

文章编号: 2095-2163(2019)05-0095-03

中图分类号: TP391.41

文献标志码: A

基于 CNN 的外观专利图像分类

苏 静^{1,2}

(1 广东工业大学 信息工程学院, 广州 510006; 2 广东海洋大学 数学与计算机学院, 广东 湛江 524088)

摘要: 针对专利图像检索系统中的图像分类问题, 研究并实现了深度学习中的卷积神经网络方法。所设计的卷积神经网络由 4 个卷积层、2 个池化层及 2 个完全连接层组成, 其中每个特征映射层只由上一层的部分特征映射层线性组合而得, 整个神经网络由反向传播算法调整权重参数和偏置项并自动完成学习任务。通过外观专利图像的分类实验表明, 该方法分类准确率可达 90%。

关键词: 深度学习; 卷积神经网络; 外观专利; 图像分类

Image classification of design patent based on CNN

SU Jing^{1,2}

(1 School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China; 2 College of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang Guangdong 524088, China)

[Abstract] For patent image classification problems of Guangdong province in image retrieval system, A convolution neural network method in the deep learning is researched and implemented. The designed convolution neural network is made of four convolution layers, two pooling layers and two complete connection layers. In the convolution neural network, each feature mapping layer is the linear combination of the last part feature mapping layers. And the whole neural network adjusts the weight parameters and the bias by the back propagation algorithm and automatically completes the study task. The patent image classification experiment shows that the average classification accuracy can reach 90% by this method.

[Key words] deep learning; convolution neural network; design patent; image classification

0 引言

在互联网技术迅猛发展的今天, 专利图像的规模不断扩大, 专利图像数据库都是海量的, 这无疑给面向大规模图像数据的分析和理解带来了重大挑战。现有基于内容的图像检索技术^[1-2]用来提取特征的训练数据大多来自人工标注样本和人工调参, 需要依靠设计者的先验知识, 很难利用大数据的优势, 极大地限制了图像检索的性能。

近年来, 深度学习^[3-4]在计算机视觉与图像检索等领域表现出强大优势。与传统模式识别方法不同的是, 深度学习无需手工设计特征, 可以直接从大量数据集中自动学习特征, 学习到成千上万的参数, 在图像检索领域应用广泛。

本文提出了一种基于 CNN 模型的家具类外观专利图像分类方法, 并对本文方法与支持向量机方法(SVM)及传统神经网络(NN)方法加以对比, 对家具类外观专利图像进行了交叉验证实验。实验结

果表明, 基于 CNN 模型的方法取得了更好的分类效果。

1 CNN 模型与算法

1.1 CNN 模型简介

卷积神经网络由 Hinton 等人于 2006 年提出, 作为一种深度学习模型受到了广泛关注, 并已成功应用在人脸识别、图像分类与检索等领域。结构上, CNN 是由多层神经网络构成, 主要包括卷积层、激励层、池化层及全相连接层^[5-6]。在每一个卷积层, 上一层的特征 Maps 被一个可以学习的卷积核进行卷积, 再通过激活函数(ReLu), 就可以得到输出特征 Map。池化层往往在卷积层后面, 计算 Map 一个区域上的某个特定特征的平均值来代表这个区域的特征, 从而降低卷积层输出的特征向量。在本实验中, 前面两层是卷积和池化的交替, 接着两层是单卷积操作, 在最后一两层(靠近输出层)是全连接的一维神经网络。网络结构如图 1 所示。

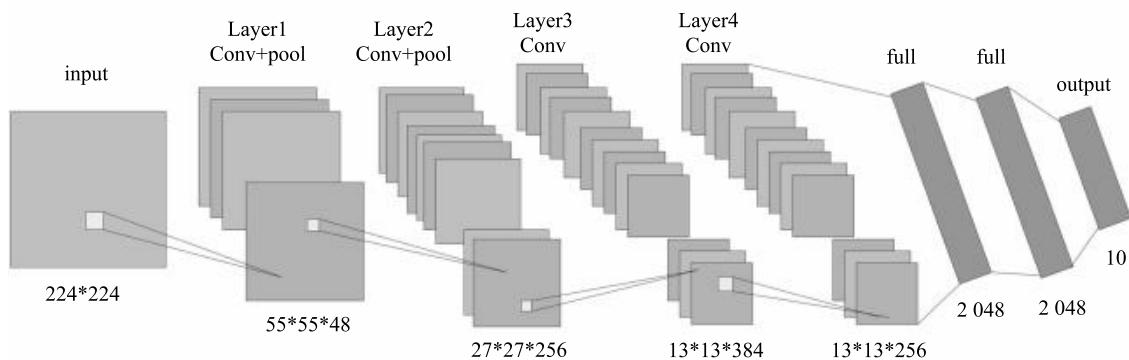


图1 CNN模型

Fig. 1 CNN model

1.2 BP 算法

CNN 通过 BP 算法^[7]对数据进行训练。在 BP 网络模型中,前面一层为输入层,最后一层为输出层,中间为隐藏层。在该神经网络中,设样本集为 $(x^{(i)}, y^{(i)})$, 其中 x 表示输入, y 表示输出的类别,网络的层数用 n_f 表示,这里 $n_f = 7$, n_k 表示第 k 层,神经网络的参数为 (W, b) , 其中 $W_{i,j}^{(k)}$ 表示第 k 层第 j 单元与第 $k+1$ 层第 i 单元之间的联接参数(即权重), $b_i^{(k)}$ 表示第 $k+1$ 层第 i 单元的偏置项。 s_k 表示第 k 层的节点数, $a_i^{(k)}$ 表示第 k 层第 i 单元的激活值。则 BP 神经网络的输出结果可以表示为:

$$h_{w,b}(x) = a^{(7)} = f(W^{(6)}a^{(6)} + b^{(6)}), \quad (1)$$

其中, $a^{(k+1)} = f(W^{(k)}a^{(k)} + b^{(k)}), k = 1, 2, \dots, 6$ 。

对于一个包含 t 个样例的数据集,其代价函数为:

$$J(W, b) = \frac{1}{t} \sum_{i=1}^t \left(\frac{1}{2} \| h_{w,b}(x^{(i)}) - y^{(i)} \|^2 \right) +$$

$$\frac{\lambda}{2} \sum_{k=1}^{n_f-1} \sum_{i=1}^{s_k} \sum_{j=1}^{s_{k+1}} (W_{ji}^{(k)})^2, \quad (2)$$

对于最后一层(输出层)的每个输出单元 i , 其残差为:

$$\delta_i^{(n_f)} = -(y_i - a_i^{(n_f)}) \cdot f'(W^{n_f-1}a^{n_f-1} + b^{n_f-1}), \quad (3)$$

第 k 层的第 i 个节点的残差为:

$$\delta_i^{(k)} = \left(\sum_{j=1}^{s_{k+1}} W_{ji}^{(k)} \delta_j^{(k+1)} \right) f'(W^{k-1}a^{k-1} + b^{k-1}), \quad (4)$$

反向传播算法步骤详述如下:

Step 1 进行前向传导计算,得到 1, 2, 3, ..., 直到第 7 层(输出层)的激活值。

Step 2 计算输出层(第 7 层)的残差,会用到如下公式:

$$\delta_i^{(n_f)} = -(y_i - a_i^{(n_f)}) \cdot f'(W^{n_f-1}a^{n_f-1} + b^{n_f-1}), \quad (5)$$

Step 3 对于 $k = 2, \dots, 5, 6$ 的各层,计算其残差:

$$\delta_i^{(k)} = \left(\sum_{j=1}^{s_{k+1}} W_{ji}^{(k)} \delta_j^{(k+1)} \right) f'(W^{k-1}a^{k-1} + b^{k-1}), \quad (6)$$

Step 4 对于 $i = 1$ to m ,

(1) 计算偏导数,具体公式为:

$$\nabla_{W^{(k)}} J(W, b, x, y) = \delta^{(k+1)} (a^{(k)})^T, \quad (7)$$

$$\nabla_{b^{(k)}} J(W, b, x, y) = \delta^{(k+1)}, \quad (8)$$

(2) 计算 $\Delta W^{(k)}$, 具体公式为:

$$\Delta W^{(k)} := \Delta W^{(k)} + \nabla_{W^{(k)}} J(W, b; x, y), \quad (9)$$

(3) 计算 $\Delta b^{(k)}$, 具体公式为:

$$\Delta b^{(k)} := \Delta b^{(k)} + \nabla_{b^{(k)}} J(W, b; x, y), \quad (10)$$

Step 5 更新权重参数,具体公式为:

$$W^{(k)} = W^{(k)} - \gamma \left[\left(\frac{1}{m} \Delta W^{(k)} \right) + \lambda W^{(k)} \right], \quad (11)$$

$$b^{(k)} = b^{(k)} - \gamma \left[\left(\frac{1}{m} \Delta b^{(k)} \right) \right]. \quad (12)$$

Step 6 使用梯度下降法来求解最小代价函数 $J(W, b)$ 的值,得到神经网络的解。

2 实验与结果分析

2.1 实验设备与数据

本实验在 Ubuntu14.04 操作系统,caffe 框架环境下完成,使用的是改进后的 Alexnet 模型。关于实验数据,使用了家具外观专利图像,其中 5 000 张为训练图像,1 000 张为测试图像。在分类方案中,根据家具种类的不同,训练图像与测试图像都被细分为 10 个类。在训练过程中,每个类都有 500 个训练样本;在测试过程中,每个类都有 100 个测试样本。每个图像的输入大小为 $224 * 224$, 格式为 PNG。部分图像样本如图 2 所示。为保证实验的客观性,实验图像为随机选取,采用交叉验证法得到实验结果的平均准确率。



图 2 部分实验样本
Fig. 2 Partial samples

2.2 实验结果分析

(1) 训练模型对分类结果的影响。不同的训练模型因网络层数和参数不同对分类结果有较大影响。对于同样的数据集, 采用 AlexNet 原模型识别准确率为 85%, 采用改进后的 AlexNet 模型识别准确率可达 90%, 仿真结果对比如图 3 所示。本设计考虑到了外观专利图像的特殊性, 对输入图像的大小、网络层数、滤波器大小都做了重新设计。实验中, 当训练次数达到 15 000 次时, 网络参数变化不大, 表示卷积网络已呈收敛状态, 分类性能达到最优。

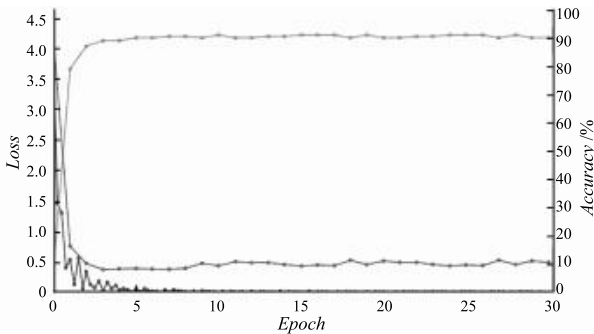


图 3 CNN 的分类准确率
Fig. 3 Accuracy of CNN classification

(2) 与 SVM、NN 方法的比较。本文将 CNN 模型的方法与 SVM 及 NN 两种分类方法进行了对比, 实验结果见表 1。可以看出, 和 SVM 方法相比, CNN 在各个类别和总体分类精度上略有提升。和 NN 方法相比, CNN 方法在 Mirror 和 Stool 两个类别分类精度相当, 而在 Swivelchair、Table、Messagechair 这 3 个类别上精度有较大提高。总地来看, CNN 方法的分类效果在三者中为最优。

表 1 CNN、NN 和 SVM 分类方法比较

Tab. 1 Comparison of CNN, NN and SVM classification methods

| Class | SVM | NN | CNN |
|--------------|------|------|------|
| Bed | 0.96 | 0.97 | 0.99 |
| Cabined | 0.94 | 0.94 | 0.96 |
| Chair | 0.87 | 0.86 | 0.88 |
| Messagechair | 0.81 | 0.78 | 0.82 |
| Mirror | 0.91 | 0.93 | 0.93 |
| Sofachair | 0.70 | 0.72 | 0.76 |
| Stool | 0.93 | 0.96 | 0.97 |
| Swivelchair | 0.93 | 0.94 | 0.99 |
| Table | 0.83 | 0.82 | 0.86 |
| Woodenchair | 0.95 | 0.96 | 0.99 |

3 结束语

本文研究了卷积神经网络在外观专利图像分类中的应用, 所设计的卷积神经网络由 4 个卷积层、2 个池化层和 2 个全连接层组成, 在外观专利图像分类实验中达到 90% 的准确率, 优于 SVM 及 NN 分类方法。但是本文只是对 Alexnet 的网络结构进行了简单的改进, 采用了均值池化方法, 针对外观专利图像的更优卷积网络模型仍有待进一步的深入研究。

参考文献

- [1] Edkins J, Graham M. Content - based image retrieval [R]. Newcastle: University of Northumbria, 1999.
- [2] DATTA R, JOSHI D, LI Jia, et al. Image retrieval: Ideas, influences, and trends of the new age [J]. ACM Computing Surveys (CSUB), 2008, 40(2) : 5.
- [3] 余凯, 贾磊, 陈雨强, 等. 深度学习的昨天、今天和明天 [J]. 计算机研究与发展, 2013, 50(9) : 1799-1804.
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6) : 84-90.
- [5] DAHL G E, SAINATH T N, HINTON G E. Improving deep neural networks for LVCSR using rectified linear units and dropout [C] // 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Vancouver, BC, Canada: IEEE, 2013 : 8609-8613.
- [6] FUKUSHIMA K. Neocognitron: A self - organizing neural network model for a mechanism of pattern recognition unaffected by shift in position [J]. Biological Cybernetics, 1980, 36(4) : 193-202.
- [7] LECUN Y, BOSER B E, DENKER J, et al. Backpropagation applied to handwritten zip code recognition [J]. Neural Computation, 1989, 1(4) : 541-551.