

基于深度学习网络的烟叶质量识别

韩东伟, 王小明*, 王新峰 (河南中烟工业有限责任公司许昌卷烟厂, 河南许昌 461000)

摘要 概述了烟叶质量和熟成度分类的主要依据, 采用自动编码器预训练方法重构的卷积神经网络构建了烟叶质量识别模型, 并采用实地采集的数据集进行了实验验证, 结果表明重构的深度训练自编码器在分类性能上达到 99.92% 的准确度。

关键词 深度学习; 智能识别; 烟叶质量识别

中图分类号 S126 **文献标识码** A **文章编号** 0517-6611(2018)10-0185-04

Identification of Tobacco Quality Based on Deep Learning Network

HAN Dong-wei, WANG Xiao-ming, WANG Xin-feng (Xuchang Cigarette Factory, Henan China Tobacco Industrial Co., Ltd., Xuchang, Henan 461000)

Abstract The main basis for the identification of tobacco leaf quality and ripening degree was summarized. The convolution neural network reconstructed by the automatic encoder pre-training method was used to construct the tobacco leaf quality identification model. The experimental data were used to verify the experimental results and the results showed that the reconstructed depth training self-encoder achieved 99.92% accuracy in classification performance.

Key words Deep learning; Intelligent identification; Tobacco quality identification

烟草的质量识别长期以来主要依赖人工分拣和分类, 效率慢且受操作人员的主观因素影响较大。深度学习网络作为一种新的机器学习技术, 在图像处理和智能识别领域得到了广泛应用, 并有着优异的性能。随着近年来深度学习技术的快速发展, 以图像处理和语音识别为主的深度学习技术的应用也得到广泛展开, 在各行各业中都得到了普遍应用, 并在农业领域取得了较大的经济与社会收益^[1]。我国的烟叶种植和烟草制品消费量均十分庞大, 将深度学习技术与图像处理相结合并在烟草工业中应用对于烟草行业的现代化建设十分关键^[2]。笔者利用深度学习网络构建烟叶的熟成度分类模型, 进行智能化的烟叶质量识别, 旨在提高烟叶分拣及分类工作的效率和精度。

1 烟叶熟成度分类

烟叶的质量和熟成度检验主要分为内部检测和外部检测。内部质量检测指的是烟叶的化学成分、燃烧测试、化学分析和通过人体感官进行的吸烟测试来实现的, 而外部质量检测主要通过人的肉眼对烟叶进行观察。由于烟叶化学成分复杂, 内部质量检测费用较高, 且费时费力^[3]。因此, 外部质量检测经常在对时间要求较高的情况下被选择来代替内部质量检测, 且烟叶的内部质量也与其外部质量密切相关, 有经验的烟厂生产人员往往仅凭肉眼观察就能分辨烟叶的质量和熟成度。烟叶的外部质量检验包括色泽、熟成度、表面纹理、大小和形状。在通过人类视觉进行烟叶质量检验时, 不可避免地受到检验人员的身体、心理和环境的影响。

烤烟烟叶的具体分级标准可能因国家和地区而异, 但烟叶分类的一般方法和考察的外部特征相似。首先, 烟叶被分类为正常烟叶和异常烟叶。一般情况下, 大部分的烟叶都属于正常烟叶, 因此对异常烟叶暂不进行讨论。其次, 正常烟

叶根据其在烟草植株的茎上生长的位置和颜色不同, 分为 X 下部(lugs)、C 中部(cutters)、B 上部(leaf)。对生长位置的判断主要基于颜色、表面纹理、烟叶的形状和叶脉特征。叶片按颜色分为 L 柠檬黄色(lemon)、F 橘黄色(orange)和 R 红棕色(red-brown)。通过生长位置和叶片颜色的组合, 得到: 下部柠檬黄色(XL)、下部橘黄色(XF)、中部柠檬黄色(CL)、中部橘黄色(CF)、上部柠檬黄色(BL)、上部橘黄色(BF)和上部红棕色(BR)。

烟叶熟成度分类参照 GB2635-1992 烤烟的规定。《烤烟》国标中将熟成度分为完熟、成熟、尚熟、欠熟和假熟 5 个等级。按照国标要求, 烤烟烟叶熟成度的判断依据包括成分、结构、密度、油分、气味、颜色、斑点、光滑程度和色度等。烟叶熟成度等级及相应外观特征见表 1。其中, 结构、颜色、斑点、光滑程度等在一定程度上都可以采用机器视觉技术来进行自动分拣、分类和评级。

表 1 烟叶熟成度等级及相应外观特征

Table 1 The ripening degree level and appearance characteristics of tobacco

等级 Level	熟成度 Ripening degree	外观特征 Appearance characteristics
1	完熟	结构疏松, 颜色多橘黄或浅红棕色, 有明显的成熟斑, 有些伴有赤星病斑
2	成熟	没有任何的含青或明显的挂灰和蒸片类的杂色, 没有大面积的光滑叶, 色度无异常
3	尚熟	外观有一定的含青, 有较大面积的光滑叶, 或有较明显的挂灰和蒸片等杂色, 颜色偏浅淡
4	欠熟	有较大程度的含青, 或者有较大面积的僵硬(光滑), 有严重的挂灰或蒸片等杂色现象
5	假熟	颜色较浅或身份较薄, 调制后表现为黄色的烟叶

2 深度学习网络

深度学习网络是基于大数据信息建立的统计学习方法^[4], 属于人工智能领域中的机器学习学科, 其目的是通过大量样本的统计学习, 建立针对某种特定问题的专家系统,

作者简介 韩东伟(1982—), 男, 河南周口人, 工程师, 硕士, 从事计算机软件开发及信息系统的管理工作。*通讯作者, 工程师, 硕士, 从事计算机软件开发及信息系统的管理工作。

收稿日期 2018-01-22

借助或不借助人类知识自动解决某一领域的特定问题。深度学习网络的主要思想:通过建立具有多个层次的神经网络,实现对输入数据的深层次表达,从而实现更好的分类与特征抽取。使用深度学习方法,可以让计算机自动学习到人工方法未必能够发现的重要特征。

深度学习作为人工智能领域和机器学习学科的一种最新的技术,近年来获得了大量关注和研究^[5]。针对深度学习的理论研究、算法设计和应用系统在许多领域广泛应用^[6-7],如语音识别、图像分类和自然语言处理等^[8]。目前最主要的深度学习包括深度置信网络(Deep belief network, DBN)^[9]、卷积神经网络(Convolutional neural network, CNN)^[10]、循环神经网络(Recurrent neural network, RNN)和堆叠自动编码器(Stacked autoencoders, SAE)等^[11]。

AE是3层神经网络,将输入表达编码为一个降维映射,并将该降维映射重新解码还原为原输入,在训练过程中最小化重构误差。

笔者采用自动编码器和卷积神经网络相结合的方法,建立一种采用无监督预训练,以提高训练深度的专门网络,通过对烤烟烟叶的图像数据识别,自动进行烤烟烟叶检验与分级。

2.1 稀疏自动编码器 稀疏自动编码器是神经网络的一种,其机制是试图在无监督过程中使其输出与输入近似。自动编码器的基本模型是一个3层网络(图1),其隐含层神经元个数往往少于输入特征数,因此输入层到隐含层相当于一个编码过程,而隐含层到输出层则视为解码。在编码过程中,自动编码器试图将输入向量 x 按一定的映射编码为 x_c ,如式(1)所示。

$$y = f(x) \quad (1)$$

当编码过程完成后,解码过程是将编码的特征向量 y 重构成新的向量 x' ,并使得 x' 尽可能地还原 x ,其过程如式(2)所示。

$$x' = g(y) \quad (2)$$

x' 的维数一般小于 x ,但其可以通过 $g(y)$ 来还原 x ,因此可视其为 x 的特征向量。自动编码器的作用就是通过特征提取将高维问题转化为低维问题,从而降低问题的复杂度,以使问题更容易求解。使 x' 接近于 x 的过程可以用反向传播算法来训练,从而获得编码 y 。该算法逻辑如图2所示。

输入层到隐含层的映射即编码过程,对任何 i 维向量 $x \in R^n$,经过输入层到隐含层的映射,获得该向量对应的低维向量 $y = f(x)$, $y \in R^d$ 。从隐含层到输入层的过程,自动编码器对该向量的对应低维向量 y 进行“解码”,得到与 x 同维度向量 x' 。通过反向传播算法训练网络,并调整 $f(x)$ 与 $g(x)$ 的参数以最小化重构误差,使存在一个足够小的正数 ζ 且 $|x' - x| < \zeta$ 。

在具体训练过程中,将电能计量器具的运行条件矩阵 X 作为自动编码器的输入矩阵, $x_i \in R^{n \times l}$ 是矩阵 X 中的第 i 条测试数据对应的条件向量,将 x_i 作为自动编码器的第 i 个输入向量,编码层的神经元数量为 d ,通过式(3)得到编码

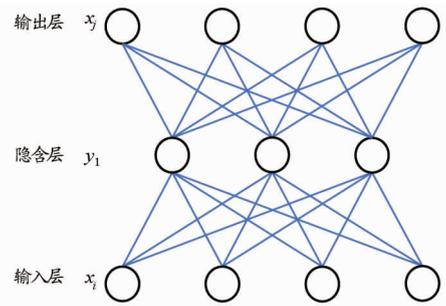


图1 自动编码器网络结构

Fig.1 The network structure of automatic coding machine

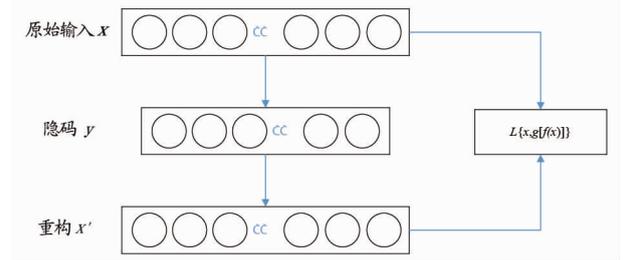


图2 自动编码器算法逻辑

Fig.2 The algorithm logic of automatic coding machine

$h_i \in R^{d \times l}$ 。

$$y_i = s_f(Wx_i + b) \quad (3)$$

式中, S_{act} 为编码器的激活函数,如sigmoid函数或者ReLU单元。 $W \in R^{d \times n}$ 是权值矩阵, $b \in R^{d \times l}$ 是偏置向量。将 y_i 输入解码层,通过式(4)重构输入向量,获得 $x'_i \in R^{n \times l}$:

$$x'_i = S_g(\tilde{W}y_i + b') \quad (4)$$

式中, S_g 为解码器的激活函数。权值矩阵一般被限制为 $\tilde{W} = W^T$, b' 则为解码层的偏置向量。

训练过程中,自动编码器通过不断调整解码层和编码层的权值矩阵和偏置向量这4个参数,即调整 $\theta = \{W, \tilde{W}, b, b'\}$,试图最小化重构误差:

$$\min_{W, \tilde{W}, b, b'} \sum_{i=1}^n \| s_g(\tilde{W}S_f(Wx_i + b) + b') - x_i \|^2 \quad (5)$$

2.2 卷积神经网络 卷积神经网络是由用于特征提取的卷积层和用于特征处理的亚采样层交叠组成的多层神经网络。网络输入是一个图像的数字矩阵,输出则是其识别结果。输入图像经过若干个“卷积”和“采样”加工后,在全连接层网络实现与输出目标之间的映射。卷积神经网络是Hubel、Wiesel等通过对猫科动物视觉系统的研究中受到启发而设计的。

卷积神经网络的卷积层由若干个特征图组成,每个特征图上的所有神经元共享同一个卷积核的参数。由卷积核对前一层输入图像进行卷积运算得到,因此有效缩减了以往全连接网络需要优化的权值个数,成功加深了训练层数。第 l 层的第 j 个特征图矩阵 x'_j 可能由前一层若干个特征图卷积加权得到,一般表示为:

$$x'_j = f\left[\sum_{i \in N_j} x_i^{l-1} * k_{ij}^l + b_j^l\right] \quad (6)$$

式中, f 为神经元激活函数; N_j 代表输入特征图的组合,*代表卷积运算, k_{ij}^l 为卷积核矩阵, b_j^l 为偏置矩阵。

池化层作用是基于局部相关性原理进行池化采样,从而在减少数据量的同时保留有用信息。采样过程可以表示为:

$$x_j^l = f(\text{down}(x_j^{l-1})) \quad (7)$$

式中,down()表示采样函数。卷积神经网络在卷积层和采样层后,通常会连接一个或多个全连接层。全连接层的结构和全连接神经网络的隐层结构相同,全连接层的每个神经元都会与下一层的每个神经元相连。对多类别图像分类任务而言,在输出层需要采用 Softmax 分类器将任务类别进行独立编码,并采用最小化交叉熵作为损失函数。Softmax 归一化概率函数定义如下:

$$z_i = z_i - \max(z_1, z_2, \dots, z_m) \quad (8)$$

$$\sigma_i(z) = \frac{\exp(z_i)}{\sum_{j=1}^m \exp(z_j)} \quad (9)$$

式中, z_i 是每一个类别线性预测的结果,由于网络最终需做归一化处理,因此减去一个最大值并不会改变最终结果,而会起到保持计算时的数值稳定性的作用。同时,根据 $\sigma_i(z)$ 来预测输入 z_i 属于每个类别的概率。这样,损失函数定义为:

$$L(z_i) = -\log \sigma_i(z) \quad (10)$$

在训练过程中,卷积神经网络仍采用输入正向传播和误差反向传播算法进行训练,最小化分类或回归误差。

3 卷积自动编码器网络结构

卷积神经网络采用权值共享和卷积操作代替了传统神经网络层与层之间的矩阵内积运算,有效加深了网络深度并提高了网络的训练效果。但对于深度网络而言,反向传播算法的梯度消失问题仍然没有彻底解决。对于更加复杂的处理对象而

言,网络深度的增加依然会导致网络的学习效能降低。

自动编码器采用最小化重构误差进行无监督训练,在堆叠自动编码器中,这个训练过程实际上成为一种预训练(pre-training)。预训练是一种广泛使用的深度网络训练方法,通过预先将网络权值调整到较为有利的位置,可以有效增加训练深度,提高网络效能。笔者以此重构一种新的网络结构,网络层间运算采用卷积操作,并在卷积操纵之后使用反卷积操作作为解码,利用最小化重构误差预调整网络权值。在预训练结束后,再采用反向传播算法微调网络参数,获得更好的训练效果。网络结构如图3所示,

此外,为自动编码器加入基于 Kullback-Leibler (KL) 散度的稀疏性限制,以在训模型练过程中尽量避免过拟合,其限制如式(11)所示。

$$\sum_{j=1}^d KL[\rho \parallel \frac{1}{n} \sum_{i=1}^n y_i] \quad (11)$$

式中, ρ_j 表示隐含层的平均活跃度, $\rho_j = \frac{1}{n} \sum_{i=1}^n y_{ij}$; $KL(\rho \parallel \rho_j)$ 表示分别以 ρ 和 ρ_j 为均值的2个变量之间的相对熵。引入稀疏性就是使尽量多的神经元参数近似于0,即权值矩阵的正则化,所以在这里希望隐含层神经元参数的一个函数,即平均活跃度 ρ_j 趋近于近似于0的常数 ρ 。KL散度的目的就是建立一个罚函数,使损失函数(loss function)因 ρ_j 偏离 ρ 的程度而获得惩罚。 $KL(\rho \parallel \rho_j)$ 的数学表达式如下:

$$KL(\rho \parallel \rho_j) = \rho \ln \frac{\rho}{\rho_j} + (1 - \rho) \ln \frac{1 - \rho}{1 - \rho_j} \quad (12)$$

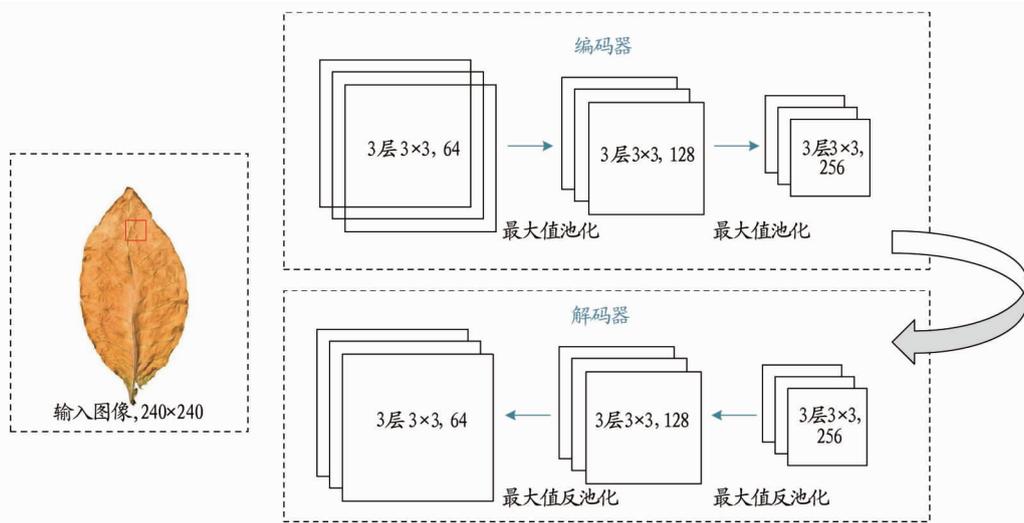


图3 卷积自动编码器网络结构

Fig. 3 The network structure of convolutional automatic encoder

则构建的自动编码器重构误差的损失函数(loss function)

$$L(\theta) = \sum_{i=1}^n \|S_g(\tilde{W}S_j(Wx_i + b) + b') - x_i\|_2^2 + \alpha \sum_{j=1}^d KL[\rho \parallel \frac{1}{n} s_j(Wx_i + b)] \quad (13)$$

4 训练试验、对比和结果

4.1 网络结构 通过使用同样数据集对最小二乘支持向量机(Least squares support vector machine, LS-SVM)、全连接多层感知机(Full-connection perceptions, FCP)和采用深度训练

归纳式如式(13)所示。式中, α 为控制稀疏限制的权重因子。

自编码器(Stacked sparse autoencoder, DRSSA)进行试验,观察其训练效果。各网络结构如图4所示。

图4a为全连接多层感知机网络结构图,该模型共含有4个隐含层,每个隐含层含有5个感知机单元,层与层之间的

感知机单元采用全连接方式。图4b为堆叠自动编码器网络结构图,共含有4个自动编码器,编码器所含单元数分别为6、5、4、3,每2层采用无监督预训练编码器输入与输出,并将上一层编码器的中间层作为下一层编码器的输入。输出层为全连接层。在训练时将样本数据的1/3作为预训练集(pre-training datasets)、其余2/3作为微调训练集(fine-tuning datasets)。

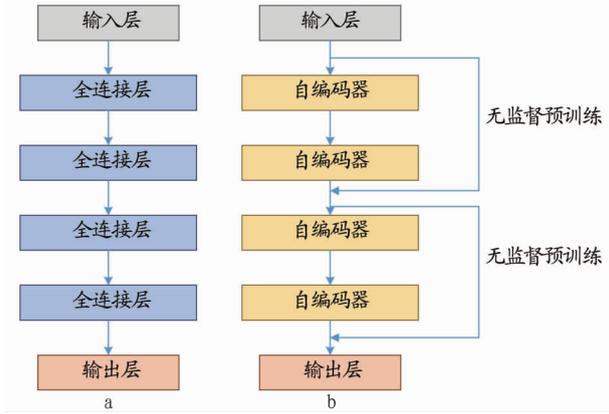


图4 全连接多层感知机与堆叠稀疏自编码器网络结构

Fig.4 Network structure of fully connected multi-layer perception machine and stack sparse self-encoder

训练学习率 $\eta = -0.001$,采用随机梯度下降法(Stochastic gradient descent)进行训练,batch size = 50。损失函数(loss function)采用均方误差(mean squared error, MSE):

$$E(f;D)_{MSE} = \frac{1}{m} \sum_{i=1}^m [f(x_i) - y_i]^2 \quad (14)$$

对于给定样例集合 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, $x_i \in R^d, y_i \in R^d$,其中 y_i 为示例 x_i 的真实标记, m 为样本数据量,学习器预测结果为 $f(x)$ 。

4.2 仿真试验 试验采用我国云南某烟草产地收集的烟草图像数据集,数据集样本数为3 000,格式均为 240×240 像素的RGB图像。

通过对该数据分别采用最小二乘支持向量机(Least squares support vector machine)、全连接多层感知机(Full-connection perceptions,FCP)和重构的堆叠稀疏自编码器(Stacked sparse autoencoder,SSA)进行试验,观察其训练效果。

从图5可以看出,LS-SVM在最初几次迭代中效果优于FCP和SSA模型,这是由于SVM在针对小样本空间的学习能力比神经网络算法要强。但随着训练样本和迭代次数的增加,神经网络的学习潜力开始展现,最终FCP和SSA的模型误差都要小于LS-SVM,说明在样本数量足够的情况下,神经网络算法对大数据的适应能力更强。此外,SSA模型学习的收敛速度更快,最终误差更小,说明SSA模型在针对烟

叶成熟度识别这一特定的问题上具有更好的性能。

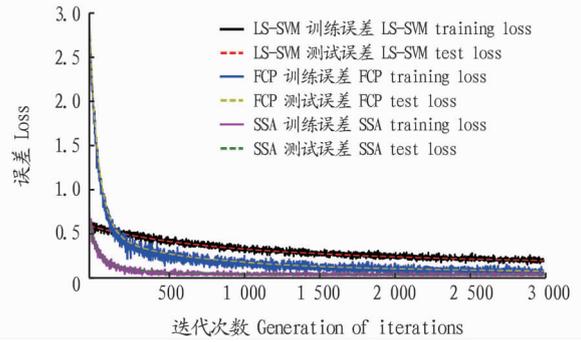


图5 LS-SVM、FCP与SSA训练效果对比

Fig.5 The training effect comparison among LS-SVM,FCP and SSA

5 结论

随着烟草行业的发展,对行业整体的自动化和智能化需求的上升,深度学习技术作为人工智能领域的前沿技术,在烟草行业中的应用还有很大的拓展空间。重构的堆叠稀疏自编码器(Stacked sparse autoencoder,SSA)模型在训练深度提高时可以有效保持训练效果,一定程度上避免误差消失,并在烟叶成熟度识别试验中表现出良好的识别性能。结果表明,重构的堆叠稀疏自编码器(Stacked sparse autoencoder,SSA)性能优于最小二乘支持向量机(Least square support vector machine)、全连接多层感知机(Full-connection perceptions,FCP)。

参考文献

- [1] 张帆,张新红,张彤.模糊数学在烟叶分级中的应用[J].中国烟草学报,2002,8(3):45-49.
- [2] 刘朝营,许自成,闫铁军.机器视觉技术在烟草行业的应用状况[J].中国农业科技导报,2011,13(4):79-84.
- [3] 王亚平.烟叶分级存在问题及改进措施[J].宁夏农林科技,2013,54(1):112-113.
- [4] 焦李成,杨淑媛,刘芳,等.神经网络七十年:回顾与展望[J].计算机学报,2016,39(8):1697-1716.
- [5] BENGIO Y. Learning deep architectures for AI(Foundations and trends(r) in machine learning) [M]. Hanover, MA: Now Publishers Inc., 2009: 1-127.
- [6] SCHMIDHUBER J. Deep learning in neural networks: An overview[J]. Neural networks, 2015,61:85-117.
- [7] 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.
- [8] 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.
- [9] DENG L. A tutorial survey of architectures, algorithms, and applications for deep learning. APSIPA Transactions on Signal and Information Processing [M]. Cambridge: Cambridge University Press, 2014.
- [10] ZHANG X Z, GAO Y S. Face recognition across pose: A review[J]. Pattern recognition, 2009,42(11):2876-2896.
- [11] LE Q V, NGIAM J, COATES A, et al. On optimization methods for deep learning[C]//Proceedings of the 28th International Conference on Machine Learning. Bellevue, Washing, USA:ICML,2011:265-272.

本刊提示 文稿题名下写清作者及其工作单位名称、邮政编码;第一页地脚注明第一作者简介,格式如下:“作者简介:姓名(出生年—),性别,籍贯,学历,职称或职务,研究方向”。