

基于 alpha 抠图算法的改良 P 图

边晓璐 李丹

四川大学锦城学院 四川 成都 610065

【摘要】针对 P 图软件的抠图功能无法有效解决前景与背景分割间隙明显的问题，提出了一种通过对抗损失进行预测以训练深度网络的算法，该算法通过上下文转换块（CS Block）组合特征，利用编码-解码（Encoder-Decoder）框架进行抠图（matting），得到前景图和 alpha 消光（matte）后的图，然后通过半监督对抗性损失（Self-Supervised Adversarial Loss）训练真实数据与 Adobe 数据集的前景和 alpha，并计算其损失，最终通过鉴别器（Discrimination）判断其真假。

【关键词】上下文转换块；半监督对抗性损失；图像抠图；鉴别器

引言

数字图像抠图技术在计算机视觉和图像处理领域中非常热门，同时也很难处理。目前，数字图像抠图技术在电影电视制作、证件照、美图方面都有一定的成就，但是效果不是特别理想，仍无法有效消除前景与背景之间颜色渐变度差异大的问题，导致轮廓看起来较为明显。

本文针对现有的 P 图软件的抠图功能无法有效解决前景与背景分割间隙明显的问题，例如：美图秀秀的抠图功能存在两个问题：第一，抠图后人物周围有线条；第二，选择无线条的抠图方式的话，人物的细节部位存在瑕疵。为此提出了一种通过对抗损失进行预测以训练深度网络的算法。

该算法存在三个受限因素。首先，通过图像 I 与前景 F 、背景 B 、透明度 α 有关的表达式： $I = \alpha F + (1 - \alpha) B$ ，可以发现存在三个未知数是无法求得确定解的，因此该算法将背景 B 作为已知量，求解 F 和 α 。所以不仅要输入一张原始的图像 I ，还要输入一张没有人物的背景图 B 。第二，需要输入非动态背景和运动信号，达到消除拍摄时手抖的因素，但是效果可能不是很理想。最后，该算法只处理人物前景，除此之外的其他类别图像或多类别图像均不适用。

1 算法实现

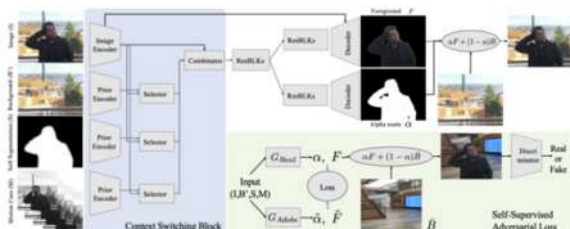


图 1 算法总流程图

1.1 上下文转换块（CS Block）

1.1.1 编码-解码框架（Encoder-Decoder）

图像编码器和优先编码器采用的结构是 $C64(k7) - C*128(k3) - C*256(k3)$ ，其中 $CN(kS)$ 代表使用步幅（stride）为 1 的 N 个 $S*S$ 滤波器的卷积层，然后进行批量标准化（Batch Normalization）以及 ReLU； $C*N(kS)$ 代表使用步幅（stride）为 2 的 N 个 $S*S$ 滤波器的卷积层，然后进行批量标准化（Batch Normalization）以及 ReLU。图像编码器的输出分辨率为 $256*128*128$ 的连通区域（blob），所以卷积层都没有偏置项。

α 解码器采用的结构是 $CU*128(K3) - CU*64(K3) - Co1(k7) - Tahn$ ，解码器的输入分辨率为 $256*128*128$ 。其中 $CU*N(kS)$ 表示使用因子为 2 的双线性向上采样，然后与步幅为 1 的带有 N 个 $S*S$ 滤波器的卷积层进行批量标准化及 ReLU。最后一层 $Co1(k7)$ 由 1 个 $7*7$ 滤波器的卷积层构成，然后是 Tahn 层。alpha 的输出范围在 $(-1, 1)$ 。

前景解码器采用的结构 $CU*128(K3) - CU*64(K3) - Co1(k7)$ ，该解码器的输入分辨率与 α 解码器相同。不同的是有一个分辨率为 $128*256*256$ 的图像输入特征的跳跃连接，该连接与 $CU*128(k3)$ 的输出相组合并传递到 $CU*64(k3)$ 。

1.1.2 转换器（Selector）

转换器的结构是 $C64(k1)$ ，“转换器”块将空间分辨率为 $512*128*128$ 的连通区域的图像特征与先验特征的并列作为输入，转换器输出分辨率为 $64*128*128$ 的连通区域。该转换器的目的是生成以图像为条件的先验特征，有助于网络不过度依赖一种功能就能将合成的数据集通过训练推广到真实图像。

1.1.3 ResBLKs

组合器的输出首先经过 K 为 7 的 ResBLKs 传递, 然后作为输入提供给 2 个单独的 K 为 3 的 Res-BLKs, 得到 alpha 和前景层。每个 ResBLKs 由 Conv256(k3)-BN-ReLU-Conv256(k3)-BN 构成, 且分辨率为 256*128*128。

1.2 半监督对抗性损失

1.2.1 消光网络 G

输入 X 和真实图像, 将用 Adobe 数据集训练的消光网络所得到的前景及透明度作为虚假的地面。然后训练对抗性损失和消光网络输出损失, 与“伪真实地面”相比较; 第二次损失赋予较小的权重, 将训练的时间间隔减少。假设公式能避免陷入局部最小, 我们用 LS-GAN 框架训练生成器 G 和鉴别器 D。

最小化生成器的更新:

$$\min_{\theta_{\text{Real}}} E_{X, B \sim p_{X, B}} [(D(\alpha F + (1-\alpha)B) - 1)^2 + \lambda \{2 \|\alpha - \tilde{\alpha}\| + 4 \|\nabla(\alpha) - \nabla(\tilde{\alpha})\| \} (1) + \|F - \tilde{F}\|_1 + \|I - \alpha F - (1-\alpha)B\|_1]$$

最小化鉴别器:

$$\min_{\theta_{\text{Disc}}} E_{X, B \sim p_{X, B}} [(D(\alpha F + (1-\alpha)B))^2] + E_{I \in p_{\text{data}}} [(D(I) - 1)^2] \quad (2)$$

1.2.2 鉴别器 (Discrimination)

鉴别器的结构是 C*64(k4)-C*I128 (k4)-C*I256(k4)-C*I512(k4), 我们使用 70*70 的 PatchGAN。C*N(kS) 表示步长为 2 的 N 个 S*S 滤波器的卷积层, 然后 leaky ReLUs 的斜率为 0.2。I 表示 leaky ReLU 之前, 除第一层以外的所有层中都存在 Instance Norm。再将 PatchGAN 应用于整个合成图像, 每 70 x 70 个 patch 进行一次卷积, 以确定其真假。

1.3 总流程

该算法的核心分为两大块, 上下文切换块和半监督对抗性损失。

首先将输入的图像 I、背景 B、柔和分割 S 和运动信号 M 分别放入编码器中, 然后通过上下文切换块选择性地不同的输入信号组合到转换器 (selector) 中, 将三个转换器与图像 I 的编码器进行结合、批处理等一系列操作, 再解码得到前景 F 和透明度 α 。

其次输入 I、B、S 和 M, 通过半监督对抗性损失 (Self-Supervised Adversarial Loss) 中的深层消光 (matting) 网络训练真实数据与 Adobe 数据集的前景 F 和透明度 α , 然后计算他们之间的损失值, 再将用真实数据训练后的 F 和 α 以及自己选的新背景 \bar{B} 带入方程 $I = \alpha F + (1 - \alpha) B$ 中, 得到新的合成图, 然后用鉴别器判断其真假。

2 实验及结果

由于该算法的要求很高, 很难有满足条件的新数据

集, 因此在网上随机找了 11 张具有代表性的图片作为数据集。为了满足该算法的要求, 用 PS 处理这 11 张图片, 将人物抠除, 留下背景并填充抠除的人物区域, 这样就得到了 11 组背景图像和原始图像, 其中图片按照人数由少到多进行编号, 图片选用了单人及多人、简单及复杂的因素进行全方面的分析。

再选择几张感兴趣的背景, 由于该算法的输出是一张图, 而不是其他类似于求准确率的算法, 所以不需要大量数据。选择了两种背景图, 风景图和城市图, 但是背景的种类和抠图并没有关系。

首先对图片进行蒙版分割, 如果分割之后为全黑, 就意味着该图像的前景无法识别, 由于本论文针对人为研究对象, 因此除此之外的其他图像均无法成功运行; 反之, 如果出现这种黑白相间的图片, 就意味着可以成功运行。



图 2 颜色对抠图效果的影响对比图

通过图 2 的 (a2) 与 (a3) 进行比较, 可以发现: (a3) 中手的部位缺失了, 同样的, 另一只手和脚都存在部分缺失, 这是由于图像的背景色与人物的阴影颜色大致相近, 导致抠图时无法确认是背景还是前景, 所以无法精确分离人物和背景。

观察图 2 的 (b1)、(b2) 和 (b3) 进行比较, 可以发现: 有光照的地方并没有存在任何部分缺失的情况, 反而是一些细节部分有所缺失, 而造成该情况的原因是颜色相近, 该算法主要通过颜色对前景图像与背景图像进行区分, 当颜色相近时, 无法识别, 自然就造成了局部缺失, 因此光照并不会影响抠图的效果。

当背景图的颜色与前景图的颜色差异很大时, 合成的新图像一眼就能看出来不是真实的, 但是采用原来的数据集就不存在这一问题, 这是由于处理数据集时通过 PS 的人工操作对人物与背景的分割达不到使用算法进行分割的精确度, 所以将原始图像中残留的背景部分带入了合成图像。而且当对多人物图像进行抠图时, 该算法抠取的是所有人物, 而不是某一个人。



图 3 细节 (发丝) 处理对比图

通过图 3 的 (a2) 与 (a3) 进行比较, 可以发现: (a3)

中头发丝部位缺失了, 造成这样的情况有两种原因。第一, 处理数据集的方式采用的是 PS 抠除人物, 那么可能是 PS 效果不好; 第二, 该算法无法处理头发丝等较为细节的部分。为了验证这个问题, 本文采用为数不多的通过拍照的方式获得背景图像和原图像, 这就排除了 PS 效果不好的因素, 通过结果验证产生该情况的真实原因。

通过图 3 的 (b1) 和 (b2) 进行比较, 可以发现: (b2) 中存在头发丝, 但是会缺失极细的部分, 因此验证了发丝等细节部位的缺失在极大程度上与 PS 有关, 不可否认的是, 还有少量因素是与该算法有关, 虽然该算法能够处理细节问题, 但对于极其细微的部分仍然无法处理。

3 同类算法对比

3.1 与人体前景的自动抠图算法对比



图 4 人体前景自动抠图算法与本文算法效果对比图

通过图 4 的 (b) 和 (c) 进行比较, 可以发现: 本文的细节处理没有其他算法的好, 这是由于本文的数据集通过人工 PS 处理。而人体前景自动抠图算法通过 RPN 网络得到前景候选框, 然后利用回归获取更精确的坐标; 同时, 对每个预测为人像前景的候选框进行前景分割 (进入二值掩码生成负载)。负载网络在 FCN 的输出层之后使用 Sigmoid 激活函数, 并用二分类交叉熵作为损失函数, 最后将属于人像前景的候选框进行组合, 获得最终结果, 有效的提高了人像与背景分离的精度。

该算法的好处在于 (1) 只需要输入一张图, 这样可选择的图片就很多, 今后在应用时的受限性很小; (2) 使用算法对人像前景和背景进行分离, 提高了轮廓的细节处理; 劣处在于 (1) 当图像包含更为复杂的背景 (如自然场景) 时, 效果不是特别好; (2) 无法对除人像前景以外的其他类别或多类别图像中的前景进行抠图。

本论文算法的好处在于使用相机拍摄照片的话, 无论背景是否复杂都能达到理想效果; 劣处在于 (1) 需要输入两张图片, 一张为有人像的图片, 另一张为只有背景的图片; (2) 只能对人像前景进行抠图。

3.2 与改进的自然图像鲁棒抠图算法对比

得到的效果与 3.1 节中的效果大致相同, 但该算法通过减少未知像素降低了后续 α 估值计算量, 然后引

入可信度, 选取可信度最高的前景、背景样本组估计未知像素的值, 并得到透明度估计值。然后构建二次目标函数优化透明度 α , 达到了改善抠图的目的。该算法的合成效果和均方误差 (MSE) 值都优于大部分 matting 算法。而且该算法的流程比本论文的算法流程简单, 但是该算法运用了大量数学公式进行推导, 提出此算法也很不容易。

4 结论

本文提出了一种通过对抗损失进行预测以训练深度网络的算法, 该算法的受限性很高, 需要有人物和没有人物的图片, 而且只能处理人像前景, 其他类别均不使用, 但是它避免使用绿屏或 trimap, 能够节省很多时间。使用人物与背景分割细节处理的很好 (即原始图像的部分背景不会残留在合成的新图像中) 的数据集的话, 会得到满意的效果。增加算法的前景与背景分离功能并能改善前景与背景颜色相近会出现缺失的问题, 将会是未来研究的重点。

【参考文献】

- [1] Soumyadip Sengupta, Vivek Jayaram, Brian Curless, Steve Seitz, and Ira Kemelmacher-Shlizerman University of Washington. Background Matting: The World is Your Green Screen.
- [2] 赵海, 雷凯茹, 朱宏博, 朴春鹤. 基于潜在半径优化策略的数字图像抠图算法 [J]. 东北大学学报 (自然科学版), 2015, 36(11): 1553-1557.
- [3] 冉清, 冯结青. 人体前景的自动抠图算法 [J]. 计算机辅助设计与图形学学报, 2020, 32(02): 277-286.
- [4] 黄睿, 王翔. 改进的自然图像鲁棒抠图算法 [J]. 计算机工程与应用, 2013, 49(12): 136-139.
- [5] 王明楸, 陈金聪, 陈东昇, 文龙, 李佐勇. 人物证件照的自动换底算法 [J]. 福建电脑, 2016, 32(06): 12-13+67.
- [6] 孟蕊, 唐好选. 基于 OneCut 和共享抠图算法的自适应衣物目标抠取 [J]. 智能计算机与应用, 2015, 5(05): 84-88.
- [7] 赵海, 雷凯茹, 朱宏博, 朴春鹤. 基于潜在半径优化策略的数字图像抠图算法 [J]. 东北大学学报 (自然科学版), 2015, 36(11): 1553-1557.
- [8] 俞克强, 黄文清, 李俊柏, 黄牧. 例样相似性的全局最优抠图算法 [J]. 工业控制计算机, 2014, 27(03): 48-49+51.
- [9] 曹森科, 方发明, 徐盈盈, 沈超敏. 基于局部区域主动轮廓与抠图的图像分割算法 [J]. 计算机工程, 2015, 41(08): 190-195.
- [10] 聂栋栋, 王丽. 改进的 KNN 抠图技术 [J]. 小型微型计算机系统, 2015, 36(06): 1316-1320.
- [11] 杨杰, 王志胜. 基于参数自适应滤波器的抠图优化算法 [J]. 计算机与数字工程, 2012, 40(02): 99-101.
- [12] 林生佑, 石教英. 基于感知颜色空间的自然图像抠图 [J]. 计算机辅助设计与图形学学报, 2005(05): 915-920.