

基于模糊深度学习网络的行人检测方法

邱 东,刘德雨

(长春工业大学 电子与电气工程学院,吉林 长春 130000)

摘 要:行人检测已经成为机器视觉研究的重要组成部分之一。深度置信网络(deep belief networks,DBN)优秀的学习能力保证了其学习得到的目标特征更加有效,有利于实现目标的准确检测。但是传统的深度置信网络模型对整体的目标进行处理,训练时间长,同时需要将所有的样本都进行预先正确的标注,这些都限制了深度置信网络的进一步发展。对此,文中提出了一种基于多特征的模糊深度置信网络的方法,该方法将经典的深度置信网络与模糊集的理论相结合,融合方向直方图特征对复杂背景下的行人进行检测识别处理。在静态行人检测库 INRIA 的测试结果表明,该方法在一定程度上减少了训练时间,同时也提高了行人检测的准确率。

关键词:机器视觉;多特征;模糊集;深度置信网络;行人检测

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2018)10-0022-05

doi:10.3969/j.issn.1673-629X.2018.10.005

A Pedestrian Detection Method Based on Fuzzy Depth Learning Network

QIU Dong,LIU De-yu

(School of Electrical and Electronic Engineering,Changchun University of Technology,Changchun 130000,China)

Abstract: Pedestrian detection has become an important part of machine vision research. Deep belief networks (DBN) has excellent learning ability to ensure that the learning of the target features more effective, which is conducive to achieving the goal of accurate detection. However, the traditional deep belief network model deals with the overall goal, which needs a long training time and all the samples need to be correctly labeled at the same time, which are limiting the further development of the deep confidence network. Therefore, we propose a method based on multi-feature fuzzy depth belief network, which combines the classical depth belief network with the theory of fuzzy sets, and the direction histogram features to detect and identify pedestrians in complex background. The test in the static pedestrian detection library INRIA shows that the method not only reduces the training time to a certain extent, but also improves the accuracy of pedestrian detection.

Key words: machine vision; multi-feature; fuzzy set; depth confidence network; pedestrian detection

0 引 言

随着机器视觉技术和人工智能的迅速发展,行人检测逐步成为军事、社会公共安全、交通和互联网发展等社会领域里的热门研究课题之一。但是行人检测问题涉及到模式识别、图像处理、计算机视觉和机器学习等多学科,同时受到行人穿着、光照、身体姿态、尺度、视角和复杂背景等因素的影响,至今也未能(也不可能)设计出一种通用的、实时的、鲁棒的检测算法。

2006 年,Hinton 教授在《Science》上提出了一种面向复杂通用学习任务的深度神经网络,指出具有大量隐藏层的神经网络具有非常好的特征学习能力,而且

网络的训练可以采用“逐层初始化”和“反向微调”技术解决,从此开启了对人工智能领域的研究热潮,深度学习的概念开始被提出。深入学习领域的大多数研究人员专注于深度网络设计和相应的快速学习算法,一些研究工作试图改进深度学习技术模型表示。孙劲光等^[1]提出数值属性的 DBN,并在 UCI 的多个数据集上进行对比验证,证明了其有效性。N. Wang 等^[2]在 2014 年提出了具有高斯线性单位的高斯限制玻尔兹曼机器(GRBM)来学习来自实值数据的表示,通过用高斯函数替换二进制值可见单元来改进 RBM。许庆勇等^[3]在 2015 年提出了一种基于多特征融合的深度

收稿日期:2017-11-14

修回日期:2018-03-20

网络出版时间:2018-05-28

基金项目:中央引导地方科技发展专项(20175000702)

作者简介:邱 东(1969-),男,博士,教授,硕导,研究方向为智能测试与优化;刘德雨(1989-),男,硕士研究生,研究方向为计算机视觉。

网络出版地址:<http://kns.cnki.net/kcms/detail/61.1450.TP.20180525.1558.036.html>

置信网络图像分类算法,通过提取样本图像中的颜色、纹理和形状特征,构成多特征融合的权重矩阵,并对特征矩阵进行归一化处理,利用构建的 4 层 DBN 分类器进行训练和分类。

文中将传统深度置信网络同模糊集理论相结合,提出一种基于多特征的模糊深度置信网络的行人检测方法。该方法在传统的深度学习模型深度置信网络的基础上引入模糊集的理论思想,一方面把深度学习的典型结构之一深度置信网络与模糊算法相结合,构建用于图像分类和识别的模型;另一方面,利用模糊受限玻尔兹曼机构建深度网络,同时改进训练过程,用于图像识别,以进一步提高行人检测的正确率。

1 基于模糊深度学习网络的行人检测方法

1.1 模糊理论

对称三角模糊数如图 1 所示。

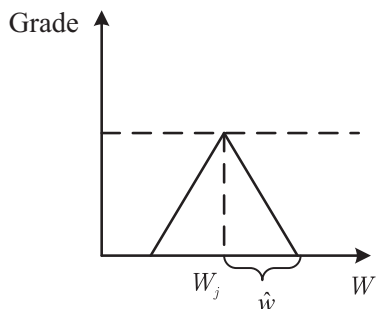


图 1 对称三角模糊数

向量 $\bar{W} = [\bar{w}_1, \bar{w}_2, \dots, \bar{w}_n]^T$ 是对称三角模糊数,其中第 j 个模糊参数 W_j 的成员函数有如下定义:

$$\bar{W}_j(w) = \max \left\{ 1 - \frac{|w - w_j|}{\hat{w}_j}, 0 \right\} \quad (1)$$

其中, W_j 为模糊数的中心; \hat{w}_j 为模糊数的宽度。

用 \bar{W}_j 表示一个模糊集,其中 $\bar{W}_j(w)$ 表示第 w 个成员值。

1.2 模糊受限玻尔兹曼机

受限玻尔兹曼机 (RBM)^[4] 是由 Hinton 和 Sejnowski 于 1986 年提出的,由一个可见层和一个隐藏层构成,可见层与隐藏层的神经元之间为双向全连接,是一种可通过输入数据集学习概率分布的随机生成神经网络。标准的受限玻尔兹曼机是由二值(布尔/伯努利)隐层和可见层单元组成,该模型是用参数 θ 表示跨层节点之间的权值和偏置的连接。权重矩阵 $W = (w_{ij})$ 中的每一个元素指定了隐藏层单元 h_j 和可见单元 x_i 之间的权重度置信网络。受限玻尔兹曼机是基于能量的概率模型,定义概率分布如下:

$$P(x, h, \theta) = \frac{e^{-E(x, h, \theta)}}{Z} \quad (2)$$

$$z = \sum_{\tilde{x}} \sum_{\tilde{h}} e^{-E(\tilde{x}, \tilde{h}, \theta)} \quad (3)$$

其中, $E(x, h, \theta)$ 为能量函数; θ 为模型的参数; z 为配分函数; \tilde{x} 和 \tilde{h} 是用来遍历和总结图单位所有可见的和隐藏的单元配置。

受限玻尔兹曼机的能量函数定义为:

$$E(x, h, \theta) = -b^T x - c^T h - h^T W x \quad (4)$$

其中, b_j 和 c_i 为偏移量; W_{ij} 为连接第 j^{th} 可视节点和第 i^{th} 隐藏节点的权重; $\theta = \{b, c, W\}$ 为系统参数。

传统受限玻尔兹曼机中代表可视节点和隐藏节点之间的参数被限制为常量,这会引发很多其他的问题。首先,它将限制表示能力,因为变量经常以某种不确定的方式相互作用。其次,它训练带有噪声的采样数据的鲁棒性也不高。2015 年, C. L. Philip Chen 等^[5]提出的模糊受限玻尔兹曼机^[2]可以很好地解决这些问题,同时通过跨层单元关系的线性化来减少误差和失真。

随着一些快速学习算法和典型深度学习算法的提出,如深度自编码器^[6]、深度置信网络^[4]和深度受限玻尔兹曼机^[7]等,受限玻尔兹曼机受到了越来越多的关注,之后受限玻尔兹曼机及其相关的深度学习结构也在降维^[8-10]、分类^[11-14]、特征提取^[15-16]等方面得到了广泛的应用。受限玻尔兹曼机通过非监督学习方式,用一定的数据集来训练网络,设置可视神经元的值匹配数据集的数据点的值,当网络训练完成以后,就可以用来对未知数据进行计算,从而进行分类。

深度置信网络是一个概率生成模型,和传统的判别模型的神经网络相对,生成模型建立一个观察数据和标签之间的联合分布。深度置信网络是由多个受限玻尔兹曼机层组成,模糊能量函数定义为:

$$\bar{E}(x, h, \bar{\theta}) = -\bar{b}^T x - \bar{c}^T h - h^T \bar{W} x \quad (5)$$

其中, $\bar{\theta} = \{\bar{b}, \bar{c}, \bar{W}\}$ 为模糊参数。

通过边缘化隐藏单元和化简式 5 得到模糊自由能量函数:

$$\bar{F}(x, \theta) = -\log \sum_{\tilde{h}} e^{-E(x, \tilde{h}, \bar{\theta})} \quad (6)$$

如果直接采用这个模糊能量函数来定义概率,那么得到的将是模型的模糊概率^[11],最终的学习优化问题转变为一个求解模糊极大似然的问题。但是,由于模糊目标函数是非线性的并且隶属函数很难计算,求解一个模糊极大似然变得非常难;此外,因为 α -分割的计算变成 NP 难问题^[17],所以这种问题很难解决。因此需要通过将模糊自由能量函数去模糊化使问题转化为一般的极大似然问题。利用面积中心法^[18]给模糊自由能量函数 $\bar{F}(x)$ 去模糊化。去模糊化之后的自

由能量函数如下:

$$F_c(x, \tilde{\theta}) = \frac{\int \theta F(x, \theta) d\theta}{\int F(x, \theta) d\theta}, \theta \in \tilde{\theta} \quad (7)$$

同时,去模糊化的概率可以表示为:

$$P_c(x, \tilde{\theta}) = \frac{e^{-F_c(x, \tilde{\theta})}}{Z}, Z = \sum_x e^{-F_c(x, \tilde{\theta})} \quad (8)$$

因此,在模糊受限玻尔兹曼机模型中,目标函数是负对数似然,表示为:

$$\Gamma(\tilde{\theta}, D) = - \sum_{x \in D} \log P_c(x, \tilde{\theta}) \quad (9)$$

其中, D 表示训练数据集。

学习过程转变为寻找最小化目标函数所对应的参数 $\tilde{\theta}$, 即 $\min \Gamma(\tilde{\theta}, D)$ 。

为了寻找到最优参数 $\tilde{\theta}$, 采用随机梯度下降方法进行极大似然学习, 即最大化输入的重构概率, 使得目标函数(9)最小。

2 模糊深度置信网络

深度置信网络是一种结合无监督学习和有监督学习方法的机器学习方法。无监督学习的优点在于, 学习使用的数据来源简单, 不需要进行大量的手工标注, 缺点在于难以得到准确的结果。而有监督学习的优缺点正好与之相对, 需要有手工标注的数据进行监督训练, 能得到更加准确的学习结果。针对传统的深度置信网络中求解最优参数计算量大及鲁棒性不高的缺点, 提出一种基于模糊深度置信网络的行人检测方法。如图 2 所示, 包括以下两个部分:

(1) 借助模糊深度置信网络提取行人特征, 从图像像素开始, 建立图像的深度置信网络, 通过逐层学习得到图像的表达;

(2) 采用监督学习提高 FDBN 分类的性能。FDBN 不仅继承了 DBN 的强大抽象能力, 而且展示了处理行人数据的诱人的模糊分类。

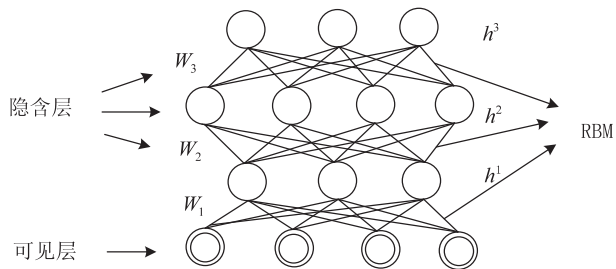


图 2 深度置信网络

FDBN 通过引入模糊集理论, 获得嵌入类先验知识与深度框架的能力。利用一个新的深度框架集成限制玻尔兹曼机(RBM)的抽象能力和模糊集的分类能力。RBM可以快速降维, 模糊集可以基于每个类的成

员函数提高深度框架的分类精度。FDBN 通过逐层贪心的非监督学习构建, 参数空间通过梯度下降监督学习微调。特征提取和分类步骤使用相同的深度框架, 在充足的非标记数据的帮助下提高了概括能力, 避免了显式的特征提取过程, 通过隐式地从训练数据中进行学习; 再者由于同一特征映射面上的神经元权值相同, 所以网络可以并行学习。同时, 局部的权值共享降低了网络的复杂性, 特别是多维输入向量的图像可以直接输入网络这一特点避免了特征提取和分类过程中数据重建的复杂度。FDBN 可以通过反向传播策略使用一个指数损失函数直接优化构建的深度框架。框架模型结构具体如图 3 所示。

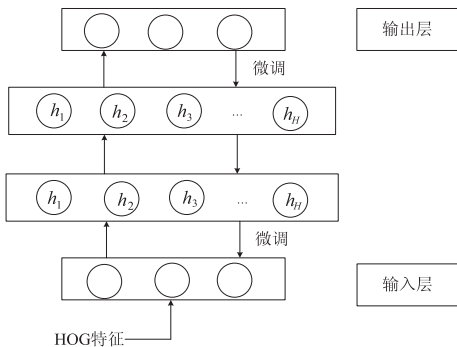


图 3 模糊深度置信网络模型

模糊深度置信网络训练过程主要有两个步骤:

(1) 根据输入的 HOG 特征自底向上地训练模糊受限玻尔兹曼机;

(2) 根据步骤 1 的训练结果, 采用 BP 算法对整个 DBN 结构进行微调, 使其更有利于分类。

由于图像特征的特性, 采用二值型模糊受限玻尔兹曼机进行特征学习将会丢失特征的属性, 因此在整个模糊深度置信网络中的第一层模糊受限玻尔兹曼机的可视层采用的是服从高斯分布的模糊受限玻尔兹曼机, 其余仍为二值的。

为了优化分类结果, 进一步完善模糊深度置信网络结构, 需要将网络的输出结果与真实的结果进行对比, 利用对比结果对模糊深度置信网络结构和参数进行微调, 以实现进一步优化分类器的目的。文中使用 BP 网络对输出值进行反向监督, 利用 BP 网络良好的反向传播能力, 将分类结果误差反向传入模糊深度置信网络, 实现整个网络模型的微调。

文中算法在 Win10 系统, 4 GB 内存, Nvidia GTX960 显卡的计算机上进行实验, 实验数据来源于 INRIA 行人数据库(大小为 64×128), 其中训练集包含 2 000 张正样本单个行人图片和 2 000 张负样本图片, 测试集为 300 张多行人图片。实验对文中提出的模糊深度置信网络行人检测方法的性能进行评估和分析, 在同一检测环境下与其他几种经典的行人检测方法的

性能进行比较。为了评估各方法的分类以及检测性能,采用最能直观反映检测性能的检测率、虚警率和检测速度等指标来衡量和分析整个系统的检测性能。性能指标的计算方法如下:

$$\text{检测率} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\%$$

(10)

$$\text{虚警率} = \frac{\text{FN}}{\text{TP} + \text{FN}} \times 100\%$$

(11)

其中,TP 表示正确分类的正样本数;FP 为被错分类的正样本数;FN 表示被误分类为正样本的负样本数。

首先对模糊深度置信网络自身结构的搭建对检测结果的影响进行检测和分析,分别搭建 3、4、5、6 层模糊置信网络结构,隐藏层的神经节点数分别设为 80、100、120、150、120、100、80,迭代次数设为 1 500。结果如表 1 所示。

表 1 不同结构的模糊深度置信网络的检测性能对比

网络结构	检测率/%	虚警率/%	检测速度/(f/s)
3 层 FDBN	91.8	9.7	0.164
4 层 FDBN	94.2	9.5	0.176
5 层 FDBN	93.9	9.5	0.182
6 层 FDBN	92.1	10.2	0.205

从表 1 中的数据可以看出,不同结构的模糊深度置信网络的检测性能并不相同,4 层深度网络的检测率最高,并且虚警率也没有任何结构比它低,并不是隐层的层数越多识别性能就越好;相反,随着整体结构越来越复杂,可能会出现过拟合的现象,从而导致识别性能的降低。综合考虑隐层数和隐层单元的识别性能,文中采用的是 4 层模糊深度置信网络,其中 3 层隐含层的神经节点数为 80、100、150 的模糊深度置信网络,设定迭代次数为 1 500。

将提出的模糊深度置信网络和文献[14-16]提出的性能较优的分类算法在相同的检测平台上进行对比,具体检测结果如表 2 所示。

表 2 不同的深度学习网络的测试结果对比

算法名称	检测率/%	虚警率/%	检测速度/(f/s)
DBN	93.4	9.6	0.178
CPSO-SVM	89.8	14.6	0.232
CS-SVC-AdaBoost	90.2	11.5	0.146
FDBN	94.3	8.9	0.185

从表 2 的数据可以看出,在相同的条件下,文中提出的模糊深度置信网络分类检测的检测率和虚警率都优于其他方法,虽然检测速度比 CS-SVM-AdaBoost 要慢一些,但是基本可以实现实时检测,满足实时性的要求。 万方数据

3 结束语

针对复杂背景下行人检测性能不稳定的问题,引入深度学习的方法进行特征提取实现对行人的识别和检测,并在此基础上将传统的深度置信网络同模糊集思想相结合,提出一种模糊深度置信网络的行人检测方法。使用模糊深度置信网络可以提取更加有效的特征,进一步提高算法的检测精度。实验结果表明,该算法在检测率、漏检率、实时性等方面都比传统的深度置信网络有所提升,可以较快地从大多数的复杂背景下检测和识别出行人目标。因为图像的质量直接影响最终的检测效果,使用提出的方法在分辨率较低的图像中难以将行人从复杂的背景中区分出来,这也是接下来需要重点解决的问题。

参考文献:

[1] 孙劲光,蒋金叶,孟祥福,等. 一种数值属性的深度置信网络分类方法[J]. 计算机工程与应用,2014,50(2):112-115.

[2] MELCHIOR J, WANG N, WISKOTT L. Gaussian - binary restricted Boltzmann machines for modeling natural image statistics[J]. PloS One,2017,12(2):e0171015.

[3] 许庆勇,江顺亮,黄 伟,等. 基于多特征融合的深度置信网络图像分类算法[J]. 计算机工程,2015,41(11):245-252.

[4] SMOLENSKY P. Parallel distributed processing:explorations in the microstructure of cognition [M]. MA, USA: MIT Press,1986:194-281.

[5] CHEN C L P, ZHANG Chunyang, CHEN Long, et al. Fuzzy restricted Boltzmann machine for the enhancement of deep learning[J]. IEEE Transactions on Fuzzy Systems,2015,23(6):2163-2173.

[6] HINTON G E, OSINDERO S, THE Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation,2006,18(7):1527-1554.

[7] SALAKHUTDINOV R, HINTON G. An efficient learning procedure for deep Boltzmann machines[J]. Neural Computation,2012,24(8):1967-2006.

[8] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006,313(5786):504-507.

[9] SCHMIDHUBER J. Multi-column deep neural networks for image classification [C]//Proceedings of IEEE conference on computing and vision pattern recognition. Providence, RI, USA:IEEE,2012:3642-3649.

[10] BENGIO Y, COURVILLE A, VINCENT P. Representation learning:a review and new perspectives[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2013,35(8):1798-1828.

[11] KLIMT G, BUCKLEY J J. Fuzzy probabilities; new approach

- and applications[M]. [s. l.]:[s. n.],2003.
- [12] LODWICK W A,KACPRZYK J. Fuzzy optimization:recent advances and applications [M]. New York, NY, USA: Springer,2010.
- [13] PEDRYCZ W,SKOWRON A,KREINOVICH V. Handbook of granular computing [M]. New York, NY, USA: Wiley, 2008.
- [14] KARNIK N N,MENDEL J M. Centroid of a type-2 fuzzy set[J]. Information Science,2001,132(1-4):195-220.
- [15] 陈秋菊,李 应. 基于优化正交匹配追踪和深度置信网的声音识别[J]. 计算机应用,2017,37(2):505-511.
- [16] ZHAO Chenglin,SUN Xuebin,SUN Songlin,et al. Fault diagnosis of sensor by chaos particle swarm optimization algorithm and support vector machine[J]. Expert Systems with Applications,2011,38(8):9908-9912.
- [17] ZHANG Yang,LIU Weiming,DENG Rufeng. Diverse cost-sensitive asymmetric SVM-boost classifier with full binary tree structure for fast pedestrian detection system[J]. Journal of Convergence Information Technology,2011,6(8):117-125.
- [18] 张 阳,刘伟铭,吴义虎. 基于深信度网络分类算法的行人检测方法[J]. 计算机应用研究,2016,33(2):594-597.

《计算机技术与发展》投稿须知

1. 《计算机技术与发展》是中国计算机学会会刊,中国科技核心期刊、中国科技论文统计源期刊、RCCSE 核心期刊。

2. 投稿时请注明是否是中国计算机学会(CCF)会员(高级会员、普通会员、学生会员)。凡作者中有 CCF 会员,则享受优先录取、优先出版待遇,第一作者为 CCF 会员者,将享受 85 折版面费优惠。

3. 论文若有国家(省、部)等基金或科研及工程项目支持,则享受优先录用待遇(请注明基金或项目名称和编号)。

4. 投稿时请写明详细通信地址、邮政编码、准确联系电话、Email 信箱等各项必备内容。稿件自留底稿,经初审通过后 40 天内通知作者终审处理意见。

5. 本刊不收取审稿费。扣除作者的稿费后,收取一定的版面费。

6. 投稿时请作者登陆本刊网站(<http://www.xactad.org>),注册成作者后按投稿指南投递。