# GAN Understanding：理解GAN

Taylor Guo, 2019年1月05日 @ innotechx上海

服务号二维码



目录

[GAN Understanding：理解GAN 1](#_Toc43022679)

[前言 1](#_Toc43022680)

[生成模型 2](#_Toc43022681)

[前言 2](#_Toc43022682)

[1 简介 2](#_Toc43022683)

[2 相关工作 5](#_Toc43022684)

[3  问题 6](#_Toc43022685)

[4  方法 6](#_Toc43022686)

[5  实现细节 8](#_Toc43022687)

[6  实验评估 8](#_Toc43022688)

[7  结论 8](#_Toc43022689)

[参考文献 9](#_Toc43022690)

[StyleGAN: 一种基于风格迁移的生成对抗网络的生成器 10](#_Toc43022691)

[前言 10](#_Toc43022692)

[生成模型 10](#_Toc43022693)

[前言 10](#_Toc43022694)

[1 简介 11](#_Toc43022695)

### 前言

生成对抗网络。

。

### 生成模型

统计分类中，包括机器学习方法，有两种主要的方法，分别是生成式方法和判别式方法。

。

### 前言

生成对抗网络。

。

### 1 简介

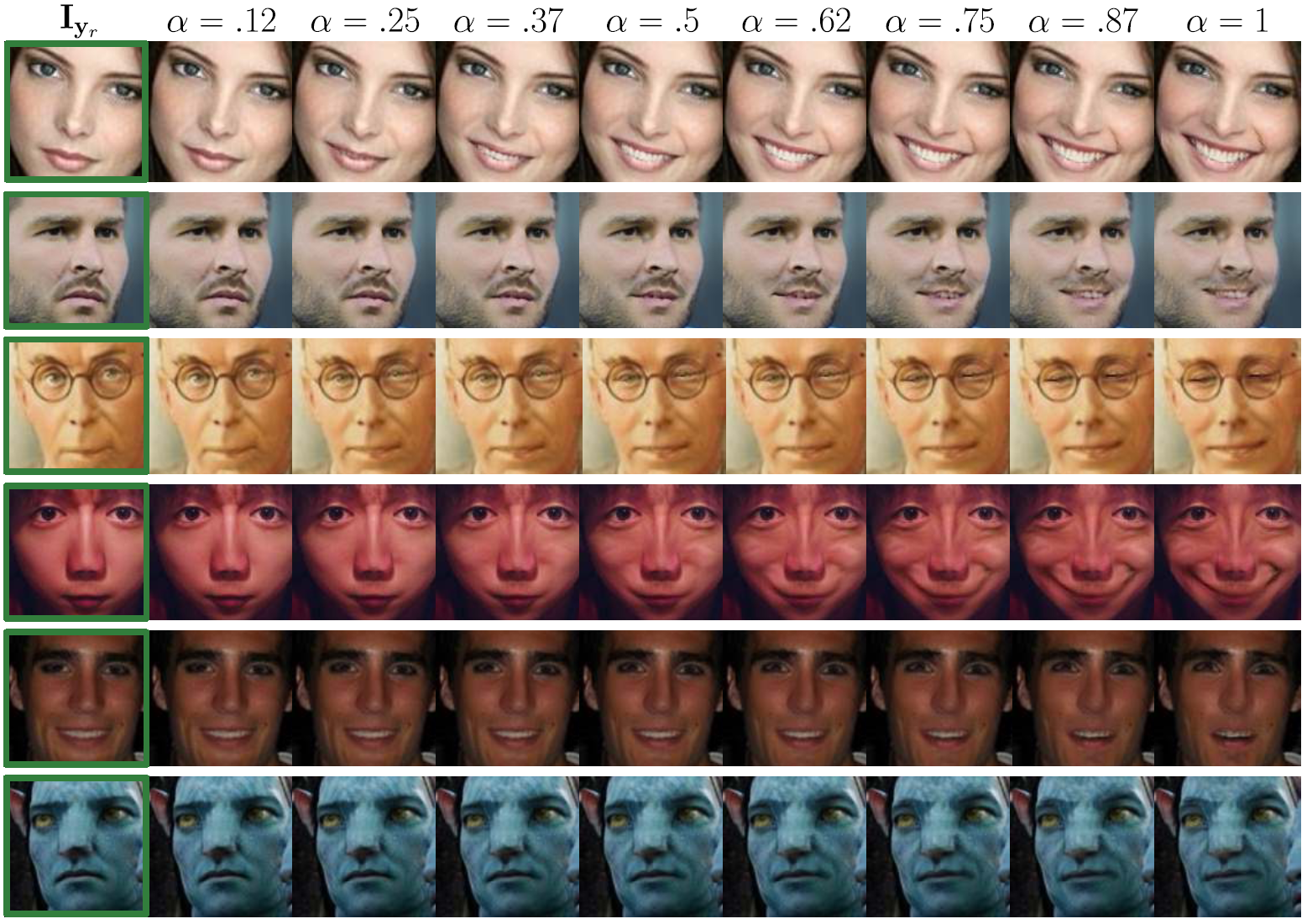


图1. 单图像人脸变换。我们提出了一种自动连贯的方法，并不局限于离散数量的表情，可以将给定图像人脸进行变换，渲染出连续的表情。在这些案例中，左上角是输入图像Iyr（绿色框里标出），参数α控制目标活动单元的激活度，比如微笑表情。另外，我们的系统也可以处理非自然的图像，比如最后一行。

DBoW3是DBoW2的增强版，这是一个开源的C++库，用于给图像特征排序，并将图像转化成视觉词袋进行表示。它采用层级树状结构，将相近的图像特征在物理存储上聚集在一起，创建一个视觉词典。DBoW3还生成一个图像数据库，带有顺序索引和逆序索引，可以使图像特征的检索和对比非常快。

DBoW3与DBoW2的主要差别：

* DBoW3依赖项只有OpenCV，DBoW2依赖项DLIB被移除；
* DBoW3可以直接使用二值和浮点特征描述子，不需要再为这些特征描述子重写新类；
* DBoW3可以在Linux和Windows下编译；
* 为了优化执行速度，重写了部分代码（特征操作都写入类DescManip）；DBoW3的接口也被简化了；
* 可以使用二进制视觉词典文件；二进制文件在加载和保存上比.yml文件快4-5倍；而且，二进制文件还能被压缩；
* 仍然和DBoW2 yml文件兼容。

图像数据库的主要组成：

* 视觉字典树
* 顺序索引
* 逆序索引

对应的DBoW3库有两个主要的类：Vocabulary和Database。视觉词典将图像转化成视觉词袋向量，图像数据库对图像进行索引。

ORB-SLAM2中的ORB Vocabulary保存在文件orbvoc.dbow3中，二进制文件在Github上：

https://github.com/raulmur/ORB\_SLAM2/tree/master/Vocabulary

### 2 相关工作

视觉字典树生成过程：

### 3  问题

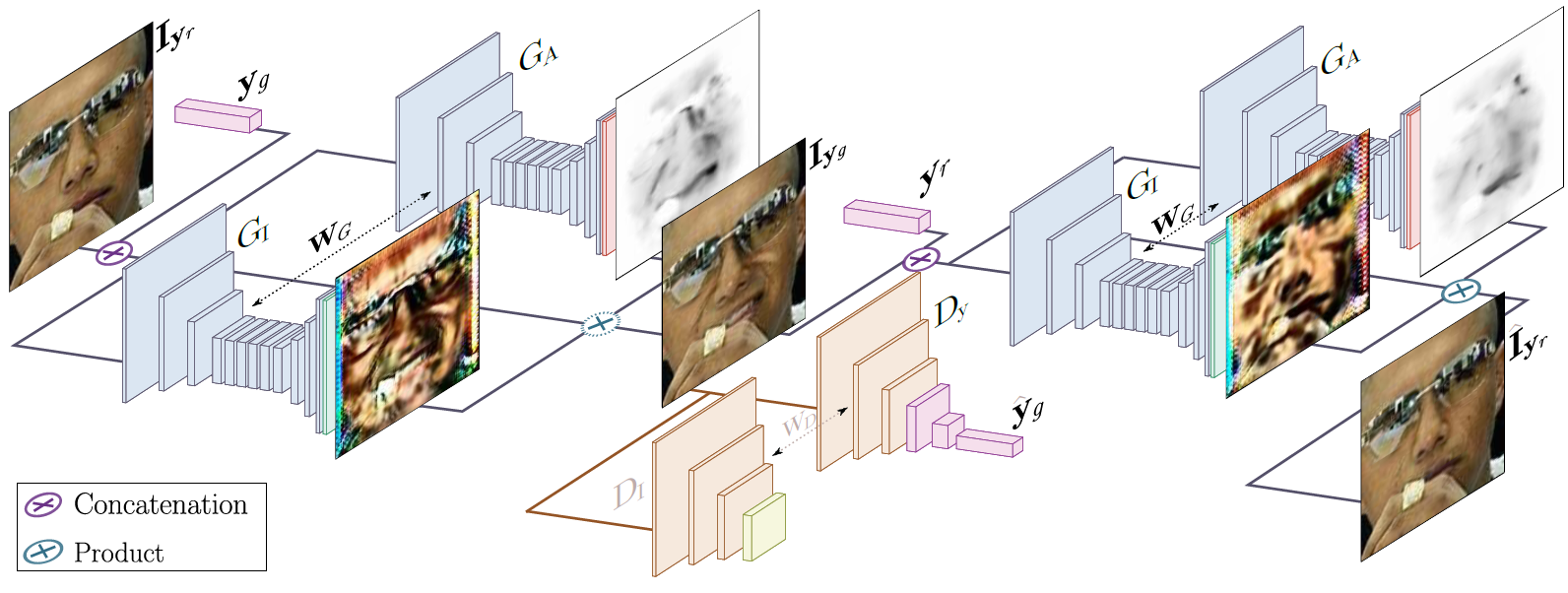


图2. 生成真实连续图像的方法。这个架构由2个主要模块组成：生成器G

### 4  方法

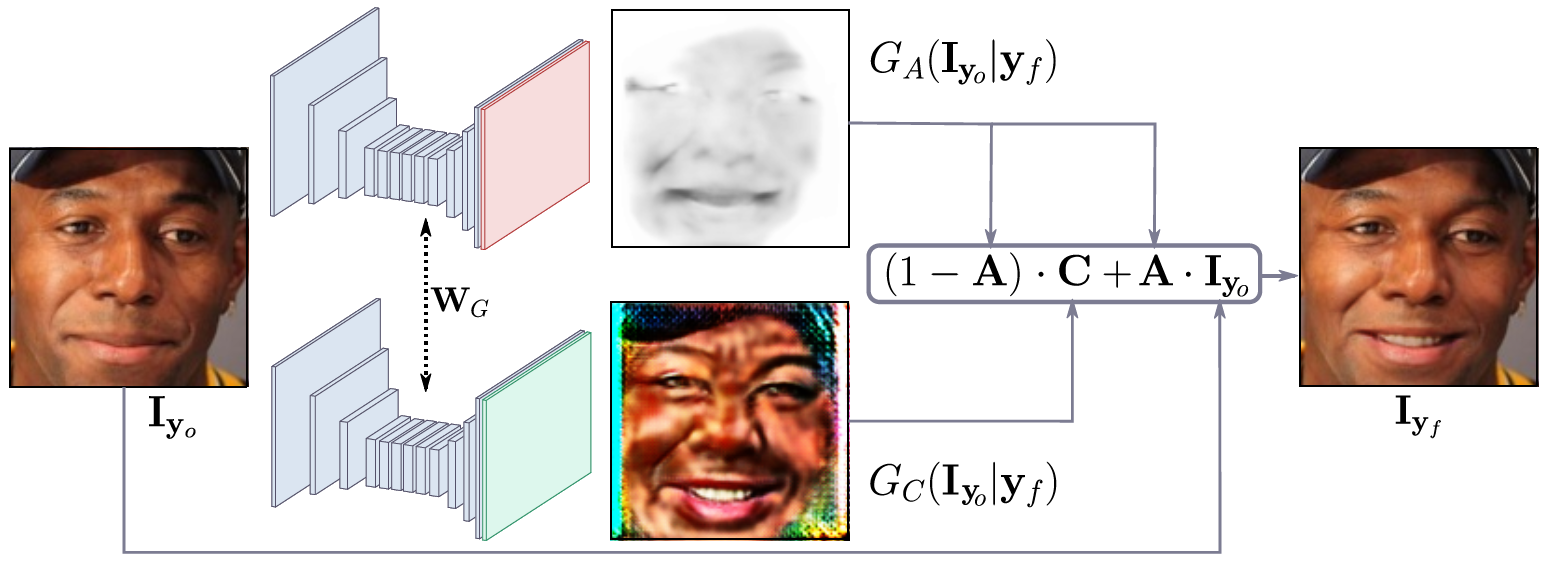


图3.

### 5  实现细节

这里只对DBoW3的库类做简单说明，这个库并不是很大。

### 6  实验评估

### 7  结论

### 参考文献

1. 《机器学习》周志华
2. 《应用多元分析》王学民
3. Bags of Binary Words for Fast Place Recognition in Image Sequences
4. Video Google: A Text Retrieval Approach to Object Matching in Videos
5. Scalable Recognition with a Vocabulary Tree
6. Interactive learning of visual topological navigation
7. Fast and incremental method for loop-closure detection using bags of visual words
8. [Bag of visual word model based on binary hashing and space pyramid](http://sci-hub.bz/10.1117/12.2245124)

## StyleGAN: 一种基于风格迁移的生成对抗网络的生成器

### 前言

生成对抗网络。

。

这种新的架构可以产生自动学习的、无监督地分离生成的图片中的高层属性和随机变量；可以直觉地对特定尺度进行控制合成过程。

### 生成模型

统计分类中，包括机器学习方法，有两种主要的方法，分别是生成式方法和判别式方法。

。

## DCGAN: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks

### 前言

最近几年，监督学习的卷积网络在计算机视觉领域大量采用。相对而言，无监督学习的卷积神经网络缺很少被关注。本文的工作希望能填补监督学习和无监督学习的这个差距。我们介绍了一种卷积神经网络，叫做深度卷积生成对抗网络。

我们的深度卷积对抗网络的生成器和判别器中学习到了从物体到场景的层级表示。另外，我们使用学习到的特征用于各种任务，发现这些特征可以用作一般图像表示。

### 1 简介

生成对抗网络为最大似然估计技术提供了另一种方法。它的学习过程，加上并没有启发式的代价函数（比如像素级均方差），用于表示学习非常有吸引力。

我们提供并评估了深度卷积生成对抗网络的网络结构，并进行了稳定训练。我们把它叫做DCGAN，深度卷积生成对抗网络。

我们使用训练好的判别器用于图像分类，比无监督的算法更有优势。

我们将GAN学习到的滤波器进行可视化，经验表示特定的滤波器可以学习到绘画特定的物体。

我们发现了生成器具有很有趣的向量数学特征，可以用于操纵生成样板的许多语义特征的质量。

### 2 相关工作

#### 2.1 无标签数据的表示学习

无监督表示学习的一种经典方法就是数据聚类（比如K-mean）