

DOI: 10.16652/j.issn.1004-373x.2021.04.038

引用格式:郭原东,雷帮军,聂豪,等.基于深度学习的智能高精度图像识别算法[J].现代电子技术,2021,44(4):173-176.

基于深度学习的智能高精度图像识别算法

郭原东, 雷帮军, 聂豪, 李 讷

(三峡大学, 湖北 宜昌 443000)

摘 要: 针对高精度训练样本缺失场景下图像识别算法泛化能力差的问题,文中提出一种改进的深度置信网络结构(DBNs)。该结构通过在DBN网络中引入随机隐退机制,使得隐含层中的部分单元失效,仅保留其连接权重,防止小样本量训练过程中产生的过拟合现象。为降低引入随机隐退机制后算法的计算复杂度,在该结构中引入基于相邻近算法的降采样机制。采用ORL开放人脸数据集进行仿真实验,结果表明,该机制可以将识别错误率由普通DBN网络的43%降低到5.0%,但计算时间有所增加。对比引入降采样算法后的网络测试结果显示,网络训练时间下降约69.9%;与AlexNet等公开网络的对比测试结果表明,该算法的识别精度可达95.2%,在计算精度与识别效率上均有一定的优越性。

关键词: 图像识别; 深度学习; 随机隐退; 图像降采样; 仿真实验; 网络测试

中图分类号: TN911.73-34; TP311

文献标识码: A

文章编号: 1004-373X(2021)04-0173-04

Intelligent high precision image recognition algorithm based on deep learning

GUO Yuandong, LEI Bangjun, NIE Hao, LI Ne

(China Three Gorges University, Yichang 443000, China)

Abstract: In allusion to the problem of the poor generalization ability of image recognition algorithm in the missing scene of high precision training samples, an improved depth belief network structure (DBNs) is proposed. In the structure, the random fading mechanism is introduced into the DBN network to make some units in the hidden layer invalid and only retain their connection weights, so as to prevent the over fitting phenomenon in the process of small sample size training. In order to reduce the computational complexity of the algorithm after introducing the random fading mechanism, a downsampling mechanism based on the adjacent near algorithm is introduced into the structure. The simulation experiments with ORL open face dataset show that the recognition error rate can be reduced from 43% to 5.0% in general DBN network, but the calculation time is increased. The network testing results after introducing the downsampling algorithm show that the network training time is reduced by about 69.9%. The comparison test results with public networks such as Alexnet show that the recognition accuracy of the algorithm can be improved up to 95.2%, which has certain advantages in calculation accuracy and recognition efficiency.

Keywords: image recognition; deep learning; random fading; image downsampling; simulation experiment; network testing

0 引 言

随着计算机计算能力的增长,深度学习算法已在图像处理、图像识别领域有了广泛应用。算法的应用除了算力的支持,还需大量的图像样本,但对于高精度的图像样本,通常存在数量的缺失带来的问题。当使用较少样本数量的图像进行深度网络训练时,容易使网络陷入局部最优点,从而产生过拟合现象。此时,训练得到的网络虽然在训练集上有较优的识别精度,但算法在测试

集上的泛化能力差,无法满足工业界的使用标准^[1-6]。近年来,学术界通过对样本缺失条件下的深度学习理论进行了众多研究,文中在这些理论的研究基础上引入了随机隐退机制^[7-10]。该方法的思想是在某些迭代过程中,随机地选择一定数量的隐含层单元使其失效,此方法的引入可以避免过拟合现象。本文还引入了降采样机制,用来降低引入随机失效方法后算法时间复杂度过高的问题。

1 理论基础

1.1 深度学习

本文使用的深度学习算法为深度置信神经网络

收稿日期:2020-07-20

修回日期:2020-09-01

基金项目:湖北省科技厅项目(2019ZYYD007)

(Deep Belief Networks, DBNs)。在该网络中,引入受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)改变算法的训练结构,并改善卷积神经网络(Convolutional Neural Networks, CNN)训练效率低的问题,在图像识别领域有较广泛的应用。该网络中的RBM结构如图1所示^[11]。

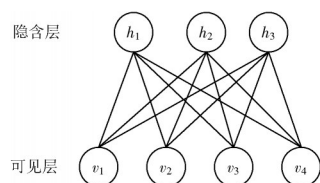


图1 RBM基本结构

其训练与学习方法如下,一个RBM的能量分布为:

$$E(v, h | \theta) = - \sum_{i=1}^n \eta_i v_i + \sum_{j=1}^m \mu_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i w_{ij} h_j \quad (1)$$

在训练中,其目的在于确定RBM的参数三元组 $\theta = (\eta_i, \mu_i, W)$ 。本文通过极大似然函数法进行学习,首先定义目标函数 $l(\theta)$,如下:

$$\theta^* = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^T \log P(v^i | \theta) \quad (2)$$

根据对比散度计算规则,对式(2)中的参数三元组计算偏导如下:

$$\begin{aligned} \frac{\partial \log(v | \theta)}{\partial w_{ij}} &= \varepsilon \left(\langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}} \right) \\ \frac{\partial \log(v | \theta)}{\partial \eta_i} &= \varepsilon \left(\langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{model}} \right) \\ \frac{\partial \log(v | \theta)}{\partial \mu_i} &= \varepsilon \left(\langle h_i \rangle_{\text{data}} - \langle h_i \rangle_{\text{model}} \right) \end{aligned} \quad (3)$$

当RBM中包含 n 个可见单元与 m 个隐含单元时,在算法迭代的过程中,其参数更新的方法如下:

$$\begin{aligned} W &= W + \varepsilon \left(P(h_1 = 1 | v_1) v_1^T - P(h_2 = 1 | v_2) v_2^T \right) \\ \eta &= \eta + \varepsilon (v_1 - v_2) \\ \mu &= \mu + \varepsilon \left(P(h_1 = 1 | v_1) - P(h_2 = 1 | v_2) \right) \end{aligned} \quad (4)$$

式中: v_i 代表可见层的单元; h_i 代表隐含层的单元。

1.2 随机隐退

当深度网络的结构较为复杂,而图像样本的数量较少时,网络的训练容易产生过拟合现象,且网络不具备泛化能力。此时,可以引入随机隐退机制,其原理如图2所示。

图2给出随机隐退机制的示意图。该机制为了避免过拟合现象,在训练的过程中,随机地删除隐含层中的神经元数量。当节点被删除后,视作已“隐退”。该节点在每次迭代中只保留当前的权重,不再进行更新。

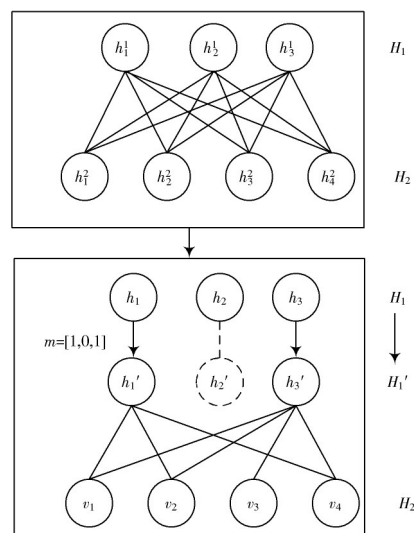


图2 随机隐退机制

随机隐退时,在隐退层引入过滤函数,过滤函数直接作用于隐含层 H_1 ,过滤后得到新的隐含层^[12-13],如下:

$$H'_1 = H_1 \cdot m \quad (5)$$

对于线性网络,其 h 层的第 i 个单元的表达式如下:

$$S_i^h = \sum_{i < h} \sum_j w_{ij}^h S_j^i \quad (6)$$

式中, $S_j^0 = I_j$, I_j 是第 j 个元素为1的单位向量。其次,在该网络中引入随机隐退机制后,该单元的输出如下:

$$S_i^h = \sum_{i < h} \sum_j w_{ij}^h \delta_i^h S_j^i \quad (7)$$

式中 δ 是一个服从伯努利分布的变量。此时,可以得到这个单元的数学期望,如下:

$$E(S_i^h) = \sum_{i < h} \sum_j w_{ij}^h p_j^i E(S_j^i) \quad (8)$$

由于神经网络是一个非线性网络,需要引入激活函数,本文使用的激活函数为Sigmoid函数,引入该函数后,网络的输出如下:

$$O = \sigma(S) = (1 + ce^{-\lambda s})^{-1} \quad (9)$$

2 方法实现

2.1 数据预处理

对于高精度的图像识别场景,训练样本通常不足,本文使用的样本库为ORL人脸库。在该人脸库中,共包含400张图片,这400张图片对应了40个不同个体,每个个体有10张不相同的图片。图3给出了其中1个个体的部分人脸图像。

首先,将ORL人脸库的图像划分为训练集与测试集,训练集与测试集中包含每个个体的各5张图像,每个集内有200个图像样本。由于本文算法引入随机隐

退与降采样机制,因此需要分别评估这两个方法对于算法性能的影响。



图3 ORL人脸库

在引入随机隐退机制时,设置网络的参数,如表1所示。

表1 引入隐退机制的网络参数

参数	参数值	备注
h_1	500	隐含层一节点数
h_2	500	隐含层二节点数
n	200	迭代次数
η	0.01	学习率
σ	50%	隐退率

图4中的虚线给出了引入50%隐退机制后,DBN网络的错误率随着迭代次数增长的变化情况。为了更优地评估隐退机制引入后的效果,使用普通的反向传播网络进行对比。由图4可以看出,普通的反向传播在迭代次数为50次时,可达到最优的错误率43%;50%隐退网络的错误随着迭代次数的进行不断降低,最终可达到5%。

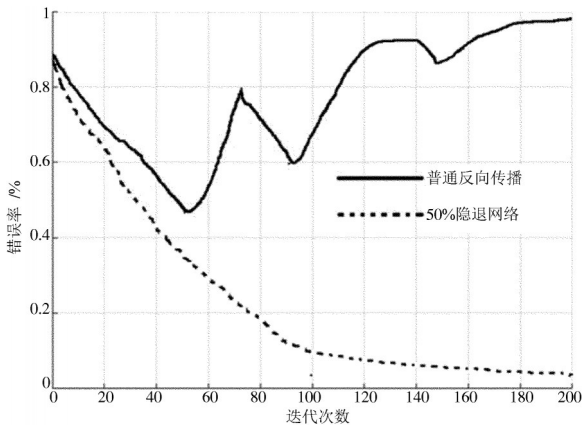


图4 算法错误率和迭代次数的关系

随后,在隐退网络上加入降采样机制,由于图像进行了降采样处理,每个输入样本包含的样本信息也随之

降低。因此,网络的学习率可以根据采样率降低为0.005,隐退率也可以降低为25%。在使用相邻近算法进行ORL库的降采样时,会消耗约1 s的计算时间。

2.2 仿真结果

通过前期的处理,共获得了3个不同网络,分别是普通的DBN网络1、引入隐退机制的DBN网络2、引入降采样与隐退机制的DBN网络3。在训练过程中,其错误率与计算时间的模型参数,如表2所示。

表2 模型参数

网络	错误率 /%	计算时间 /s
1	43	531
2	5.0	722
3	5.2	217

由表2可以看出,普通的DBN网络在本文的图像训练样本缺失的高精度图像识别场景中,错误率与训练的时间均较高。当在该网络中引入随机隐退机制后,网络的错误率有明显的改善,由43%降低到5.0%。但该机制的引入另一方面降低了网络的训练效率,网络的训练时间由531 s提升到722 s。然后通过再次引入降采样机制,使得网络的性能与训练效率均有明显的改善。虽然网络的错误率由5.0%略微增加到5.2%,但网络的训练时间由722 s降低到217 s,降低了69.9%。最终,文中将本文方法与现有的一些图像识别方法在OLP数据集上的计算结果进行了对比,算法对比结果如表3所示。

表3 算法对比结果 %

网络	识别率
AlexNet	94.1
LaNet-5	93.6
2DDCCA	93.5
本文方法	95.2

由表3的结果可以看出,本文算法的识别率可达到95.2%,相较当前识别率较优的AlexNet网络,可以提升约1.1%,具有在工业界进行应用与推广的潜力。

3 结 论

高精度的图像识别算法在训练前首先需要搜集大量的训练样本,这是一个耗时、耗力的过程。本文设计的算法中引入了随机隐退机制,使得深度学习算法对于样本数量的需求大幅度降低。从算法的测试结果可以看出,本文算法较大地开拓了深度学习算法的适用范围,对于算法在工业界的深入应用具有较强的推动作用。

参 考 文 献

- [1] DUAN M, LI K, YANG C, et al. A hybrid deep learning CNN-ELM for age and gender classification [J]. Neurocomputing, 2018, 75(2): 448-461.
- [2] LU N, WU Y D, FENG L, et al. Deep learning for fall detection: 3D-CNN combined with LSTM on video kinematic data [J]. IEEE journal of biomedical & health informatics, 2018 (7): 13-21.
- [3] WANG Ping, JIANG Lei. Deep learning - based object classification through multimode fiber via a CNN-architecture SpeckleNet [J]. Applied optics, 2018(3): 233-242.
- [4] ZHANG M D, LIAO G S, HE X P, et al. Unambiguous forward - looking SAR imaging on HSV - R using frequency diverse array [J]. Sensors, 2020, 20(4): 1169-1183.
- [5] LI D, LI R, ZHANG S. A deep learning image recognition framework accelerator based parallel computing [C]// The 2nd International Conference. Beijing: ACM, 2018: 101-110.
- [6] HUANG H W, LI Q T, ZHANG D M. Deep learning based image recognition for crack and leakage defects of metro shield tunnel [J]. Tunnelling and underground space technology, 2018, 77(7): 166-176.
- [7] HUANG Hongwei, LI Peng. Deep learning based image recognition for crack and leakage defects of metro shield tunnel [J]. Tunnelling & underground space technology, 2018(1): 31-40.
- [8] YANG X, ZENG Z, TEO S G, et al. Deep learning for practical image recognition: case study on kaggle competitions [C]// The 24th ACM SIGKDD International Conference. Hangzhou: ACM, 2018: 42-55.
- [9] AIZEZI Y, JIAMALI A, ABUDUREXITI R, et al. Research on image recognition method based on deep learning algorithm [C]// IEEE International Conference on Advanced Video & Signal Based Surveillance. Boston: IEEE, 2018: 312-320.
- [10] MEZGEC S, EFTIMOV T, BUCHER T, et al. Mixed deep learning and natural language processing method for fake-food image recognition and standardization to help automated dietary assessment [J]. Public health nutrition, 2018(6): 1-10.
- [11] POSTALCIOGLU S. Performance analysis of different optimizers for deep learning based image recognition [J]. International journal of pattern recognition & artificial intelligence, 2019(8): 1120-1133.
- [12] HUANG L Q, LI J, HAO H, et al. Micro - seismic event detection and location in underground mines by using Convolutional Neural Networks (CNN) and deep learning [J]. Tunnelling and underground space technology, 2018, 81 (11): 265-276.
- [13] 蔡昌许. 基于布隆过滤器的WSN链路层地址隐藏方法[J]. 计算机工程与设计, 2016, 37(5): 1208-1211.

作者简介: 郭原东(1995—), 男, 贵州黔南人, 硕士研究生, 研究方向为深度学习、图像识别、基于人体的行为分析等。