基于CNN的图像识别技术在火灾自动报警系统的应用研究

马　奇[1]，冯欧阳[2]

（1. 同济大学建筑设计研究院（集团）有限公司，上海 200092；
2. 国网浙江省电力有限公司瑞安市供电公司，浙江 温州 325205）

摘要：卷积神经网络（CNN）在图像处理方面相对于神经网络有着更好的表现，基于 CNN 能够更加准确便捷识别图像类信息；同时消防安全体系的构成在工业生产、电力发电以及民用建筑等领域愈加重要，其重要性不可忽视。研究的重点是如何利用 CNN 的图像处理和识别的优势，将其应用于图像型探测器，从而实现构建更加优化的火灾自动报警系统。

关键词：卷积神经网络；图像识别；火灾自动报警系统

Application of image recognition technology based on CNN in automatic fire alarm system

MA Qi[1], FENG Ouyang[2]

（1. Tongji University Architectural Design and Research Institute (Group) Co., Ltd.,Shanghai 200092, China;
2. State Grid Zhejiang Electric Power Co., Ltd., Rui'an Power Supply Company, Wenzhou 325205, China）

Abstract: Convolutional neural network (CNN) has a better performance than the neural network in image processing, which can identify image information more accurately and conveniently based on CNN. At the same time, the composition of the fire safety system is becoming more and more important in industrial production, electric power generation, and civil construction, and its importance can not be ignored. The focus of research is on how to use the advantages of convolutional neural networks in image processing and recognition, and apply them to image detectors, so as to realize the construction of a more optimized automatic fire alarm system.

Key words: CNN; image recognition; automatic fire alarm system

0　引言

由于出色的图像处理和识别能力，卷积神经网络（CNN）广泛用于图像识别领域。通常情况下，感温、感烟类传感器不太适合于关键环境（大空间、隧道和无人值守的变电站等），在发生警报时需要人的参与来确认火灾，包括访问火灾位置。此外，此类系统通常无法提供有关火灾大小、位置和燃烧程度的信息。研究基于CNN的深层特征监控网络的早期火灾探测技术，应用小波变换提取图像的光谱特征，然后将这些特征输入到CNN的不同层阶段，可以克服这些局限性。本文选择了最简单的小波2D Haar（该方法并不局限于Haar），因为它足以描述不同频率的火焰信息。关键在于CNN中的卷积层和池层可以被视为光谱分析的一种有限形式。因此，这两层可以通过2D Haar变换进行推广，以实现光谱分析。

1　火灾自动报警系统

1.1　火灾自动报警系统基本概念

火灾自动报警系统主要是由以下几个部分组成：①触发装置；②火灾报警装置；③联动输出装置；④具有其他辅助功能的装置。能够在火灾初期阶段，通过触发装置，将火焰、烟雾等燃烧产生的现象通过探测器转化为电信号，传输到火灾报警控制器，及时向控制中心传达火灾的位置，使得火灾在早期被及时发现，帮助人们处置火情，有效地阻止火势蔓延，减少生命财产的损失。

1.2　现行的火灾自动报警系统框架

火灾自动报警系统框图如图1所示。根据《14X505-1火灾自动报警系统设计规范》图示可知，该系统主要由火灾预警系统、消防联动控制系统、火灾探测报警系统、电源、管理中心、传输设备和相关设备设施构成。图像型火灾探测器是火灾探测器系列的组成部分，如何将CNN技术应用于该型探测器，从而实现有效的火灾识别、探测是本文研究的重点内容。

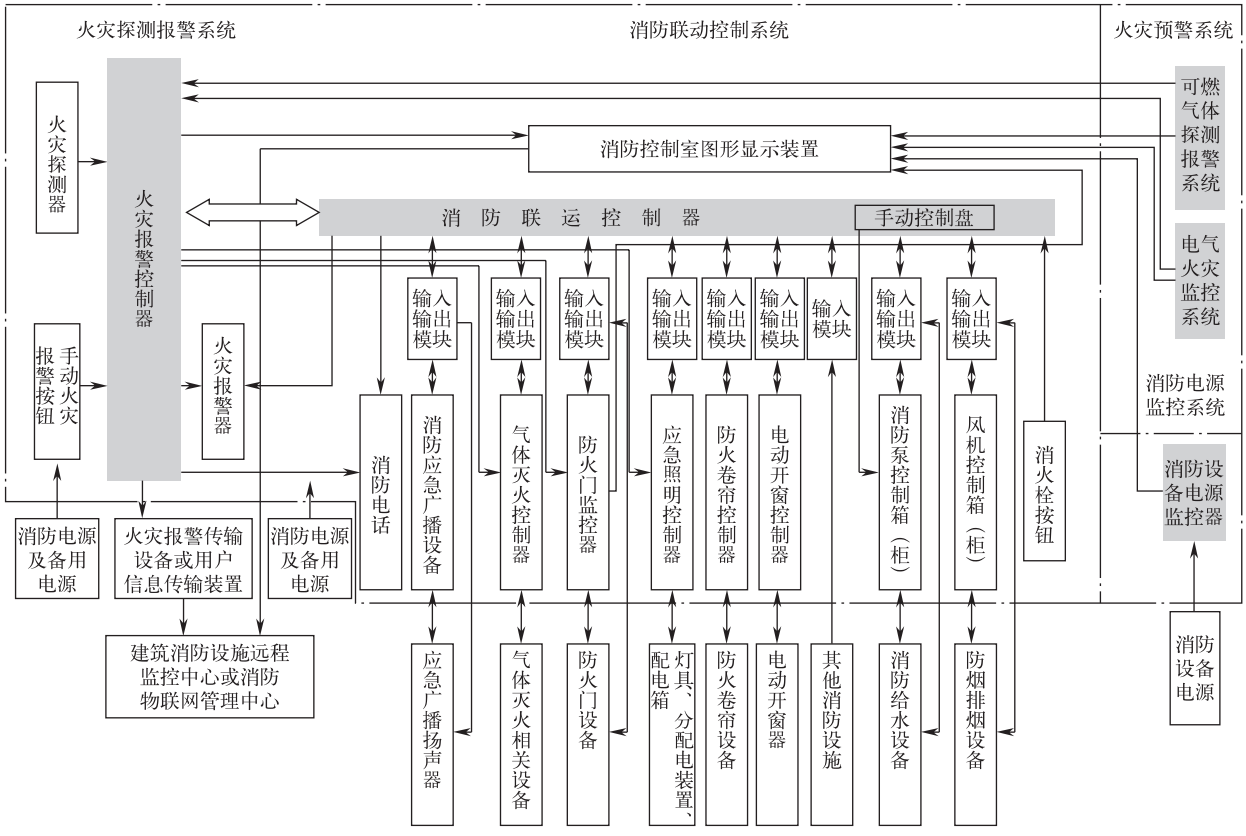

图 1　火灾自动报警系统框图

Fig. 1 Block diagram of automatic fire alarm system

1.3　图像型火灾探测器及其应用

图像型火灾探测器是火灾探测器的一种类型，主要用于室外、隧道、室内高大空间和无人值守的变电站等区域。

图像型火灾探测器拥有诸多优点：①能够灵活地接入报警系统；②有多重探测维度（信号强度、时间分辨率和三维空间尺寸）；③可以快速确认火灾，无须现场确认；④具备视频监视功能；⑤视觉和图像一体化处理。图像型火灾探测器如图2所示。

图 2　图像型火灾探测器

Fig. 2 Image fire detector

2　卷积神经网络（CNN）

2.1　卷积神经网络基本概念

卷积神经网络（Convolutional Neural Network，CNN）是一种前馈神经网络，其人工神经元可以响应一部分覆盖范围内的周围单元，对于大型图像处理有出色表现。它由一个或多个卷积层和顶端的全连通层（对应经典的神经网络）组成，同时包括关联权重和池化层（Pooling Layer）。这一结构使得CNN能够利用输入数据的二维结构。与其他深度学习结构相比，CNN在图像方面能够给出更好的结果。这一模型也可以使用反向传播算法进行训练。相比较其他深度、前馈神经网络，CNN需要考量的参数更少，使之成为一种颇具吸引力的深度学习结构。其基本结构如图3所示。

图 3　卷积神经网络结构

Fig. 3 Convolutional neural network structure

2.2　卷积层

卷积层的作用是负责提取图像中的局部特征，可以产生一组平行的特征图（Feature Map），通过在输入图像上滑动不同的卷积核并执行一定的运算而组成。此外，在每一个滑动的位置上，卷积核与输入图像之间会执行一个元素对应乘积并求和的运算，将感受的信息投影到特征图中的一个元素。这一滑动的过程可称为步幅Z_s，是控制输出特征图尺寸的一个因素。卷积核的尺寸要比输入图像小得多，且重叠或平行地作用于输入图像中，一张特征图中的所有元素都是通过一个卷积核计算得出的，也即一张特征图共享了相同的权重和偏置项。

2.3　线性整流层

线性整流层使用线性整流（ReLU）作为这一层神经的励磁函数。它可以增强判定函数和整个神经网络的非线性特性，而本身并不会改变卷积层。相较于其他一些用于增强网络的非线性特性的函数，如双曲正切函数或者Sigmoid函数，ReLU函数可以将神经网络的训练速度提升数倍，不会对模型的泛化准确度造成显著影响。

2.4　池化层

池化层用来大幅降低参数量级（降维），实际上是一种非线性形式的降采样。有多种不同形式的非线性池化函数，其中最大池化（Max Pooling）最为常见。它是将输入的图像划分为若干个矩形区域，对每个子区域输出最大值。直觉上，这种机制能够生效的原因在于，一个特征的精确位置远不及它相对于其他特征的粗略位置重要。池化层会不断地减小数据的空间大小，因此参数的数量和计算量也会下降，这在一定程度上控制了过拟合。通常来说，CNN的网络结构中的卷积层之间都会周期性地插入池化层。池化操作提供了另一种形式的平移不变性。因为卷积核是一种特征发现器，通过卷积层可以很容易地发现图像中的各种边缘。但是卷积层发现的特征往往过于精确，即使高速连拍一个物体，照片中物体的边缘像素位置也不大可能完全一致，通过池

化层可以降低卷积层对边缘的敏感性。池化层每次在一个池化窗口（Depth Slice）上计算输出，然后根据步幅移动池化窗口。目前最常用的池化层如图4所示，步幅为2，池化窗口为2×2的二维最大池化层。每隔2个元素，从图像划分出2×2的区块，然后对每个区块中的4个数取最大值，这会减少75%的数据量，即

$$f_{X,Y}(S)=\max_{a,b=0} S_{2X+a,2Y+b} \quad (1)$$

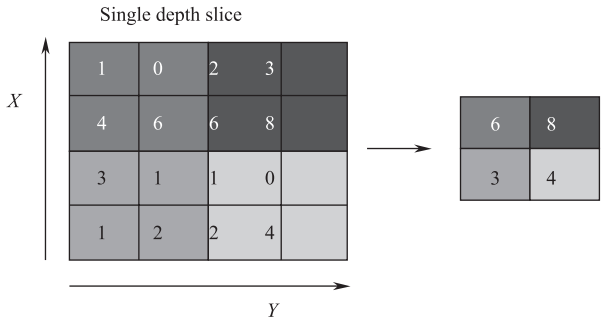

图4 常用的池化层

Fig. 4 Common pooling layer

除了最大池化之外，池化层也可以使用其他池化函数，例如“平均池化”甚至“L2-范数池化”等。平均池化的使用曾经较为广泛，但是最近由于最大池化在实践中的表现更好，平均池化已经不太常用。由于池化层过快地减少了数据的大小，目前文献中的趋势是使用较小的池化滤镜，甚至不再使用池化层。RoI池化（Region of Interest）是最大池化的变体，其中输出大小是固定的，输入矩形是一个参数。池化层是基于Fast-RCNN架构的卷积神经网络的一个重要组成部分。

2.5 完全连接层

完全连接层起到输出想要结果的作用，在经过几个卷积和最大池化层之后，神经网络中的高级推理通过完全连接层来完成。和常规的非卷积人工神经网络一样，完全连接层中的神经元与前一层中的所有激活都有联系。因此，它们的激活可以作为仿射变换来计算，先乘以一个矩阵，然后加上一个偏差偏移量（向量加上一个固定的或者学习来的偏差量）。

3 火灾自动报警系统中应用CNN技术

3.1 技术应用的基本框架

本文是将CNN应用于火灾自动报警系统中的图像型火灾探测器，相较于基于神经网络的图像识别而言，CNN具备强大的图像识别能力，优势明显，能够更好地实现火情的辨别和报警处置，其基本作用流程图如图5所示。

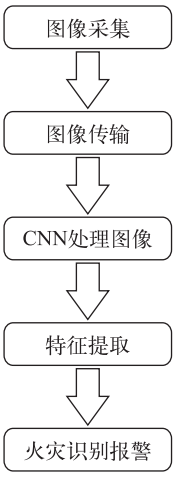

图5 基本作用流程图

Fig. 5 Flow chart of basic functions

3.2 模型实现

（1）VGG模型结构

VGG网络是牛津大学Visual Geometry Group团队研发，该项目的主要目的是证明增加网络深度，在一定程度上提高网络的准确度。VGG有，A、B、C、D、E五种模型，其中的E模型VGG19是参加ILSVRC 2014挑战赛使用的模型，获得了ILSVRC定位第一名和分类第二名的成绩。整个过程证明，通过把网络深度增加到16～19层，能够提高网络性能。VGG网络跟之前学习的LeNet网络和AlexNet网络有很多相似之处。本文搭建的VGG19模型分成了五个大的卷积层和三个大的全链层；同时，为了减少网络训练参数的数量，整个卷积网络均使用3×3大小的卷积。参考的VGG网络有五种模型结构。A、B、C、D、E模型均是由5个stage和3个全链层和一个softmax分类层组成，其中每个stege有一个

max-pooling层和多个卷积层。每层的卷积核个数从首阶段的64个开始，每个阶段增长一倍，直到达到512个。A是最基本的模型，8个卷基层和3个全连接层，一共11层；A-LRN忽略；B在A的基础上，在stage1和stage2基础上分别增加了1层3×3卷积层，一共13层；C在B的基础上，在stage3、stage4和stage5基础上分别增加了一层1×1的卷积层，一共16层；D在B的基础上，在stage3、stage4和stage5基础上分别增加了一层3×3的卷积层，一共16层；E在D的基础上，在stage3、stage4和stage5基础上分别增加了一层3×3的卷积层，一共19层；模型D就是经常说的VGG16网络，模型E则为VGG19网络。

（2）利用Tensorflow构建VGG19网络

VGG团队用Tensorflow搭建VGG16和VGG19网络，在使用网络前，需要下载已经训练好的参数文件，根据该模型对VGG19网络做了一些微调以适应训练需求，增加了精调训练。在原版的基础上做了一些改动，增加了入参num_class，该参数代表分类个数，在有N个种类的图片需要训练时，这个值必须设置成N。

3.3 实现方式

主程序逻辑是将图片转为三维BGR矩阵，然后使用Tensorflow的队列来自动获取每个batch的数据，队列可以把图片及对应的标签准确地取出来，同时还自动打乱顺序，之后将其加入到TensorBoard中。整个VGG19网络一共训练100个epochs，每个epoch有100个迭代，同时使用交叉熵和梯度下降算法精调VGG19网络的最后三个全链层fc6、fc7和fc8。

3.4 输入图像的识别和转换

训练数据集使用了来自挪威的数据集，通常模型训练需要的正样本和负样本数量要相等，并且数据集越多越好，但在本次训练中，所有图片均是从网络抓取而来，且训练集中的fire的图片和non-fire图片是不相等的，分别为223张和445张，测试集中的fire图片和non-fire的图片则相等，均为50张。

3.5 模型准确度评估

使用精确率（Precision）、召回率（Recall）和两者的调和平均F_1值，相较于准确率而言，可以在正负样本不平衡的情况下较好地衡量网络性能。精确率（Precision）是针对预测结果而言的，表示的是预测为正的样本中有多少是对的。有两种值：一种是把正类预测为正类（TP）；另一种是把负类预测为正类（FP）。召回率（Recall）是针对原来的样本而言，表示的是样本中的正例有多少被预测正确了。也有两种值：一种是把原来的正类预测成正类（TP）；另一种就是把原来的正类预测为负类（FN）。

F_1为精确率和召回率的调和平均，当两者值较高时，F_1值也较高。

True Positive（真正, TP）：将正类预测为正类数。

True Negative（真负，TN）：将负类预测为负类数。

False Positive（假正, FP）：将负类预测为正类数 →→ 误报 （Type I error）。

False Negative（假负，FN）：将正类预测为负类数 →→ 漏报 （Type II error）。

精确率（Precision）的计算公式为

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

召回率（Recall）的计算公式为

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

两者的调和平均F_1计算公式为

$$F_1 = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{4}$$

4 结束语

基于CNN的图像识别技术属于成像型探测，模拟“人眼”，具有火焰检测和视频监控功能，可从不同角度（诸如火焰的光学特性、跳动频率、形状、变化趋势、烟雾的运动模型、烟雾色调特征、静态形状特性和边缘模糊特征等）

进行综合判断，便于值班人员快速反应，节省时间。

基于CNN图像型火灾探测器具备非常稳定可靠的报警能力，只要报警，几乎可以确认发生了火灾，还能有效避免诸如辐射、电晕、电弧、电动机电枢、*X*射线和核辐射干扰对探测器的影响。

由于具备了火灾探测和监视两大功能，只需要一套系统的成本，可以大幅度地节约资金，同时，在有火情出现后，会自动提示操作人员发生了报警，应该采取相应的处置，节约了一定的人力成本。

经过训练和测试，通过网络途径获取了大量的火灾图像，对不同的火灾图片进行识别和测试。经过测试发现，仍然存在一些细微的不足之处，测试结果显示，此次研究测试的网络精确率（Precision）为60.0%，召回率（Recall）为95.0%，F_1值为72.6%。经过分析，原因是由于图片量不够大，且火灾往往伴随着大量烟雾，因此性能指标不是很理想，但是在图片样本池扩大后，这些指标可以得到优化和加强，问题可以解决。将CNN的图像识别技术应用到火灾自动报警系统中，可以带来一些良好的效益。

参考文献

[1] 高敬鹏，江志烨，赵娜．机器学习[M]．北京：机械工业出版社，2020．

[2] 中国建筑标准设计研究院．《火灾自动报警系统设计规范》图示：14X505-1[M]．北京：中国计划出版社，2014．

[3] 陈学军．基于空间变换的图像型火灾探测器坐标定位方法[J]．仪表技术与传感器，2019(11)：122-126．

[4] 舒日勇．浅谈公路隧道分布式智能图像型火灾探测器[J]．中国交通信化，2018(12)：137-138．

[5] 李征．浅议数字图像处理在火灾报警系统中的应用[J]．科学与信息化，2022(3)：66-68．

[6] 高娟，曲娜，林丽华，等．图像型火灾探测系统设计[J]．河南科技，2018(11)：12-13．

作者简介

马奇（1992—），男，工程师，主要研究方向：建筑电气设计。

冯欧阳（1992—），男，工程师，主要研究方向：建筑电气设计。

（收稿日期：2022-08-11）