

SealGAN: 基于生成式对抗网络的印章消除研究

李新利¹ 邹昌铭¹ 杨国田¹ 刘禾¹

摘要 发票是财务系统的重要组成部分. 随着计算机视觉和人工智能技术的发展, 出现了各种发票自动识别系统, 但是发票上的印章严重影响了识别准确率. 本文提出了一种用于自动消除发票印章的 SealGAN 网络. SealGAN 网络是基于生成式对抗网络 CycleGAN 的改进, 采用两个独立的分类器来取代原本的判别网络, 从而降低单个分类器的分类要求, 提高分类器的学习性能, 并且结合 ResNet 和 Unet 两种结构构建下采样-精炼-上采样的生成网络, 生成更加清晰的发票图像. 同时提出了基于风格评价和内容评价的综合评价指标对 SealGAN 网络进行性能评价. 实验结果表明, 与 CycleGAN-ResNet 和 CycleGAN-Unet 网络相比较, SealGAN 网络不仅能实现自动消除印章, 而且还能更加清晰地保留印章下的发票内容, 网络性能评价指标较高.

关键词 印章消除, 生成式对抗网络, SealGAN, CycleGAN, 评价指标

引用格式 李新利, 邹昌铭, 杨国田, 刘禾. SealGAN: 基于生成式对抗网络的印章消除研究. 自动化学报, 2021, 47(11): 2614–2622

DOI 10.16383/j.aas.c190459

SealGAN: Research on the Seal Elimination Based on Generative Adversarial Network

LI Xin-Li¹ ZOU Chang-Ming¹ YANG Guo-Tian¹ LIU He¹

Abstract Invoice is an important part of the financial system. With the development of computer vision and artificial intelligence technologies, various automatic invoice identification systems have been developed. However, the seals on the invoice often affect the identification success rate. In this paper, the SealGAN network structure is proposed, which can automatically eliminate the seal of invoice. SealGAN network is an improvement based on generative adversarial network CycleGAN. The two discriminant networks are replaced with two independent classifiers in the SealGAN, which reduces the classification requirement of each classifier, and the learning performance of classifiers can be improved. In addition, it combines the ResNet and UNet structures to construct a downsampling-refining-upsampling generation network, which can generate clearer images of invoice. And the network comprehensive evaluation index, including style evaluation and content evaluation, is proposed to evaluate the performance of network. Experimental results demonstrate that the SealGAN network can not only eliminate the seal automatically, but also retain the invoice content under the seal clearly, and the network performance evaluation index is higher than that of the CycleGAN-ResNet and CycleGAN-Unet.

Key words Seal elimination, generative adversarial networks, SealGAN, CycleGAN, evaluation index

Citation Li Xin-Li, Zou Chang-Ming, Yang Guo-Tian, Liu He. SealGAN: Research on the seal elimination based on generative adversarial network. *Acta Automatica Sinica*, 2021, 47(11): 2614–2622

发票是财务系统的重要组成部分, 发票的识别和处理是财务人员的首要工作. 人工智能已被广泛应用在人脸识别、车辆识别、物体检测等各个领域, 而基于人工智能的发票识别, 可为财务人员节省大量的时间. 目前发票自动识别根据使用场景不同, 可分为移动客户端、云端和本地客户端. 移动端客户端发票识别是通过手机端的离线 SDK, 集成到公司财务系统的 APP 内, 自动检测发票的边线并拍照识别, 提取发票上的关键信息; 云端发票识别是

通过调用云平台的 API 接口, 将拍照或者扫描的发票图像传送到云平台上进行识别; 而本地客户端识别是识别软件与扫描仪硬件配合使用进行识别. 三种方式本质都是基于识别软件进行的. 由于发票种类繁多, 格式不固定, 有的发票存在手写的內容, 且不同人手写字的风格不一, 因此基于传统数字图像处理的识别软件, 相应识别准确率较低. 随着卷积神经网络的兴起, 包括表格线定位、手写文字识别等问题得到很好的解决. 文献 [1] 采用卷积神经网络实现增值税发票批量识别, 通过对发票图像进行采集、预处理、字符切割, 基于 CNN 神经网络进行字符识别, 同时融入人工修改, 提高识别率. 但是发票上的印章对发票识别准确率依旧具有一定影响, 且印章是每张发票必不可少的, 因此如何消除印章

收稿日期 2019-06-18 录用日期 2020-03-12

Manuscript received June 18, 2019; accepted March 12, 2020

本文责任编辑 金连文

Recommended by Associate Editor JIN Lian-Wen

1. 华北电力大学控制与计算机工程学院 北京 102206

1. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206

也是现在 OCR (Optical character recognition) 领域的研究热点. 目前大多数研究者对于发票印章的研究主要是印章识别与验证^[2-4]. 针对发票印章的去除问题, 若通过印章定位, 而后直接挖除印章则会丢失印章下的发票内容; 若采用过滤颜色的方式来消除, 会导致发票中与印章颜色相近的文字变得模糊不清, 而且不同发票的印章颜色不同. 文献^[5]提出一种基于彩色图像的二次分割算法来去除印章, 但是对于发票的要求较高, 需要保证印章的颜色为红色或者蓝色, 票据不能出现明显的扭曲和变形, 字体还要求是黑色或灰色, 其实用性较差. 文献^[6]针对印章覆盖、发票折痕等干扰因素影响发票号码分割的问题, 采用基于数字结构特征的识别算法, 通过对噪声粘连区域进行修复, 解决干扰因素对数字分割的影响, 实现发票号码识别. 但是如何判断噪声区域最小连通面积的阈值是算法关键, 当发票数量增多, 印章颜色深度和位置不统一时, 消除印章干扰则很困难.

各种图像处理软件也可用在印章消除上, 如 PhotoShop, 但需要基于软件进行手动处理, 大量发票的印章消除则会新增大量工作量, 反倒得不偿失. 因此研究如何自动消除发票印章, 对提高发票识别准确率具有重要意义. 生成式对抗网络 GAN (Generative adversarial network) 是 Goodfellow 等在 2014 年提出的一种生成式网络^[7]. 在 GAN 提出之前, 常见的生成式网络有最大似然估计法、近似法^[8]、马尔科夫链法^[9]等. 这些方法的基本步骤是, 先对样本数据进行分布假设, 然后利用数据样本直接对假设分布的参数进行拟合, 这会导致学习到的生成式模型受到限制^[10]. 而 GAN 不同于上述的生成网络, 该方法采用对抗学习方式, 先通过判别器学习生成分布和真实分布之间的差异, 再驱使生成器去缩小差异. GAN 相比于其他的生成网络, 对数据的分布不做显性的限制, 从而避免了人工设计网络分布^[11]. GAN 目前常用于实现图像的风格迁移以及超分辨率图像的生成. 本文基于生成对抗网络提出用于消除印章的 SealGAN 网络. SealGAN 借鉴了 CycleGAN 网络的循环结构, 采用两个独立的分类器来取代判别网络, 并针对印章的特点去设计生成网络, 实现发票图像的带印章到不带印章的风格迁移, 从而达到消除发票印章的效果.

1 生成式对抗网络和 CycleGAN 网络

1.1 生成式对抗网络

生成式对抗网络结构如图 1 所示, 包括一个生成网络 (G) 和一个判别网络 (D). 生成网络用于生成目标图像, 而判别网络用于判断图像是生成的虚

假图像还是真实目标图像. 在训练时, 生成网络会在损失函数的牵引下让生成的图像尽可能地逼近目标图像, 判别网络则将生成图像和真实目标图像区分开, 由此构成对抗博弈的训练方式. 通过竞争方式, 生成网络生成的图像会越来越真实, 而判别网络的判别能力也会越来越强^[7].

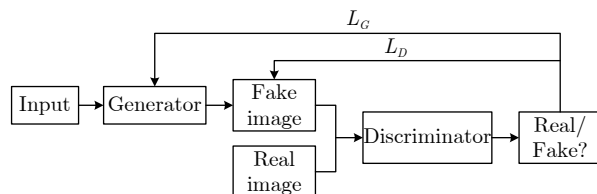


图 1 生成式对抗网络结构

Fig. 1 Generative adversarial networks structure

图 1 中 L_G 是更新生成网络 G 的损失, L_D 是更新判别网络 D 的损失, 如下式所示:

$$L_G = -\frac{1}{m} \sum_i \log(D(G(z^i))) \quad (1)$$

$$L_D = -\frac{1}{m} \sum_i [\log(D(x^i)) + \log(1 - D(G(z^i)))] \quad (2)$$

式中 z^i 表示的是第 i 个输入, x^i 表示对应的真实图像. 生成网络和判别网络的损失函数均为交叉熵损失, 只是判别网络的目标是让真实图像的判别结果趋近于 1, 生成图像结果趋近于 0, 而生成网络则希望生成图像的判别结果趋近于 1^[7].

1.2 CycleGAN 网络

生成式对抗网络的概念提出后, 便成为学术界研究的热点. 在此基础上, 一些改进网络结构相继被提出, 如引入卷积神经网络的 DCGAN^[12], 用于生成超分辨率图像的 SRGAN、LapGAN^[13-14], 用于风格迁移的 pix2pix、cVAE-GAN^[15-16]. 但是大部分生成式对抗网络都要求待处理图像和目标图像一一匹配. 针对带印章的发票, 则很难获取印章图像与其对应的无章图像. 2017 年, 伯克利大学的研究团队在 ICCV 上提出了 CycleGAN 网络, 其最大特点是不需要待处理图像和目标图像一一匹配, 只需要两个分布不同的数据集即可实现数据集之间的相互风格迁移.

CycleGAN 的基本思想是采用两个基础的生成式对抗网络, 采用循环的方式训练, 其网络结构如图 2 所示^[17].

假设有两个不同分布的数据集 A 与 B , CycleGAN 中包含两个基础 GAN 网络, 一个负责把 A

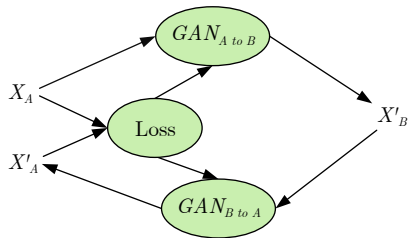


图 2 CycleGAN 的网络结构

Fig.2 GycleGAN networks structure

类数据转化成 B 类数据, 记作 $GAN_{A \text{ to } B}$, 另一个负责把 B 类数据转化成 A 类数据, 记作 $GAN_{B \text{ to } A}$. 设原始的 A 类图像为 X_A , 通过 $GAN_{A \text{ to } B}$, 可以得到原始图像对应的 B 类风格图像 X'_B . 如果采用传统图像对模式, 则存在 X_A 对应的真实图像 X_B , 这时只需将 X'_B 和 X_B 做均方差之类的损失计算,

通过反向传播更新生成网络参数即可驱使生成式网络生成内容相匹配的图像. 当无法获取图像对时, CycleGAN 则把生成的 X'_B 输入 $GAN_{B \text{ to } A}$, 生成相应的 A 类图像 X'_A , 换言之就是试图把生成的 X'_B 还原成原图像 X_A , 此时计算 X_A 和 X'_A 之间的损失即可反向传播去更新两个生成对抗网络. 这种循环的训练模式巧妙地实现了不存在图像对时的两类数据集之间相互转化.

2 基于 CycleGAN 改进的 SealGAN 网络

2.1 SealGAN 网络结构

SealGAN 网络结构如图 3 所示, $GEN_{A \text{ to } B}$ 和 $GEN_{B \text{ to } A}$ 是生成式对抗网络中抽取出来的生成网

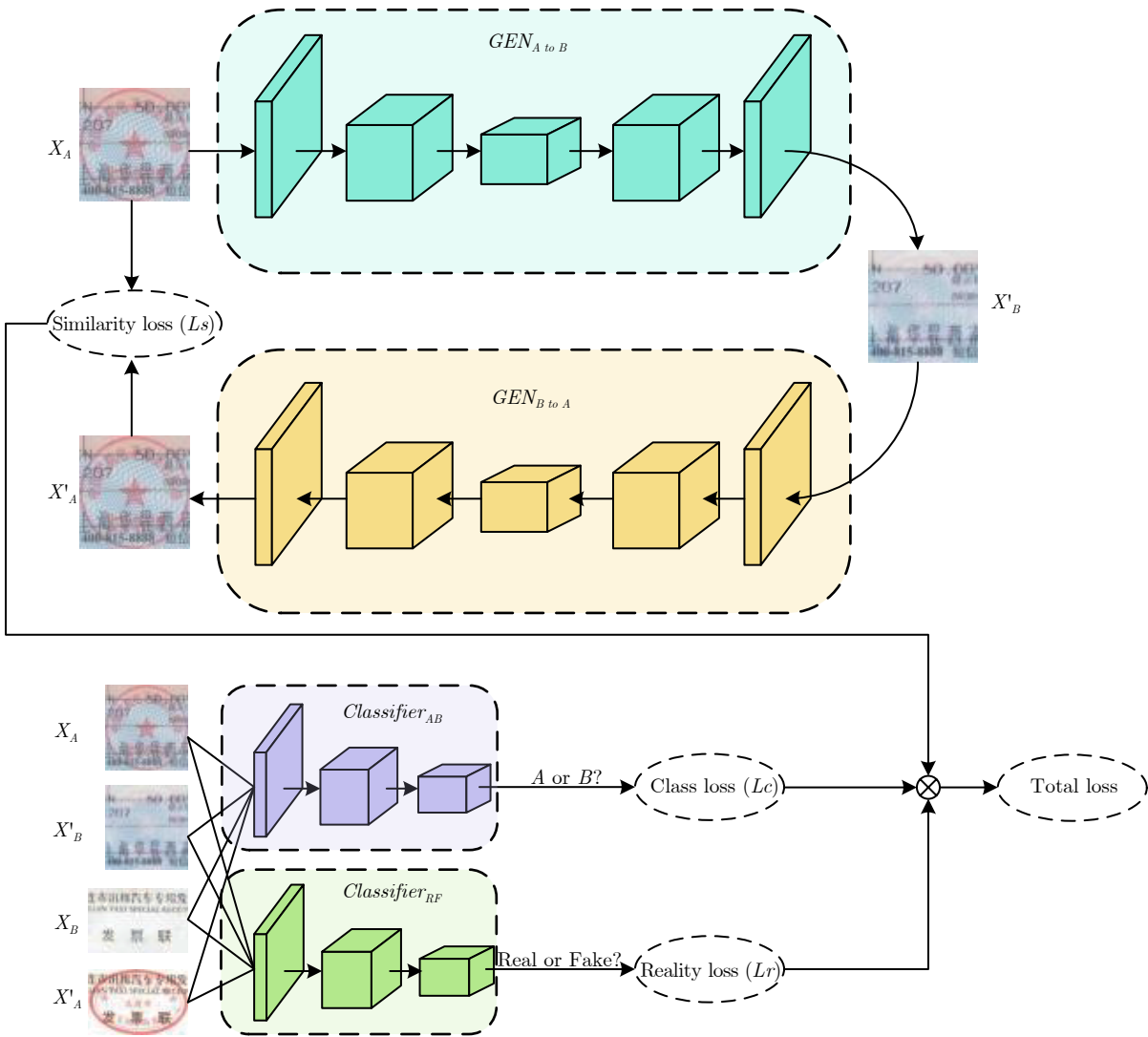


图 3 SealGAN 网络结构

Fig.3 SealGAN networks structure

络, 采用循环训练的方式. 与 CycleGAN 网络相比, 本文所提出的 SealGAN 中, 每个生成网络并没有配备相应的专属判别网络, 而是采用原始图像和所有生成图像单独训练了两个分类器, 其中 $Classifier_{AB}$ 用于区分 A 类和 B 类数据, $Classifier_{RF}$ 用于区分真实图像和生成图像, 然后将两个分类器所计算出的损失加入到总损失中来更新整体网络的参数.

在 CycleGAN 网络中, 采用 $GAN_{A \rightarrow B}$ 和 $GAN_{B \rightarrow A}$ 两个基础的生成式对抗网络进行循环训练, 且有各自的判别网络. 以 $GAN_{A \rightarrow B}$ 为例, 它的判别网络既需要判断输入图像是属于 A 类还是 B 类, 还需要判断输入图像是生成图像还是原始真实图像. 其实, 生成对抗网络中的判别网络是一个分类器, 对于分类器而言, 分类要求越高则分类效果越差, 而且两个判别网络虽然相互独立, 但是分类要求相似, 采用的训练样本交集较大. 因此本文所提出的 SealGAN 网络采用两个独立的分类器取代两个判别网络, 一个分类器用于区分 A 类和 B 类数据, 另一个分类器用于区分生成图像和真实图像. 相当于用两个独立的分类器来取代原本的判别网络, 网络整体参数数量没有太大变化, 但是对于单个分类器其分类要求降低, 分类效果更好.

2.2 SealGAN 网络中的生成网络

CycleGAN 中有两种生成网络结构, 残差网络 ResNet 结构和 UNet 结构. ResNet 结构用于解决层数较深的网络难以训练的问题. 在 ResNet 结构提出之前, 研究人员在提升网络性能上往往是单纯的叠加网络层数. 尽管网络层数的增加能够提高网络的特征容量, 理论上能更好地对样本进行拟合. 但是实际训练中, 当网络层数过大时, 参数的训练

难度也大幅提升, 过分引入的卷积层和非线性激活不仅无法提升网络性能, 而且还难以维持浅层网络所达到的效果, 从而出现了模型退化问题^[18]. ResNet 结构的核心是采用了跨层连接思想, 引入了捷径连接的概念^[19]. 残差网络结构如图 4 所示, 由于捷径连接的存在, 在训练过程中, 拟合的不再是目标分布, 而是目标分布和上一层网络输出之间的残差. 当浅层网络能够达到较好效果时, 残差趋近于零. 尽管在拟合残差时也引入了卷积和非线性激活, 但是要想拟合零只需要把所有参数都趋向零即可, 网络训练易于实现. 因此, ResNet 结构的引入能够在加深网络层数, 提高网络拟合能力的同时, 也能保证达到浅层网络的效果, 从而消除网络退化问题. 图 4 中 $F(x)$ 为待拟合的残差函数, $H(x)$ 为待拟合的完整目标函数.

UNet 结构常用于图像分割领域, 其网络结构如图 5 所示. UNet 结构包括下采样和上采样的过

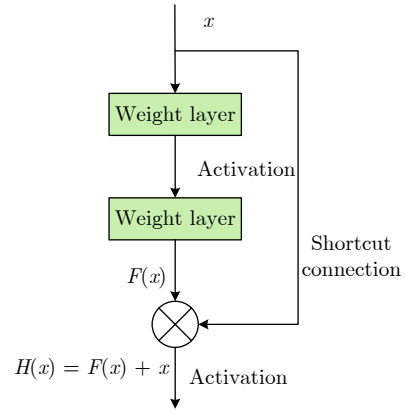


图 4 残差网络结构

Fig.4 Residual networks structure

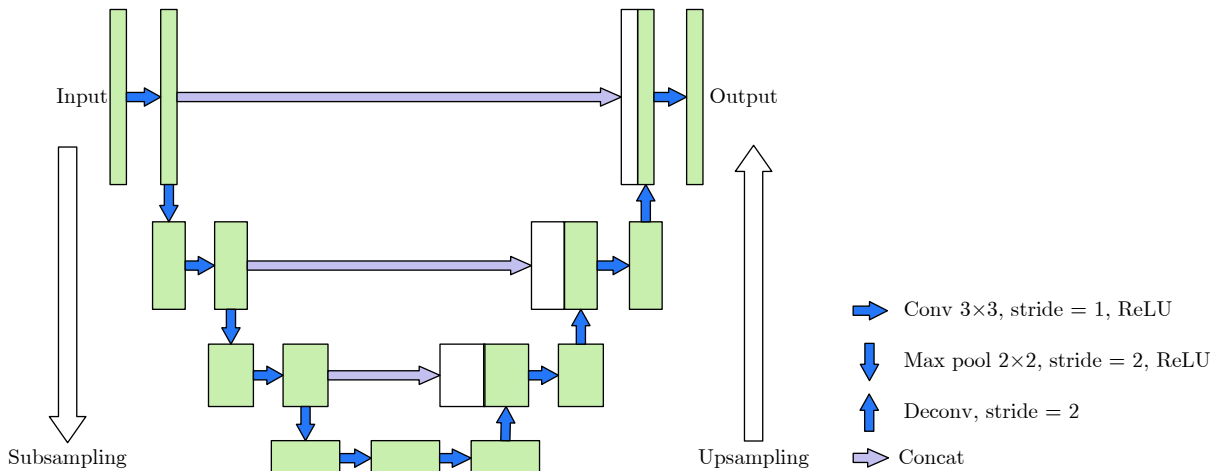


图 5 UNet 网络结构示意图

Fig.5 Schematic diagram of UNet structure

程. 在上采样过程中, 先把下采样和上采样中尺度相同的层拼接在一起, 再进行卷积和反卷积操作. 上采样层和下采样层之间的连接是 UNet 结构的亮点, 这些连接使得在上采样过程中可保留了更多的原始图像的纹理信息^[20].

本文提出的 SealGAN 网络中, 生成网络结合 ResNet 和 UNet 两种结构, 搭建了一个下采样-精炼-上采样的网络结构, 如图 6 所示. 由于在消除印章的同时, 也希望能尽可能地保留印章下的内容, 而 UNet 结构的优势正是能更多地保留原始图像纹理信息, 因此在 SealGAN 网络的生成网络中, 上下采样部分采用了 UNet 结构, 而精炼部分采用了 ResNet 结构以保证网络的效果.

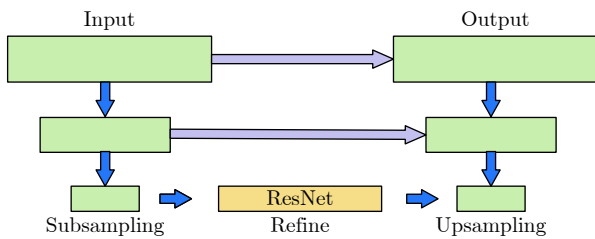


图 6 SealGAN 生成网络结构示意图

Fig. 6 Schematic diagram of SealGAN generative networks structure

3 实验结果与讨论

3.1 网络参数设置

输入图像的尺寸统一缩放为 256×256 , 并把图像像素归一化到 $[-1, 1]$ 区间. 为了保证消除印章后能够尽可能保留印章下的内容, 下采样通道数应多于上采样通道数, 以保证在上采样时能够有足够多的原始图像信息. 多次实验发现, 下采样的通道数与上采样的通道数的比例设置为 1.5 : 1 时效果最佳. 在生成网络中, 除了输出层, 其余层的激活函数均采用 LeakyReLU.

$$\sigma_{\text{LeakyReLU}}(x) = \begin{cases} x, & \text{if } x > 0 \\ 0.2x, & \text{else} \end{cases} \quad (3)$$

输出层采用 tanh 激活函数, 保证输出结果在 $[-1, 1]$ 区间内.

$$\sigma_{\text{tanh}}(x) = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4)$$

分类器网络前四层的激活函数采用 LeakyReLU, 最后一层没有激活. 生成网络以及分类器的相关参数如表 1 所示.

3.2 损失函数以及权重

损失函数包括两个分类器的损失 $Loss_{AB}$ 、 $Loss_{RF}$ 和一个循环回来的图像之间的像素损失 $Loss_G$. 设 z 为原始图像, G 为生成网络, D 是分类器, A 为带印章图像, B 对应的无印章图像. 原始的生成式对抗网络所使用的损失均为交叉熵损失, 在最近的研究中发现, 采用交叉熵损失存在梯度消失和模型坍塌的风险^[21]. 相比于交叉熵损失, 最小二乘损失效果更佳^[22]. 因此, SealGAN 采用了最小二乘损失替换了交叉熵损失. 除此以外, 在计算生成图像和原始图像的内容损失时, 也不再使用常见的 L_2 损失, 而是采用 L_1 损失, 因为 L_2 损失驱动下的生成图像更为模糊. 在训练分类器时, 损失函数如式 (5)、(6) 所示. 当训练生成网络时, 损失函数如式 (7)、(8)、(9) 所示. 其中, $Loss_{AB}$ 和 $Loss_{RF}$ 的权重为 1, $Loss_G$ 的权重为 10 时, 生成效果最佳.

$$Loss_{AB} = \sum \{(D(z_A))^2 + (D(z_B) - 1)^2\} \quad (5)$$

$$Loss_{RF} = \sum \{(D(z_A))^2 + (D(z_B))^2 + (D(G(z_A)) - 1)^2 + (D(G(z_B)) - 1)^2\} \quad (6)$$

$$Loss_{AB} = \sum \{(D(G(z_A)) - 1)^2 + (D(G(z_B)))^2\} \quad (7)$$

表 1 生成网络和分类器参数表

Table 1 Parameters of the generative network and classifier

生成网络			分类器
下采样	精炼	上采样	
7×7 conv, 96	Residual_block(3×3, 384) ×7	4×4 deconv, 256, ×2	4×4 conv, 64, /2
4×4 conv, 192, /2		4×4 deconv, 256, ×2	4×4 conv, 128, /2
4×4 conv, 384, /2		4×4 deconv, 256, ×2	4×4 conv, 256, /2
4×4 conv, 384, /2		4×4 deconv, 256, ×2	4×4 conv, 512, /2
4×4 conv, 384, /2		4×4 deconv, 128, ×2	4×4 conv, 1
4×4 conv, 384, /2		4×4 deconv, 64, ×2	
4×4 conv, 384, /2		7×7 conv, 3	

$$Loss_{RF} = \sum \{(D(G(z_A)))^2 + (D(G(z_B)))^2\} \quad (8)$$

$$Loss_G = \sum \{|z_A - G_{BtoA}(G_{AtoB}(z_A))| + |z_B - G_{AtoB}(G_{BtoA}(z_B))|\} \quad (9)$$

3.3 网络性能综合评价

为了更好地评价网络的性能, 本文提出针对印章消除的综合评价指标 Evaluation-score (ES), 包含风格评价指标 Class-score (CS_1) 和内容评价指标 Content-score (CS_2), 如下式所示.

$$ES = \alpha \times CS_1 + \beta \times CS_2 \quad (10)$$

式中的 α 和 β 是超参数, 可基于需求重要性进行调整, 本文实验中 α 和 β 均为 1.

风格评价指标需要搭建一个分类网络, 采用的训练样本是真实的发票样本, 包含带印章和不带印章的样本. 本论文所搭建的分类网络借鉴 GoogleNet 网络结构^[23], 损失函数采用余弦损失函数, 目标是让带印章样本的输出趋向于 0, 不带印章样本的输出趋向于 1. 采用余弦损失函数能够更好地增大类间距离, 公式如式 (11) 所示^[24]. 将生成的图像通过训练好的分类网络, 所得到的结果即为 CS_1 , 得分越高, 则消除印章的效果越好.

$$L_{cos} = \frac{1}{N} \sum_i -\log \frac{e^{s(\cos(\theta_{i,i})-m)}}{e^{s(\cos(\theta_{i,i})-m)} + \sum_{j \neq i} e^{s(\cos(\theta_{i,j})-m)}} \quad (11)$$

内容评价指标则是利用生成图像和原带章图像之间的均方根误差计算的, 均方根误差越小, 意味着保留的内容越多. 定义的内容评价指标如下:

$$CS_2 = \frac{1}{\log \left(1 + \sqrt{\frac{\sum (x_i - y_i)^2}{N}} \right)} \quad (12)$$

3.4 结果与讨论

训练图像一共有六千余张发票图, 其中带印章样本和不带印章样本的比例约为 1 : 1, 印章样本包含了出租车票印章、增值税发票印章、高速公路收费票据印章、定额发票印章以及各类企业单位加盖公章. 在训练网络过程中, 为选择最佳的训练集和测试集比例, 对 CycleGAN-ResNet 网络和 CycleGAN-UNet 网络, 以及本文所提的 SealGAN 网络等三种网络, 采用不同的训练集和测试集划分比例, 包括 7 : 3、8 : 2 和 9 : 1 三种, 进行对比实验, 并基于网络性能综合评价指标 ES 进行评价. 三种网络

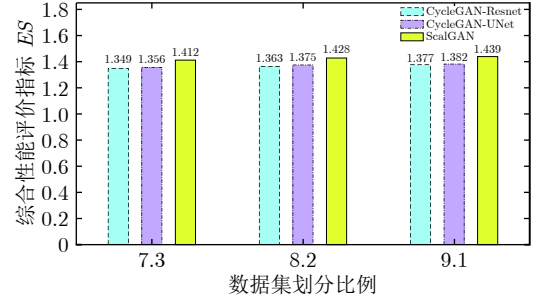


图 7 三种网络在不同数据集合划分比例下的性能指标
Fig.7 Performance indices of three networks under different division proportion of data set

不同数据集合划分比例的综合评价指标如图 7 所示.

由图 7 可知, 随训练集和测试集比例增大, 网络综合性能评价指标也在提升. 为了让每种网络效果达到最佳, 因此在论文实验中采用训练集和测试集比例为 9 : 1 分布进行实验对比.

本文分别使用了文献 [5] 的二次分割方法、采用 ResNet 和 UNet 生成网络的 CycleGAN 网络, 以及 SealGAN 网络, 进行发票印章消除实验. 训练模型时, 针对三种常见的优化器 Momentum、RMS-Prop 和 Adam 进行实验对比, 最终选择稳定性最佳的 Adam 优化器. 每种网络均训练了 20 个 Epoch, Batch Size 均为 4. 学习率的初始值是 0.0002, 每训练一个 Epoch, 学习率下降 10 %. 在 CycleGAN 模型中, 针对生成器和判别器不同的损失权重比, 包括 1 : 1, 5 : 1, 10 : 1 和 15 : 1, 通过网络综合评价指标 ES 进行测试对比, 可知 10:1 为最佳损失权重比. 基于上述的参数设定进行实验, 图 8 是四种方法对发票印章消除的效果.

由图中可知, 采用二次分割的方法效果最差, 由于方法的核心思想是对颜色进行阈值分割来获取印章部分再进行修改, 所以当发票颜色复杂时, 对印章部分的定位不完整, 导致印章消除不全, 而且对于获取到的印章部分采用直接填充背景方法, 对发票内容造成影响. 采用 ResNet 生成网络的 CycleGAN, 生成的结果比较模糊, 色彩与原图相比, 色差巨大; 除此以外, 内容丢失也很严重. 采用 UNet 生成网络的 CycleGAN, 由于 UNet 在进行上采样时融合了下采样层的特征, 保证了上采样过程中保留更多的原始信息, 因此色彩问题得到解决, 但是依旧有明显的内容丢失, 特别是当内容的颜色与印章的颜色相近时. 基于 SealGAN 网络的印章消除, 采用了下采样-精炼-上采样的生成结构, 而且上采样过程也融入了下采样层的原始信息, 因此色差问题基本不存在, 生成的结果相比于前两个网络的实验结果要清晰很多. 除此以外, 由于 SealGAN 并不



图 8 基于二次分割、CycleGAN-ResNet、CycleGAN-UNet 和 SealGAN 的印章消除对比
Fig.8 Comparsion of effect of the seal elimination based on re-segmentation, CycleGAN-ResNet, CycleGAN-UNet and SealGAN

是简单地用两个基础 GAN 网络搭建的循环结构,而是用两个专门的分类器来替代判别器,其中一个分类器专门区分真实样本和生成样本,另一个专门区分带章样本和不带章样本. 这种替换使得网络能够更好地学习到哪些部分属于印章,哪些部分不属于印章,因此 SealGAN 网络可更好地保留非章内容.

三种不同网络消除发票印章的评价如表 2 所示,可知 SealGAN 具有最高得分.

表 2 三种网络性能评价指标

Table 2 Performance evaluation indices of three kinds of network

网络类型	CS_1	CS_2	ES
二次分割	0.340	0.991	1.331
CycleGAN-ResNet	0.678	0.699	1.377
CycleGAN-UNet	0.703	0.679	1.382
SealGAN	0.699	0.740	1.439

4 结语

本文针对发票上消除印章的问题,分析了 CycleGAN 网络的优缺点,并基于此进行相应改进,提出

SealGAN 网络结构,设计了包括风格评价和内容评价的网络综合评价指标. SealGAN 网络用两个独立的分类器来取代原本的判别网络,可降低单个分类器的分类要求,提高网络的学习效果,同时生成网络结合 ResNet 和 Unet 两种结构,实现下采样-精炼-上采样的过程,能够生成更加清晰的图像. 基于发票印章消除实验,与原始的 CycleGAN 进行对比可看出,本文所提出的 SealGAN 无论是在模糊程度、生成色彩等图像细节上,还是内容的保留程度上都要明显优于原始的 CycleGAN 网络,且综合评价指标高于 CycleGAN 网络,实验结果表明了所提出的 SealGAN 网络的有效性. 未来可进一步提高非章内容的保留效果,尤其是当印章与印章下的发票内容颜色很相近时,并可针对发票复印件进行印章消除研究.

References

1 Hu Ze-Feng, Zhang Xue-Xi, Li Xian-Zhao. Batch invoice recognition system based on convolutional neural network (CNN). *Industrial Control Computer*, 2019, **32**(5): 104-105, 107 (胡泽枫, 张学习, 黎贤钊. 基于卷积神经网络的批量发票识别系统研究. *工业控制计算机*, 2019, **32**(5): 104-105, 107)

2 Ou-Yang Huan, Fan Da-Zhao, Li Dong-Zi. Invoice seal identific-

- ation based on multi-feature fusion decision. *Computer Engineering and Design*, 2018, **39**(9): 2842–2847
(欧阳欢, 范大昭, 李东子. 多特征融合决策的发票印章识别. 计算机工程与设计, 2018, **39**(9): 2842–2847)
- 3 Chung W H, Wu M E, Ueng Y L, Su Y H. Seal imprint verification via feature analysis and classifications. *Future Generation Computer Systems*, 2019, **101**: 458–466
- 4 Xie R N, Mao W H, Shi G Z. Electronic invoice authenticity verifying scheme based on signature recognition. In: Proceedings of International Conference on Advanced Algorithms and Control Engineering (ICAACE), 2019, 1–8
- 5 Ji Jing-Jing, Lou Zhen. Filtering of color seal on bank notes based on re-segmentation. *Modern Electronics Technique*, 2014, **37**(22): 5–9
(季婧婧, 娄震. 基于二次分割的银行票据彩色印章的滤除. 现代电子技术, 2014, **37**(22): 5–9)
- 6 Cui Wen-Cheng, Ren Lei, Liu Yang, Chao Hong. Invoice Number recognition algorithm based on numerical structure characteristics. *Journal of Data Acquisition and Processing*, 2017, **32**(1): 119–125
(崔文成, 任磊, 刘阳, 邵虹. 基于数字结构特征的发票号码识别算法. 数据采集与处理, 2017, **32**(1): 119–125)
- 7 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, WardeFarley D, Ozair S, et al. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems, 2014, 2672–2680
- 8 Rezende D J, Mohamed S, Wierstra D. Stochastic back-propagation and approximate inference in deep generative models. In: Proceedings of International Conference on Machine Learning (ICML), Beijing, China, JMLR:W&CP. 2014.
- 9 Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, **18**(7): 1527–1554
- 10 Wang Kun-Feng, Gou Chao, Duan Yan-Jie, Lin Yi-Lun, Zheng Xin-Hu, Wang Fei-Yue. Generative adversarial networks: The state of the art and beyond. *Acta Automatica Sinica*, 2017, **43**(3): 321–332
(王坤峰, 苟超, 段艳杰, 林懿伦, 郑心湖, 王飞跃. 生成式对抗网络GAN的研究进展与展望. 自动化学报, 2017, **43**(3): 321–332)
- 11 Lin Yi-Lun, Dai Xing-Yuan, Li Li, Wang Xiao, Wang Fei-Yue. The new frontier of AI research: generative adversarial networks. *Acta Automatica Sinica*, 2018, **44**(5): 775–792
(林懿伦, 戴星原, 李力, 王晓, 王飞跃. 人工智能研究的新前线: 生成式对抗网络. 自动化学报, 2018, **44**(5): 775–792)
- 12 Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. In: Proceedings of International Conference on Learning Representations (ICLR), 2015. San Diego, CA.
- 13 Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, et al. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, 105–114. Honolulu, HI, USA, Institute of Electrical and Electronics Engineers.
- 14 Denton E, Chintala S, Szlam A, Fergus R. Deep generative image models using a Laplacian pyramid of adversarial networks. In: Proceedings of the 28th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, Neural information processing systems foundation, 2015, 1486–1494
- 15 Isola P, Zhu J Y, Zhou T, Efros A A. Image-to-image translation with conditional adversarial networks. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, Institute of Electrical and Electronics Engineers. 2017, 5967–5976
- 16 Bao J M, Chen D, Wen F, Li H Q, Hua G. CVAE-GAN: Fine-grained image generation through asymmetric training. In: Proceedings of IEEE International Conference on Computer Vision (ICCV), Venice, Italy, Institute of Electrical and Electronics Engineers. 2017, 2764–2773
- 17 Zhu J Y, Park T, Isola P, Efros A A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of IEEE International Conference on Computer Vision (ICCV), Venice, Italy, Institute of Electrical and Electronics Engineers. 2017, 2242–2251
- 18 He K, Sun J. Convolutional neural networks at constrained time cost. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, Institute of Electrical and Electronics Engineers. 2015, 5353–5360
- 19 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, Institute of Electrical and Electronics Engineers. 2016, 770–778
- 20 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Proceedings of 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, Springer International Publishing Switzerland. 2015, 234–241
- 21 Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial network. In: Proceedings of the International Conference on Machine Learning, Sydney, Australia. 2017, 298–321
- 22 Mao X D, Li Q, Xie H R, Lau, R Y K, Wang Z. Least squares generative adversarial networks. In: Proceedings of IEEE International Conference on Computer Vision (ICCV), Venice, Italy, Institute of Electrical and Electronics Engineers. 2017, 2813–282
- 23 Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, Institute of Electrical and Electronics Engineers. 2015, 1–9
- 24 Wang H, Wang Y T, Zhou Z, Ji X, Gong D H, Zhou J C, Li Z F, Liu W. CosFace: Large margin cosine loss for deep face recognition. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, USA, Institute of Electrical and Electronics Engineers. 2018, 5265–5274



李新利 华北电力大学控制与计算机工程学院副教授. 主要研究方向为模式识别与智能系统, 图像处理, 燃烧过程检测技术. 本文通信作者.

E-mail: lixinli@ncepu.edu.cn

(**LI Xin-Li** Associate professor at the School of Control and Computer Engineering, North China Electric Power University. Her research interest covers pattern recognition and intelligent system, image processing, combustion process detection technology. Corresponding author of this paper.)



邹昌铭 华北电力大学控制与计算机工程学院硕士研究生. 主要研究方向为深度学习, 图像处理.

E-mail: 1172227195@ncepu.edu.cn

(**ZOU Chang-Ming** Master student at the School of Control and Computer Engineering, North China Electric Power University. His research interest covers deep learning and image processing.)



杨国田 华北电力大学控制与计算机工程学院教授. 主要研究方向为智能机器人, 计算机视觉, 火力发电精细化燃烧与优化控制.

E-mail: ygt@ncepu.edu.cn

(**YANG Guo-Tian** Professor at the School of Control and Computer Engineering, North China Electric Power University. His research interest covers intelligent robot, computer vision, fine combustion and optimal control of thermal power plant.)



刘 禾 华北电力大学控制与计算机工程学院教授. 主要研究方向为图像处理, 计算机视觉, 模式识别.

E-mail: lh@ncepu.edu.cn

(**LIU He** Professor at the School of Control and Computer Engineering, North China Electric Power University. His research interest covers image processing, computer vision, pattern recognition.)