

# 基于迁移学习的糖尿病视网膜病变眼底图像合成研究

粘伟杰, 金海龙

燕山大学电气工程学院, 河北 秦皇岛

收稿日期: \*\*\*\*\*

## 摘要

文章研究了利用生成对抗网络(GAN)和迁移学习在小数据集上生成糖尿病视网膜病变(DR)眼底图像的有效方法。传统的图像生成技术在高维数据下表现不佳, 难以准确学习数据分布。自2014年GAN被提出以来, 众多研究致力于其优化与应用, 然而在小样本情况下仍面临训练难度和过拟合问题。文章通过迁移学习方法, 利用在大数据集上预训练的GAN模型, 结合条件生成对抗网络(CGAN), 在小数据集上重新训练部分参数, 从而生成高清晰度的DR图像, 并提高生成图像的类别控制能力。研究表明, 采用迁移学习的方法显著提升了图像生成质量, 实验结果显示, 基于迁移学习的GAN和CGAN在小数据集上的FID分数明显优于无迁移学习的对照组, 表明生成的图像在质量和多样性上有显著改善。这项研究不仅为医学图像生成提供了新的思路, 也为解决小数据集和不平衡问题提供了有效的解决方案。

## 关键词

深度学习, 生成对抗网络, 迁移学习, 糖尿病视网膜病变

# Image Synthesis of Diabetic Retinopathy Based on Transfer Learning: A Study

Weijie Nian, Hailong Jin

School of Electrical Engineering, Yanshan University, Qinhuangdao Hebei

Received: \*\*\*\*\*

## Abstract

This paper investigates an effective method for generating fundus images of diabetic retinopathy (DR) on small datasets using Generative Adversarial Networks (GANs) and transfer learning. Traditional

image generation techniques struggle to accurately learn data distributions in high-dimensional data, leading to poor performance. Since the introduction of GANs in 2014, numerous studies have focused on their optimization and applications; however, they still face challenges such as training difficulties and overfitting in small sample cases. This study employs transfer learning by utilizing a GAN model pre-trained on a large dataset, combined with Conditional Generative Adversarial Networks (CGAN), to retrain some parameters on a small dataset. This approach generates high-quality DR images and enhances the class control capability of the generated images. Research shows that using transfer learning significantly improves the quality of image generation. Experimental results indicate that the FID scores of GANs and CGANs based on transfer learning are notably superior to those of the control group without transfer learning, demonstrating significant improvements in the quality and diversity of the generated images. This research not only provides new insights for medical image generation but also offers effective solutions for addressing small datasets and imbalance issues.

## Keywords

Deep Learning, Generative Adversarial Network, Transfer Learning, Diabetic Retinopathy

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

图像生成属于计算机视觉领域其中一个研究方向, 该方向主要利用各种方法寻找数据分布, 并建立一个能生成属于该分布的数据模型。传统图像生成算法包括变分自动编码、玻尔兹曼向量机、近似推理马尔可夫链、自回归模型等。由于图像数据的高维特点, 所建立的模型通常难以学习到数据的分布, 导致模型不能生成近似样本分布的图像。再有, 所建立的模型难以执行难解的推断和归一化的分布。面对这些计数上的困难, 可以采用近似计算或者设计不用去做复杂计算的模型两种方法[1]。基于这两种方法, 研究人员开展了一系列相关研究, 但效果并不十分理想。直到 2014 年, Goodfellow 等人引入的生成对抗网络(Generative Adversarial Networks, GAN)取得了突破性的进展[2]。之后研究人员在此方向开展了大量研究工作, 产生了许多 GAN 的版本。主要分为无条件合成、条件合成和交叉合成[3]。

无条件合成, 即在用 GAN 生成图像时没有加入任何信息, 从输入的随机向量生成近似图像。包括 DCGAN、WPGAN 和 PGGAN, 其中[4][5]利用 DCGAN 生成医疗图像达到了可以欺骗医疗专家的效果。而 PGGAN 采用渐进增长方法生成了高清晰的逼真图像。前两种方法可用于合成最大  $256 \times 256$  分辨率的图像, 后一种方法可生成  $1024 \times 1024$  分辨率的高清图像。但大部分利用上述方法合成核磁共振图像, 原因可能是这种图像开源数据较多并且采用核磁共振设备扫描时间长[3]。如果要利用这种方法进行数据增强, 需要为每一个类别训练一个 GAN, 即单样本合成。Frid-Adar 等人用三个 DCGAN 合成三种肝损伤图像, 生成的样本被证明有助于提高肝损伤分类器的分类效果[5]。林志鹏等人将 4 类宫颈细胞图像分为有无病变 2 个类别, 然后训练两个基于 DCGAN 的模型分别生成有病变图像和无病变图像, 然后将训练好的图像用于二分类任务增强, 通过使用 GAN 扩充后的图像集训练的分类器, 准确率达到了 97%, 明显优于使用传统仿射变换扩充的数据集(准确率为 93%)[6]。Bermudez 等人利用 DCGAN 生成逼真的脑核磁共振图像, 但与真实的图像存在一定差异[7]。

条件合成, 通过在 GAN 中加入限制信息(可以是图像类别、文字描述、位置分割图等), 使之生成与

信息相关的图片或者其他内容。这种可以生成带标签的医疗图像很具有挑战性并且使用价值也更大。如 Jin 等人利用肺的 CT 数据集成带病灶的肺部 CT 图片, 提高了肺部 CT 图片的分割效果[8]。Mok 等人用 CGAN 生成更多脑瘤图像, 利用分割图作为条件输入模型中, 由粗到细地生成脑部核磁图像, 提高了脑瘤分割效果[9]。

交叉合成, 为了诊断病情, 医生需要给病人拍摄多种图像, 比如核磁共振和 CT, 这增加了病人的医疗费用也耗费很多时间, 而且多次拍摄 CT 图像对人体有伤害, 为解决这个问题, 很多研究[10]-[12]利用 CycleGAN 进行交叉合成, 即通过核磁图像生成对应的 CT 图像。

除了 GAN 外, 为了获得更多数据来让数据平衡, 也有很多传统图像处理方法用于数据增强, 比如水平翻转、垂直翻转、随机旋转、随机剪裁等等, 但这些方法都不能生成不同的逼真的样本, 这些方法被证明对数据增强有一定效果。但很多情况下, 我们的数据太过不平衡, 不同类别的样本不在一个数量级上, 如目前已知公开的最大 DR 数据集为 EyePACS [13], 正常的眼底图像(0 级)要比 3、4 级 DR 高出两个数量级, 如此差异巨大的数据集采用传统方法就不足以彻底解决数据不平衡的问题[14]。生成对抗网络(Generative Adversarial Network, GAN)可以生成不存在的逼真图像[2], 但是 GAN 相对不易训练, 而且对算力要求高, 因此小数据集难以训练 GAN [15]。因此, 如何在小数据集和相对低的算力条件下, 用 GAN 生成医学图像(如糖尿病视网膜病变图像)还需进一步开展研究。

借鉴在小数据集分类任务中, 很多工作利用迁移学习达到了很好的效果[16], 那么如果能利用其他研究者在相似任务训练好的 GAN 模型用到目标任务上, 则可能对 GAN 的训练有帮助, 受此启发, 考虑利用其他研究中训练好的 GAN, 再用我们的数据集重新训练。然而, 还存在很多问题, 因为要生成高清晰的图像, GAN 的网络结构通常很大, 而大的网络结构需要用大量数据才能充分训练, 用小数据集训练这样的 GAN 容易过拟合[17], 所以选择拿部分网络进行训练需要研究。由文献[18]可知, 靠近图像一侧的层是通用层, 这些层内的信息基本是直线、曲线、纹理、色彩等所有图像共有信息, 而且这些层需要大量数据训练。受此启发, Zhao 等人在预训练的 GP-GAN 基础上, 采用迁移学习, 把在大数据上训练好的 GAN 模型, 固定靠近图像一侧层的网络参数, 更改生成器输入侧和分类器输出侧的网络结构, 这样需要训练的参数大大减少, 并用目标数据集训练这一部分参数, 为了更好地把预训练的通用层应用到其目标任务上, 开发了 AdaFM 技术(对冻结的参数进行加权处理, 增加的权重和偏差是可训练的, 因此通过训练可以更好地将参数应用于目标任务), 使得冻结的层数就变得更灵活了, 并且他们的实验显示冻结相对多的层会得到更好的效果, 但过多效果会变得更差。这也说明远离图像侧的层是用于具体任务的, 通过迁移学习和 AdaFM 的运用取得很好的效果[17]。

然而, 虽然预训练的分类器很多, 预训练的 GAN 也越来越多, 但利用预训练的 GAN 的工作并不多[19]。公开的预训练 GAN 有 SNGAN (Miyato 等人, 2018)、GP-GAN (Mescheder 等人, 2018)和 BigGAN (Brock 等人, 2019)等。

以上这些能生成图像的 GAN 都不能指定生成图像的类别, 事实上, 我们缺少的主要是带标签的数据, 不带标签的数据在网上有很多。目前条件下生成对抗网络(Conditional Generative Adversarial Networks, CGAN) [20]可以生成带标签的数据, 但目前很少有公开的在大数据上预训练的 CGAN。另外, 由于 CGAN 需要在图像中增加信息, 通常的做法是生成一个和图像大小相同的标签嵌入层, 但这样一来就需要改变网络结构, 特别是靠近图像的那一层, 这层的参数就无法迁移, 导致很难迁移现有的 GP-GAN、SNGAN 等模型用于 CGAN。因此, 本文主要针对如何把迁移学习应用到条件生成对抗网络上进行研究。

研究思路: 在迁移学习中(微调迁移学习), 浅层网络用于图像通用特征(线条、色彩、纹理等)的提取[16], 那么能否将标签信息加到浅层网络后面, 即保留浅层网络的结构和参数不动, 修改其后面的网络结

构并训练其参数, 这一点是本文研究重点。

## 2. 方法

### 2.1. 研究背景概述

能够生成高清图像(1024 像素)的生成对抗网络[21][22]通常网络很深、训练时间长而且所需设备要求较高。随着 GAN 技术的发展, Brock 等人的工作[22]和 Karras 等人的 PGAN 生成了高清晰的图像, Karras 等人在训练 GAN 时, 由小到大逐渐在网络输出端增加更大网络层, 使得输出的图像尺寸一点点变大, 图像的清晰度也逐渐变好, 达到了生成高清晰图像的效果[23]。2019 年, Karras 等人在 PGAN 基础上, 在网络不同层引入噪声和风格信息, 生成了不同风格的高清人脸图像, 并且合成的图像可通过风格信息(将输入通过映射网络变为控制变量对 AdaIN 进行调整)进行控制[21]。在此基础上, 2020 年, Zhao 等人在预训练的 GAN 的基础上, 利用迁移学习技术和 Karras 等人所采用部分方法(加入噪声和类型信息的类型模块), 并给固定参数的网络层中加入了可训练的权重和偏差(AdaFM), 在小数据集(花的数据集)上训练, 生成了清晰的花的图像。

然而, Tero Karras 所采用的方法计算量非常大且所需设备要求高, 由于 Tero Karras 等人属于英伟达公司, 模型训练采用的设备一般研究者并不具备。并且生成的人脸图像, 虽然非常清晰而且可以变换风格(肤色、发型、面孔、年龄等等), 但生成的图片与眼底图像还是有很大不同, 生成的脸部图片很少有斑点, 而且人脸面部信息位置相对固定。而医学图像, 如糖尿病视网膜病变的眼底图像不同, 病灶基本为斑点且位置不固定, 真实的 DR 图像拍摄环境、角度和设备等等都不同, 所以 DR 图像相对复杂。而且, 以上方法生成的图像类别不可控。

在以上研究基础上, 本研究采用迁移学习思想, 利用 Zhao 等人的预训练 GAN 网络(和分类任务不同, GAN 的输出都是 1 和 0, 所以不需要更改网络结构), 在小数据集(DR 数据集)上重新训练 GAN 网络的部分参数(生成器的浅层参数和深层中的 AdaFM 中参数, 以及分类器的深层参数), 使之能生成清晰的 DR 眼底图像。然后, 再研究条件生成对抗网络与迁移学习的结合, 难点在于训练好的 GAN 很难加入标签信息, 而标签信息通常是和输入叠加在一起输入网络中, 导致输入的样本尺寸发生变化, 网络结构也因此必须改变。即标签信息的加入会改变网络结构, 网络结构改变后, 原来网络训练好的参数就不能再利用了, 使得迁移学习难以应用。本研究提出把标签信息加入到分类网络中间层, 保留浅层网络结构(主要用于图像基本特征提取, 需要大量数据训练)。实验结果显示, 生成的图像具有更多的类别信息(病灶特征), 但图像清晰度也略有降低。但要远远好于不加载预训练参数的网络。实验证明了在小数据集上, 迁移学习在条件生成对抗网络中的应用显著提高了训练效果。

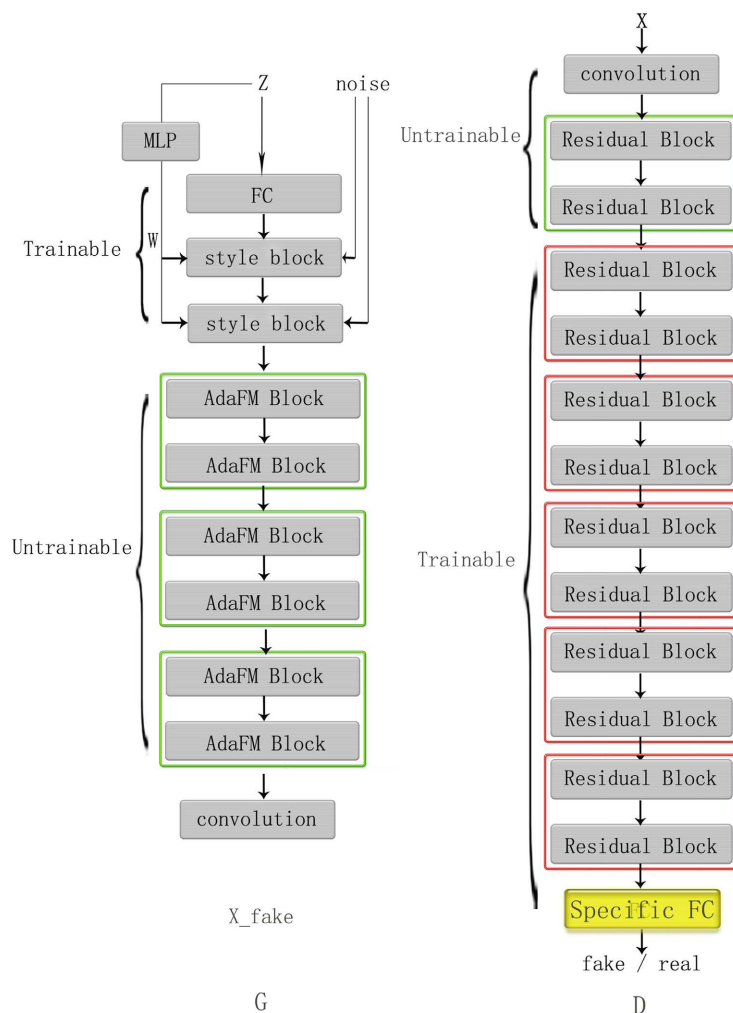
### 2.2. 基于迁移学习的糖尿病视网膜病变生成对抗网络

#### 2.2.1. 模型结构

本研究所采用的模型是利用在公开 Imagenet 大数据集上训练好的 GAN 模型[22] (如图 1), 固定生成器 G 深层网络和分类器 D 浅层网络参数, 并在本任务采用的数据集上训练, 生成了 DR 眼底图像。

图 1 中 G 为生成器模型,  $z$  为输入随机向量, 经过 FC 层后, 变成某一尺寸的常数, 这一点与传统 GAN 不同, 常数的主要作用就是给网络一个起点, 一个初始偏差。 $z$  主要作用是通过 MLP (8 层全连接层)后作为风格控制参数来调控 style block, FC 的作用是解除耦合, 即消除各个  $w$  间的相互关联和影响, 否则各个  $w$  间相互缠绕, 就如同一个  $w$ , 网络很难调整这一个  $w$  控制不同的风格。打个比方, 用一个开关控制不同灯, 按照想要的方式亮和灭, 是很难达到的。所以, 要接触各个  $w$  间的相互缠绕, 而这是通过网络学习得到的, 因为为了更好地生成图像, 网络需要  $w$  间彼此独立。





**Figure 1.** Structure of GAN

**图 1.** 生成对抗网络结构图

noise 的作用是给网络加入噪声，通过风格调控与噪声相结合调控输出图片。

图 1 G 中的类型模块(style block)如图 2 所示：

其中 AdaIN (Adaptive Instance Norm 类)是一种标准化：

$$\text{AdaIN}(x, w) = \sigma(w) \left( \frac{x - \mu(x)}{\sigma(x)} \right) + \mu(w) \quad (1)$$

$\mu$  表示均值， $\sigma$  表示标准差，通过  $w$  控制 AdaIN 输出的均值和方差，公式中  $\frac{x - \mu(x)}{\sigma(x)}$  先标准化输入，目的是去除风格类型信息，然后通过  $\sigma(w)$  和  $\mu(w)$  加入我们想要图像获得的风格。而之前加入的 noise 相当于一个定位器，告诉网络(比如人脸检测中的头发、皱纹等等)风格加入的位置。通常浅层网络从 noise 中学到的相对少，因为越往深层网络，越是细节信息，而要生成的细节信息需要通过 AdaIN 和 noise 结合起来加入到网络中，这是通过网络对比假图像和真图像之间的对比损失函数，然后通过反向传播算法来更新相关参数来达到的。也就是说，某一层风格  $w$  控制图片中的某尺度下的某些特征有多少，网络所要学习的的就是如何加入这些信息，使得假图和真图更接近，以便使得损失函数减少。

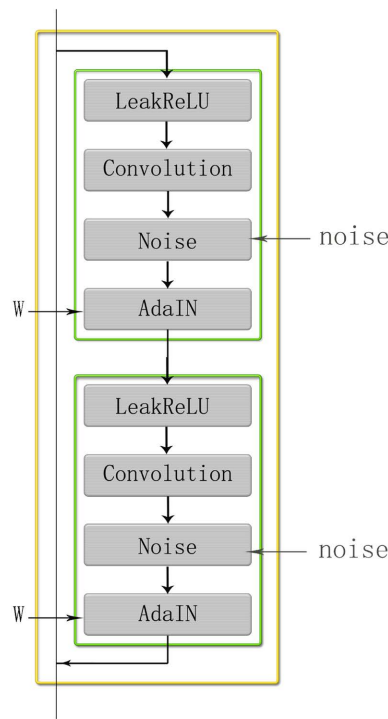


Figure 2. Structure of style module  
图 2. 类型模块结构图

图 1G 中的 AdaFM Block 是利用论文[24]中的网络, 该网络卷积层中加入可训练的权重和偏差, 采用了其中的风格混合思想, 使得网络更好地适应目标任务。训练时采用分阶段训练, 先训练浅层网络, 再训练深层网络, 目的同 SGAN 相同, 即由粗到细的训练网络生成更清晰的图像。

图 1D 为分类器, 浅层网络提取的是图像通用的信息(纹理、线条等), 这些信息通常需要大数据来训练, 通过迁移学习, 利用大数据上训练好的网络, 来缩短训练时间, 同时解决小数据集过拟合的问题。网络的输出端不同于传统分类网络(需要调整输出类别), 由于生成对抗网络标签只有 1 和 0, 所以无需调整。

### 2.2.2. 实验过程

生成器和分类器交替训练:

步骤一, 训练生成器:

生成器输入为随机向量, 长度为 100。形状为: batch size, 100, 1, 1。输出为和真实图像尺寸相同的假图像, 对应的标签为 0。真实图像对应的标签为 1。

先固定分类器, 把生成器生成的图像输入分类器, 得到输出 pred (为真实图像的概率), 然后计算损失, 对简单的二分类问题采用损失函数选择二进制交叉熵, 为了通过定时保存训练的参数和观察生成器输出的样本质量来判断学习停止时间。如果生成器生成的图像不再变得更好, 那么停止训练, 手动调整学习率, 下载生成图像最好的模型参数, 再继续训练。

网络训练时采用固定学习率 0.001, 刚开始训练网络的浅层 style block 部分, 主要考虑网络浅层控制图像的粗糙部分, 然后再训练深层网络中的可训练参数(即加入的 AdaFM)。

步骤二, 训练分类器:

采用与生成器相同的学习率, 固定网络浅层(如图 1D 所示), 训练网络深层。然后重复步骤一和二,

并不断观察生成图像的质量,直到生成的图像质量不再变好,就停止训练。并保持生成器和分类器参数,视情况调整超参数,下载刚刚训练好的模型,继续训练,直至生成的图像不能再变好为止。

## 2.3. 基于迁移学习的糖尿病视网膜病变条件生成对抗网络

### 2.3.1. 生成器模型结构

本研究所采用的模型是利用在公开大数据集上训练好的 GAN 模型,更改部分模型结构和参数,并加入分类标签信息,获得条件生成对抗网络(CGAN)。即把 Zhao 和 Choe 等人在论文[22][24]中训练的 GAN 模型,该模型如图 1G 所示(原本用于生成各种花的网络),保留卷积层参数,在生成器和分类器输入端分别插入嵌入层(注意分类器输出层不动),再用本研究采用的数据训练该网络。

由于条件生成对抗网络需要在 GAN 的生成器和分类器中分别增加标签信息,而分类器中增加标签信息面临的问题是需要更改网络结构,而更改网络结构就不能用迁移学习了。即使只更改分类器的输入层,也会影响很大,因为分类器的输入层是用于提取图像基本特征(纹理、线条、色彩等)的,而这样的网络参数需要很大的数据才能训练好,考虑到这一点,本研究采用把嵌入层叠加到分类器通用层(浅层)的后面,如图 1D 中的 yembedded block 中。这样既避免了更改网络结构所导致的无法利用迁移学习,又将普通 GAN 转换为 CGAN。

本研究所采用的生成器网络模型如图 3(b)所示,图 3(a)为原网络模型。

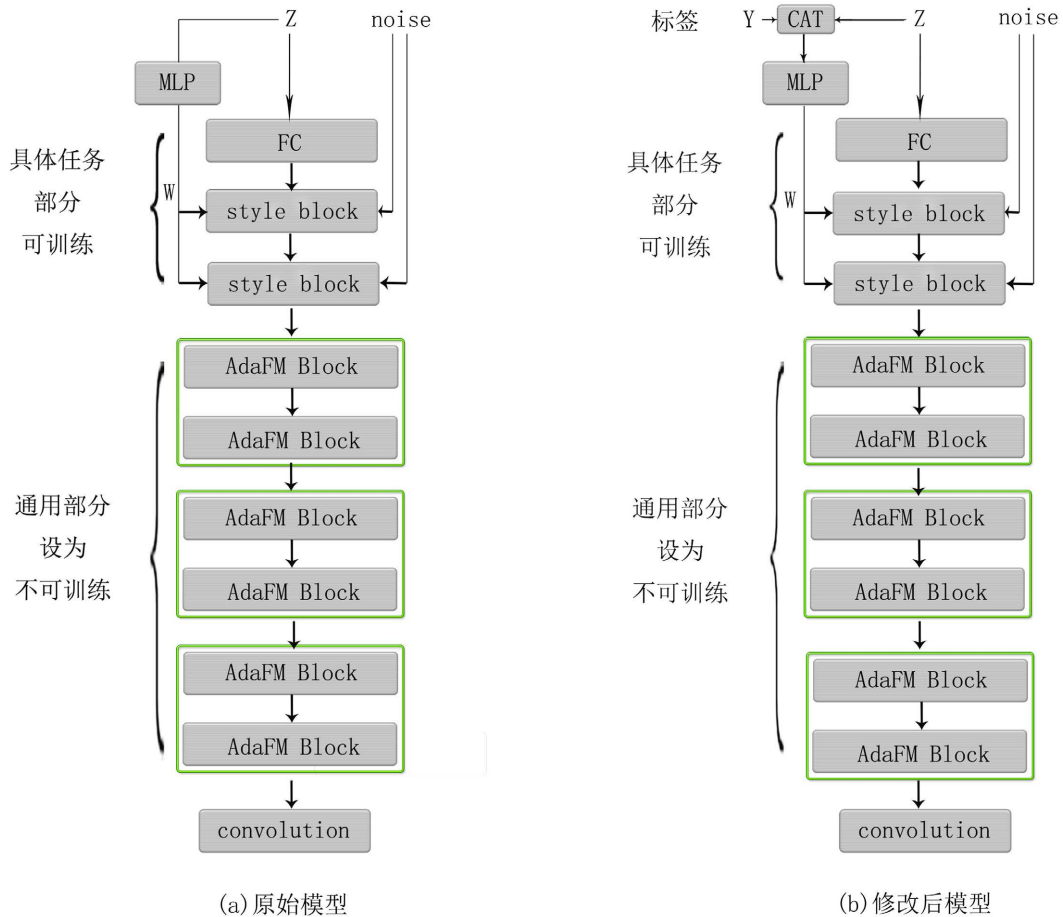
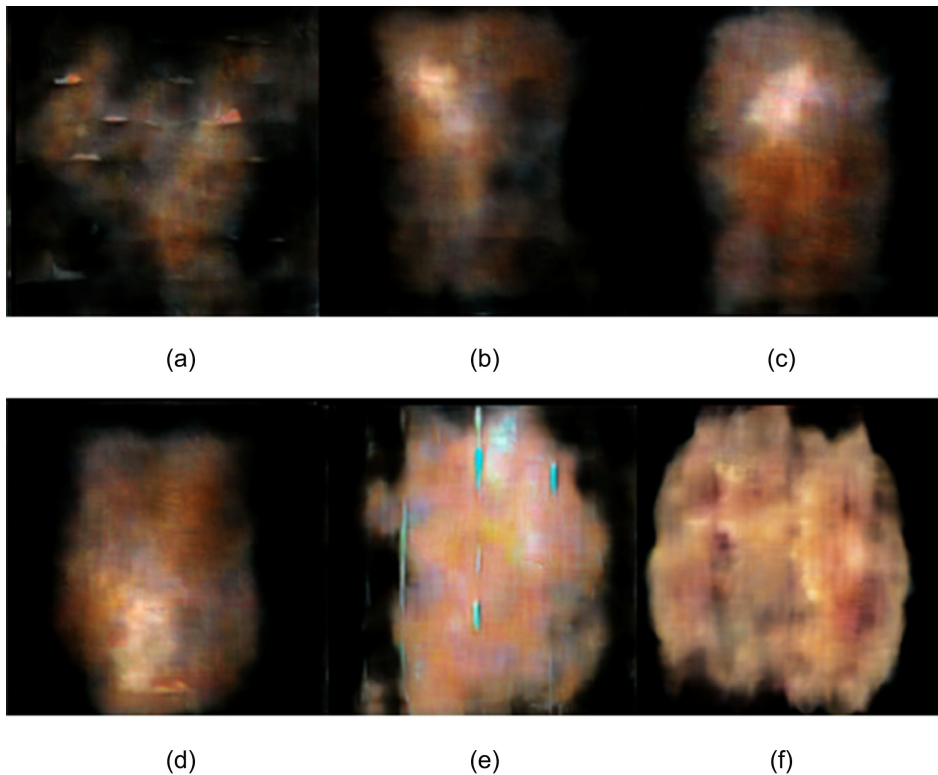


Figure 3. Structures of GAN and CGAN

图 3. 生成对抗网络和条件生成对抗网络结构图

虽然仅仅在输入层中加入了类别标签和随机变量叠加在一起, 但对整个网络影响非常大, 简言之, 会使得输出面目全非, 因为这导致了输入样本分布发生了变化, 因此需要重新训练部分网络参数。

在图 3(b)中, 即本研究所用的生成器输入端, 类别标签先经过嵌入层, 变为和标签尺寸相当的标签嵌入层  $Y$ , 嵌入层与输入随机向量  $Z$  有 Pytorch 中的 `cat` 函数, 叠加在一起输入网络 MLP(多层全连接层, 用于将标签信息和输入随机向量  $Z$  进行变换来控制网络模型输出)。图 3(b)中 MLP、FC、style block 为可训练层, 其中 MLP 是用于图像微调, 而 FC 和 style block 是用于控制图像生成大的轮廓(如眼底图像的圆形)等低频特征, 因此训练时先训练 FC 和 style block, 再训练 MLP。如果同时训练这三个部分, 则很难生成眼底图像, 即生成的图像会出现各种形状(如图 4 所示), 而不是由方形逐渐变成圆形, 细节信息逐渐增加的过程。最后再训练 AdaFM 中的参数, AdaFM 中卷积参数是固定不训练的, 而通过 AdaFM 加的权重和偏差是可以训练的, 通过这种方式, 更好地使用原网络适应目标任务(训练过程生成的图片如图 7 所示)。



**Figure 4.** Training process of GAN without pre-trained  
**图 4.** 模型整体训练的训练过程图

### 2.3.2. 分类器模型架构

要利用迁移学习, 网络结构不能变。但考虑到条件生成对抗网络必须要增加嵌入层, 因此需要更改网络结构, 这给迁移学习带来了障碍。且考虑到分类器中的浅层网络是提取图像通用特征的, 因此设计保留浅层网络结构, 而把标签嵌入到其后的网络中, 如图 5(b)中的黄色 yembeded block 层。即分类器的标签嵌入层是叠加到固定不训练层的后一层上, 并相应更改了其下一层的网络结构。在相应卷积层的输入通道加一, 并把标签生成和该层输入尺寸相匹配, 然后叠在一起输入到接下来的网络中。网络的输出层不动, 即输出通道数为 1。



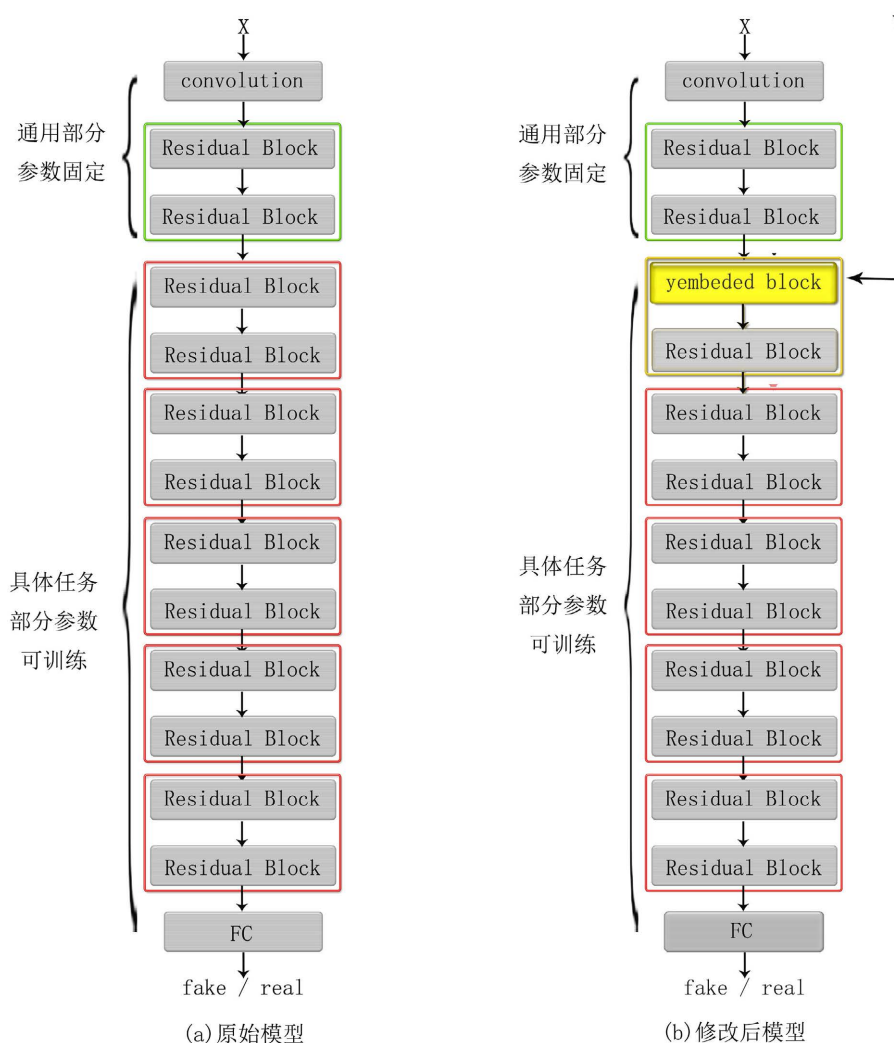


Figure 5. Structure of classifier of CGAN  
图 5. 条件生成对抗网络分类器结构图

## 2.4. 训练过程

数据来源于网上公开数据集 Kaggle 失明检测公开竞赛数据集, 数据集是从多家医院收集的, 其中的图像和标签都含有噪声, 图像是很多成像设备拍摄的, 有的图像包含伪影、曝光不够或曝光过度、不聚焦和尺寸不同等缺陷。这和真实数据集很像。数据集中包括 35,126 张眼底图像, 各个类别数量为(严重程度递增): 25,810, 2443, 5292, 873, 708。其中 33,545 张(前三类)是不影响视力的眼底图像, 其它 1581 张(后两类)是对视力造成伤害的 DR 眼底图像, 显然数据存在分布不均的情况。

由于数据不平衡, 用 DataLoader 进行取样时, 网络会得到更多无症状的 DR 图像, 那么网络会更加倾向于生成无症状的眼底图像。因此, 为了减少这一弊端采用过采样方法, 即按照放回抽样的方式, 从各个类别中均匀抽取样本, 根据样本分布  $d$  设计取样器权重  $w$ 。

$$d = [25810, 2443, 5292, 873, 708] \quad w = [0.1, 1, 0.5, 3, 4]$$

首先, 导入数据, 由于数据各个类别是在不同文件夹下, 用 torchvision.datasets.ImageFolder 方法导入数据。然后, 建立模型, 并根据模型参数名导入预训练的参数, 设置优化器为 Adam。接着, 根据需要固

定的参数，将预训练好的模型参数进行加载。然后固定分类器靠近输入端的两层神经网络。把生成器生成的图像和对应类别标签(输入一)和真实图像和类别标签(输入二)，分别输入分类器。

3. 实验结果与分析

从头开始训练的 GAN 训练过程如图 6 所示：

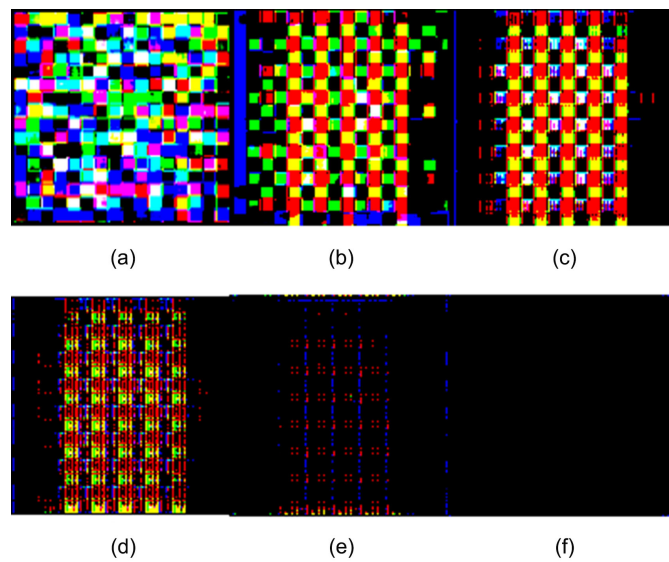


Figure 6. Output images of GAN without pre-trained  
图 6. 无预训练的生成对抗网络训练过程输出图像

从图 6 可见，数据量小的网络得不到有效训练，图 6(a)为网络开始训练 GAN 生成的图像，然后继续训练(从图 6(a)到图 6(f))，网络最后生成的是黑图(如图 6(f))，然后网络生成的图片不再改变了。  
基于迁移学习的 GAN 训练过程，生成器输出图像变化如图 7 所示：

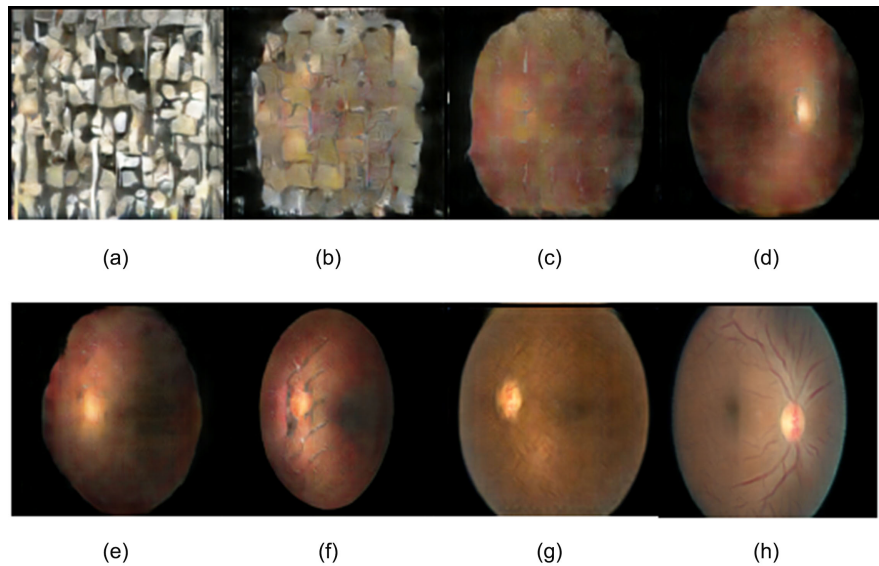
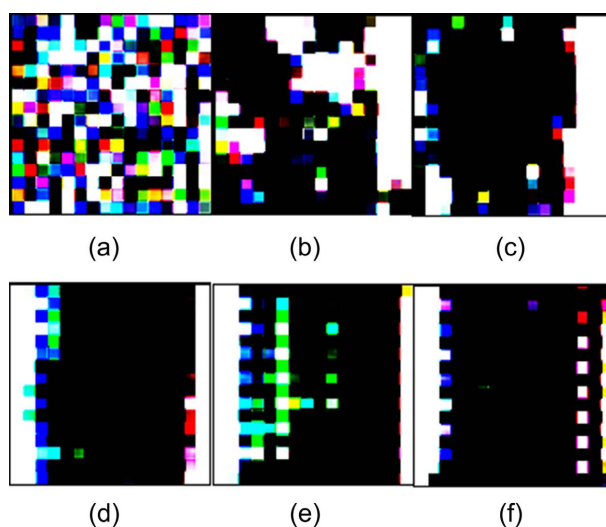


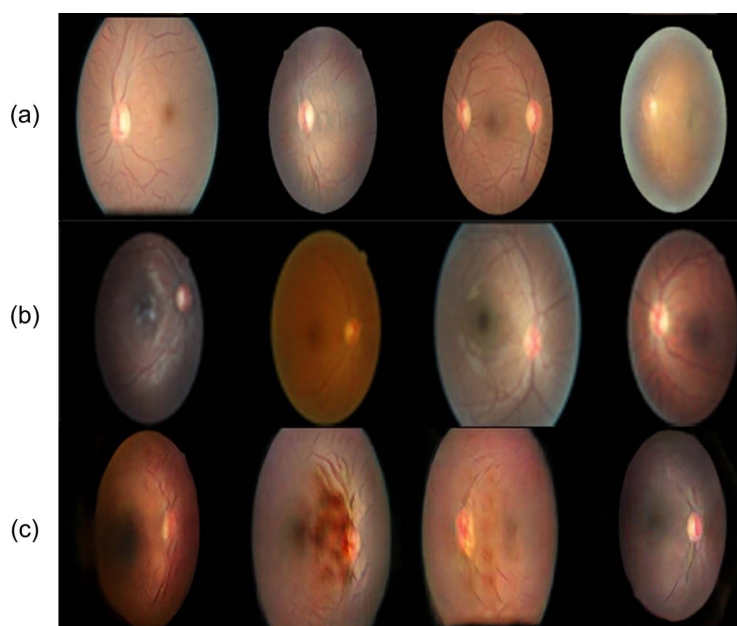
Figure 7. Output images of pre-trained GAN  
图 7. 预训练的生成对抗网络训练过程输出图像

对比基于迁移学习的条件生成对抗网络和从零开始训练的条件生成对抗网络可见：迁移学习明显加快了网络训练进度。从刚开始训练基于迁移学习的 GAN 网络输出的图像可见，图 7 中开始是一些花的图案，随着训练的增加，生成器网络输出图像形状一点点变圆，图案由花的绿色和白色一点点变为眼底图像的黄红等颜色，而且眼底图像的视盘很快就出现了，网络的血管开始并没有出现，原因在于训练过程，即刚开始网络训练浅层参数，不包括 AdaFM Block，设置训练 1000 个批次后再将其加入训练，而且随着训练轮次增加学习率也相应降低，这使得生成器网络生成的图像细节信息更为丰富。

如图 8，从零开始训练的 CGAN，从形状上难以达到圆形，而且基本的图像特征也难以生成，最后生成的图像在类似图 8(e)和图 8(f)间左右徘徊，这反映了数据小造成的模型无法变得更好了。



**Figure 8.** Output images of CGAN without pre-trained  
**图 8.** 无预训练的条件生成对抗网络训练过程输出图像



**Figure 9.** Comparison between real images and generative images  
**图 9.** 生成图像与真实图像对比

图 9 中, 图 9(a)为基于迁移学习的生成对抗网络生成的图片, 图 9(b)为真实图片, 图 9(c)为基于迁移学习的条件生成对抗网络生成的图片。

从图 9(a)中可见, 第三幅图像有两个视盘, 这与真实图像显然不同, 但图 9(a)中血管生成的还是很明显, 虽然与真实图像有一点点区别, 但和真实眼底图像比较接近。图 9(b)是真实眼底图像, 有些图像血管和病灶也不是很明显, 这也是医生判断不准确的原因, 因为数据集来源于不同医疗机构, 数据质量并不是很理想, 反映成像设备不同, 并且拍摄角度和环境也各不相同。这也给深度学习带来挑战, 深度学习模型需要根据不同场景和设备拍摄的图像学习分类 DR 病情。

图 9(c)是基于迁移学习的条件生成对抗网络生成的图像, 和图 9(a)相比有更明显的细节病灶, 这反映了加入条件信息提高了对生成器的可控性。通过迁移学习和条件生成对抗网络的结合, 不但生成了 DR 图像, 也对生成图像中类别增加了一定的控制。

不论是使用迁移学习和生成对抗网络结合, 还是迁移学习与条件生成对抗网络结合, 都避免了图 6 和图 8 中模型没有学到如何生成眼底图像的问题; 而通过采用预训练模型, 正如图 7 和图 9 所示, 生成模型能够成功生成逼真的眼底图像。从从头训练的 GAN 和 CGAN 生成图像的变化可见, 小数据集难以训练 GAN, 这反映了大数据对深度学习的重要性, 也反映小数据集需要迁移学习来充分利用数据, 来训练那些专用于特定任务的网络层, 即生成器的深层和分类器的浅层网络需要大数据来训练。

非预训练的生成对抗网络和条件生成对抗网络难以在小数据集上生成似真的眼底图像, 从生成图像的过程可见, 无迁移学习网络难以学习, 训练经过相同轮次训练后生成的图像与真实图像相去甚远, 而且难以继续提高图像质量。研究认为, 由于数据集太小, 难以训练过多的参数, 所以, 不论训练多少轮最后都难以生成逼真的图像。而在迁移学习基础上, 大大减少了所需训练的参数, 这样小数据集训练较少的参数可以达到较好的效果。

通过 FID 度量比较基于迁移学习和无迁移学习 GAN 生成图片质量和多样性如表 1:

**Table 1.** Score of FID  
**表 1.** FID 分数

方法	FID 分数
无迁移学习 GAN	293.9
迁移学习 GAN	86.3
无迁移学习 CGAN	287.7
迁移学习 CGAN	124.5

由表 1 可见, 无论是 GAN 还是 CGAN, 迁移学习大大提高了小数据集上生成对抗网络生成图片的质量和多样性, 基于迁移学习的 GAN 生成的眼底图片质量要远远好于无迁移学习 GAN 生产的图片质量。

#### 4. 结论

深度学习在糖尿病视网膜病变(DR)自动筛查领域的应用还面临一些困难: 医学数据通常是小数据集, 而且数据存在不平衡现象。因此, 本研究利用生成对抗网络、迁移学习对这项任务进行研究: (1) 研究利用迁移学习搭建生成对抗网络(GAN)生成 DR 眼底图像, 建立了基于迁移学习的生成对抗网络, 实现了在小数据集上生成逼真的糖尿病视网膜病变眼底图像。这种方法避免了 GAN 需要大数据来训练的弊端; (2) 研究利用预训练的生成对抗网络构造条件生成对抗网络, 增加对生成 DR 眼底图像的控制, 提出了基于迁移学习的条件生成对抗网络, 实现了条件生成对抗网络和迁移学习成功结合, 这增加了对生成图像的

控制, 使得生成眼底图像有更多的病灶信息。

## 参考文献

- [1] 陈佛计, 朱枫, 吴清潇, 等. 生成对抗网络及其在图像生成中的应用研究综述[J]. 计算机学报, 2021(44): 347-369.
- [2] Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., et al. (2014) Generative Adversarial Networks. *27th International Conference on Neural Information Processing Systems*, Bangkok, 8-13 December 2014, 2672-2680.
- [3] Singh, N.K. and Raza, K. (2021) Medical Image Generation Using Generative Adversarial Networks: A Review. In: *Studies in Computational Intelligence*, Springer, 77-96. [https://doi.org/10.1007/978-981-15-9735-0\\_5](https://doi.org/10.1007/978-981-15-9735-0_5)
- [4] Chuquicuma, M.J.M., Hussein, S., Burt, J. and Bagci, U. (2018) How to Fool Radiologists with Generative Adversarial Networks? A Visual Turing Test for Lung Cancer Diagnosis. 2018 *IEEE 15th International Symposium on Biomedical Imaging*, Washington, 4-7 April 2018, 240-244. <https://doi.org/10.1109/isbi.2018.8363564>
- [5] Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J. and Greenspan, H. (2018) Gan-Based Synthetic Medical Image Augmentation for Increased CNN Performance in Liver Lesion Classification. *Neurocomputing*, **321**, 321-331. <https://doi.org/10.1016/j.neucom.2018.09.013>
- [6] 林志鹏, 曾立波, 吴琼水. 基于生成对抗网络的宫颈细胞图像数据增强[J]. 科学技术与工程, 2020, 20(28): 11672-11677.
- [7] Plassard, A.J., Davis, L.T., Newton, A.T., Resnick, S.M., Landman, B.A. and Bermudez, C. (2018) Learning Implicit Brain MRI Manifolds with Deep Learning. *Medical Imaging 2018: Image Processing*, Houston, 11-13 February 2018, Article 105741. <https://doi.org/10.1117/12.2293515>
- [8] Jin, D., Xu, Z., Tang, Y., et al. (2018) CT-Realistic Lung Nodule Simulation from 3D Conditional Generative Adversarial Networks for Robust Lung Segmentation. <https://ui.adsabs.harvard.edu/abs/2018arXiv180604051J>
- [9] Mok, T.C.W. and Chung, A.C.S. (2019) Learning Data Augmentation for Brain Tumor Segmentation with Coarse-to-Fine Generative Adversarial Networks. In: *Lecture Notes in Computer Science*, Springer, 70-80. [https://doi.org/10.1007/978-3-030-11723-8\\_7](https://doi.org/10.1007/978-3-030-11723-8_7)
- [10] Chartsias, A., Joyce, T., Dharmakumar, R. and Tsafaris, S.A. (2017) Adversarial Image Synthesis for Unpaired Multi-Modal Cardiac Data. In: *Lecture Notes in Computer Science*, Springer, 3-13. [https://doi.org/10.1007/978-3-319-68127-6\\_1](https://doi.org/10.1007/978-3-319-68127-6_1)
- [11] Maspero, M., Savenije, M.H.F., Dinkla, A.M., Seevinck, P.R., Intven, M.P.W., Jurgensliemk-Schulz, I.M., et al. (2018) Dose Evaluation of Fast Synthetic-CT Generation Using a Generative Adversarial Network for General Pelvis MR-Only Radiotherapy. *Physics in Medicine & Biology*, **63**, Article 185001. <https://doi.org/10.1088/1361-6560/aada6d>
- [12] Yang, H., Sun, J., Carass, A., Zhao, C., Lee, J., Xu, Z., et al. (2018) Unpaired Brain MR-to-CT Synthesis Using a Structure-Constrained CycleGAN. In: *Lecture Notes in Computer Science*, Springer, 174-182. [https://doi.org/10.1007/978-3-030-00889-5\\_20](https://doi.org/10.1007/978-3-030-00889-5_20)
- [13] Dugas, E., Jared, Jorge and Cukierski, W. (2020) Diabetic Retinopathy Detection. <https://www.kaggle.com/c/diabetic-retinopathy-detection>
- [14] Zhou, Y., Wang, B., He, X., et al. (2020) DR-GAN: Conditional Generative Adversarial Network for Fine-Grained Lesion Synthesis on Diabetic Retinopathy Images. *IEEE Journal of Biomedical and Health Informatics*, **26**, 56-66.
- [15] 马崇慕, 唐娉, 赵理君, 等. 深度学习图像数据增广方法研究综述[J]. 中国图象图形学报, 2021(26): 487-502.
- [16] Pan, S.J. and Yang, Q. (2010) A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, **22**, 1345-1359. <https://doi.org/10.1109/tkde.2009.191>
- [17] Zhao, M., Cong, Y. and Carin, L. (2020) On Leveraging Pretrained GANs for Generation with Limited Data. <https://ui.adsabs.harvard.edu/abs/2020arXiv200211810Z>
- [18] Mo, S., Cho, M. and Shin, J. (2020) Freeze the Discriminator: A Simple Baseline for Fine-Tuning GANs. <https://ui.adsabs.harvard.edu/abs/2020arXiv200210964M>
- [19] Giacomello, E., Loiacono, D. and Mainardi, L. (2020) Brain MRI Tumor Segmentation with Adversarial Networks. 2020 *International Joint Conference on Neural Networks*, Glasgow, 19-24 July 2020, 1-8. <https://doi.org/10.1109/ijcnn48605.2020.9207220>
- [20] Mirza, M. and Osindero, S. (2014) Conditional Generative Adversarial Nets. *Computer Science*, **2014**, 2672-2680.
- [21] Karras, T., Laine, S. and Aila, T. (2019) A Style-Based Generator Architecture for Generative Adversarial Networks. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 4396-4405. <https://doi.org/10.1109/cvpr.2019.00453>



- [22] Brock, A., Donahue, J. and Simonyan, K. (2018) Large Scale GAN Training for High Fidelity Natural Image Synthesis. <https://ui.adsabs.harvard.edu/abs/2018arXiv180911096B>
- [23] Karras, T., Aila, T., Laine, S., *et al.* (2018) Progressive Growing of GANs for Improved Quality, Stability, and Variation. *6th International Conference on Learning Representations*, Vancouver, 30 April-3 May 2018, 10-15.
- [24] Choe, J., Oh, S.J., Lee, S., Chun, S., Akata, Z. and Shim, H. (2020) Evaluating Weakly Supervised Object Localization Methods Right. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 13-19 June 2020, 3130-3139. <https://doi.org/10.1109/cvpr42600.2020.00320>