|  |  |  |
| --- | --- | --- |
| 南农大  **专业文献综述** | | |
|  | 题 目: | 卷积神经网络验证码识别技术综述 |
|  | 姓 名: | 邱日 |
|  | 学 院: | 信息科学技术学院 |
|  | 专 业: | 计算机科学与技术 |
|  | 班 级: | 计科151 |
|  | 学 号: | 19215116 |
|  | 指导教师: | 顾兴健 职称: 讲师 |
| 2018 年9月14日 | | |
| 南京农业大学教务处制 | | |

**卷积神经网络验证码识别技术综述**

作者：邱日 指导老师：顾兴健

摘要： 当下验证码在互联网被广泛应用着。使用验证码的主要目的是通过人机交互来减少机器自动化攻击的可能性。本文首先简要介绍了几种不同类型的验证码，并归纳总结国内外验证码识别技术的常用方法。其次重点介绍了卷积神经网络在基于文本的验证码识别领域中的应用和最新研究成果。也指出了当前验证码的不足和未来的发展方向。这对于改进验证码以营造良好的网络环境有一定的指导作用。

关键字：验证码; 识别; 卷积神经网络; 分割

**An Overview on CNN-based CAPTCHA Recognition**

Nanjing Agricultural University, Network Engineering, Jiangsu Nanjing 210095

Author：Ri Qiu Tutor：XingJian Gu

**Abstract:** Currently, verification code is widely used on the Internet. The main purpose of use of the verification code is to reduce the possibility of machine automation attacks through human-computer interaction. This paper first briefly introduces several different types of verification codes, and summarizes the common methods of domestic and foreign verification code recognition technology. Secondly, the application and latest research results of convolutional neural network in the field of text-based verification code identification are introduced. It also points out the shortcomings of the current verification code and the future development direction. This has a certain guiding role in improving the verification code to create a good network environment.

**Key words:** CAPTCHA; recognition; convolutional neural network; segmentation

引言：随着网络的飞速发展，利用网络犯罪的情况也越来越常见。验证码作为规避一些恶意事件的手段得以发展起来。不少网站为了防止用户利用机器人自动注册、登录、灌水，都采用了验证码技术。所谓验证码，就是将一串随机产生的数字或符号，生成一幅图片，图片里加上一些干扰像素。由用户肉眼识别其中的验证码信息，输入表单提交网站验证，验证成功后才能使用某项功能。而近年来计算机视觉领域机器学习前沿的不断突破，特别是卷积神经网络的发展，使得验证码识别准确率提高不少。

1. 验证码的发展概述

验证码是互联网时代区分人类与计算机的一种较为高效的图灵测试手段，验证码（CAPTCHA）是“Completely Automated Public Turing test to tell Computers and Humans Apart”[1]的缩写。验证码现已经广泛应用在各个领域，如网站用户身份的验证、反爬虫场景等。在这些场景中，验证码的作用是区分人与计算机，从而避免可能带来的网络安全问题，如对违规批量注册、信息大规模非法采集、恶意注册和访问、暴力破解密码[2]等恶性事件。所以，验证码可以一定程度上做到维护网络环境。研究如何识别验证码可以有针对性地改进现有验证码和设计新型验证码，从而可以有效提高验证码的安全性，维护网络秩序。

* 1. 验证码产生的背景

相当多的网络应用场景中，特别是与计算机交互的时候，有必要判断访客是计算机还是人类，验证码应运而生。验证码（CAPTCHA）一词是由John Langford、 Nicholas J. Hooper 和 Luis Von Ahn[3]三人于2000年在卡耐基梅隆大学引入。到目前为止，国际上已经有了大量关于如何安全有效地设计验证码的方法，有一些已经过时，一些仍在广泛使用中。它涉及到用来证明是人还是机器的一系列协议的实现，即人类交互式证明Human Interactive Proof (HIP)。验证码可以有效预防一些自动程序伪装成人类用户的恶意活动，如信用卡欺诈、非法访问用户信息等。验证码的可靠性在于：它可以由人类用户快速理解和识别，那些很难完成既定任务的对象就被判定为机器。验证码是图灵测试的一个较好的实现。当人类用户和机器同样面对测试题时，两者都尝试让系统判定他/它为人类。基于他们的回答，验证码系统需要判定哪个回答是语义明确且准确无误的（即人类），哪一个是模棱两可或回答错误（即机器）的。

* 1. 验证码的应用领域

验证码可以被应用在如下领域：

1. 免费电子邮箱服务. 防止机器每分钟自动注册大量的电子邮箱账号；
2. 反搜索引擎爬虫. 防止被搜索引擎索引到；
3. 蠕虫病毒. 防止从恶意电子邮件中感染病毒。
4. 防止字典爆破攻击. 防止黑客字典爆破攻击密码系统;
5. 防治博客“水军”. 比如有效防治微博僵尸粉丝现象；
6. 保护正常的网站注册.
   1. 验证码分类

第一个验证码系统是在 1997 年，由Alta Vista的研究员们提出的。目前，验证码的使用范围非常广，由于网络安全性要求不同，验证码的设计样式也多种多样。

根据验证码设计机制的特点我们可将其分为不同类型。总体上说，这是一种涉及图像处理方法和密码学等领域的综合技术。比较常见的验证码有：基于字母和数字的文本验证码、基于图片信息的验证码，基于视频内容的验证码，基于音频的验证码和基于数学知识运算的验证码等。

目前，各大网站多数使用基于文本的图片验证码。一般而言，基于文本的图片验证码，通常是将一串随机产生的字符生成图片，然后随机加入若干像素干扰、形变干扰以及颜色干扰等等。本文主要探讨图片验证码的识别。

* + 1. 文本验证码

文本验证码是目前网络中使用最频繁的一种验证码，其主要特点是图像中的内容是由英文字母和数字随机组合而成。这种验证码的实现相对容易，并且人类识别通过率高，使用很方便。为了提离验证码的安全性，人们在验证码图像中加入了复杂的背景、干扰线和噪声。有的验证码使用了多种字体，有的验证码字符个数不定，有的验证码采用了空白字符策略，更有的验证码使用了扭曲粘连字符的措施。

* + 1. 图像验证码

在认知科学中有大量的研究表明，人类对于图像的记忆能力要远远优于对文字的 记忆能力。这也就表示基于图像的验证码具有深远的研究意义。

图像验证码是一种内容更为复杂的验证码。图像验证码中的内容不再是简单的英文字符与数字的姐合，而是更加复杂的图形，例如动物、风景、物体等等，从而图像验证码包含更多的信息。使用图像验证码的主要原因是计算机在理解图像内容上存在着＂语义鸿沟＂。图像验证码利用语义鸿沟，能使验证码具有更高的安全性。但同时由此降低了用户体验，使用户在识别验证码时用时过长，识别率低于文本验证码。 图像验证码的验证模式一般是要求用户识别出验证码图像中的物体，然后根据要求点击相应的一个或者多个物体来进行验证。

图像验证码的使用率比文本验证码低，其原因主要是用户体验度低，人类识别通过率低，同时这种验证码还存在着两个缺点。第一，图像验证码中的图像是从图像数据库里选取的，图像数据库里面的图像必须经过人工分类保存，工程量巨大，图像的数量和种类难保证。如果计算机大量收集图像验证码，通过对其进行分类学习便可成功破解图像验证码。如果改进图像数据库，使验证码图像中出现多种物体，那么验证码生成难度将大大提高，但目前这还比较少见。

* + 1. 语音验证码

基于语音的验证码应用范围较为狭窄，现在并不常见，当碰到视觉障碍的用户时，语音验证码是个不错的选择。语音验证码形式多是将一段有效语音中掺杂一些噪音，然后需要用户输入其中听到的有效语音，其存在形式多数是数字或者字母等。另一种现在比较主流的语音验证码是基于通信的，用户需要在听到的语音中获得相关信息并完成测试。该类型验证码容易受到背景中的噪音、通信质量以及录制语音人员口音等影响，所以用户友好度并不好，很容易出现输入错误的状况。

* 1. 国内外对验证码识别研究现状

识别文本验证码的思路，在国内外主要有两种。第一，基于模板匹配的识别算法，这类算法出现在早期，主要是通过提取字符特征点并与模板字符进行匹配来完成验证码的识别，其算法的正确识别率较低。第二，基于机器学习的识别算法，这类算法一般是通过图像预处理，字符分割，字符识别三个步骤进行验证码的识别，其正确识别率较高，因此大多数识别算法都采用这种机器学习的方法进行验证码的识别。但是对于字符站连扭曲的验证码，这种识别算法的流程存在着缺陷，即一旦字符分割错误，字符识别也就会发生错误，并且会影响下一个字符的识别，造成错误的累积。

研究员Suphanne Sivakorn已经能够分别以70.78% 和83.5%的准确率攻克reCAPTCHA和Facebook的验证码[4]，另外Vu Duc Nguyen 也声称他们的算法能够解决所有的六位字符的动态验证码，其准确率从16% 到100%不等 [5]。Goodfellow[6] 、Hong[7] 、Bursztein和Stark都使用了深度学习技术，并且都取得了不错的效果。其中Goodfellow从直接从谷歌街景地图图片上识别预测数字，他们使用了Jeff Dean主导研制的谷歌第一代的深度信念网络，并使用了多个计算机来协同分布工作，以避免单机内存不足的限制。这项技术后来被用于验证码的识别上。他们的预测成功率达到了最高99.8%。Hong是先对验证码图片做旋转和分割出单字符的操作，然后使用了两层全连接层，Bursztein使用了预处理、单字符分割然后是基于KNN的识别技术，之后他使用了不同的CNN来检测来自百度、维基百科、reCAPTCHA和雅虎的验证码。Stark 的场景是在很有限测试数据条件下识别验证码，他们使用了一种叫激活学习的技术来从新的训练数据下得到反馈以提高总体效果。

* 1. 研究的目的及意义

验证码的设计是基于人工智能领域的开放性问题，验证码的破解意味着一个AI 问题得到解决，因此，验证码研究领域以“设计—破解—再设计” 的状态向前发展。总体上说验证码的发展是一个螺旋形上升的过程。验证码运行机制如下：一个专门的服务器负责产生和评估验证码测试，用户使用某个需要验证的网络服务时，服务器提供给用户一个测试，用户做完后将测试结果提交给服务器，服务器根据结果进行评估，决定该用户能否通过测试。测试能被绝大多数人类用户通过，而拒绝几乎所有的机器程序。一般要求人类用户通过率大于 90%， 机器通过率小于0.01%。

1. 验证码识别过程

验证码识别可分为以下几个阶段：图片预处理、语义分割、字符识别。在预处理阶段，验证码图片转化为灰度图像，这之后，为了恢复图像中的原始单词，起干扰作用的线和点被移除了。语义分割阶段图片被裁减以识别其中的每个字符，当字符提取好后，通过传统的图像处理方式或是卷积神经网络的方法将字符与数据库中的进行匹配，若字符被正确识别了，那基于文本的验证码便被攻克了。

* 1. 图片预处理

在验证码识别过程中，传统的验证码图片预处理方法有：图像灰度法、阈值处理算法、迭代法和大律法等等。验证码图片去除噪声的方法主要包括：基于空间域的滤波模型、基于形态学的去噪法以及基于连通域的去噪法。

* + 1. 二值化

通过阈值把灰度图像或者直接将彩色图像处理成黑白图像的过程，称为图像 的二值化。处理后，图像中只有像素值为 0 和 255 的像素点，不再涉及像素的多 级值，该后续的处理过程变得简单，和数据的处理量也小得多。对图像进行二值化处理的关键是阈值的选择与确定。二值化主要包括全局阈值法、局部阈值法和动态阈值法。

* + 1. 去除冗余背景并降噪

网络中很多验证码图像使用了复杂背景。文献[8]己经证明，这种给验证码图像添加背景的策略是不安全的，即使背景图像是随机选择的，仍然有很多算法可将背景去除。虽然验证码图像中添加了背景，但是为了让人能正确识别出验证码字符，验证码字符与背景之间必然存在着差异，这便为去除背景干扰提供了理论依据。去除验证码图像背景干扰的算法有基于颜色直方图[9]的算法，分水岭[10]算法等等。

降噪的主要目的是去除图像中的噪声，降噪方法有很多如：平滑、低通滤波等。平滑降噪。比如平滑降噪具体方法是通过统计每个像素点周围像素值的个数来判断将改点置为和值。如果一个点周围白色点的个数大于某一阈值则将改点置为白色，反之亦然。通过平滑降噪已经可以将剩下的噪声点全部去除了。这里需要注意的是对二值图像进行降噪时应注意强度，当验证码笔画较细时，降噪强度过大可能会破坏验证码本身的信息，这可能会影响到后面的识别效果。

* + 1. 验证码分割

验证码的分割这里即图像分割。它指的是根据灰度、颜色、纹理和形状等特征把图像划分成若干互不交迭的区域。并且划分使这些特征在同一区域内呈现出相似性，而在不同区域间呈现出明显的差异性[11]。基于图像局部特征的分割方法大致可分为四大类。这四大类即基于阈值、基于区域、基于边缘检测和基于人工智能领域的方法等[12]。图片切割还有很多算法如投影法、CFS以及滴水法等。投影法适用于字符垂直方向上没有粘连和重合的情况，CFS能够很好的切割垂直方向有粘连但是没有粘连的字符，水滴法可以分割粘连字符。

* + 1. 字符识别

传统方法该阶段主要利用一些常用算法提取图像特征。经过训练之后，系统将对分割后的验证码内容做出判定。一般是先得到样本特征，进行训练，然后再分类。这里简要介绍两个算法，一是SVM算法，二是模板匹配字符识别算法。

SVM学习问题可以表示为凸优化问题[13]，因此可以利用已知的有效算法发现目标函数的全局最小值。而其他分类方法（如基于规则的分类器和人工神经网络）都采用一种基于贪心学习的策略来搜索假设空间，这种方法一般只能获得局部最优解。SVM应该算是用的较多的分类方法，一般大多适合于二分类问题。

模板匹配字符识别算法是图像识别中的经典算法之一，该算法的核心思想是：通过比较待识别字符图像的字符特征和标准模板的字符特征，计算两者之间的相似性。相似性最大的标准模板的字符即为待识别的字符。该方法首先要建立标准模板库，其中标准模板库中的字符的大小是一样的。然后应将待识别的字符规格化，其大小应该和模板库中的字符一样。最后将待识别的字符和标准模板库中的所有字符进行匹配，计算相似度。模板匹配字符识别算法适用于印刷字体、字体规范的字符等，但是对字符变形、弯曲、字符旋转等情况的抗干扰能力差。

1. 卷积神经网络

在过去的五六年,视觉前沿由于卷积神经网络（CNN）这个新成员的加入而不断飞速发展。卷积神经网络通常也被称为深度学习。深度学习技术其实较早的时候就已经被提出了，但是由于大数据集的缺乏使得它表现不佳。近年来，计算机视觉开始往更大型[14-16]数据集上演进，更大的数据集使得这些老技术被重新定义并且被优化。因此，最近这几年计算机视觉的所有前沿领域都在更快发展。特别是基于深度学习的文本识别获得了不俗的表现，在识别不同种的字符方面几乎要达到完美了。

在计算机视觉领域，传统的验证码识别技术被分为预处理（去噪等）、定位、分割（字符分割）、字符识别等模块。各模块独立存在，这样意味着一旦某个模块出现错误，将导致整个验证码识别系统的识别率降低，而基于卷积神经网络的方法则可以更好地解决该问题。深度神经网络有很多网络层，不同层可以学习到不同等级的特征，整个网络统一训练来完成该系统最终的验证码识别这一目标，不同的神经网络有其独特的使用目的。

* 1. 卷积神经网络介绍

卷积神经网络（Convolutional Neural Network, CNN）是一种前馈神经网络，它的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。一个卷积神经网络主要由一个输入层和一个输出层及许多隐藏层组成。卷积神经网络在图像和语音识别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，卷积神经网络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。

* 1. 网络结构

本节将介绍卷积神经网络中一些常用的网络结构。卷积神经网络是一种建立在传统神经网络基础上的深度学习网络。在结构上，它主要由单个或多个卷积层和池化层（也叫下采样层）组合而成.

* + 1. 卷积层

对图像（不同的数据窗口数据）和滤波矩阵（一组固定的权重：因为每个神经元的多个权重固定，所以又可以看作一个恒定的滤波器filter）做内积（逐个元素相乘再求和）的操作就是所谓的卷积操作，也是卷积神经网络的名字来源。

* + 1. 池化层

在卷积神经网络中，池化层往往跟在卷积层的后面，池化层的作用有两个： 一是降卷积层输出的特征向量的维度。 二是减少过拟合现象。过拟合现象的存在是因为结果对于输入的某些误差过于敏感，通过max-pooling或mean-pooling可以减少噪声。

池化实际上是一种形式的降采样。有多种不同形式的非线性池化函数，而其中“最大池化（Max pooling）”是最为常见的。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，这种机制能够有效的原因在于，在发现一个特征之后，它的精确位置远不及它和其他特征的相对位置的关系重要。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，这在一定程度上也控制了过拟合。通常来说，CNN的卷积层之间都会周期性地插入池化层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[17]，甚至不再使用池化层[18]。

* + 1. 激活函数层

损失函数层（loss layer）用于决定训练过程如何来“惩罚”网络的预测结果和真实结果之间的差异，它通常是网络的最后一层。各种不同的损失函数适用于不同类型的任务。例如，[Softmax](https://zh.wikipedia.org/wiki/Softmax%E5%87%BD%E6%95%B0)交叉熵损失函数常常被用于在K个类别中选出一个，而[Sigmoid](https://zh.wikipedia.org/wiki/%E9%82%8F%E8%BC%AF%E5%87%BD%E6%95%B8)交叉熵损失函数常常用于多个独立的二分类问题。

* + 1. Dropout层

Dropout是指在模型训练时随机让网络某些隐含层节点的权重不工作，不工作的那些节点可以暂时认为不是网络结构的一部分，但是它的权重得保留下来（只是暂时不更新而已），因为下次样本输入时它可能又得工作了。 训练神经网络模型时，如果训练样本较少，为了防止模型过拟合，Dropout可以作为一种技巧供选择。

* + 1. 全连接层

全连接层（fully connected layers，FC）在整个卷积神经网络中起到“分类器”的作用。如果说卷积层、池化层和激活函数层等操作是将原始数据映射到隐层特征空间的话，全连接层则起到将学到的“分布式特征表示”映射到样本标记空间的作用。在实际使用中，全连接层可由卷积操作实现：对前层是全连接的全连接层可以转化为卷积核为1x1的卷积；而前层是卷积层的全连接层可以转化为卷积核为h乘以w的全局卷积，h和w分别为前层卷积结果的高和宽。

* 1. 卷积神经网络在验证码识别中的应用

卷积神经网络识别文本型验证码，其实就是字符的识别，相对于传统的方法，可以减少人工提取特征的环节。比较流行的一种搭建结构是这样, 从下到上的顺序, 首先是输入的图片(image), 经过一层卷积层 (convolution), 然后在用池化(pooling)方式处理卷积的信息。然后在经过一次同样的处理后, 把得到的第二次处理的信息传入两层全连接的神经层 (fully connected),这也是一般的两层神经网络层,最后再接上一个分类器(classifier)进行分类预测.

1. 验证码识别与网络安全

验证码是维护网络安全的重要组成部分。网络在线投票系统、免费邮箱服务等系统中都使用验证码进行认证，同时验证码还能阻止网络爬虫抓取重要页面，防止接收蠕虫邮件和垃圾邮件，防止密码系统遭受字典攻击等等。验证码在计算机安全领域如此重要，但其生成系统却简单高效并且成本低廉，其认证过程拥有良好的交互体验，从计算机程序的角度考虑其识别难度大识别准确率低。总而言之，验证码作为一种安全工具，既具有较高的安全性，又具有良好的用户体验，在网络上的应用必然越来越广。

1. 总结和展望

本文先介绍了验证码的发展历程，介绍了验证码的产生背景和分类，总结了国内外对验证码的研究现状。其次，文章又说明了验证码识别的一般过程，包括预处理、分割和传统的字符识别方法。之后本文简述了近年来卷积神经网络的发展，以及该领域的相关知识。在说明卷积神经网络在验证码识别中的应用后，总结了传统与机器学习方法的异同。最后谈到验证码识别与网络安全的关系，并对全文作了总结并提出对验证码前景的展望。

验证码技术的不断发展体现在：设计与识别两个方面都在不断进步。两个方面虽然看似完全对立，但是其实两者相辅相成，随着攻击手段即验证码识别的提高，新型验证码也将不断被设计出来。研究验证码识别技术有助于我们对现有验证码的有针对性的改进。

参考文献：

[1]周正,文亚飞,鲍文平.基于深度学习的人工智能用于识别破解字符型验证码[J].通信技术,2017,50(11):2572-2576.

[2]汪洋,许映秋,彭艳兵.基于KNN技术的校内网验证码识别[J].计算机与现代化,2017(02):93-97.

[3] Ahn L. von, M. Blum and J. Langford. 2004. Telling Humans and Computer Apart Automatically. Communications of the ACM[J]. 47(2): 57-60.

[4]S. Sivakorn, I. Polakis, and A. D. Keromytis, “I am Robot: (Deep) learning to break semantic image CAPTCHAs,” [J]Proceedings - 2016 IEEE European Symposium on Security and Privacy, EURO S and P 2016, pp. 388–403, 2016.

[5] Nguyen V D, Chow Y W, Susilo W. Breaking an Animated CAPTCHA Scheme[C]// International Conference on Applied Cryptography and Network Security. Springer-Verlag, 2012:12-29.

[6] Goodfellow, Ian J. Multi-digit number recognition from street view imagery using deep convolutional neural networks. [J] arXiv preprint arXiv:1312.6082 (2013).

[7] Hong, Colin. Breaking Microsoft’s CAPTCHA. [J] (2015).

[8] Yan J, Ahmad A S E. Breaking Visual CAPTCHAs with Naive Pattern Recognition Algorithms[C] Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual. IEEE, 2007:279-291.

[9]量化的均匀与非均匀对提取颜色直方图的影响及比对研究[J].于翔,沈美.青海大学学报(自然科学版). 2015(01)

[10]基于改进分水岭算法的运动目标行为理解[J].张惊雷,张云飞.计算机工程与设计. 2015(07)

[11] 韩思奇,王蕾.图像分割的阈值法综述[J].系统工程与电子技术,2002,24(6):91-94,102.

[12] 钱智明.图像分割方法研究[D].北京.国防科学技术大学,2010.

[13] 殷光，陶亮.一种 SVM 验证码识别算法[J].计算机工程与应用，2011，47（18）：188-190

[14] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C] International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.

[15] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: Common Objects in Context[J]. 2014, 8693:740-755.

[16] Russell B C, Torralba A, Murphy K P, et al. LabelMe: A Database and Web-Based Tool for Image Annotation[J]. International Journal of Computer Vision, 2008, 77(1-3):157-173.

[17] Graham B. Fractional Max-Pooling[J]. Eprint Arxiv, 2014.

[18] Springenberg J T, Dosovitskiy A, Brox T, et al. Striving for Simplicity: The All Convolutional Net[J]. Eprint Arxiv, 2014.

**专业文献综述成绩评阅表**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 学院 | 信息科技学院 | 专业 | 计算机科学与技术 | 姓名 | 邱日 |
| 题目 | 卷积神经网络验证码识别技术综述 | | | | |
| 指导教师意见(包括选题是否恰当、文字表达水平、论文的难度和创新性、参考文献质量、格式是否规范等方面，请使用钢笔书写或打印)： | | | | | |
| 论文评定成绩：    指导教师签名：  2018 年 9 月 20 日 | | | | | |