

6-23-2021

Research of Super-resolution Processing of Invoice Image Based on Generative Adversarial Network

Xinli Li

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

Changming Zou

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

Guotian Yang

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

Liu He

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

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Research of Super-resolution Processing of Invoice Image Based on Generative Adversarial Network

Abstract

Abstract: Automatic identification of invoices can effectively improve financial efficiency. But low-resolution invoice image reduces the accuracy of automatic identification, an *ESRGAN (Encoder Super-resolution Generative Adversarial Network)* network for super-resolution processing of invoice images is proposed. The *ESRGAN* network is based on a conditional generative adversarial network. An auxiliary encoder is designed to guide the network to generate a more realistic super-resolution image. Based on the actual invoice image, the *ESRGAN* network and the conventional image processing, *SRCNN (Super-resolution Convolutional Neural Networks)* network and *SRGAN (Super-resolution Generative Adversarial Network)* network. The model is evaluated through two evaluation indicators of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). The experimental results show that the images processed based on *ESRGAN* super-resolution are better on visual effects and evaluation indicators.

Keywords

invoice image, super-resolution, generative adversarial networks, *ESRGAN*, evaluation indicator

Recommended Citation

Li Xinli, Zou Changming, Yang Guotian, Liu He. Research of Super-resolution Processing of Invoice Image Based on Generative Adversarial Network[J]. *Journal of System Simulation*, 2021, 33(6): 1307-1314.

基于生成式对抗网络的发票图像超分辨率研究

李新利, 邹昌铭, 杨国田, 刘禾

(华北电力大学 控制与计算机工程学院, 北京 102206)

摘要: 发票自动识别可有效提高财务工作效率。为避免低分辨率的发票图像影响自动识别的准确性, 提出了一种用于对发票图像进行超分辨率处理的 *ESRGAN (Encoder Super-resolution Generative Adversarial Network)* 网络。*ESRGAN* 网络是基于带条件的生成式对抗网络, 设计了辅助编码器, 引导网络生成更加真实的超分辨率图像。基于实际发票图像, 将 *ESRGAN* 网络与常规图像处理、*SRCNN (Super-resolution Convolutional Neural Networks)* 网络和 *SRGAN (Super-resolution Generative Adversarial Network)* 网络进行对比实验, 并通过峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似性(Structural Similarity, SSIM)评价指标进行模型评价。实验结果表明基于 *ESRGAN* 超分辨率处理的图像在视觉效果和评价指标上均具有良好的效果。

关键词: 发票图像; 超分辨率; 生成式对抗网络; *ESRGAN*; 评价指标

中图分类号: TP391.4 文献标志码: A 文章编号: 1004-731X (2021) 06-1307-08

DOI: 10.16182/j.issn1004731x.joss.20-0095

Research of Super-resolution Processing of Invoice Image Based on Generative Adversarial Network

Li Xinli, Zou Changming, Yang Guotian, Liu He

(School of Control and Computer Engineering, North China Electric Power University, Beijing, 102206, China)

Abstract: Automatic identification of invoices can effectively improve financial efficiency. But low-resolution invoice image reduces the accuracy of automatic identification, an *ESRGAN (Encoder Super-resolution Generative Adversarial Network)* network for super-resolution processing of invoice images is proposed. The *ESRGAN* network is based on a conditional generative adversarial network. An auxiliary encoder is designed to guide the network to generate a more realistic super-resolution image. Based on the actual invoice image, the *ESRGAN* network and the conventional image processing, *SRCNN (Super-resolution Convolutional Neural Networks)* network and *SRGAN (Super-resolution Generative Adversarial Network)* network. The model is evaluated through two evaluation indicators of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). The experimental results show that the images processed based on *ESRGAN* super-resolution are better on visual effects and evaluation indicators.

Keywords: invoice image; super-resolution; generative adversarial networks; *ESRGAN*; evaluation indicator

引言

发票是财务系统中的重要组成部分, 基于人工智能的发票识别, 可提高财务人员工作效率。随着

硬件设备以及互联网技术的高速发展, 发票处理更多是利用拍照、扫描的方法生成数字图像进行提交。在财务处理流程中, 提交的发票数字图像需要

收稿日期: 2020-02-28 修回日期: 2020-05-02

第一作者: 李新利(1973-), 女, 博士, 副教授, 研究方向为模式识别、图像处理、燃烧过程检测技术。E-mail: lixinli@ncepu.edu.cn

进行多次传输,其中难免会被反复压缩。除此以外,有些发票图像中可能会包含多张发票,尤其是小面积的出租车发票,在进行发票内容识别时,往往出现发票图像分辨率过低,导致发票内容识别错误或漏识别,进而增加财务人员工作量。因此对发票图像进行超分辨率处理,对后续发票内容自动识别具有重要意义。

对图像进行超分辨率处理的常规方法是对低分辨率图像进行插值操作,生成一张相应的高分辨率图像,然后再对模糊的高分辨率图像进行锐化处理,使得图像变得清晰。在图像锐化处理上,常用经典锐化算子和图像进行卷积操作,锐化算子包括拉普拉斯算子、Sobel 算子等^[1]。这种方法的核心是利用图像中局部区域的梯度信息来凸显图像的边缘,从而让图像看起来更加的清晰。但人为设置卷积核的算法缺乏普适性,提取的特征有限,而且图像处理方法感受野狭小,只根据局部像素信息进行超分辨率处理,无法考虑稍远距离的像素信息。在深度卷积神经网络兴起后,科研人员开始研究如何利用深度卷积神经网络来实现图像的超分辨率处理。2014 年提出了 SRCNN(Super-resolution Convolutional Neural Networks)网络,利用深度卷积神经网络对芯片图像进行超分辨率处理^[2-3]。深度卷积神经网络中包含大量的卷积核,每个卷积核参数均由反向传播更新得到,而不是人为设定,可保证提取特征的丰富度。除此以外,深度卷积神经网络会进行多层的卷积操作,保证了充足的感受野。

GAN (Generative Adversarial Network) 是 Goodfellow 等在 2014 年提出的一种生成式网络^[4]。GAN 采用了对抗学习的方式,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 对数据分布不做显性的限制,从而避免了人工设计网络分布的需要^[5-8]。GAN 目前常用于实现图像的风格迁移以及超分辨图像的生成。2016 年 Twitter Cortex 提出了 SRGAN (Super-resolution Generative Adversarial Network)模型,用于将低分辨率图像转成高分辨率图像^[9];2017 年加

利福尼亚大学提出了 pix2pix 模型^[10],Jun-Yan Zhu 等在同年提出了 CycleGAN 模型^[11],这 2 个模型都是基于生成式对抗网络来实现图像的风格迁移,常用于语义分割;2018 年提出能够实现从一个风格到多个风格迁移的 StarGAN 和 AttGAN^[12-13]以及 Attentive GAN 网络用于从单幅图像中去除雨滴^[14]等。

本文基于生成对抗网络提出用于发票图像超分辨率处理的 ESRGAN(Encoder Super-resolution Generative Adversarial Network)网络。ESRGAN 网络是基于带条件的生成式对抗网络,该网络设计了辅助编码器,让相匹配的低分辨率图像和高分辨率图像所得到的编码距离最小,反之则编码距离最大。利用对低分辨率图像的编码结果作为引导条件,引导网络生成更加真实的超分辨率结果。基于实际发票图像,采用 ESRGAN 网络进行超分辨率处理,并与常规图像超分辨率处理算法、SRCNN 网络和 SRGAN 网络进行效果对比。

1 基于带条件生成式对抗网络的超分辨率网络—ESRGAN

1.1 基于生成式对抗网络的超分辨率网络—SRGAN

引入深度卷积神经网络的 SRCNN 网络并不是最佳的超分辨率处理方法,它的不足源自于其损失函数,即生成的超分辨率图像和真实高分辨率图像的像素均方差损失。在现实中,每一张低分辨率图像所对应的高分辨率图像其实是有无穷多种可能的。网络在运行过程中,会计算出每个位置上不同分辨率的可能性概率。由于网络参数更新是由每一轮损失驱动的,网络自然会保守地将每个位置的像素生成为所有像素的可能性加权和,从而导致生成的图像和真实图像相比会模糊一些。

生成式对抗网络可以改善这个情况。生成式对抗网络的核心思想是用生成图像和真实图像来训练一个判别器,判别器只需要判断输入的图像是否

为真实图像, 而判别器的结果又可以用来更新生成器的参数, 让生成器生成更加真实的图像。在生成式对抗网络中, 判别器相当于一个可以训练的损失函数, 而不再局限于用点对点的像素损失。这样生成的图像可能像素的均方差损失会变大, 但效果会更加清晰真实。SRGAN 网络核心思想即是利用生成式对抗网络来进行超分辨率处理^[9]。

1.2 带条件的生成式对抗网络—cGAN

在引入生成式对抗网络后, 虽然生成的图像变得更加清晰, 看起来更加真实。但是每一张低分辨率图像都对应着无穷多个高分辨率图像的可能性。SRGAN 为了保证图像视觉效果真实性而趋向于其中某一个可能地高分辨率图像, 尽管生成的结果在视觉效果上十分真实, 但是内容可能和目标内容不符合, 因此 SRGAN 生成的结果往往会出现一些比较突兀的内容。

2017 年提出带条件的生成式对抗网络(cGAN)^[15], 其核心思想是在生成器和判别器的输入中加入引导条件, 引导条件会引导生成器生成特定的图像。

原始 GAN 的训练目标:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}} [\log D(x)] + E_{z \sim p_z} [\log(1 - D(G(z)))] \quad (1)$$

加入条件后, GAN 的训练目标变为

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}} [\log D(x|y)] + E_{z \sim p_z} [\log(1 - D(G(z|y)))] \quad (2)$$

式中: y 表示加入的引导条件。

1.3 基于带条件生成式对抗网络的超分辨率网络—ESRGAN

针对 SRGAN 网络在进行图像超分辨率处理上会生成一些比较突兀不合理的内容, 是因为一个低分辨率图像会对应着很多高分辨率图像, 原始的生成式对抗网络在生成超分辨率图像时, 会选择众多高分辨率图像可能性的其中一种, 而这种可能性往往不完全和现实相符, 与现实不符的部分会显得十

分突兀。带条件的生成式对抗网络 cGAN 思想能有效缓解这个问题。在 cGAN 的思想中, 可以使用一个表征目标图像的条件加入到输入中以引导生成式对抗网络的生成结果。

本文借鉴了 cGAN 的思想, 改进了 SRGAN 超分辨率网络, 提出了一种全新的超分辨率网络 ESRGAN。在借鉴 cGAN 的思想时, 存在一个明显的难点, 即表征目标高分辨率图像的条件应该如何选取。在超分辨率图像处理中, 为了普适性, 目标高分辨率图像应该是无穷多的, 而且各张高分辨率图像并没有种类之分, 如何选择表征条件十分重要。

对于超分辨率网络, 应当在网络训练完成后, 只需要输入一张低分辨率图像即可生成相应的高分辨率图像。因此, 只能在低分辨率图像上提取相应的引导特征。针对网络训练集中, 每张低分辨率图像都有相应的高分辨率图像, 本文提出的 ESRGAN 网络设计了一个辅助编码器, 用于对所有图像进行编码处理, 包括低分辨率图像和高分辨率图像, 目标是让相匹配的低分辨率图像和高分辨率图像所得到的编码距离最近, 不相匹配的低分辨率图像和高分辨率图像所得到的编码距离最远。当训练好辅助编码器后, 让低分辨率图像通过辅助编码器, 获取相应的编码, 把得到的编码作为引导条件用于引导生成式对抗网络生成超分辨率图像。基于 ESRGAN 网络的超分辨率图像处理流程如图 1 所示。

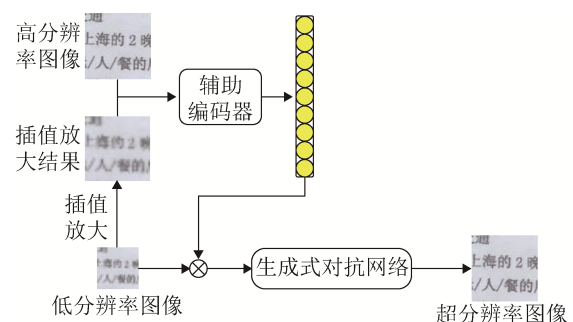


图 1 基于 ESRGAN 网络的超分辨率图像处理流程
Fig. 1 Super-resolution image processing flow based on ESRGAN network

辅助编码器主要包括卷积层和全链接层，如图 2 所示，全链接层输出即为相应的引导条件。在使用辅助编码器之前需要对其进行独立训练。首先把低分辨率图像插值放大，然后把放大后的图像以及所有高分辨率图像作为训练样本，对辅助编码器进行分类训练。把低分辨率放大样本和相应的高分辨率样本归为同一类。在进行分类训练时，需要在辅助编码器的结尾增加一层全连接层以保证正常训练，倒数第二层全链接层输出为引导条件。

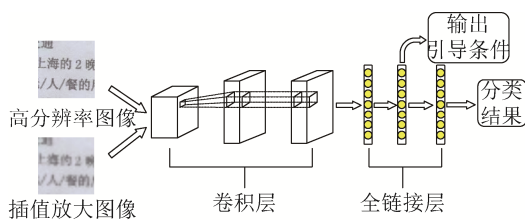


图 2 辅助编码器结构
Fig. 2 Auxiliary encoder structure

在训练辅助编码器时，采用的损失函数是余弦损失，有利于拉近相匹配的低分辨率图像和高分辨率图像的编码结果，拉远不相匹配的低分辨率图像和高分辨率图像的编码距离。余弦损失函数为^[16]

$$L_{\cos} = \frac{1}{N} \sum_i -\log \frac{e^{s(\cos \theta_{i,i-m})}}{e^{s(\cos \theta_{i,i-m})} + \sum_{j \neq i} e^{s(\cos \theta_{i,j})}} \quad (3)$$

式中： $\cos \theta_{i,j}$ 为第 i 个低分辨率图像的编码结果和第 j 个高分辨率图像的编码结果的余弦角度。

当辅助编码器训练完成后，无论是训练生成对抗网络还是最终生成超分辨率图像，均不再需要原始高分辨率图像，只需要将低分辨率图像插值放大后通过训练好的辅助编码器获取相应的引导条件即可。

在引导条件与低分辨率图像融合时，采用的是串联拼接，其融合过程如图 3 所示。原始低分辨率图像为 3 通道图像，本文所设计辅助编码器输出的引导条件为 128 维，因此在进行特征融合后，生成 131 维通道的特征矩阵。

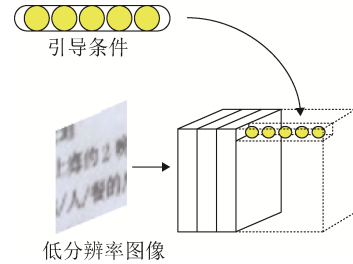


图 3 引导条件与低分辨率图像融合示意图
Fig. 3 Schematic diagram of guidance conditions and low resolution image fusion

2 基于 ESRGAN 网络的发票图像超分辨率处理

2.1 网络的参数设置

基于本文所提的 ESRGAN 网络，进行发票图像超分辨率处理实验。实验所采用的数据集均裁剪自真实的发票图像，裁剪位置随机决定，裁剪尺寸为 128×128 。裁剪出来的图像作为高分辨率图像，相应低分辨率图像制作是先采用高斯滤波对裁剪图像进行模糊处理，再进行尺寸收缩操作，以模拟发票图像在传输过程中被反复缩放的情况，且低分辨率图像长宽为原来的 $1/4$ 。图像在输入网络之前，将像素归一化至 $[-1, 1]$ 区间，以便输入像素的均值为零，有利于加速模型的训练以及训练的稳定性。同时，由于生成器输出层的激活函数为 Tanh ，将像素归一化至 $[-1, 1]$ 也便于损失的计算。在生成器中，除了输出层采用 Tanh 激活函数，其余层均采用 ReLU 激活函数。在判别器与辅助编码器中，最后一层均不采用激活函数，其余层采用 Leaky ReLU 激活函数。ESRGAN 网络结构的参数如表 1 所示。

2.2 损失函数及其权重

ESRGAN 网络中包含 4 个损失函数，分别是生成器损失函数 $Loss_g$ 、判别器损失函数 $Loss_d$ 、辅助编码器损失函数 $Loss_e$ 以及超分辨率图像和原高清图像之间的像素损失 $Loss_p$ 。假设低分辨率图像为 a ，相应的高分辨率图像为 A ，4 个损失函数为

$$Loss_g = \sum D(G(a))^2 \quad (4)$$

$$Loss_d = \sum \{(D(G(a)) - 1)^2 + (D(A))^2\} \quad (5)$$

$$Loss_e = \sum (E(a) - E(A))^2 \quad (6)$$

$$Loss_p = \sum |G(a) - A| \quad (7)$$

经过多次实验, 最终采用的最佳权重如下: 像素损失的权重为 10, 其余损失的权重为 1。

表 1 ESRGAN 网络结构的具体参数
Tab. 1 Parameters of the ESRGAN network

生成器	判别器	辅助编码器
3×3, conv, 64, ReLU	3×3, conv, 64, Leaky ReLU	3×3, conv, 64, Leaky ReLU
3×3, conv, 64, ReLU	3×3, conv, 64, /2, Leaky ReLU	3×3, conv, 64, /2, Leaky ReLU
3×3, conv, 64, ReLU	3×3, conv, 128, Leaky ReLU	3×3, conv, 128, Leaky ReLU
Res_Block(3×3,64, ReLU)×4	3×3, conv, 128, /2, Leaky ReLU	3×3, conv, 128, /2, Leaky ReLU
3×3, conv, 64, ReLU	3×3, conv, 256, Leaky ReLU	Dense, 1024, Leaky ReLU
3×3, conv, 256, ReLU	3×3, conv, 256, /2, Leaky ReLU	Dense, 128
3×3, deconv, 128, ×2, ReLU	3×3, conv, 512, Leaky ReLU	
3×3, conv, 64, ReLU	3×3, conv, 512, /2, Leaky ReLU	
3×3, deconv, 32, ×2, ReLU	Dense, 1024, Leaky ReLU	
3×3, conv, 3, Tanh	Dense, 1	

2.3 网络性能评价指标

对于不同的图像超分辨率处理方法, 不能仅凭视觉效果进行模型评价, 应该采用合理的、统一的评价指标来衡量每种模型的好坏。在图像超分辨率的研究中, 常用峰值信噪比评价指标(PSNR)和结构相似性评价指标(SSIM)进行模型评价^[17-18]。

峰值信噪比评价指标 PSNR 公式为

$$PSNR = 10 \lg \left(\frac{MAX^2}{MSE} \right) \quad (8)$$

式中: MAX 为每个像素可能的最大像素值, 此处是 255; MSE 为图像与图像之间的对应像素的均方误差。 $PSNR$ 指标越大表示生成图像效果越好。

$$MSE = \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M (f_{ij} - f'_{ij})^2 \quad (9)$$

结构相似性评价指标 SSIM 公式为

$$SSIM(x, y) = l(x, y)c(x, y)s(x, y) \quad (10)$$

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (11)$$

$$c(x, y) = \frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (12)$$

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \quad (13)$$

式中: μ_x 和 μ_y 为图像的像素均值; σ_x 和 σ_y 为图像像素的标准差; σ_{xy} 为图像间的协方差; c_1 , c_2 和 c_3 为 3 个小数值常数, 用以避免分母为 0 的情况。 $l(x, y)$ 表示图像间的亮度比较, $c(x, y)$ 表示图像间的对比度比较, $s(x, y)$ 表示图像间的结构比较。 $SSIM$ 指标越大表示生成图像效果越好。

2.4 实验结果与讨论

实验数据集裁剪自真实的发票样本, 包含出租车发票、增值税发票以及高速公路收费发票等。所有高低分辨率图像数据集大小约为 6 万张, 为了保证测试结果的可靠性, 其中 70% 作为训练集, 30% 作为测试集。在本实验中, 除了采用本文所提出的 ESRGAN 网络进行图像超分辨率处理外, 还用了基于常规图像超分辨处理方法(双线性插值放大+Sobel 算子、双线性插值放大+拉普拉斯算子)、SRCNN 网络、SRGAN 网络进行对比。

在训练 SRCNN 和 SRGAN 2 个网络时, 网络结构均复现文献[6,13]中的结构。其中 SRGAN 网络中, 像素损失与对抗损失的权重比为 1 000: 1^[9]。训练次数与本文提出的 ESRGAN 网络相同, 均为 20 轮, 在优化器的选择上采用稳定性最佳的 Adam 优化器, Batch size 均为 16, 学习率初始值为 0.002, 每训练一轮, 学习率降低至上一轮的 90%。实验所生成的图像效果如图 4, 5 所示。

由实验结果可以看出常规图像超分辨率处理的效果与相应的高分辨率图像视觉差距较大, 采用 Sobel 算子进行锐化处理后, 强行突出了图像里的边缘, 视觉效果比较突兀。而采用拉普拉斯算子进行锐化尽管看着不突兀, 但是锐化程度很弱, 没有明显的提升。在引入深度卷积神经网络后, SRCNN

整体效果比常规图像处理要好很多,但是仍然比较模糊。从图 5 可以看到 SRCNN 的生成结果中,有些字显示是一片糊状,这与 SRCNN 中人为设定的像素均方差损失有关。基于原始生成式对抗网络的 SRGAN 超分辨率网络,很好地缓解了 SRCNN 生成结果模糊的情况。尽管生成的字体轮廓和线条变得清晰,但是生成字的线条有一些是弯曲的,与真实情况不符,这也是生成式对抗网络自身的特点,能

很好地还原图像真实性,但无法保证生成图像的内容真实性。本文所提出的 ESRGAN 网络保持了 SRGAN 的优势,能够生成较为清晰图像。此外,ESRGAN 还引入了 cGAN 的思想,采用一个辅助编码器来生成相应的编码向量,该向量用以引导生成式对抗网络去生成尽可能符合真实内容的图像,很好地缓解了生成图像的内容真实性问题,由图 5 可知其视觉效果最佳。

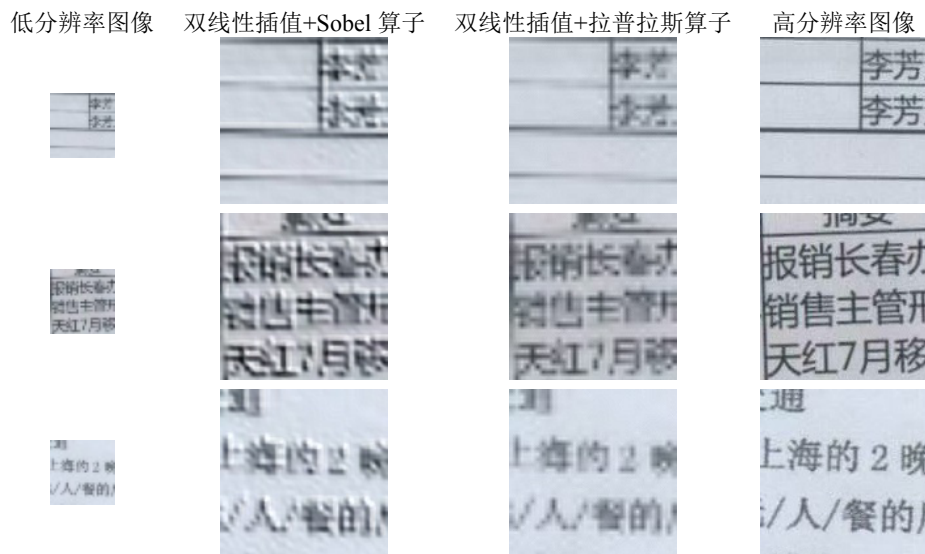


图 4 基于双线性插值+Sobel 算子和双线性插值+拉普拉斯算子的图像超分辨率处理结果

Fig. 4 Results of super-resolution image processing based on bilinear interpolation+Sobel operator and bilinear interpolation + Laplace operator

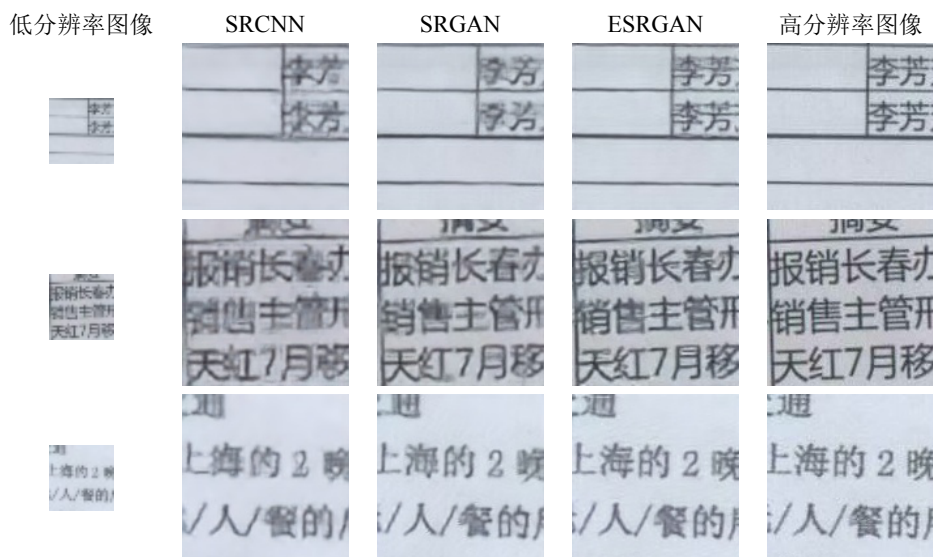


图 5 基于 SRCNN, SRGAN 和 ESRGAN 网络的图像超分辨率处理结果

Fig. 5 Results of super-resolution image processing based on SRCNN, SRGAN and ESRGAN networks

5 种图像超分辨率处理算法的评价指标如表 2 所示。从表 2 中可以看出本论文提出的 ESRGAN 网络无论是 PSNR 评价指标还是 SSIM 评价指标均优于其他方法。

表 2 不同图像超分辨率处理的评价指标
Tab. 2 Evaluation indicators of different image super-resolution processing algorithm

超分辨率方法	PSNR	SSIM
双线性插值+Sobel 算子	32.008 4	0.628 1
双线性插值+拉普拉斯算子	33.111 3	0.812 8
SRCNN	33.651 3	0.865 5
SRGAN	34.326 8	0.908 9
ESRGAN	35.000 2	0.945 8

3 结论

针对发票图像超分辨处理, 本文在分析了深度卷积神经网络的超分辨率模型 SRCNN、基于生成式对抗网络的超分辨率模型 SRGAN, 以及带条件的生成式对抗网络 cGAN 的算法及特点, 提出一种全新的超分辨率模型 ESRGAN, 并设计辅助编码器, 用于对所有图像进行编码处理, 使得相匹配的低分辨率图像和高分辨率图像所得到的编码距离最近, 不相匹配的低分辨率图像和高分辨率图像所得到的编码距离最远, 将得到的编码作为引导条件用于引导生成式对抗网络生成超分辨率图像。基于实际发票图像, 采用五种超分辨率处理算法, 包括双线性插值+Sobel 算子、双线性插值+拉普拉斯算子、SRCNN 网络、SRGAN 网络和本文所提出的 ESRGAN 网络, 进行发票图像超分辨率处理, 采用峰值信噪比(PSNR)和结构相似性(SSIM)2 个评价指标进行模型评价。实验结果表明基于 ESRGAN 超分辨率处理的图像无论是在视觉上还是在评价指标上, 均具有良好效果。

参考文献:

- [1] Rafael C G, Richard E W. 数字图像处理[M]. 3 版. 北京: 电子工业出版社, 2014: 97-101.
Rafael C G, Richard E W. Digital Image Processing [M]. 3rd ed. Beijing: Publishing House of Electronic Industry, 2014: 97-101.
- [2] Chao D, Chen C L, He K M, et al. Learning a Deep Convolutional Network for Image Super-Resolution[C]// European Conference on Computer Vision (ECCV). Switzerland: Springer International Publishing, 2014: 184-199.
- [3] 范明明, 池源, 张铭津, 等. 基于深度学习的芯片图像超分辨率重建[J]. 模式识别与人工智能, 2019, 32(4): 353-360.
Fan Mingming, Chi Yuan, Zhang Mingjin, et al. Super-resolution Reconstruction of Chip Images Based on Deep Learning[J]. Pattern Recognition and Artificial Intelligence, 2019, 32(4): 353-360.
- [4] Goodfellow I J, Poutget-Abadie J, Mirza M, et al. Generative Adversarial Nets[C]// The 27th International Conference on Neural Information Processing Systems. Montreal. CANADA: NIPS, 2014: 2672-2680.
- [5] Rezende D J, Mohamed S, Wierstra D. Stochastic Back-propagation and Approximate Inference in Deep Generative Models[C]// International Conference on Machine Learning (ICML). Beijing: ICML, 2014: 1-9.
- [6] Hinton G E, Osindero S, Teh Y W. A Fast Learning Algorithm for Deep Belief Nets[J]. Neural Computation (S0899-7667), 2006, 18(7): 1527-1554.
- [7] 王坤峰, 苟超, 段艳杰, 等. 生成式对抗网络 GAN 的研究进展与展望[J]. 自动化学报, 2017, 43(3): 321-332.
Wang Kurfeng, Gou Chao, Duan Yanjie, et al. Research Progress and Prospect of Generative Adversarial Networks[J]. Acta Automatica Sinica, 2017, 43(3): 321-332.
- [8] 林懿伦, 戴星原, 李力, 等. 人工智能研究的新前线: 生成式对抗网络[J]. 自动化学报, 2018, 44(5): 775-792.
Lin Yilun, Dai Xingyuan, Li Li, et al. A New Frontier for Artificial Intelligence Research: Generative Adversarial Networks[J]. Acta Automatica Sinica, 2018, 44(5): 775-792.
- [9] Ledig C, Theis L, Huszar F, et al. Photo-realistic Single Image Super-resolution Using a Generative Adversarial Network[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 105-114.
- [10] Isola P, Zhu J Y, Zhou T, et al. Image-to-Image Translation with Conditional Adversarial Networks[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 5967-5976.
- [11] Zhu J Y, Park T, Isola P, et al. Unpaired Image-to-Image

- Translation Using Cycle-consistent Adversarial Networks[C]// IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 2242-2251.
- [12] Choi Y, Choi M, Kim M, et al. StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation[C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, UT, USA: IEEE, 2018: 8789-8797.
- [13] He Z L, Zuo W M, Kan M N, et al. AttGAN: Facial Attribute Editing by Only Changing What You Want[J]. IEEE Transactions on Image Processing (S1057-7149), 2019, 28(11): 5464-5478.
- [14] Qina R, Tan R T, Yang W H, et al. Attentive Generative Adversarial Network for Raindrop Removal From a Single Image[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, UT, USA: IEEE, 2018: 2482-2491.
- [15] Mirza M, Osindero S. Conditional Generative Adversarial nets[C]// The 18th Annual Conference of the International Speech Communication Association (INTERSPEECH). Stockholm, Sweden: ISCA, 2017: 2814-2818.
- [16] Wang H, Wang Y T, Zhou Z, et al. CosFace: Large Margin Cosine Loss for Deep Face Recognition[C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, UT, USA: IEEE, 2018: 5265-5274.
- [17] Wang Y J, Li J H, Lu Yi, et al. Image Quality Evaluation Based on Image Weighted Separating Block Peak Signal to Noise Ratio[C]// 2003 International Conference on Neural Networks and Signal Processing (ICNNSP'03). Nanjing: IEEE, 2003: 994-997.
- [18] Wang Z, Bovik A C, Sheikh H R, et al. Image Quality Assessment: from Error Visibility to Structural Similarity[J]. IEEE Transactions on Image Processing (S1057-7149), 2004, 13(4): 600-612.