基于CNN的验证码识别技术应用综述

作者：邱日

班级：计科151 学号：19215116

摘要： 当前的验证码在因特网上广泛使用，其目的主要是防止减少网络黑客暴力破解的可能。 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。然后，介绍了CNN的基本原理和最近的研究及应用，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

**Author：**Ri Qiu

Nanjing Agricultural Universit， Jiangsu Nanjing 210095

**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

验证码作为规避一些恶意事件的手段随着网络上不法行为的增多而发展起来。但是，近年来计算机视觉领域机器学习前沿也在不断突破，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“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。

1.1 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。

1.2验证码的应用领域

验证码可以被应用在如下领域：

1）免费电子邮件服务。

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.

1.3验证码分类

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：基于数字字母的验证码、基于图片的验证码、基于用户交互输入的验证码、基于数学常识的验证码等。

基于图片文本的验证码用的较多。通常，基于文本的图像验证码通常从一系列随机生成的字符生成图片，然后随机地添加许多像素干扰，变形干扰，颜色干扰等。 本文主要讨论图像验证码的识别。

1.3.1 文本验证码 基于文本的验证码用的比较广，这种验证码的实现相对容易，人识别通过率高，使用方便。 为了提高验证码的安全性能，人们想办法在验证码中添加干扰的背景或者是干扰线等。一些验证码使用各种字体，一些验证码字符是可变的，一些验证码使用空白字符策略，一些验证码使用措施来扭曲粘合字符。

1.3.2 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。这也意味着基于图像的验证码具有深远的研究意义。

图像验证码是更复杂的验证码。 例如，动物，风景，物体等，使图像验证码包含更多信息。 然而，同时，用户体验减少，使得用户花费太长时间来识别验证码，并且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低于文本验证码的使用率。图像验证码要求数据库存储大量的数据库，需要大量的人工操作。若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。

1.3.3 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。

1.4国际上对验证码领域的研究

国内外识别文本验证码的方法主要有两种。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。其次，基于机器学习识别算法，这种算法一般采用图像预处理，字符分割和字符识别三个步骤来识别验证码。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，字符识别也会产生错误并影响对下一个字符的识别，从而导致错误的累积。

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使用预处理，单字符分割和基于 KNN的识别技术，之后他使用不同的 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。

1.5研究的目的及意义

验证码的设计基于人工智能领域的开放性问题。 因此，验证码研究领域在“设计 - 破解 - 重新设计”的状态中取得了进展。 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：机器自动生成验证码，操作者给出结果， 通常，人类用户通过率大于90％，机器通过率小于0.01％。

2.验证码识别过程

验证码识别时首先要预处理，然后对语义进行分割。分割出单个字符后进行识别。 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。

2.1图片预处理

在验证码识别过程中，传统的验证码图像预处理方法是： 图像灰度法，阈值处理算法，迭代法和宏观法。

2.1.1去除冗余背景并降噪 在二值化之后，还要解决的问题是，网络中的许多验证码图像使用复杂的背景。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。

2.1.2验证码分割 验证码的分割这里即图像分割。它指的是基于诸如灰度，颜色，纹理和形状的特征将图像划分成彼此不重叠的区域。并且划分使得这些特征在相同区域中显示出相似性，并且在不同区域之间显示出明显的差异[11]。基于图像的局部特征的分割方法可大致分为四类。这四个类别基于阈值，基于区域，基于边缘的检测和基于人工智能的方法[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。

2.1.3字符识别 经过训练之后，系统将采用诸如SVM算法[13]等对分割后的验证码内容做出判定。通常，首先获得样本特征，训练样本特征，然后进行分类。

3.卷积神经网络

在过去的五六年中，由于增加了卷积神经网络（CNN）的新成员，视觉前沿迅速增长。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，近年来计算机视觉的所有前沿领域都发展得更快。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。

3.1卷积神经网络介绍

卷积神经网络（CNN）是前馈神经网络，其人工神经元可以响应覆盖区域的一部分中的周围单元。 对于大型图像处理有出色表现。卷积神经网络主要由输入层和输出层以及许多隐藏层组成。

3.2网络结构

本节描述了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基于传统神经网络的深度学习网络。在结构上，它主要由单个或多个卷积层和池化层（也称为下采样层）组成。

3.2.1 卷积层 对于图像（不同的数据窗口数据）和滤波器矩阵（一组固定的权重：因为每个神经元的多个权重是固定的，所以它可以被视为恒定滤波器。也是卷积神经网络的名字来源。

3.2.2池化层 在卷积神经网络中，池化层倾向于遵循卷积层。一个是减少的卷积输出的特征向量的维数。 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。 有许多不同形式的非线性池函数，“最大池化”是最常见的。直观地说，这种机制起作用的原因是在发现一个特征之后，它的确切位置远不如它与其他特征的相对位置重要。 汇集层将不断减小数据空间的大小，因此参数的数量和计算量也会减少，这也会在一定程度上控制过度拟合。 通常，CNN的卷积层被周期性地插入到池化层中，因为池化层太快地减小了数据的大小。 文献中的当前趋势是使用较小的池化过滤器或甚至使用池化层。

3.2.2 损失函数层 损失函数层中各种不同的损失功能可用于不同类型的任务。 例如，Softmax交叉熵损失函数通常用于选择K类别中的一个，而Sigmoid交叉熵损失函数通常用于多个独立的分类问题。

3.2.3 Dropout层 网络中权重可能对最终结果不起作用，但这些权重必须保留（暂时不会更新），因为它可能在下次输入样品时再次起作用。 在训练神经网络模型时，Dropout层可以在训练样本较少的时候有效地防止实验结果过度拟合。在实际使用中，全连接层可以通过卷积操作实现：完全连接到前层的完全连接层可以转换成1x1的卷积核的卷积; 前层是卷积层。

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。这是一种比较流行的建筑结构。在一层卷积之后，以汇集方式处理卷积信息。然后，在相同处理之后，将所获得的第二处理信息发送到两层完全连接的神经层。这也是一般的双层神经网络层，最后分类器用于分类预测。

4.验证码识别与网络安全

验证码在网络安全中的作用主要体现在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。 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。

5.总结和展望

本文首先介绍了验证码的开发过程，介绍了验证码的背景和分类，总结了国内外验证码的研究现状。其次，文章还解释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。本文简要介绍了近年来卷积神经网络的发展及该领域的相关知识。在解释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，这将让我们的公共平台更加安全。

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. Breaking Visual CAPTCHAs with Naive Pattern Recognition Algorithms[ C] Computer Security Applications Conference，2007. ACSAC2007. Twenty- Third Annual. IEEE，2007:279-291.

[9] 于翔，沈美.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].青海大学学报(自然科学版). 2015(01)

[10] 张惊雷，张云飞.基于改进分水岭算法的运动目标行为理解[J].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.