优化卷积神经网络在复杂验证码识别中的运用

马　阳[1]　李克昌[2]
1 江西开放大学　江西南昌　330000
2 南昌广播电视台网络传输中心　江西南昌　330000

摘　要：作为一种有效的网络安全防护方法，网络验证代码的使用越来越广泛。验证代码的识别，不但能从验证代码的反认证的角度来设计出更安全、更易于使用的认证代码，同时也能在短时间内检测出认证代码的安全缺陷，从而提高服务器的工作效率和用户的安全。近年来，深度学习技术已被广泛地应用于各个领域，但侧重于深度学习的深度学习，特别是卷积神经网络在各种验证码的识别方面的研究和应用还比较欠缺，本文主要针对数字、字母、汉字等常用的验证码，利用卷积神经网络进行验证图像的识别。
关键词：验证码图片；字符信息；预处理

Optimization of convolutional neural network in complex verification code recognition

Yang Ma[1] Kechang Li[2]
1.Jiangxi Open University Nanchang, Jiangxi Province 330000, China
2.Nanchang Radio and Television Network Transmission Center Nanchang City, Jiangxi Province 330000, China

Abstract: The use of network validation code has become increasingly widespread as an effective method for network security protection. Identifying validation codes not only enables the design of more secure and user-friendly authentication codes from the perspective of anti-certification, but also allows for the detection of security vulnerabilities in authentication codes within a short period of time, thereby improving server efficiency and user security. In recent years, deep learning technology has been widely applied in various fields, but research and application of convolutional neural networks, which focus on deep learning, in the recognition of various types of verification codes are still relatively lacking. This paper focuses on commonly used verification codes such as digits, letters, and Chinese characters, and utilizes convolutional neural networks for recognition of validation images.
Keywords: CAPTCHA image; character information; pre-processing

验证码是一种能将电脑与人区别开来的一种图灵式考试，通常会在用户登录网站、网站注册、查询信息、网站发布等情况下使用。在这种情况下，人们需要的是真实的人类而非电脑程式。图像文字的辨识对于人而言非常容易，但对于电脑程式而言则更困难。因此，有了这个验证代码可以保证一个安全的网络。因此，大部分的网页都会选择用图像进行身份认证，而最近几年，也有了类似的代码，比如 SMS，比如动画代码，比如逻辑判断代码，比如简单的逻辑判断代码，比如数字字和字的组合。有些情况下，验证码的出现也会对使用者的使用造成一定的负面影响，而为了保证安全，我们可以通过对验证码进行身份验证，从而让使用者更加安全、更加人性化，从而为使用者创造一个安全的安全环境。于是，很多人就开始研究如何破译，目前最常用的就是利用光文字辨识（OCR），因为大部分的验证代码都是由数字、字母和汉字组成，因此它的辨识也可以称为文字辨识。首先对图像进行二值化、去噪、细化、缩放、分割、旋转等方法进行图像的预处理。第 2 个步骤是训练模式，现在常用的是 K 近邻、支持向量机（SVM）、SVM（SVM）、“神经网”等。第 3 个步骤是对第二阶段所获得的模型进行预测和辨识图像的扭曲、重叠、斜线、变形等干扰，是一种通用的图像自动识别方法。验证代码的工作过程很简单，比如将用户的问题发送到服务器，与用户的答案进行匹配，如果回答正确，就会被认为是一个人，或者是一个自动的程序，为了增强网站的抵抗能力，防止服务被滥用，保证网络的安全性，发现验证码的设计缺陷，识别出网站的验证码，减少网络信息的威胁。目前有关验证代码识别的研究已经取得了很大的成果。如何对验证码的字符或图像进行形态识别、对图像像素进行处理、对错乱的验证代码进行归类、从验证码中抽取可

读的文字信息等。

一、卷积神经网络

卷积神经网（CNN）是基于传统的神经网络的一种深度学习方法，它包括多个卷积层和终端的全连通层，并包含相关权值和池化层，从而使 CNN 可以有效地利用二维的输入资料。深度学习的诞生与发展并非一蹴而就，它经过了一个漫长而又曲折的发展历程，一开始的神经网络仅仅是通过一种简单的线性加权和法来实现输入与输出之间的转化，不过这个过程中必须要有人工设定的权值，人工的因素越多，就越难获得最佳的结果。然而，感知机的建模仅仅能够解出线性的可分性问题，而无法求解线性的不可解问题。20 世纪八十年代，分布的表示和逆向传递的方法开始出现。而分布式表示理论的关键在于，在真实的环境下，要用多个神经元来表示真实的知识和观念，并且每个神经元都可以参与多种特性的表达，从而增强了模型的表征，使得神经网络在处理线性非分割问题上更加有效。而逆向传递的方法也极大地减少了网络的学习难度，目前为止，逆向传递是目前最常用的学习方法。计算机的运算能力不断提高，云计算和 GPU 的发展，使得运算能力在神经网络领域的应用越来越广泛。

1.1 局部感知

一般人们对外部世界的认识都是由局部到整体的，而影像又与各区域的象素密切相关。卷积神经网络利用稀疏连接技术对空间相关信息进行挖掘，并利用局部感知域对边缘、方向线段等主要视觉特征进行提取。再将该区域的特征信息进行综合，从而获得该区域的整体特征。

1.2 权值共享

一般情况下，在图像中不同部位的统计特征是相似的。这样，所学到的特性可以扩展到其他的图像中。卷积神经网络采用局部链接和权重共享的方法，可以大大降低训练网络的参数数量，提高迭代的速度。

1.3 池化

从原理上讲，对不同的滤光片进行卷积处理，获得多个卷积后的图象特征，而用这种图象特征进行分类法则需要大量的运算。采用非线性下取样技术，即“池”运算，减少了上层网络的运算复杂性，提高了对位移、缩放等转换的鲁棒性，并在一定程度上避免了过度拟合。

二、验证码识别

当前主流的网页大多采用文字为基础的图像验证。一般情况下，基于文字的图像验证码，往往是在一组随机生成的文字中，添加一些像素干扰、形变干扰和色彩干扰。在进行深神经网络的学习时，为了保证模型的精度，必须要有足够的样本来保证。针对数量庞大的图像，我们可以使用程式来自动地生成，而对一个简单的数字编码，则采用一系列的随机字符或字符，在图像中加入一组干涉的象素。比如 Python, Python 中的 PIL （Python Imaging Library）中的图像、图像 Draw、图像 Font、ImageFilter 等多种方式来产生普通的验证代码。当然，也有很多第三方的代码库，比如 python 中的 captcha，它的产生方式非常的简洁。这种算法能够在较快的速度下获得大量的验证代码，从而能够很好地适应大规模的学习。与此同时，验证代码可以在产生的过程中将它们的内容进行录制，并根据它们的内部信息来给它们起一个名字，从而在产生的过程中就会对它们进行标识。

2.1 常见验证码分类

2.1.1 问答验证码

文本验证代码一般是以问答的方式呈现的，比如：向用户提问，让用户回答，给出古诗词上的一句，让使用者把下一句话写出来。由于所有的验证问题和解答都是预先存储在资料库中的，所以攻击者很容易利用爬虫和规则表达式来破解。

2.1.2 静态图验证码

静态验证码是一种应用非常广泛的认证代码。这种验证代码通常需要使用者在验证图像中输入数字或字符，通过扭曲、干扰、变形等方式增加了验证代码的自动识别难度。受限于验证码设计者的经验，其难度也是不同的。

2.1.3 短信邮件验证码

手机的普及，移动互联网的发展，使得手机的使用越来越广泛。伺服器向使用者的移动电话传送验证码，并在指定的时限内，请使用者输入确认的验证码。这样的验证码主要是为了确认自己的行为。

2.2 卷积神经网络验证码识别方法

众所周知，不同的验证码图像具有很大的特性和差异，因此，在进行验证码的识别时，首先要对各种类型的验证码图像进行预处理、字符分割和特征抽取。因此，传统的验证码识别技术更多的是依靠图像预处理、字符分割和特征抽取，而不注重对其进行分类。对于不同的验证码图像，分别进行预处理、降噪、字符分割和特征抽取等步骤，这一工作既复杂，也不利于推广。

优化卷积神经网络在复杂验证码图片识别中的应用方法

人们总是会把前因后果和历史背景放在第一位，而不是从现在开始，我们的决定就会变得更有说服力。也就是说，我们的思想是有记忆力的。但由于传统的神经网络仅能够对现有的数据进行一定程度的加工，从而限制了传统的神经网络应用范围。而递推神经网络则可以很好地解决这个问题。使网路模式像人那样具备储存的能力，目前的输出是以先前的输出与现时的输入来确定的。由于这种性质，它被大量地用于求解某些时序问题。例如，在处理语言、机器翻译、语音识别、视频序列标记等领域，都有很好的应用前景。因为每一列的象素都可以看作是一种时间顺序，因此这一节我们将使用一种基于长期和长期的递归神经网络模型来进行辨识。

三、预处理复杂验证码图片

对复杂的验证码图像进行预处理，剔除与识别字符无关的象素点，如噪声、干扰背景等。在保留了原验证码图像的亮度、色差信息的前提下，扫描二值化图像，扩大和缩小图像连通面积，将中间象素点与相邻象素点的象素值进行对比，得到图像空间与结构要素的交叉操作。得到了连接区域的尺寸，选择了连接区域内的结构象素，并将其作为结构单元的中心。通过对结构象素中的噪声进行抑制，使噪声点向连通域的边界运动，如果所有的结构象素点都能被包含在连通区域内，则可以保持象素点。通过对图像连通区域内的直线干涉线进行检测，以大到小的次序连接域象素，设置直线干扰线中的象素点数目，判断出象素点数目在预定值以上的线为干扰线，并将验证码图像的直线干扰线剔除。选取影像中的任何一个点作为起点，通过对起点附近的象素点进行遍历，对每个象素点进行加权，对验证码中的剪裁线进行筛选。基于该方法，对相邻点进行检索，并将其归入待删除的曲线线组，并依据干扰线曲率的特点，判断出宽线为干扰线，消除了图像曲线的干扰线。

3.1 生成复杂验证码图片特征库

对复杂的验证码进行分割，抽取特征矢量，并建立了特征模板库。初步裁剪验证码图片，从上到下，从左到右依次扫描图像连通域中各列的像素点数，确定图像的上、下、左、右边，再对图像进行倾斜、规范化处理，计算出整体图像的平均象素高度，并根据平均高度设置统一标准的文字高度，将字符调整为相同大小和位置级别。考虑到图像预处理后出现的空白，采用文字之间的空隙，以图像的左上方为中心，画出四条直线，组成一个长方形的方框，将验证码图像中的字符进行分割，并将所要识别的各个特征块进行分割。对每一行和每一列的象素点进行累积，得出验证码影像的投影，以所扫描的象素点为投影值，剔除投射值较小的一行，用横向投射的方法，排除文字以外的一行的干扰，用竖线投射，获得有谷线的投射图，以谷线为分割点，以谷点为切入点，将文字与背景分开，设置投影分割阈值，统计黑字象点，并对累积数值超出阈值的行和列进行初步分割。然后，利用连通域分割法，对图像中的连通域进行标注，并对每个区域的像素进行计算，在保留面积大的情况下，将具有相同标志的像素点视为要被分割的字符，然后由左到右对其进行分割。

3.2 基于优化卷积神经网络识别复杂验证码图片

20 世纪六十年代，生物学家维塞尔曾研究过，当猫咪的大脑中的神经元被投射到电脑前的准确方位时，大脑中的神经细胞会对某些特殊的光线产生最大的反应，而对其它的神经则基本没有任何影响。他们叫这种感觉。由于感觉区只能在一定范围内发挥功能。因此，这可以很好地反映出图像的地方相关性。受到这个理念的启发，福田信介设计了 CNN 的首个实施 CNN 的工程。CNN 的基础架构包括输入层、卷积层、池化层、全连接层和输入层。卷积是一系列的卷积核心，它能够从多个卷积核心中抽取出一些特殊的特征。卷积核心是一种加权（3x3 或 5x5 的矩阵）（例如 2 D 图像）。CNN 的卷积层能够对图像进行特征提取，第一级卷积通常提取出边缘、线条、角落等特征，之后的各个卷积层会提取先前特征的特征，而卷积操作则具备稀疏权重、共享参数、等变表示等特征，从而提高了卷积操作的效率。在池化层中，池化功能利用特定的节点附近的总的统计特性来替代这个地点的网络。池化层是对从卷下层中抽取的特征进行二次抽取，常用的池化方法是最大池，最大池提供了邻近矩形区的最大值，平均池化池提供了邻近矩形区的平均值，池化可以使输入值接近不变，而平移值不变则是在对一个小的平动时，大部分的输出不会变化。在卷积和池中，各神经元与前一层的各神经元相连。在此基础上，将最终一级完全连通的输出数据传输到输出层次，利用软映射逻辑进行划分。由该网络模式所预言的类型在软件最大的输出级中具有最大的权值。在分类中，分类函数的选取、激活函数的选取、网络层数的选取以及卷积核的个数和尺寸的选取都具有一定的意义。通过对图像进行特征提取，并将其输入到最优卷积神经网络中，实现图像的图像信息的识别。在此基础上，采用最优卷积神经网络，对字符的轮廓进行训练，以图像特征矢量为学习样本，使卷积层能够学习到验证码图像的特征矢量，并对其进行循环，重构卷积层，得到字符特征的标记概率分布，然后由变换层产生一系列的特征矢量，将标记概率分布转化为特定的字符，从而消除验证码图像中的空白。通过对错误信号进行动态反馈，消除了神经网络输出的误差。深度学习是一种以学习资料为基础的机器学习方法，它与其他特殊的任务方法不一样。其受功能与构造的启发，即所谓的神经网络。深度学习具有极大的灵活性和广阔的应用范围，它能让我们了解到更多的概念和层次，并能根据较少的抽象概念来处理更多的抽象表示。换句话说，机器学习仅注重于实际问题的求解，并在对大量的数据进行统计、并根据结果进行预测；在对大量的书籍进行分析之后，深入研究的学生可以从中抽取出样本的特性，并让机器“记住”样本的共性，这样，机器就可以拥有自己的“思想”，并根据样本的特性来判断。近年来，深度学习成为了许多人工智能爱好者们的热门话题，而深度学习在自然语言处理、图像识别等领域中也得到了越来越多的重视。

四、结束语

综上所述，本论文利用卷积神经网络对复杂的验证码图像进行了识别，并对其进行了降噪。但是，这种设计仍然有一些缺陷，可以通过采用模板匹配的方法来简化验证码图片的破译过程，并利用更多的机器学习技术来改善其描述的客观性。

参考文献：

[1] 张爱民 . 优化卷积神经网络在复杂验证码图片识别中的应用 [J]. 信息工程大学学报 , 2021(006):022.

[2] 刘欢 , 邵蔚元 , 郭跃飞 . 卷积神经网络在验证码识别上的应用与研究 [J]. 计算机工程与应用 , 2016, 52(18):7.

[3] 周文凯 . 基于卷积神经网络的极验验证码识别系统研究与设计 [D]. 东华大学 .

江西省教育厅科学技术研究项目 GJJ2210602 基于四元数特征提取的验证码识别算法研究