

周报 (2020-10-12)

赖泽强

2020 年 10 月 12 日

目录

1 本周工作	1
2 下周计划	1
3 论文总结	1
3.1 Bring Old Photos Back to Life	2
3.1.1 相关工作	2
3.1.2 方法	2
3.1.3 重要公式解读	2
3.1.4 主要贡献	4

1 本周工作

1. 阅读论文 “Bring Old Photos Back to Life”。
2. 学习 VAE。
3. 复习生成式模型相关知识，如 GDA，朴素贝叶斯等。
4. 复习概率论，概率模型相关知识。

2 下周计划

1. 阅读并尝试运行 “Bring Old Photos Back to Life” 的代码。
2. 继续学习 “Bring Old Photos Back to Life” 论文里一些不懂的概念和数学知识。
3. 阅读与 “Bring Old Photos Back to Life” 相关的论文。

3 论文总结

主要只读了一篇论文。

3.1 Bring Old Photos Back to Life

这个工作 [1][2] 的目标任务是修复老照片，即恢复老照片原来的样子。

具体来说，包括修复两大方面损伤：

1. 非结构化损伤 (Unstructured defects)：包括胶片噪声，模糊，褪色等。
2. 结构化损伤 (Structured defects)：包括折痕和污渍等。

3.1.1 相关工作

之前也有一些工作尝试对老照片进行修复，但这些工作存在两个主要问题：

1. 首先是 Mixed degradation issue。老照片的损伤通常是多方面的，但是之前的工作要么是只针对某一方面进行修复，要么是使用一种方法对付所有方面。
2. 其次，目前大部分的模型，包括深度学习模型，都是使用合成的数据对进行训练，这种模型对于真实数据的效果会大打折扣。CycleGAN 虽然使用了真实数据，但是它使用的不匹配的数据进行训练。目前没有工作同时使用真实数据和配对数据（存疑，需要调研下）。

3.1.2 方法

这篇论文同时使用真实老照片，合成老照片及其对应的 ground truth 作为训练数据。

如图1所示，文中所述模型主要分为三部分：

- 首先是两个 VAE[3] 负责将老照片和好照片分别映射到两个隐空间。
 - 合成老照片和真实老照片公用一个 VAE。
 - VAE 以非监督的形式单独训练，训练好后固定不变。
- 然后是一个翻译网络 T，负责将老照片的隐空间映射到好照片的隐空间。
 - 为了处理 Structured defects（通常需要更多的上下文），作者引入了 Partial nonlocal block[4]。

3.1.3 重要公式解读

VAE

论文所用的 VAE Loss 前两项和传统的一致，最后一项叫 least-square loss (LSGAN)[5]，用于解决 VAE 中的 over-smooth issue。

$$\begin{aligned}\mathcal{L}_{\text{VAE}_1}(r) = & \text{KL}(E_{\mathcal{R},\mathcal{X}}(z_r | r) \parallel \mathcal{N}(0, I)) \\ & + \alpha \mathbb{E}_{z_r \sim E_{\mathcal{R},\mathcal{X}}(z_r | r)} [\|G_{\mathcal{R},\mathcal{X}}(r_{\mathcal{R} \rightarrow \mathcal{R}} | z_r) - r\|_1] \\ & + \mathcal{L}_{\text{VAE}_1, \text{GAN}}(r)\end{aligned}\tag{1}$$

为了更好缩小真实老照片和合成老照片的 domain gap，作者还引入了一个 Adversarial loss 训练一个 Discriminator, $D_{r,x}$ 越大，则 Discriminator 认为这个 feature 越接近真实老照片的 feature。

$$\begin{aligned}\mathcal{L}_{\text{VAE}_1, \text{GAN}}^{\text{latent}}(r, x) = & \mathbb{E}_{x \sim \mathcal{X}} [D_{\mathcal{R},\mathcal{X}}(E_{\mathcal{R},\mathcal{X}}(x))^2] \\ & + \mathbb{E}_{r \sim \mathcal{R}} [(1 - D_{\mathcal{R},\mathcal{X}}(E_{\mathcal{R},\mathcal{X}}(r)))^2]\end{aligned}\tag{2}$$

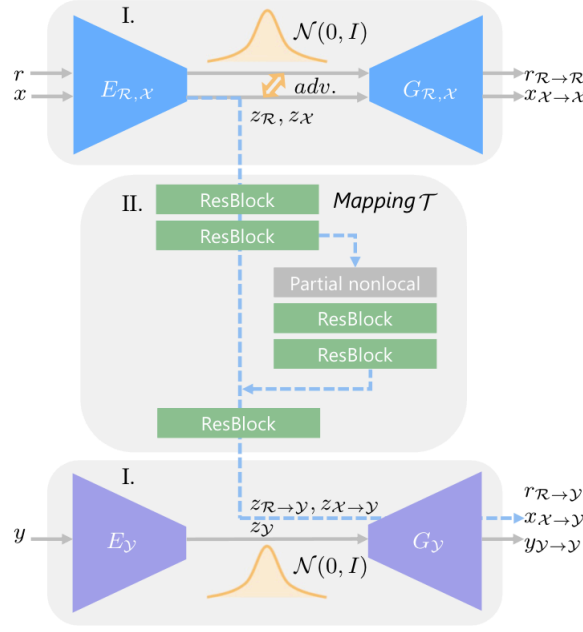


图 1: 网络结构

下面的就是第一个 VAE 总的 Loss:

$$\min_{E_{\mathcal{R}, \mathcal{X}}, G_{\mathcal{R}, \mathcal{X}}} \max_{D_{\mathcal{R}, \mathcal{X}}} \mathcal{L}_{\text{VAE}_1}(r) + \mathcal{L}_{\text{VAE}_1}(x) + \mathcal{L}_{\text{VAE}_1, \text{GAN}}^{\text{latent}}(r, x)$$

Mapping

$$\mathcal{L}_{\mathcal{T}}(x, y) = \lambda_1 \mathcal{L}_{\mathcal{T}, \ell_1} + \mathcal{L}_{\mathcal{T}, \text{GAN}} + \lambda_2 \mathcal{L}_{\text{FM}}$$

$$\mathcal{L}_{\mathcal{T}, \ell_1} = \mathbb{E} \|\mathcal{T}(z_x) - z_y\|_1$$

这个 Loss 用于训练翻译网络 T，第一部分是隐向量的 L1 Loss，第二部分也是 least-square loss (LSGAN)，第三部分是 feature matching loss 用于 stabilize the GAN training. 后两部分目前没看懂。

Multiple degradation restoration

在基础的 Mapping 网络中，Resnet Block 使用的 CNN filter 只能关注到 local 的 feature，但是一些 structured defect，比如褶皱等，必须要有更大范围的上下文信息才能够进行修复（你不可能靠周围的褶皱修复褶皱）。

因此，作者引入了一个 nonlocal block，之所以叫 partial 是因为 nonlocal block 的原文里使用的是所有像素的信息，这里作者只关注非褶皱像素（作者还引入了一个 defect region detection network 标出一个 defect mask，即下面的 m，m=1 表示是 defect region）。

下面这个公式是 nonlocal block 原文里给出的公式，简单来说 nonlocal block 就是将某个像素的特征值就是其他像素特征值的加权平均。公式里的 \mathbf{f} 就是计算权值的一个函数， \mathbf{g} 则是计算特征值的函数， $C(x)$ 是一个正则项。

$$\mathbf{y}_i = \frac{1}{\mathcal{C}(\mathbf{x})} \sum_{\forall j} f(\mathbf{x}_i, \mathbf{x}_j) g(\mathbf{x}_j)$$

下面的公式都是本篇论文里的公式：

- $s_{i,j}$ 相当于计算权值，从公式里不难看出，这个权值对于 defect region 里的像素是 0，即我们不关注 defect region 里的像素的特征值。
- $f_{i,j}$ 用于计算两个像素之间的像素度，显然，越相似的像素我们应该越关注。
- 第三个公式几乎就是 nonlocal block 原文中的本文形式；
- 最后一个公式表示，对于 defect region 里的像素，我们采用 nonlocal block 计算出来的特征值，对于其他像素，特征值保持不变。

$$s_{i,j} = (1 - m_j) f_{i,j} / \sum_{\forall k} (1 - m_k) f_{i,k}$$

$$f_{i,j} = \exp \left(\theta (F_i)^T \cdot \phi (F_j) \right)$$

$$O_i = \nu \left(\sum_{\forall j} s_{i,j} \mu (F_j) \right)$$

$$F_{\text{fuse}} = (1 - m) \odot \rho_{\text{local}} (F) + m \odot \rho_{\text{global}} (O)$$

3.1.4 主要贡献

- 提出了一种比较有效的，在 domain translation 问题中，使用未配对数据的方法。
- 这个方法有机会使用到别的 domain translation 问题上，比如降噪，可以结合魏学长之前的工作进行扩展。

References

- [1] Ziyu Wan et al. *Bringing Old Photos Back to Life*. 2020. arXiv: 2004.09484 [cs.CV] (cit. on p. 2).
- [2] Ziyu Wan et al. *Old Photo Restoration via Deep Latent Space Translation*. 2020. arXiv: 2009.07047 [cs.CV] (cit. on p. 2).
- [3] Diederik P Kingma and Max Welling. *Auto-Encoding Variational Bayes*. 2014. arXiv: 1312.6114 [stat.ML] (cit. on p. 2).
- [4] Xiaolong Wang et al. *Non-local Neural Networks*. 2018. arXiv: 1711.07971 [cs.CV] (cit. on p. 2).
- [5] Xudong Mao et al. *Least Squares Generative Adversarial Networks*. 2017. arXiv: 1611.04076 [cs.CV] (cit. on p. 2).