHONG J, HE W Y, TAN S H. Research on human-computer interaction gesture recognition based on PCA dimensionality reduction and machine learning algorithm[J]. Machine Tool & Hydraulics, 2020, 48(6): 181-186. (in Chinese)
- [12] YANG J, ZHANG D D, FRANGI A F, et al. Two-dimensional PCA: a new approach to appearance-based face representation and recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(1): 131-137.

- [13] 胡娜,马慧,湛涛. 融合LBP纹理特征与B2DPCA技术的手指静脉识别方法[J]. 智能系统学报,2019,14(3):533-540.
HU N, MA H, ZHAN T. Finger vein recognition method combining LBP texture feature and B2DPCA technology[J]. CAAI Transactions on Intelligent Systems, 2019, 14(3): 533-540. (in Chinese)
- [14] WANG Y L, ZHANG H X, ZHANG G W. cPSO-CNN: an efficient PSO-based algorithm for fine-tuning hyperparameters of convolutional neural networks[J]. Swarm and Evolutionary Computation, 2019, 49: 114-123.
- [15] 马芳武,韩丽,吴量,等. 基于遗传与粒子群算法的隔振平台减振性能优化[J]. 吉林大学学报(工学版),2020,50(5):1608-1616.
MA F W, HAN L, WU L, et al. Damping optimization of heavy-loaded anti-vibration platform based on genetic algorithm and particle swarm algorithm[J]. Journal of Jilin University(Engineering and Technology Edition), 2020, 50(5): 1608-1616.
- [16] 王晨阳,段倩倩,周凯,等. 基于遗传算法优化卷积长短记忆混合神经网络模型的光伏发电功率预测[J]. 物理学报,2020,69(10):149-155.
WANG C Y, DUAN Q Q, ZHOU K, et al. A hybrid model for photovoltaic power prediction of both convolutional and long short-term memory neural networks optimized by genetic algorithm[J]. Acta Physica Sinica, 2020, 69(10): 149-155. (in Chinese)
- [17] 曾宇,户文成. 贝叶斯优化卷积神经网络公共场所异常声识别[J]. 应用声学,2020,39(3):409-416.
ZENG Y, HU W C. Recognition of abnormal sound in public places based on Bayesian optimal convolutional neural network[J]. Journal of Applied Acoustics, 2020, 39(3): 409-416. (in Chinese)
- [18] 崔佳旭,杨博. 贝叶斯优化方法和应用综述[J]. 软件学报,2018,29(10):3068-3090.
CUI J X, YANG B. Survey on Bayesian optimization methodology and applications[J]. Journal of Software, 2018, 29(10): 3068-3090. (in Chinese)
- [19] WU J, CHEN X Y, ZHANG H, et al. Hyperparameter optimization for machine learning models based on Bayesian optimization[J]. Journal of Electronic Science and Technology, 2019, 17(1): 26-40.
- [20] 李文宽,刘培玉,朱振方,等. 基于卷积神经网络和贝叶斯分类器的句子分类模型[J]. 计算机应用研究,2020,37(2):333-336,341.
LI W K, LIU P Y, ZHU Z F, et al. Sentence classification model based on convolution neural network and Bayesian classifier[J]. Application Research of Computers, 2020, 37(2): 333-336, 341. (in Chinese)
- [21] 生龙,马建飞,杨瑞欣,等. 基于特征交换的CNN图像分类算法研究[J]. 计算机工程,2020,46(9):268-273.
SHENG L, MA J F, YANG R X, et al. Research on CNN image classification algorithm based on feature exchange[J]. Computer Engineering, 2020, 46(9): 268-273. (in Chinese)
- [22] 冯玉芳,殷宏,卢厚清,等. 基于改进全卷积神经网络的红外与可见光图像融合方法[J]. 计算机工程,2020,46(8):243-249,257.
FENG Y F, YIN H, LU H Q, et al. Infrared and visible light image fusion method based on improved fully convolutional neural network[J]. Computer Engineering, 2020, 46(8): 243-249, 257. (in Chinese)
- [23] MATTEO P, LUIGI C, GIUSEPPE P. A light CNN for detecting COVID-19 from CT scans of the chest[J]. Pattern Recognition Letters, 2020, 140: 95-100.

编辑 金胡考