

受损图片分类的双并行交叉降噪卷积神经网络

王 达^{1,2}, 周雪峰², 徐智浩², 蔡奕松², 陈 洲^{1,2}

WANG Da^{1,2}, ZHOU Xuefeng², XU Zhihao², CAI Yisong², CHEN Zhou^{1,2}

1. 沈阳工业大学 信息科学与工程学院, 沈阳 110870

2. 广东省智能制造研究所 广东省现代控制技术重点实验室, 广州 510070

1. School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China

2. Guangdong Key Laboratory of Modern Control Technology, Guangdong Institute of Intelligent Manufacturing, Guangzhou 510070, China

WANG Da, ZHOU Xuefeng, XU Zhihao, et al. Double parallel cross denoising convolution neural network for damaged image classification. Computer Engineering and Applications, 2018, 54(18): 147-153.

Abstract: In order to deal with the classification problem of a large number of disturbed images, a double parallel cross-over noise reduction convolution model is proposed. The model consists of two parts of parallel cross-network structure, corresponding to improved self-coding and parallel cross-convolutional neural network. At the same time, the model uses the method of batch normalization and improvement of traditional activation function in the model training. It is proved by experiments that the model has the characteristics of strong noise reduction, good robustness, strong generalization ability and high accuracy in comparison with the same model. Secondly, it avoids overfitting and speeds up convergence. In the case of different degrees of picture damage, it can also accomplish the task of image target recognition classification well.

Key words: image classification; autoencoder; convolution neural networks; double parallel cross

摘 要: 应对大量受干扰图像的分类问题, 提出了一种双并行交叉降噪卷积模型, 该模型由两部分并行交叉网络结构组成, 分别对应改进的自编码方式和并行交叉卷积神经网络, 同时在该模型训练的过程中, 使用批量正则化和改进激活函数的方法。经实验验证, 与同类模型相比, 该模型首先具有降噪能力强、鲁棒性好、泛化能力强和准确率高的特点, 其次避免过拟合, 加快收敛速度。在图片不同程度受损的情况下, 它也可较好地完成图像目标识别分类任务。

关键词: 图像分类; 自编码器; 卷积神经网络; 双并行交叉

文献标志码: A **中图分类号:** TP183; TP389.1; TP911.73 **doi:** 10.3778/j.issn.1002-8331.1802-0090

1 引言

图像识别与分类技术作为计算机视觉中的一个重要的研究领域, 近些年广泛应用于国防和民用的许多领域。但由于识别背景、环境及目标的复杂性, 机器视觉图像目标的高效、准确识别具有一定的困难。近些年计算机与互联网技术的进步, 使得以神经网络为基础的深度学习得到发展, 运用深度学习的主要目的是建立多层神经网络, 从而实现模仿人脑机制来分析和解释图

像^[1-2]。深度学习通过组合浅层特征形成更加抽象的高层特征, 以发现数据的更深层次分布式特征表示。目前应用在图像识别、场景识别和物体追踪等方面, 并表现出极大的应用价值。但也存在以下问题: 对于受到干扰的图像信息, 识别效果与精度达不到目标; 需要处理大量的训练样本并提取高层次的特征, 因此计算模型复杂, 计算量极大^[3-4]; 由于没有捕捉到图像中物体的特征信息, 整体特征的突然变化对于图像目标的识别非常

基金项目: 国家自然科学基金(No.51405091); 广东省科学院优秀青年科技人才基金(No.rcjj201504); 广州市珠江科技新星专项(No.201610010054)。

作者简介: 王达(1993—), 男, 硕士研究生, 主要研究领域为医疗机器人, 深度学习, E-mail: 379566200@qq.com; 周雪峰(1982—), 男, 博士, 副研究员, 主要研究领域为机器人技术; 徐智浩(1989—), 男, 博士, 助理研究员; 主要研究领域为深度学习, 多目标优化; 蔡奕松(1992—), 男, 助理工程师, 主要研究领域为机器人力控制; 陈洲(1992—), 男, 硕士研究生, 主要研究领域为机器人技术。

收稿日期: 2018-02-07 **修回日期:** 2018-03-27 **文章编号:** 1002-8331(2018)18-0147-07

不利。

针对图像识别领域存在的问题,发展出如卷积神经网络(Convolutional Neural Network, CNN)有效的方法,与普通的神经网络相比, CNN的特点是增加了卷积层和池化层。其中,卷积层的特点提取相邻像素之间的共同特性效果好,具有权值共享特性,能够大量减少神经网络训练过程中的参数数量;池化层的特点则是可以保证位移不变^[5-6]。CNN充分利用隐藏层逐层深入地学习图像的抽象信息,更全面直接地获取图像特征,因数字图像是以矩阵来描述,卷积神经网络能更好地从局部信息块出发,进而描述图像的整体结构^[7]。虽然CNN可以有效地处理信息,但由于真实数据存在许多干扰,其网络模型通常缺少抗噪能力,鲁棒性不强,且层数较多计算量大。

为应对各种受干扰的图像分类问题,本文改进传统自编码方式与CNN模型,结合两种改进方法,模拟人眼的双目结构,提出一种双并行交叉降噪卷积模型(Double Parallel Cross Denoising Convolution model, DPCDC),同时在该模型的训练过程中讨论不同激活函数以及处理数据的方法对其精度的影响。

2 相关工作

2.1 自编码方法

自动编码器是对原始信号进行编码,同时解码恢复数据,从而实现编码解码的整个过程^[8],避免了传统方法不加生成约束的情况下,容易形成直接将输入向量复制到输出向量,或者只作微小的改变以产生较小的重构误差,使得模型表现效果不佳^[9]。自动编码器属于多层前传神经网络是深度学习中常用的模型之一,基于人工神经网络具有网络层次结构的特点。主要由三个部分组成:编码器、解码器、隐含层。工作方式:原始输入 x 经过加权、映射之后得到 y ,再对 y 反向加权映射输出 z 。通过反复迭代训练两组的 w 和 b ,使得误差函数最小,即尽可能保证 z 近似 x ,完成对 x 的重构,具体公式如下:

$$y = f\theta(x) = s(wx + b) \quad (1)$$

$$z = g\theta'(y) = s'(w'y + b') \quad (2)$$

其中 w 与 w' 为两组权值, b 与 b' 为两组偏置, $s()$ 与 $s'()$ 分别表示不同的激活函数, $f\theta()$ 、 $g\theta'()$ 分别表示 x 到 y 与 y 到 z 的映射关系。

2.2 CNN模型

利用深度特征实现图像的分类与识别,目前主要使用CNN模型。对于CNN的基本结构为:输入层、卷积层、池化层、全连接层及输出层构成。其中卷积层由多个特征面组成,每个特征面由多个神经元组成,它的每一个神经元通过卷积核与上一层特征面的局部区域相

连,卷积核是一个权值矩阵(如对于二维而言可为 3×3 或 5×5 矩阵)^[10]。CNN的卷积层通过卷积操作提取输入的不同特征,设第 l 卷积层,第 $l-1$ 层为池化层或者输入层,则第 l 层的计算公式为:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l\right) \quad (3)$$

x_j^l 为 l 层的第 j 个特征图,等式右边表示 $l-1$ 层所有关联的特征图 x_i^{l-1} 和第 l 层的第 j 个卷积核 k_{ij}^l 做卷积运算并求和,其中 b_j^l 表示偏置,最后通过 $f()$ 激活函数计算。

池化层也称为取样层,紧跟在卷积层之后,同样由多个特征面组成,它的每一个特征面唯一对应其上一层的一个特征面,不会改变特征面的个数取样层旨在通过降低特征面的分辨率来获得具有空间不变性的特征^[11]。设第 l 为池化层,第 $l-1$ 层为卷积层,则第 l 层的计算公式为:

$$x_j^l = f(\beta_j^l \text{down}(x_j^{l-1}) + b_j^l) \quad (4)$$

其中 $\text{down}()$ 为下采样函数,可将 $n \times n$ 像素里的所有像素相加,下采样作用于图像不重合的区域,因此所得的特征图大小为原来的 $1/n$ 。 β 表示为权重,对于 2×2 的mean pooling而言,权值为 $1/4$ 。

在CNN结构中,深度越深、特征面数目越多,网络能够表示的特征空间也就越大、网络学习能力也越强,然而也会使网络的计算更复杂,极易出现过拟合的现象。

2.3 激活函数

在传统的CNN中,激励函数一般使用饱和非线性函数(saturating nonlinearity)如sigmoid函数、tanh函数等。激活函数,相比较于饱和非线性函数,不饱和非线性函数能够解决梯度爆炸、梯度消失问题,同时其也能够加快收敛速度^[12]。Jarrett等人^[13]探讨了卷积网络中使用纠正非线性函数,能够显著提升卷积网络的性能,文献[14]也验证了这一结论。而在一般网络结构中基本使用的不饱和非线性函数的激活函数为线性整流函数(Rectified Linear Unit, relu),又称修正线性单元,是一种人工神经网络中常用的激活函数,通常指代以斜坡函数及其变种为代表的非线性函数。relu函数的计算公式如下所示:

$$f(x) = \max(0, x) \quad (5)$$

3 深度学习优化策略

3.1 总体模型设计思路

为应对各种大量受干扰的图像分类问题,同时降低训练模型的复杂程度,本文模拟人眼在观察物体时,通过双目和两条视神经来进行的,两条视神经所传递的视觉信息经过交叉传导,形成交叉混合信息以供大脑分析的方法,对现有的自编码方法和CNN模型进行改进,结合这两种改进模型设计出了一如图1所示,优化的双

并行交叉降噪卷积模型,同时在模型训练时,使用批量正则化和改变传统激活函数的方法来提高效果避免过拟合的情况发生。



图1 双并行交叉降噪卷积框图

该模型主要通过4个方面来分析受干扰数据,提升精度。第一,将改进的自编码方式和并行交叉卷积神经网络进行结合,形成双并行交叉网络结构的双并行交叉降噪卷积模型;第二,传入数据进行噪声训练,通过改进的降噪自编码器处理图片,使模型有较强的鲁棒性,更好的泛化能力;第三,将数据通过并行交叉卷积神经网络对图片进行分类处理操作,提高提取特征的能力,使模型进而提高了识别精度;第四,在训练该模型时,改变传统的激活函数以及引入批量正则化的方法,从而进一步地提升该模型的收敛速度以及精度。最终形成的双并行交叉降噪卷积模型具有模型泛化能力强,大幅度降低传统深度学习的复杂度的特点。以下将具体介绍每一部分。

3.2 优化自编码方法

基于自编码原理,设计一种结合卷积神经网络改进式的并行交叉降噪自编码器,如图2所示。结合卷积神经网络的特点,重复的卷积核能够在检测特征时可以不考虑位置信息,并且共享权值可以极大地减少学习的自由参数个数。其中模型顶层和底层的2条线路用来提取输入图像的高级特征,中间一条执行线路,用来提取输入图像差异性特征信息,首先将数据进行不同破坏率的变化后传入两条并行交叉支路,分别通过一对卷积层处理后,再将数据传输给下一层的池化层,与此同时也将数据传输给交叉输入邻域层用来计算两条线路上每个特征位置与邻域的差异;通过图中nn层改变图片的尺寸,再将数据通过反向卷积处理传到全连接层还原数据,独立完成特征的提取和映射。其特点是融合卷积神经网络既可极大地减少学习的自由参数个数,又可大幅降低深度神经网络的复杂度,但是网络识别精度却没有降低,提高学习效率与实验效果,同时具有较强的鲁棒性,从而增强模型的泛化能力。

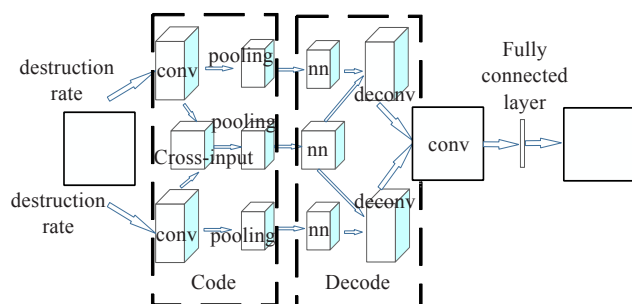


图2 并行交叉卷积降噪自编码器

在模型中,交叉输入邻域层用来计算两条线路上每个特征位置与邻域的差异^[15]。对于交叉层(cross-input)的计算, $f_i(x, y)$ 和 $g_i(x, y)$ 分别为两张输入图像的点 $p(x, y)$ 在第 i 通道特征映射的值,通过一个 $n \times n$ 的块大小来计算差异,如式(6):

$$k_i(x, y) = f_i(x, y)I(n, n) - N(g_i(x, y)) \quad (6)$$

其中, $I(n, n)$ 是单位矩阵, $N(g_i(x, y))$ 表示 $g_i(x, y)$ 的 $n \times n$ 邻域。

通过交叉熵损失函数(Cross-entropy cost function)来衡量预测值与实际值。与二次代价函数相比,它能更有效地加速模型训练。对于整个模型参数的更新,通过损失函数的值进行反向推导,从而改变模型的权值与偏置。在第 L 层的损失函数可表示为:

$$l(x, z) = - \sum_{i=1}^d [x_i \ln z_i + (1 - x_i) \ln (1 - z_i)] \quad (7)$$

其中,原始输入 x_i 破坏为 \hat{x}_1 和 \hat{x}_2 , 设 \hat{x}_i 为两个破坏量的平均值,向量 y_i 为 x_i 的有损压缩量, z_i 为重构向量,其目标函数式:

$$O = \arg \min \frac{1}{n} \sum_{i=1}^n l(x_i, g_{\theta'}[f_{\theta}(\hat{x}_i)]) \quad (8)$$

其中, f_{θ} 与 $g_{\theta'}$ 是 x_i 到 y_i , y_i 到 z_i 的映射关系。在目标函数最优解的过程中,通过使用目标函数不断迭代更新使得误差值最终达到收敛状态。

3.3 改进CNN模型

本文设计了一种改进的共享并行交叉的卷积神经网络,如图3所示,在训练过程中,并行交叉的卷积层同时对一个图片不同的位置进行特征值提取,在模型网络中有3条执行线路,顶层和底层的2条线路用来提取输入图像的高级特征,中间一条执行线路,用来提取输入图像差异性特征信息。

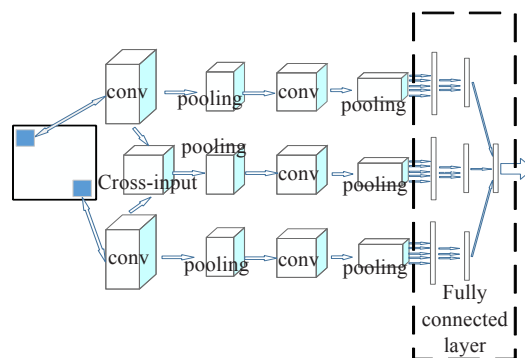


图3 并行交叉卷积网络

这种模型在相应层实现了参数共享,可有效减少整个网络参数的个数,大幅降低深度神经网络的复杂度,但是网络识别精度却没有降低。因为图像的空间联系是局部的,每个神经元只需要感受局部的图像区域,然后在最高层将这些感受到不同局部的神经元综合起来就可以得到全局信息。不同的卷积核参数值不同,这样

就实现多种特征信息的提取,但是大幅降低了网络整体的参数数量,能在网络参数减少的情况下,保证整个深度神经网络的识别精度,同时这种模型也可以提高提取特征的能力。

设 X_L 是前一层的输出值, w_L 是权值矩阵, g_L 是变换函数(或激活函数),在前向传播中,输出函数可记为:

$$z = g_L(X_L; w_L) \quad (9)$$

在模型中,交叉输入邻域层用来计算两条线路上每个特征位置与邻域的差异,通过公式(6)计算。

对于模型的损失,通过 Log-loss 误差函数求出。对于整个模型参数的更新,通过损失函数的值进行反向推导,从而改变模型的权值与偏置。其在第 L 层的损失函数可表示为:

$$l(x_i, y_i) = -\frac{1}{n} \sum_{i=1}^n (y_i - \ln(g_L(x_i; w_L))) + \lambda \sum_{k=1}^L \sum_{j=1}^k \|w_j\|^2 \quad (10)$$

式中, x_i 为样本, n 为样本总数, y_i 为样本标签值,采用 L_2 正则化 (l_2 -norm), λ 为正则化项系数,用于抑制过拟合。可使用批量正则化,效果为减少参数的人为选择,可以取消 L_2 正则项参数,或者采取更小的 L_2 正则项约束参数。整个网络的目标函数为:

$$O = \arg_{w_L} \min \frac{1}{n} \sum_{i=1}^n l(g_L(x_i; w_L), y_i) w_L \quad (11)$$

在求取目标函数最优解的过程中,输出层通过式(11)进行更新(同理,可对偏置进行更新),即:

$$\frac{\partial l}{\partial w_L} = -\frac{1}{n} (Y_L - g_L(X_L; w_L)) \times g'_L(X_L; w_L) + \lambda w_L \quad (12)$$

在求取目标函数最优解的过程中,通过不断使用公式(12),本文使用梯度下降法进行迭代,使得误差值最终达到收敛状态。

3.4 模型性能优化技巧

为避免过拟合以及提高精度,本文使用批量正则化(batch normalization)方法,在数据进入激活函数之前,先对数据进行归一化操作^[16],再选取较好的激活函数进行计算。

对于批量正则化处理方法,所谓批量正则化中的批量是指随机梯度下降中的批量样本,正则化是指让每一

层激活值的每一维度的概率分布转变为均值为0,方差为1的稳定概率分布。通过数据进入激活函数之前,先对数据进行归一化操作。其特点是可以把过大或过小的数据进行归一化,可使整体使用一个较高的学习率,避免传统方法需迁就小数据的维度;其次是归一化后使得更多的权重分界面落在了数据中,降低了过拟合的可能性,因此使用一些防止过拟合但会降低速度的方法。

假设某一层的输入为 $x = (x(1), x(2), \dots, x(k))$, 共 k 维,一批样本集合为 $B = \{x_1, x_2, \dots, x_m\}$, 批量正则化的公式如下:

$$u_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (13)$$

$$\delta_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - u_B)^2 \quad (14)$$

$$\hat{x}^{(k)} = \frac{x^{(k)} - u_B}{\delta_B} \quad (15)$$

$$y^{(k)} = r^{(k)} \hat{x}^{(k)} + \beta^{(k)} \quad (16)$$

其中, $x^{(k)}$ 表示输入 x 的第 k 维, u_B 表示对样本集合 B 的期望, δ_B^2 表示对样本集合 B 的方差, $\hat{x}^{(k)}$ 表示输入 x 的正则化结果, $y^{(k)}$ 表示对 $x^{(k)}$ 批量正则化后的结果。 $r^{(k)}$ 、 $\beta^{(k)}$ 表示待学习的参数,通过训练迭代求解,以增加输出概率分布的稳定性。

对于激活函数,本文选取 relu6、sigmoid、elu 激活函数替代传统激活函数 relu 进行对比实验,经实验验证,基于本文方面 relu6 与 elu 相对于 relu 表现效果更好,收敛更快。relu6、sigmoid、elu 函数计算公式如下所示:

$$f(x) = \max(0, 6, x) \quad (17)$$

$$f(x) = 1 / (1 + \exp(-x)) \quad (18)$$

$$f(x) = \begin{cases} x, & x > 0 \\ \alpha(\exp(x) - 1), & x \leq 0 \end{cases} \quad (19)$$

结合以上各部分具体描述,以及两个并行交叉网络结构,形成图4所示的双并行交叉降噪卷积模型,该模型首先具有降噪能力强,鲁棒性好,泛化能力强和准确率高的特点,其次避免过拟合,加快收敛速度。

对于整体的损失计算如下:

$$\text{loss} = \alpha l(x, z) + \beta l(x_i, y_i) \quad (20)$$

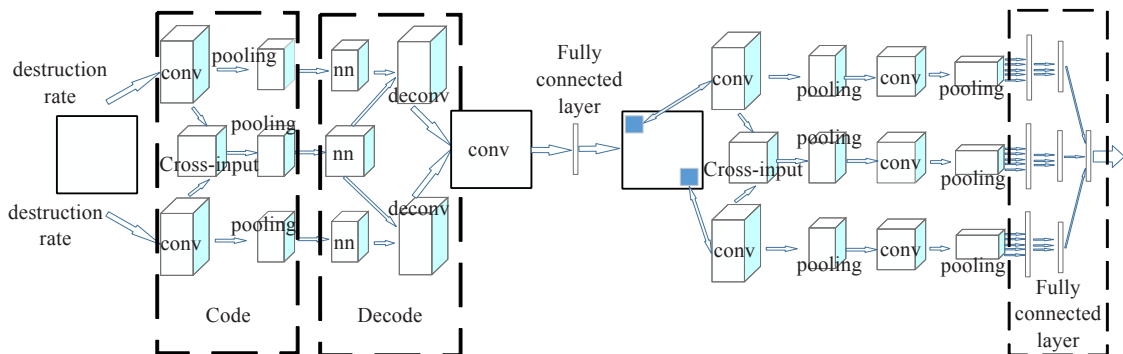


图4 双并行交叉降噪卷积模型

其中, α , β 为降噪和分类权重, $loss$ 为整体的损失。在目标函数(20)中,最优解的过程通过使用目标函数不断迭代更新使得误差值最终达到收敛状态。

$$O = \arg \min \frac{1}{n} \sum_{i=1}^n loss \quad (21)$$

对于权值 w 和阈值 b 的更新使用公式(22)与(23),其中 η 为学习率。

$$w = w' - \eta \times \frac{\partial loss}{\partial w'} \quad (22)$$

$$b = b' - \eta \times \frac{\partial loss}{\partial b'} \quad (23)$$

其中,准确率的计算公式如公式(24):

$$a = \frac{g_z}{g} \times 100\% \quad (24)$$

其中, a 为准确率, g_z 为一批测试中识别正确的个数, g 为一批测试总个数。

4 实验与结果分析

4.1 实验准备

本文实验使用 Windows 系统,基于 tensorflow1.2 机器学习框架通过 spyder 软件进行实验,其中图 5 为软件效果图展示,使用的数据集为 minst 数据集,为手写数字数据库,训练库有 55 000 张手写数字图像,测试库有 10 000 张,每个图像 $784=28 \times 28$ 。其中表 1、2 为模型两部分的各层参数,通过在图片中人为加入噪声模拟受损图片。由于数据集是识别数字 0~9 的手写图片,所以解决的问题为图片分类问题,需要分类数为 10 种类型。

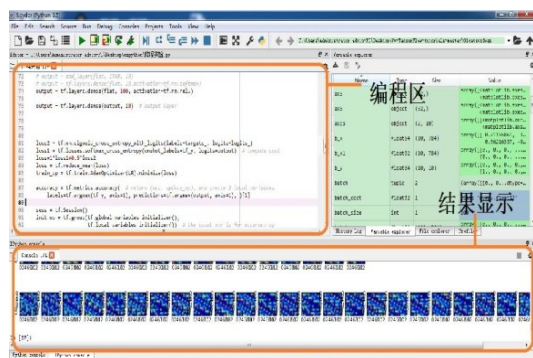


图5 软件截图效果显示

表1 自编码网络

层	类型	特征图	核大小
0	卷积	$28 \times 28 \times 64$	$3 \times 3 \times 64$
1	池化	$14 \times 14 \times 64$	$2 \times 2 \times 64$
2	NN	$28 \times 28 \times 64$	$2 \times 2 \times 64$
3	卷积	$28 \times 28 \times 64$	$3 \times 3 \times 64$
4	卷积	$28 \times 28 \times 1$	$3 \times 3 \times 1$
5	全连接	$28 \times 28 \times 1$	$28 \times 28 \times 1$

4.2 激励函数效果分析

本文对于模型激励函数的选取,通过在卷积层与全

表2 卷积分类网络

层	类型	特征图	核大小
0	卷积	$28 \times 28 \times 16$	$5 \times 5 \times 16$
1	池化	$14 \times 14 \times 16$	$2 \times 2 \times 16$
2	卷积	$14 \times 14 \times 32$	$5 \times 5 \times 32$
3	池化	$7 \times 7 \times 32$	$2 \times 2 \times 32$
4	全连接	$1\ 568 \times 1$	$1\ 568 \times 1$
5	全连接	100×1	100×1
6	全连接	10×1	10×1

连接层进行实验分析,选取激活函数为使用最广泛的 relu 与 relu6、sigmoid、elu 进行 4 组对比实验,图 6、7 为在不同激活函数情况下准确率的变化。

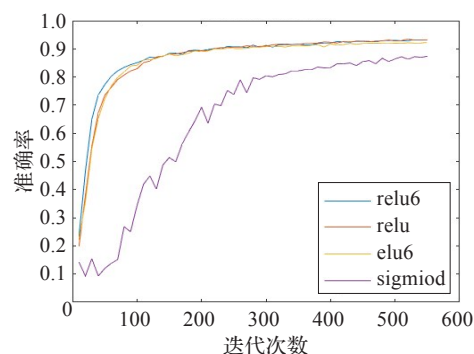


图6 卷积层不同激活函数准确率变化曲线

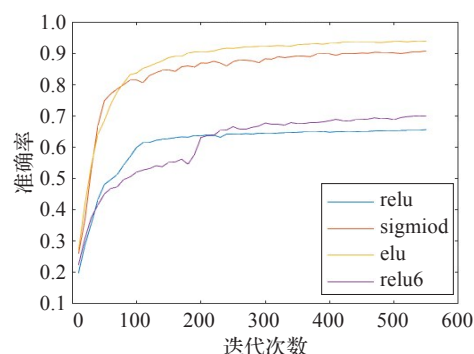


图7 全连接层不同激活函数准确率变化曲线

根据实验表明,对于本实验,在卷积层进行不同激活函数实验时,relu6 虽然在准确率上与 relu 相似,但在收敛速度上 relu6 明显优于 relu;在全连接层进行不同激活函数实验时,在最终准确率上 elu 的效果明显优于 relu 的效果。经过实验分析,模型在卷积层使用 relu6 激活函数,在全连接层处使用 elu 激活函数效果最佳。以此模型选取在卷积层使用 relu6 激活函数,全连接层使用 elu 激活函数。

4.3 降噪效果分析

如图 8、9 所示,图 8 为原始图片加噪声效果,图 9 为传统降噪方法与本文降噪效果对比展示。在图 8 中显示,原始图的特征显示为红色部分。对比发现,传统编码降噪方法虽然可以去除图片中的噪声干扰,还原部分原始图片中红色部分,但效果不是特别明显,改进的降噪自编码器的效果明显优于传统自编码器的效果,可以

有效地处理图片中的干扰,虽然无法全部排除干扰,但效果明显优于传统方法,对于传统方法几乎无法还原图中数字3与4,改进的方法则可以相对地还原出原始效果。

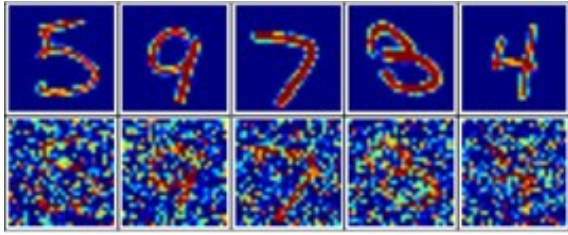


图8 原始图片加噪声效果

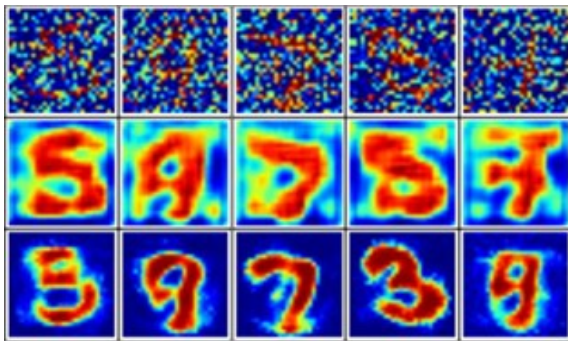


图9 降噪效果对比展示

4.4 模型效果展示及精度分析

图10为数据集随机中选出个体展示,图11与图12为三条线路上卷积后的部分可视化效果图展示,图13与图14为三条线路上池化后的部分可视化效果图展示。

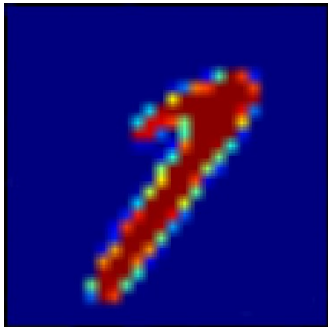


图10 数据集随机选出个体

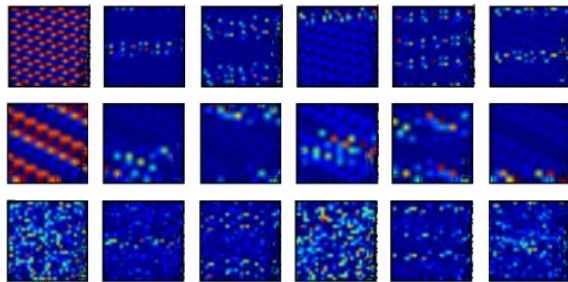


图11 降噪阶段卷积后部分效果

在不同情况下分别进行5组实验取平均值,选取10个不同图片损坏程度,当图片损坏程度为10时,说明图片几乎无法分类。分别表示测试图片集不同受损情况。建立表3至表5,选取BP与CNN与本文双并行交

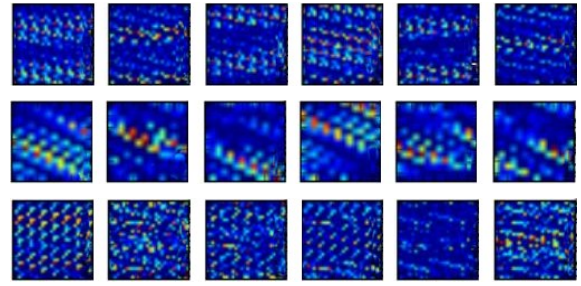


图12 分类阶段卷积后部分效果

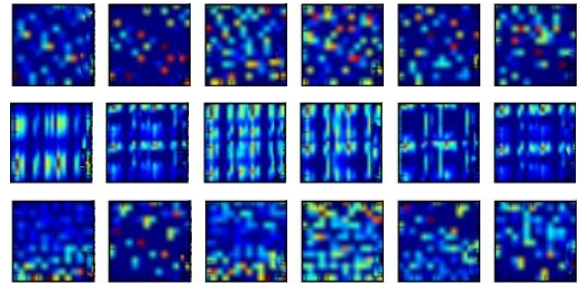


图13 降噪阶段池化后部分效果

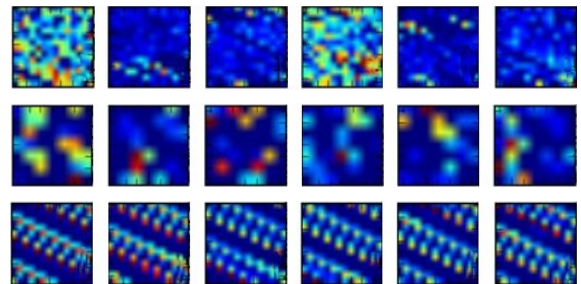


图14 分类阶段池化后部分效果

叉降噪卷积模型(DPCDC)进行对比实验,同时记录不同迭代次数下准确率的情况。

表3 测试图片无损失

模型	50	100	300	550
BP	63.1	78.3	83.2	84.8
CNN	56.9	66.0	82.7	90.0
DPCDC	59.6	73.8	85.7	90.3

表4 测试图片损失率为2

模型	50	100	300	550
BP	38.7	45.1	52.0	53.4
CNN	41.3	54.7	73.4	81.6
DPCDC	60.0	70.3	81.1	88.4

表5 测试图片损失率为5

模型	50	100	300	550
BP	19.0	24.3	27.4	28.1
CNN	21.9	25.6	33.1	35.9
DPCDC	20.5	38.4	43.1	58.4

实验结果表明,对于图片分类问题BP精度在任何情况下都无法达到CNN与DPCDC的精度。在图片受到干扰时CNN模型精度明显下降,当图片受损率达到程度5时,分类图片的精度与BP相差不大。而本文

DPCDC 模型当图片无损失时,识别精度与 CNN 相似,当损失继续加大时,虽然精度开始下降,但效果明显优于 CNN 模型,损率达到程度 5 时也可以满足简单的分类处理。

图 15 为在损失率为程度 5 时,有无批量正则化对比实验,当模型中不使用批量正则化时,5 次实验中有 1 次出现图中红色线无批量正则化的实验结果。实验结果说明当无批量正则化操作时,实验结果容易陷入过拟合,影响准确率。

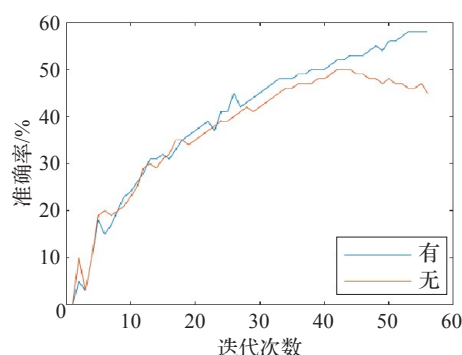


图 15 有无批量正则化对比实验

5 结论

为了应对各样大量受干扰的图像分类问题,改进传统编码方式与 CNN 模型,结合两种改进的方法,提出了一种双并行交叉降噪卷积模型。首先传入数据进行噪声训练,通过改进的降噪自编码器处理图片,使模型具有较强的鲁棒性,更好的泛化能力;再将数据通过并行交叉卷积神经网络对图片进行分类处理操作,提高提取特征的能力,使模型进而提高了识别精度。双并行交叉降噪卷积模型优点在于增强模型的泛化能力,也可大幅度降低传统深度学习的复杂度,对于受损图片进行识别,效果明显优于其他传统方法。而在训练该模型时,改变了传统的激活函数以及引入了批量正则化的方法,从而进一步地提升了该模型的收敛速度以及精度。实验结果证明了模型的合理以及有效性。

参考文献:

- [1] Choi J Y, Yong M R, Plataniotis K N. Color local texture features for color face recognition[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2012, 21(3): 1366-80.
- [2] McCulloch W S, Pitts W A. A logical calculus of the ideas immanent in nervous activity[J]. Bulletin of Mathematical Biophysics, 1943, 5(4): 115-133.
- [3] 卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. 数据采集与处理, 2016, 31(1): 1-17.
- [4] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems, Curran Associates Inc, 2012: 1097-1105.
- [5] Lee H, Grosse R, Ranganath R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations[C]//International Conference on Machine Learning, 2009: 609-616.
- [6] Huang G B. Learning hierarchical representations for face verification with convolutional deep belief networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2012: 2518-2525.
- [7] 王伟凝, 王励, 赵明权, 等. 基于并行深度卷积神经网络的图像美感分类[J]. 自动化学报, 2016, 42(6): 904-914.
- [8] Vincent P, Larochelle H, Lajoie I, et al. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion[J]. Journal of Machine Learning Research, 2010, 11(12): 3371-3408.
- [9] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders[C]//International Conference on Machine Learning, 2008: 1096-1103.
- [10] Haykin S, Kosko B. Gradient based learning applied to document recognition[M]. [S.l.]: Wiley-IEEE Press, 2009.
- [11] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [12] Xu B, Wang N, Chen T, et al. Empirical evaluation of rectified activations in convolutional network[J]. Computer Science, 2015.
- [13] Jarrett K, Kavukcuoglu K, Ranzato M, et al. What is the best multi-stage architecture for object recognition[C]//IEEE International Conference on Computer Vision, 2010: 2146-2153.
- [14] Nair V, Hinton G E. Rectified linear units improve restricted boltzmann machines[C]//International Conference on International Conference on Machine Learning, 2010: 807-814.
- [15] Jiang Y G, Wu Z, Wang J, et al. Exploiting feature and class relationships in video categorization with regularized deep neural networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99): 1.
- [16] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//International Conference on International Conference on Machine Learning, 2015: 448-456.
- [17] Nair V, Hinton G E. Rectified linear units improve restricted boltzmann machines[C]//International Conference on Machine Learning, 2010: 807-814.
- [18] Yoo H J. Deep convolution neural networks in computer vision[J]. IEEE Transactions on Smart Processing & Computing, 2015, 4(1): 35-43.