

基于卷积神经网络多层特征提取的目标识别^①

江彤彤, 成金勇, 鹿文鹏

(齐鲁工业大学 信息学院, 济南 250353)

摘要: 目标识别一直是人工智能领域的热点问题. 为了提高目标识别的效率, 提出了基于卷积神经网络多层特征提取的目标识别方法. 该方法将图像输入卷积神经网络进行训练, 在网络的每个全连接层分别进行特征提取, 将得到的特征依次输入到分类器, 对输出结果进行比较. 选取经过修正线性单元 *relu* 函数激活的低层全连接层作为特征提取层, 比选取高层全连接层特征提取的识别率高. 本文构建了办公用品数据集, 实现了基于卷积神经网络多层特征提取的办公用品识别系统. 选择 AlexNet 卷积神经网络模型的 *relu6* 层作为特征选取层, 选择最优训练图像数量和最优分类器构建系统, 从而证明了该方法的可行性.

关键词: 卷积神经网络; 特征提取; 深度学习; 识别; 分类器

引用格式: 江彤彤, 成金勇, 鹿文鹏. 基于卷积神经网络多层特征提取的目标识别. 计算机系统应用, 2017, 26(12): 64-70. <http://www.c-s-a.org.cn/1003-3254/6082.html>

Target Recognition Based on Multilayer Feature Extraction of Convolution Neural Network

JIANG Tong-Tong, CHENG Jin-Yong, LU Wen-Peng

(School of Information, Qilu University of Technology, Jinan 250353, China)

Abstract: Target recognition has been the hot issue in the field of artificial intelligence. In order to enhance the efficiency of target recognition, this paper proposes a method based on multilayer feature extraction of convolutional neural network. By inputting images into convolutional neural network for training, this method implements feature extraction at each full connection layer of network, inputs the features obtained into classifier, and then compares the output results. The lower full connection layer activated by *relu* function is selected as feature extraction layer, whose recognition rate is higher than that in higher full connection layer. This paper builds up office supplies dataset, and realizes the office supplies identification system based on the multilayer feature extraction of convolutional neural network. The layer *relu6* of AlexNet is selected feature extraction layer, and the optimal training image quantity as well as the optimal classifier construction system is chosen, which verifies the feasibility of this method.

Key words: convolution neural networks; feature extraction; deep learning; recognition; classifier

近年来, 深度学习方法因为具有更高的灵活性和普适性得到迅速发展. 卷积神经网络 (Convolutional neural networks, CNN) 是一项实现深度学习应用广泛的技术, 应用在多个领域并取得优异效果, 其中包括手写字符识别^[1,2], 人脸识别^[3], 人眼检测^[4], 行人检测^[5], 机器人导航^[6]等.

卷积神经网络首次应用在 20 世纪 80 年代末, 美国纽约大学教授 LeCun 等人^[7]提出用手写数字识别作为美国银行识别票据的工具, 并在识别系统中应用 CNN. 2012 年, 加拿大多伦多大学教授深度学习领军人物 Alex Krizhevsky^[8]构建深层的卷积神经网络在大尺度图像识别数据集 ImageNet 上测试识别率在 84.7% 左右.

① 基金项目: 国家自然科学基金 (61502259); 山东省自然科学基金 (ZR2011FQ038)

收稿时间: 2017-03-03; 修改时间: 2017-03-20; 采用时间: 2017-03-29

本文提出基于卷积神经网络多层特征提取的目标识别方法. 针对卷积神经网络在全连接层的每层所提取的特征输入到分类器得到的识别效果进行分析, 在卷积神经网络提取特征所选取的特征层, 并非在全连接的最后一层效果最好, 实验结果表明 relu6 层效果突出. 使用多种分类器分别在数据集 Caltech101 中结合卷积神经网络提取的特征做分类处理, 实验结果表明提取特征一致时, 与支持向量机分类器相结合的图像识别方法效果最好, 更适合处理数量大的图像, 并且有效的提高了识别的速度和识别精度. 对不同训练图像个数输入卷积神经网络训练的识别效果进行分析, 随训练图像数量增加, 识别率增加, 增加速度渐缓. 以上实验说明基于卷积神经网络多层特征提取的目标识别方法是有效的、可行的. 使用此方法针对办公用品进行分类识别获得识别率, 并与摄像头连接进行实时识别.

1 相关技术

1.1 卷积神经网络模型

如今应用广泛的 CNN 模型中, 比较成熟的是针对手写体的 LeNet-5 模型如图 1 所示. 特征提取阶段的

卷积神经网络的结构可以概括为由卷积层和子采样层两类结构交替构成的, 该网络的输入是 32×32 的图像, 如图 1 所示, 将输入层 C1、S2、C3、S4、C5 构成目标子块的特征提取部分, 输入层读入目标, 子块中的每一个像素看作神经元, C1 层是一个由 6 个 28×28 特征图构成的卷积层, 即 6 个特征图分别对应 6 个卷积滤波器, 每个神经元与输入图像的一个 5×5 的邻域相连接. 每个卷积层后面跟随着一个降采样层, S2 层是子采样层, 是由 6 个大小为 14×14 的特征图组成, 进行 2×2 的平均化降采样的操作特征图中的每个神经元与 C1 中相对应特征图的 2×2 区域相连, S2 层每个神经元的结果由 C1 层相邻的 4 个神经元相加后乘以可训练权值参数, 再加上一个可训练偏置参数, 最后通过 sigmoid 函数计算得到, 降采样层进行局部平均和降采样的操作, 降低特征图的分辨率, 同时降低了网络输出对于位移和变形的敏感程度^[9]. 卷积神经网络采用的是权值共享, 这样一来通过一个特征图上的神经元共享同样的权值就可以大大减少自由参数, 可以用来检测相同的特征在不同角度表示的效果从而得到更好的鲁棒性^[10].

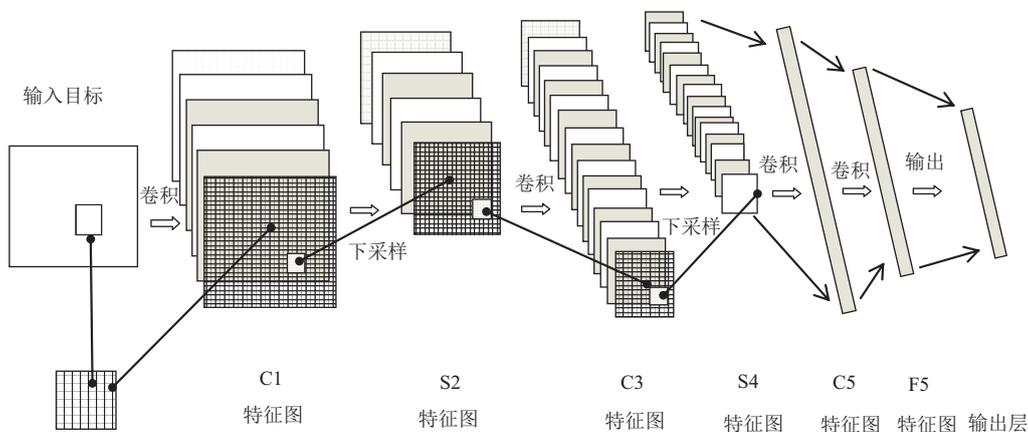


图 1 LeNet-5 模型图

1.2 分类器

目前识别领域的分类器有很多种, 其中支持向量机分类器和 softmax 分类器针对图像识别有突出表现.

Vapnik^[11]等提出的支持向量机 (Support vector machine, SVM) 是解决分类问题的一个重要方法, 该方法将高位特征空间的点积运算替代为原空间的核函数, 降低了计算的复杂性. 支持向量机能保证所得极值解

就是全局最优解, 减少了神经网络容易陷入局部最小的可能, 因此它在解决小样本、非线性及高维模式识别中均表现出许多特有的优势^[7].

Softmax 分类器是在深度学习中经常被提到的分类器, 该分类器是在多分类问题中 logistic 回归模型的扩展. Softmax 分类器对于诸如 MNIST 手写数字分类等问题上有很好的效果, 该问题就是识别不同的单个

数字和图像,可以应用于车牌识别、目标检测等方向。

2 基于 CNN 多层特征提取的目标识别

2.1 问题的提出与总体方案

多分类 SVM 分类器是具有较强泛化能力的通用学习算法,不但可以处理非线性、高维数的分类问题,而且能够解决神经网络中经常出现的局部极小值的问题,被广泛应用于图像识别领域并取得良好的效果。但多分类 SVM 算法本质是通过二次规划获得支持向量的,当分类样本数量 n 较大时,算法进行的 n 阶矩阵运算不仅计算量巨大,并且占机器内存大,运行速度缓慢。需要应用多分类 SVM 分类器的鲁棒性和泛化能力的优点去解决大量图像识别问题,结合提取特征降维可以减少计算量,提高计算速度。

选择卷积神经网络提取特征的原因,一方面,图像可以直接输入到卷积神经网络中训练,比其他网络的预处理工作简单。另一方面卷积神经网络的卷积和下采样工作是图像进行平移操作,根据平移不变性,提取出的特征向量与原始数据不仅不会过度拟合而且结果相同,这样省去对样本进行变造的过程,使得图像中的平移部分对于最后的特征向量是没有影响的。除此之外,卷积神经网络的卷积和下采样具有可变性和灵活性,可以改变特征向量大小的特点,当分类器有拟合现象可以通过调节特征向量维数解决。

两种方法相结合取长补短,SVM 解决了神经网络的局部极小值的问题,卷积神经网络提取特征解决了 SVM 算法的处理大样本问题,同时减少拟合现象提高运行速率。选取合适的 CNN 模型和特征提取层是分类识别的重要问题。根据多种论文中提出的一般选择分类层的前一层作为特征提取层的问题,提出在不同全连接层都进行特征提取的方法与上述结论进行对比。因此本文提出基于 CNN 多层特征提取的目标识别方法。

本文设计的基于卷积神经网络多层特征提取的目标识别系统整体框图如图 2 所示。

2.2 CNN 多层特征提取模型

CNN 多层特征提取模型包括训练 CNN 模型和特征提取两个部分。

2.2.1 CNN 模型的训练

本文采用的是由 Alex Krizhevsky 等人^[8]提出的大型的深度卷积神经网络,其中总共有 23 层,包含八个学习层——五个卷积层、三个全连接层的卷积神经网络,

网络的构建过程如图 3 所示,最后一个全连接层的输出被送到一个 Softmax 层,通过 Softmax 分类层分 1000 类标签,该框架 ImageNet ILSVRC-2012 大赛中依靠 top-5 测试误差率 15.3% 取得了胜利。

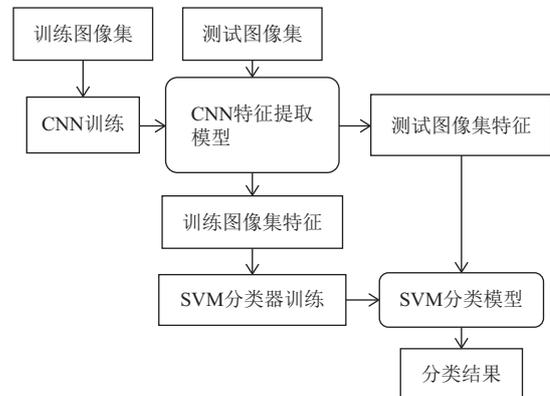


图 2 多层特征提取目标识别系统流程图

第一个卷积层利用 96 个大小为 $11 \times 11 \times 3$ 、步长为 4 个像素的核,来对大小为 $224 \times 224 \times 3$ 的输入图像进行滤波。第二个卷积层需要将第一个卷积层经过归一化和下采样后的输出作为自己的输入,且利用 256 个大小为 $5 \times 5 \times 48$ 的核对其进行滤波。第三、第四和第五个卷积层彼此相连,没有任何介于中间的下采样层与归一化层。第三个卷积层有 384 个大小为 $3 \times 3 \times 256$ 的核被连接到第二个卷积层的输出。第四个卷积层拥有 384 个大小为 $3 \times 3 \times 192$ 的核,第五个卷积层拥有 256 个大小为 $3 \times 3 \times 192$ 的核。全连接层都各有 4096 个神经元。全连接层有三层,在三层全连接层中选取一层得到 4096 张特征图排列为一条列向量,得到最终的特征向量。

2.2.2 特征提取

AlexNet 共 8 层学习层,包括 5 个卷积层和 3 个全连接层。第 1、2 层卷积后进行局部相应归一化。第 1、2、5 层有重叠池化。卷积层和全连接层的每层输出都有 relu。是一种经典的达到速度快、错误率低的一种卷积神经网络。

AlexNet 模型使用 Nair 和 Hinton^[12]提出的修正线性单元 (relu) 作为激活函数,式 (1) 是 relu 函数,训练带 relu 的深度卷积神经网络比带 tanh 单元的同等网络要快好几倍,式 (2) 是 tanh 函数。式 (3) 是传统的非线性的 Sigmoid 函数,双侧抑制,增大中央区信号,同时抑制两侧信号。relu 模型是近似生物神经激活函数,其

特点是单侧抑制、相对宽阔的兴奋边界以及稀疏激 活性。

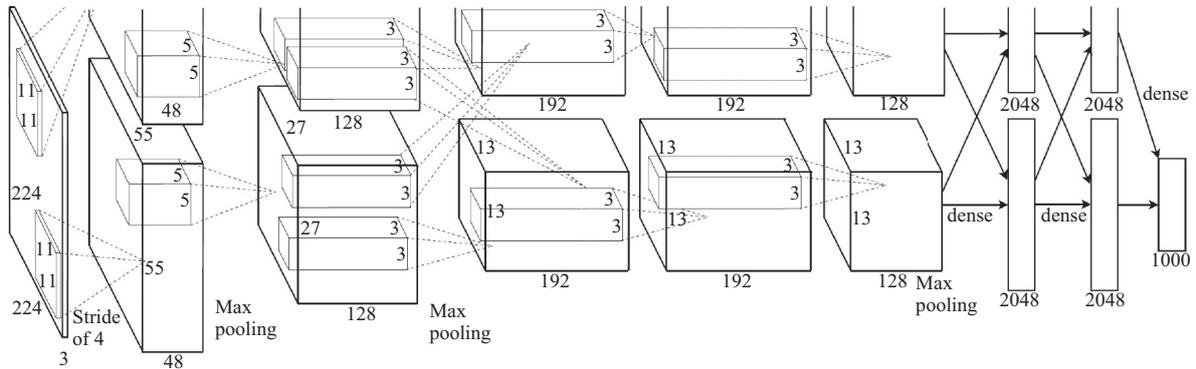


图3 Alex Krizhevsky 等人提出的大型的深度卷积神经网络 AlexNet 模型图^[8]

$$\max(0, x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases} \quad (1)$$

$$\Theta(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} \quad (3)$$

根据公式(1)(2)(3)所得, relu 函数是分段函数, 在大于 0 时导数不变, 反向传播梯度, 而且不容易出现梯度发散问题, sigmoid 和 tanh 函数的导数比较复杂, 在两端接近目标时, 导数变小. 根据神经网络训练的原理, 导数变小会导致误差反向传播, 收敛慢. relu 函数右侧不发散, 所以多处隐含层输出为 0, 网络因此变得稀疏, 降低过拟合的可能性, 提高卷积效率. 因此经过 relu 函数激活后的全连接卷积层进行特征提取的识别率明显比全连接卷积层的识别率高.

选取全连接层作为特征提取层, AlexNet 卷积神经网络模型中分为三个全连接层从前往后分别命名为 fc6、fc7、fc8, 在 fc6、fc7 卷积层后面都有 relu 层以提高速率, 这两层也包含在特征提取层的范围内.

2.3 分类识别

多分类 SVM 分类器作为一种具有较强泛化能力的通用学习算法, 被广泛应用于图像识别领域并取得良好的效果. 但使用多分类 SVM 解决小样本问题有较好的效果, 在解决数量大、类型多的图像集时, 比较不进行特征提取的多分类 SVM 识别率, 使用卷积神经网络特征提取后进行多分类 SVM 识别的效果最佳, 为了验证本文提出的卷积神经网络多层特征提取的有效性, 选择 SVM 分类器对特征进行训练和测试.

3 实验

3.1 数据集

Caltech 101 数据集创造在 2003 年 9 月, 编写由 Fei-Fei Li, Marco Andreetto 和 Aurelio Ranzato 在加州理工学院. Caltech 101 数据集包括总共 9146 个图象如图 4 所示, 101 个区别对象类别, 以及一个另外的背景或凌乱类别. 每个对象类别平均包含在 40 个和 800 个图象之间.

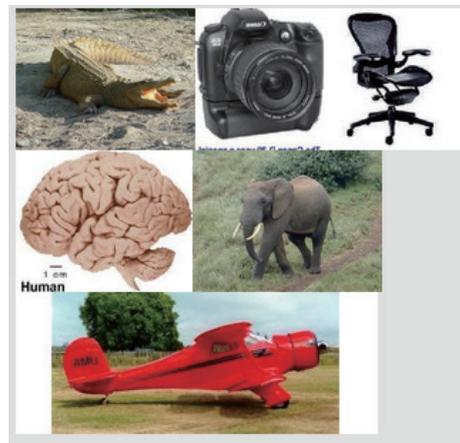


图4 Caltech101 数据集

办公用品数据集是由 15 类物品, 共 450 个图像, 有 15 类办公室中常见物品, 每类物品 30 张图像, 图像由百度图库和淘宝网提供, 如图 5 所示.

3.2 实验结果与分析

Caltech 101 数据集中每类图像个数的限制, 取 102 类图像做实验时, 取这 102 类中图像个数最小的作为整个实验的数据集, 每类 31 张.

在使用 CNN 提取特征前, 将图像处理为 image

data store 格式, 该格式存储每个图像的位置, 提取每个图像的名字作为标签, 将输入 CNN 网络的数据集中所有彩色图像的大小重置为 227*227*3.



图 5 办公用品自建数据集

本文将提出基于 AlexNet 卷积神经网络模型特征提取, 用 matlab 自带的多分类支持向量机进行分类的目标识别方法, 简称为 ACNSVnet, 用 Softmax 分类器进行分类的方法称为 ACNSOFTnet. 在 Caltech 101 数据集中, 固定训练个数, 单独使用 Softmax 分类器和支持向量机分类的识别率分别只有 56% 和 10% 左右. 由于图像数据集的数量过大、标签种类多单独使用这两种分类器运算复杂、容易判断错误导致运算时间很慢、识别率低、效率低, 因而效果不佳. 现用 ACNSVnet 方法得到的识别率约在 87.91% 左右, 去除数据集中凌乱数据集得到的识别率为 89.10%, 该方法不仅提高了识别率, 也增加了速度, 整体增加图像识别系统的效率.

表 1 对比了在 AlexNet 卷积神经网络模型中, 固定每个类别的训练图像个数为 21 张, 测试图像为 10 张, 进行三折交叉验证取平均值得到的从不同层中提取特征所得到的实验效果, 其中识别率按照式 (4) 计算, 选取 fc6、relu6、fc7、relu7、fc8 五个全连接层作为特征提取层得到的实验结果.

$$\text{识别率} = \frac{\text{正确分类图像数目}}{\text{总图像数目}} \times 100\% \quad (4)$$

表 1 特征提取层的识别率和时间

特征提取层	fc6	relu6	Fc7	relu7	Fc8
识别率(%)	84.71	87.25	83.04	84.02	79.8
时间(s)	326	365	412	380	414

根据以往经验和各种论文中提到的基于卷积神经网络进行特征提取应在卷积神经网络分类层前一层的全连接层中进行, 该实验结果表明并非在最后一层全连接层进行特征提取得到的效果最优, 最后一层全连

接层特征提取分类的结果为 79.8%, 前两层全连接层的特征提取分类结果远远高于最后一层的结果分别为 84.71% 和 83.04%. 根据实验可以发现前两个全连接的识别率相差不大. 出现这种情况的原因是因为随着全连接层数越往后的增加, 导致过度拟合, 从而识别率下降, 所以当构建全连接层数较多的卷积神经网络模型时, 特征提取层应选取在低层的全连接层中.

实验结果表明, 选取经过 relu 函数激活后的全连接卷积层 (relu6 和 relu7) 进行特征提取的识别率 87.25% 和 84.02% 比全连接卷积层的识别率 (fc6 和 fc7) 高. 这里的原因很简单, 前面已经提到过 relu 是激活函数可以提高卷积神经网络的效率. relu 函数是分段函数, 在大于 0 时导数不变, 反向传播梯度, 而且不容易出现梯度发散问题, 右侧不发散, 所以多处隐含层输出为 0, 网络因此变得稀疏, 降低过拟合的可能性, 提高卷积效率.

表 2 对比了每个类别训练图像个数在 fc6 层和 relu6 层的识别率和运行时间, 这里受到 Caltech 101 数据集中每类图像个数的限制, 每类取 5 张递增到 25 张图像作为训练集的数据进行实验.

表 2 fc6 层和 relu6 层不同训练图像个数的识别率和时间

训练图像个数	5	8	10	12	15	18	20	23	25
fc6识别率(%)	72.13	80.02	80.10	81.32	82.78	84.54	84.90	85.66	86.11
运行时间(s)	304	318	322	323	336	351	365	366	384
relu6识别率(%)	77.41	80.14	82.63	83.44	84.62	87.10	87.25	87.87	87.91
运行时间(s)	297	299	306	324	329	348	365	378	379

实验结果表明, 在 relu6 层中提取特征, 25 张训练图像范围内, 随训练数据的增大, 识别率越高. 当训练图像在 15 张时识别率为 84.62%, 而在 18 张部分识别率为 87.10%, 在这期间识别率提高了三个百分点, 训练图像大于 18 张的实验识别率都平均保持在 87%-88% 之间, 识别结果稳定. 实验结果说明, 随着训练图像的增加, 运行时间基本也是逐渐增加的. 同时, 通过对比 fc6 层和 relu6 层的结果表明, 在训练集个数相同情况下在 relu6 层取特征得到的识别率总是高于 fc6, 运行时间相对 fc6 层相对短. 所以综上三个结果, 为了提高识别系统的效率, 选择在 relu6 层提取特征, 选取每类训练图像个数在 18-25 张左右的图像, 这一方面提高了识别率, 另一方面缩短系统运行的时间.

分类器除了支持向量机还有 Softmax、BP 神经网

络、经常使用的 libsvm, 使用卷积神经网络对数据集 Caltech101 提取特征, 将特征输入到上述分类器中其结果分别如表 3, 表 4, 表 5 所示。

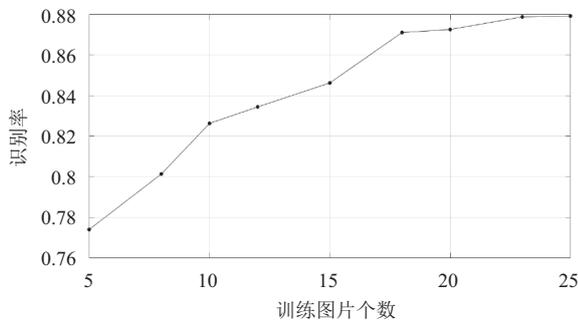


图 6 relu6 层提取特征训练图像个数与识别率的关系

表 3 Softmax 分类器的特征提取层识别率

特征提取层	fc6	relu6	fc7	relu7	fc8
识别率(%)	85.1	87.06	82.84	84.51	80.1

表 4 libsvm 分类器的特征提取层识别率

特征提取层	fc6	relu6	fc7	relu7	fc8
识别率(%)	1.56	1.37	1.56	9.8	32.451

表 5 BP 神经网络分类器的特征提取层识别率

特征提取层	fc6	relu6	fc7	relu7	fc8
识别率(%)	1.47	4.37	1.56	12.3	11.8

相比较而言, matlab 自带的多分类支持向量机在 relu6 层获得最优异的识别率, Softmax 分类器的识别率更加稳定结果与多分类支持向量机区别不大。而 libsvm 识别率较低, BP 神经网络由于网络复杂, 根据不同图像集需要进行参数调整, 训练速度非常慢, 识别率也不理想。根据实验结果表明, 选择 matlab 自带的多分类支持向量机为最优选择, Softmax 分类器可以作为代替该分类器的另一个比较好的选择, 而 libsvm 和 BP 神经网络在识别率和时间上都不占优势, 在这类实验中不是最优选择。

根据上述三个实验得到的结论, 使用自建办公用品数据集进行实验, 在五类不同特征提取层中获得的识别率。预处理同上, 根据上述结论为了提高识别系统的效率, 选取 matlab 自带多分类支持向量机作为分类器, 选取每类训练图像个数在 18-25 张左右的图像, 该实验选择 25 张训练图像, 5 张图像进行测试, 并进行三折交叉验证取平均值得到的从不同层中提取特征所得到的实验效果如表 6 所示。

表 6 SVM 分类器的特征提取层在自建办公用品

特征提取层	数据库识别率				
	fc6	relu6	fc7	relu7	fc8
识别率(%)	89.33	92	89.33	91.33	85.33
时间(s)	35	34	39.9	40.3	41.23

根据实验结果发现, 选取经过 relu 函数激活后的全连接卷积层 (relu6 和 relu7) 进行特征提取的识别率明显比全连接卷积层的识别率 (fc6 和 fc7) 高, 与第一个实验结果一致, 所以当构建全连接层数较多的卷积神经网络模型时, 特征提取层应选取在低层的全连接层中。

构建办公用品识别系统分为两步分别为训练和识别。训练部分选取 AlexNet 卷积神经网络模型的 relu6 层作为特征选取层, 输入 375 张办公用品图像 (每类 25 张训练图像) 到卷积神经网络模型进行训练并提取特征, 将得到的特征提取集输入到 matlab 多分类支持向量机分类器中进行训练。剩余 75 张 (每类 5 张测试图像) 对识别系统进行测试, 得到识别率。使用 c# 作为接口将电脑连接摄像头实时截图所拍物品, 将所截图输入到识别系统进行识别得到识别结果。针对办公室机器人图像识别方向的研究做了一个简单的系统, 有助于快速实时的定位物品。

4 结束语

本文提出基于卷积神经网络多层特征提取的目标识别的方法, 对图像数据集进行分类, 针对选取特征提取层效果最优的问题进行了研究。分类结果表明在卷积神经网络中提取特征所选取的特征层, 并非在全连接的最后一层或分类层的前一层效果最好。因为全连接层的作用是在一个更加抽象的层次上学习, 随着全连接层数越往后的增加, 导致过度拟合, 提取的特征从而识别率下降, 所以当构建全连接层数较多的卷积神经网络模型时, 特征提取层应选取在低层的全连接层中。在本次实验中经过 relu 函数激活后的全连接卷积层 relu6, 进行特征提取的识别率 87.25% 远远高于在最后一层全连接层 fc8 提取特征的识别率 79.8%, relu6 层较 relu7 层是低层的全连接层。实验结果表明当提取的特征一致时, 选取每类训练图像个数在 18-25 张左右的图像, 与支持向量机分类器相结合图像识别方法效果最好, 在办公用品数据集中识别率最高能达到 92% 的识别率。以上实验说明基于卷积神经网络多

层特征提取的目标识别的方法是有效的、可行的。本文构建了办公室用品数据集,并连接摄像头实现了针对办公室用品的实时目标识别系统。本文提出的方法可以推广应用于目标识别的各个领域。

参考文献

- 1 Wei XH, Lu SJ, Wen Y, *et al.* Recognition of handwritten Chinese address with writing variations. *Pattern Recognition Letters*, 2016, 73: 68–75. [doi: [10.1016/j.patrec.2015.12.018](https://doi.org/10.1016/j.patrec.2015.12.018)]
- 2 Singh P, Verma A, Chaudhari NS. On the performance improvement of devanagari handwritten character recognition. *Applied Computational Intelligence and Soft Computing*, 2015, 2015: 193868.
- 3 Zhang L, Song YY, Liu L, *et al.* Dissociable roles of internal feelings and face recognition ability in facial expression decoding. *NeuroImage*, 2016, 132: 283–292. [doi: [10.1016/j.neuroimage.2016.02.049](https://doi.org/10.1016/j.neuroimage.2016.02.049)]
- 4 徐欢, 高雁凤, 郑恩辉, 等. 一种用于驾驶员疲劳检测的人眼检测方法. *计算机工程*, 2013, 39(9): 254–257, 276.
- 5 江源远, 胡进, 苏芮琦, 等. 基于积分通道特征的车辆前方行人检测方法研究. 2016 中国汽车工程学会年会论文集. 上海. 2016.
- 6 Charalampous K, Kostavelis I, Gasteratos A. Robot navigation in large-scale social maps: An action recognition approach. *Expert Systems with Applications*, 2016, 66: 261–273. [doi: [10.1016/j.eswa.2016.09.026](https://doi.org/10.1016/j.eswa.2016.09.026)]
- 7 LeCun Y, Bottou L, Bengio Y, *et al.* Gradient-based learning applied to document recognition. *Proc. of the IEEE*, 1998, 86(11): 2278–2324. [doi: [10.1109/5.726791](https://doi.org/10.1109/5.726791)]
- 8 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Proc. of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, Nevada, USA. 2012. 1097–1105.
- 9 任陈俊, 仲朝亮, 刘士荣, 等. 基于卷积神经网络的场景运动目标分类研究. 中国自动化学会控制理论专业委员会. 第三十四届中国控制会议论文集 (C 卷). 杭州, 中国. 2015.
- 10 Ijjina EP, Chalavadi KM. Human action recognition using genetic algorithms and convolutional neural networks. *Pattern Recognition*, 2016, 59: 199–212. [doi: [10.1016/j.patcog.2016.01.012](https://doi.org/10.1016/j.patcog.2016.01.012)]
- 11 Vapnik VN. *The Nature of Statistical Learning Theory*. New York, USA: Springer, 1995.
- 12 Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. *Proc. of the 27th International Conference on Machine Learning*. Haifa, Israel. 2010.