

基于PCANet的图像识别及其改进

邢子扬

(上海大学理学院信息与计算科学 200444)

摘要:作为一种结合了主成分分析和卷积神经网络的特征提取算法, PCANet 在图像识别领域有着强大的功能和广泛的应用, 它具有训练时间短、识别准确率高、模型简单等优点。PCANet 算法的结构由三个阶段组成, 前两个阶段十分相似, 都是对输入矩阵分块采样、去均值化后, 利用 PCA 的方法构造PCA 滤波器, 然后对每个采样块进行卷积处理, 在第三阶段中, 我们对输入矩阵进行哈希编码和直方图化处理来提取图像特征, 至此, 利用 PCANet 来提取图像特征的过程就完成了, 最后我们采用 SVM 来对图像进行分类。虽然 PCANet 具有诸多优势, 但它仍然有上升的空间。在PCANet算法中, 滤波器的选择对特征提取的效果有着至关重要的作用, 所以我们尝试用傅里叶滤波器和小波滤波器来替换PCA 滤波器, 以此来提高算法对图像识别的准确性, 经过试验后, 我们确实发现更换滤波器后的算法有着更高的准确性。

关键词: PCANet; 图像识别; 傅里叶分析; 小波分析

第一章 引言

在图像识别领域, 深度学习是一种非常普遍和流行的做法。深度学习的思想是发现多个层次的表示, 希望更高层次的特征能够表示数据更抽象的语义。图像分类深度学习成功的一个关键因素是卷积结构的使用。卷积深度神经网络体系结构主要由三层组成——卷积层、池化层、全连接层。卷积层和池化层可交替出现多次, 最后由全连接层结尾。PCANet 就是结合了卷积神经网络的结构和主成分分析的思想, 所构建出的一个简单地深度学习网络结构。但以 PCA 的方法构建出的滤波器自然而然的也继承了 PCA 的一些局限性, 比如, 几何意义上, PCA 是要找到一个方向, 使得在这个方向上所有数据点的投影差异最大化, 即方差最大。但是, 如果仅有某一个点差距很大, 这一个点就可能贡献出非常大的方差, 从而导致分析出现错误。所以我们考虑换一个角度来看待图像特征, 我们想要通过傅里叶分析和小波分析把图像信息从空间域转换到频率域, 从而构造出傅里叶滤波器和小波滤波器, 并把它们带入到 PCANet 的整体框架中, 形成基于傅里叶滤波器和小波滤波器的改进算法^[1]。

第二章 PCANet 结构与原理

2.1 第一阶段

2.1.1 图片预处理

假设我们的训练集中有 N 张图片, 每张图片的尺寸是 $m \times n$, 首先, 我们要对这每一张图片进行块采样。采样时, 对于图片的每一个像素周围都要进行一次块采样, 我们设这个采样块的尺寸 $k_1 \times k_2$, 每个采样块表示为 $x_{i,1}, x_{i,2}, \dots, x_{i,m-n}, m-n = m - \lceil k_1/2 \rceil$, $n-n = n - \lceil k_2/2 \rceil$, $\lceil z \rceil$ 表示大于或等于 z 的最小整数。然后把收集到的所有采样块进行级联处理, 以第 i 张图片为例, $X_i = [x_{i,1}, x_{i,2}, \dots, x_{i,m-n}]$, 然后对采样后得到的数据进行零均值化处理, $\bar{X}_i = [x_{i,1}, x_{i,2}, \dots, x_{i,m-n}]$ 。将训练集中所有图片进行如此操作后就得到了训练样本矩阵。

2.1.2 构造滤波器

$$X = [\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N] \in \mathbb{R}^{k_1 k_2 \times N \tilde{m} \tilde{n}} \quad (2.1)$$

如果图像的某一区域与滤波器检测的特征十分相似, 那么当滤波器经过该区域时, 就会得到一个很高的值, 反之, 如果图像的某一区域与滤波器检测的特征相差很远时, 那么得到的数值很低。由此我们发现, 图像的滤波器在对整个图像滑动过滤时, 每经过图像一个区域得到的值越高, 则该区域与滤波器检测的特定特征相关

度越高。所以, 滤波器的构造是该算法的重中之重。

我们假设在第一阶段中, 滤波器的数量是 L_1 , 我们接下来会通过 PCA 的方法构造出第一阶段的 L_1 个滤波器, 具体方法如下: 对矩阵 X 进行主成分分析, 即提取矩阵 X 的前 L_1 个最大特征值对应的特征向量, 以此构成映射矩阵。在这些主成分中, 保留了前 L_1 个最重要, 最易于识别的特征的信息, 去除了其他识别度较低, 有效信息较少的特征的信息。这样既能有效进行分类, 又能降低计算的难度。滤波器的表示如下:

$$W_1 = \text{matk}_{,k} (q_1(X_1^T)) \in \mathbb{R}^{k_1 \times k_2}, l = 1, 2, \dots, L_1 \quad (2.2)$$

其中, $\text{matk}_{,k}(v)_{l2}$ 表示将向量 $v \in \mathbb{R}^{k_1 k_2}$ 转换为矩阵 $W \in \mathbb{R}^{k_1 \times k_2}$, $q_l(X_1^T)$ 表示 X_1^T 的第 l 个主特征向量。得到 L_1 个滤波器后, 分别用每个滤波器对原始图像进行卷积映射。映射的结果就是第一层的输出矩阵。

2.2 第二阶段

第二层的操作与第一层几乎完全相同, 这里唯一需要注意的是, 在卷积之前, 要对原始图像进行边缘补零操作, 否则映射图像会与原始图像的尺寸不同。首先将第一层的输出作为第二层的输入, 我们假设在第二层中的滤波器数量为 L_2 , 然后如同第一层中的操作一样, 对输入矩阵进行分块采样、级联、零均值化。

$$Y = [Y_1^T, Y_2^T, \dots, Y_{L_1}^T] \in \mathbb{R}^{k_1 k_2 \times L_1 N \tilde{m} \tilde{n}} \quad (2.3)$$

然后同样进行主成分分析, 用 PCA 的方法的到第二层的 L_2 个滤波器, 并用这些滤波器进行卷积映射。

$$W_2 = \text{matk}_{,k} (q_l(Y_1^T)) \in \mathbb{R}^{k_1 \times k_2}, l = 1, 2, \dots, L_2 \quad (2.4)$$

在第一层中, 我们用 L_1 个滤波器对原始图片进行卷积映射, 会得到 L_1 个输出矩阵, 在第二层中, 我们又用 L_2 个滤波器对每一个输出矩阵进行了卷积映射, 所以最终每张图片会有 $L_1 \times L_2$ 个输出矩阵。

2.3 第三阶段

首先是对第二阶段的每个卷积的结果做二值化, 每一组得到 L_2 张二值图片, 对这 L_2 张二值图片进行编码, 得到十进制的图像矩阵。接下来对每张图片做 histBlock 到 vector 的变换, 假设对原图为 28×28 的图选取 8×8 的 histBlock, 假设 overlap 系数 0.5, 那么原图就会有 36 个 histBlock, 将 histBlock 变换为 vector, 得到 64×36 的矩阵。对这个 64×36 的矩阵做直方图统计。最后将这个矩阵转化为 vector, 这样就完成了一张图片的 PCANet 的特征提取^[1]。

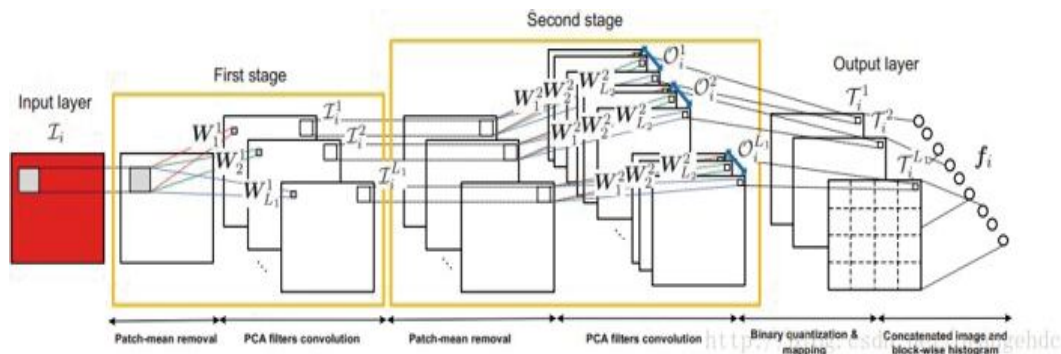


图 2.1 PCANet 模型

第三章 傅里叶分析与小波分析

我们考虑到 PCA 方法的一些缺陷, 设想通过傅里叶变换和小波变换, 把图像的特征信息, 从空间域转换到频率域, 然后从频率的角度来对图像信息的特征进行提取, 于是, 我们通过把 PCA 滤波器替换为傅里叶滤波器和小波滤波器对 PCANet 的算法进行了改进^[2]。

3.1 傅里叶分析

傅里叶变换可以表示出图片在不同频率下的信息, 研究人员发现, 不同的频率可以捕捉图像中不同程度的信息。例如, 高通滤波器将仅选择高频信号以获得边缘等结构化信息, 而低通滤波器将选择低频信号, 从而生成过度平滑和模糊的图像。接下来我们介绍用傅里叶分析构造滤波器的方法。

我们令 $\omega_k = 2k\pi/n$, $k \in F_n$, $F_n = \{-\lfloor \frac{n-1}{2} \rfloor, \dots, \lfloor \frac{n}{2} \rfloor\}$ 。[x]表示x的整数部分。

$$C(\omega_k) = \frac{1}{\sqrt{n}}(1, \cos\omega_k, \dots, \cos(n-1)\omega_k)^T, \quad (3.1)$$

$$S(\omega_k) = \frac{1}{\sqrt{n}}(1, \sin\omega_k, \dots, \sin(n-1)\omega_k)^T, \quad (3.2)$$

为了避免傅里叶基的重复, 我们把 ω_k 中的 k 限制在 F_{n+} 之内, F_{n+} 表示 F_n 中的

非负元素。然后我们令

$$c_k = \frac{1}{\sqrt{n}}(\omega_k, X \square \square 1, \quad k = 1, \dots, k1k2, \quad (3.3)$$

$$s_k = \frac{1}{\sqrt{n}}(\omega_k, X \square \square 1, \quad k = 1, \dots, k1k2, \quad (3.4)$$

这样, 我们就会得到 $k1k2$ 个 c_k 和 s_k , 我们从中选择出最大的 L 个 c_k 或 s_k , 它们所对应的向量 $C(\omega_k)$ 或 $S(\omega_k)$, 就是我们所构造出的傅里叶滤波器。

3.2 小波分析

小波分析通常会用到低通滤波器和高通滤波器。低通滤波器, 可以将输入信号的高频部份滤掉而输出低频部份。高通滤波器, 与低通滤波器相反, 滤掉低频部份而输出高频部份。对于许多信号, 低频成分相当重要, 它常常蕴含着信号的特征, 而高频成分则给出信号的细节或差别。人的话音如果去掉高频成分, 听起来与以前可能不同, 但仍能知道所说的内容; 如果去掉足够的低频成分, 则听到的是一些没有意义的声音。在小波分析中经常用到近似与细节。

小波滤波器与傅里叶滤波器的构造方法类似, 本文所采用的小波变换的低通滤波器和高通滤波器如下所示:

$$\mathbf{h} = \begin{bmatrix} \frac{1+\sqrt{3}}{4} & \frac{3+\sqrt{3}}{4} & \frac{3-\sqrt{3}}{4} & \frac{1-\sqrt{3}}{4} \end{bmatrix} \quad (3.5)$$

$$\mathbf{g} = \begin{bmatrix} \frac{1-\sqrt{3}}{4} & \frac{\sqrt{3}-3}{4} & \frac{3+\sqrt{3}}{4} & \frac{-1-\sqrt{3}}{4} \end{bmatrix} \quad (3.6)$$

然后, 通过构造如下矩阵并与数据矩阵相乘得到矩阵 A :

$$A = \begin{pmatrix} h_0 & h_1 & h_2 & h_3 & \dots & \dots & \dots \\ g_0 & g_1 & g_2 & g_3 & \dots & \dots & \dots \\ 0 & 0 & h_0 & h_1 & h_2 & h_3 & \dots \\ 0 & 0 & g_0 & g_1 & g_2 & g_3 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \ddots \end{pmatrix} \quad (3.7)$$

$$C = AX \quad (3.8)$$

然后, 取矩阵 C 每一行的 1 范数, 对最大的前 L 个 1 范数所对应的矩阵 A 的行向量转置, 就得到了 L 个小波滤波器 [4]。

第四章 实验及结果

分类器的选择可以是多样的, 本文采用了 SVM 来进行测试集的图像识别与分类。本文利用手写体数据集对基于 PCANet 的图像识别算法以及分别将滤波器改为傅里叶滤波器和小波滤波器的改进算法进行了实验, 本文在实验室, 选取了 5000 个图片作为训练集, 选取了 500 个图片作为测试集。

实验一: 无噪声情况下识别准确率测试及各个数字图片正确与错误个数

实验二: 有噪声情况下识别准确率测试及各个数字图片正确与错误个数

第五章 总结

从实验结果不难看出, 在无噪声的情况下, PCANet 算法、傅里叶改进算法和小波改进算法的识别准确率分别为 92.20%、92.40% 和 93.40%, 准确率依次升高, 而在有噪声的情况下, 这三种算法的准确率依次变为 87.20%、70.40% 和 77.40%。说明本文的两种改进算法在无噪声的情况下表现出了比 PCANet 更为良好的特征提取能力, 然而, 在有噪声的情况下, 傅里叶滤波器和小波滤波器的性能都大幅下降, 而 PCA 滤波器虽然也受到一定的影响, 但却远比其他两者的效果要好, 在有噪声时, PCA 滤波器表现出了很好的稳定性。

参考文献:

- [1]Chan, T.-H., Jia, K., Gao, S., Lu, J., Zeng, Z., and Ma, Y. Pcanet: A simple deep learning baseline for image classification? IEEE transactions on image processing, 24(12):5017 - 5032, 2015.
- [2]Yifei Li ,Zheng Wang, Kuangyan Song , Yiming Sun, "FrequentNet: A New Deep Learning Baseline for Image Classification", arXiv:2001.01034v1 [cs.CV] 4 Jan 2020
- [3]Costen, N. P., Parker, D. M., and Craw, I. Effects of high-pass and low-pass spatial filtering on face identification. Perception psychophysics, 58(4):602 - 612, 1996.
- [4]Antoine, J.-P., Murenzi, R., V anderghynst, P., and Ali, S. T. Two-dimensional wavelets and their relatives. Cambridge University Press, 2008.