

改进 GAN 网络在生成短视频的应用研究

于晓明 黄 铎

陕西科技大学电子信息与人工智能学院 西安 710021

(494636031@qq.com)

摘 要 在研究生成对抗网络(GAN)生成动态图像时,经常出现前后帧图像内容中的部分物体颜色不一致和生成的细节不自然等问题。针对当前生成视频的不理想问题,采用的主要方案是分别对 GAN 网络中的生成器和判别器进行改进,具体表现在两个方面:一方面是在生成器中对视频的前景和背景分别建模,并且使用多重空间自适应归一化(Multi Spatially-Adaptive Normalization, M-SPADE)算法;另一方面是在判别器的选取上使用双视频判别器(DVD-GAN),然后在 Kinetics-600 数据集进行训练,实验后的结果分别比对 F-Vid2Vid, WC-Vid2Vid 等生成方法。实验结果证明了对 GAN 网络改进的方法在处理生成短视频的前后帧颜色不一致的问题和细节上有着不错的效果,生成的图像相对的更加清晰。

关键词: 生成对抗网络;合成短视频;前景-背景图像建模;多重空间自适应归一化;双视频判别器

中图分类号 TP391

Research on Application of Improved GAN Network in Generating Short Video

YU Xiao-ming and HUANG Hua

College of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China

Abstract In the study of the dynamic image generated by GAN, there are many problems, such as inconsistent colors of some objects and unnatural details of generated images. In order to solve the problem of unsatisfactory video generation, the main schemes adopted are to improve the generator and discriminator of GAN network respectively, which are shown in two aspects. On the one hand, the foreground and background of the videos are modeled separately in the generator and the Multi Spatial-Adaptive Normalization (M-Spade) algorithm is used. The other aspect is the use of dual video discriminator (DVD-GAN) on discriminator selection, which trained on Kinetics 600 dataset. The experimental results are compared with F-VID2VID, WC-VID2VID and other generation methods. The results show that the method of combining the two methods has a great effect on the problem of color inconsistency before and after the short video and the details processing, and the generated images are relatively clearer.

Keywords Generative adversarial networks, Synthesis short video, Foreground-background image modeling, Multispace adaptive normalization, Dual video discriminator

1 引言

近年来,随着对计算机视觉的深入研究,生成图像的发展也越来越迅速,因此对 GAN 网络模型的研究也始终处于最前沿^[1],许多生成动态图像也以 GAN 网络为方法基础。

在 Kinetics-600 数据集上训练 GAN 网络能够产生比以前工作更复杂、质量更高的视频样本,可以更好地获得模型结构。但是生成模型很容易受到两个方面的困扰,一是数据复杂性,二是计算量的增加。基于上述问题,许多研究生成视频的学者提出了不同的见解,希望通过修改网络结构来达到目的,如 Jia 等^[2]与 Brock 等^[3]都对改进训练稳定性提出看法;Nitish 等^[4]重新设计回归模型;Yamaguchi 等^[5]在生成模型中使用归一化流等。

当前,研究生成视频正处于蓬勃发展时期,许多优化后的 GAN 网络模型在合成视频领域中取得了很好的成果。Wang

等^[6]首次提出了视频到视频的合成问题(Vid2Vid 框架),提出了把光流约束加入到生成器中、把光流信息加入到判别器中和对输入的图像序列进行前景-背景分开建模的方法。Wang 等^[7]对 Vid2Vid 框架进行延伸(F-Vid2Vid 模型),通过使用一个新的网络权重生成模块来实现 Vid2Vid 框架所不具备的泛化能力。Mallya 等^[8]对 Vid2Vid 框架进行优化后提出使用导航图像的方法去实现图像的时间一致性(WC-Vid2Vid 模型)。Clark 等^[9]提出在判别器中使用双视频判别器(DVD-GAN),然后利用判别器的计算分解把复杂的视频扩展成更长和质量更高的视频。

上述方法在合成视频领域中取得了很好的成绩,但也存在以下不足之处。生成模型不能保证目标对象会随着时间的推移具有一致的外观;同时,输入的标签图信息在合成过程中会出现消失的情况;网络模型主要基于输入图像,如果输入出现失败,整个网络也将失败。基于此,本文在 vid2vid 框架的

基金项目:陕西省科技厅项目(2014KRM80);咸阳市科技局项目(2013K15-07)

This work was supported by the Science and Technology Department of Shaanxi Province, China(2014KRM80) and Project of Xianyang Science and Technology Bureau, China(2013K15-07).

通信作者:黄铎(1792156660@qq.com)

基础上选取对输入图像序列中的图像进行前景-背景分开建模和在判别器中使用 DVD-GAN 相结合的方式。在相同的实验下,本模型与对比模型相比,效果更优。

2 生成视频原理

生成视频建模是一个经常探讨的问题。因此生成视频的主要研究方向一个是无监督视频合成,即在训练生成任何随机的视频;另一个是有监督模型,包括图片序列以内容形式传输到另一个视频为条件的生成^[10-11]、每帧进行分割掩模等;还有一种处于两者之中,比无监督更加结构化,但是在建模领域上更加具有挑战性。生成视频主要采用半监督模式,这种模式可以生成给定物体类别的视频。

研究生成视频的主要目的是学习一个映射函数,可以使一系列的输入语义图像 S_1^T ($S_1^T = \{S_1, S_2, \dots, S_T\}$) 生成相对应的输出图像序列 \tilde{S}_1^T ($\tilde{S}_1^T = \{\tilde{S}_1, \tilde{S}_2, \dots, \tilde{S}_T\}$), X_1^T ($X_1^T = \{X_1, X_2, \dots, X_T\}$) 表示对应实际目标图像序列,那么对应的映射函数应该满足条件概率,如式(1)所示,对应的转换图像过程如图 1 所示。

$$P(\tilde{S}_1^T | S_1^T) = P(X_1^T | S_1^T) \quad (1)$$


图 1 图像转换过程

Fig. 1 Image conversion process

对视频转换问题的建模采用 Markov 假设因式分解的条件分布,如式(2)所示:

$$P(\tilde{S}_1^T | S_1^T) = \prod_{t=1}^T P(\tilde{S}_t | \tilde{S}_{t-L}^-, S_{t-L}^-) \quad (2)$$

其中, \tilde{S}_t 表示当前的源序列帧, S_{t-L}^- 表示经过的 L 帧图像, \tilde{S}_{t-L}^- 表示过去 L 帧后生成的帧。最后通过训练卷积神经网络(CNN)对条件分布 $P(\tilde{S}_1^T | S_1^T)$ 进一步建模。

$$\tilde{S}_t = F(\tilde{S}_{t-L}^-, S_{t-L}^-) \quad (3)$$

3 相关理论

3.1 生成对抗网络

GAN 网络由生成器 Generator(G)网络和判别器 Discriminator(D)网络组成^[1],是由生成器 G 和判别器 D 两者相互之间进行博弈的生成模型。GAN 网络生成视频的结构如图 2 所示。

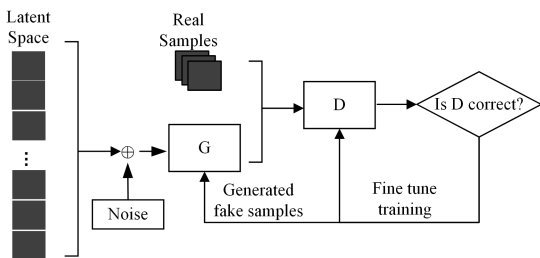


图 2 GAN 网络生成视频结构图

Fig. 2 Structure diagram of GAN network generated video

使用 GAN 网络生成模型对输入视频序列进行相对应匹

配, $\tilde{S}_1^T = G(S_1^T)$, 其中 G 为输入序列到相对应的输出序列的生成器,所以可以通过对极大极小优化问题的求解来训练生成器。

3.2 方法介绍

Pirsiavash 等^[12] 提出在 G 网络中将前景和背景进行分开建模,然后通过一个 Mask 加权将前景和背景相结合。采用分开建模的原因主要有两点:一是背景在实际情况中极少发生变化,在视频的多帧图像中背景信息都是相同的,因此对背景建模只需要考虑“空间”特点而忽略“时间”维度即可;二是由于前景信息特点相对复杂,涉及到的不仅有人还有其他物体的运动变化,对“时间”和“空间”维度都需要考虑。因此对前景-背景分开建模可以在极大程度上保证生成的连续帧中的物体轮廓图像被识别为同一物体。

M-SPADE 模块结构最初是由 Park 等^[13] 提出,并用到的 GauGAN 网络中。与生成图像的 GAN 网络的不同在于,其并不是简单地将卷积层、归一化层以及非线性层堆叠在一起构成生成模型,而是输入语义分割的 Mask 图。所以 M-SPADE 模块的主要思想是使用语义标签、光流扭曲和引导图像作为条件输入,保证了生成的视频具有真实性,能更好地处理对细节的刻画,从而展示物体的长期一致性。因此在本文中,对输入的语义图先进行前景-背景建模,建模后的图像经过 M-SPADE 模块优化后可以生成最终的图像。

在 D 网络中使用 DVD-GAN 模型中的空间判别器(Spatial Discriminator, D_S)和时间判别器(Temporal Discriminator, D_T),利用两个判别器的强大计算效率,可以把视频缩放到更长和更高分辨率。综上所述,在 G 网络中对前景-背景分开建模可以更加方便 D 网络中的 DVD-GAN 网络进行判别,进而使识别速率得到提高。

3.3 数据集介绍

Kinetics-600 数据集由 50 万段 10 s 高分辨率的视频组成, Kinetics-600 的特征具有多样化,因此能够缓解 Overfitting(过拟合)对实验的干扰,不会出现对已知数据可以预测的很好而对未知数据很难预测的现象。

Kinetics-600 里面包括 600 个类,每个类至少有 600 个视频,总共约 5×10^5 万个视频,在对数据集进行选取时,因为硬件条件的限制,并不会全部使用,通常使用最优的视频片段,即对某一个类进行实验,本次实验选取的是人物动态视频。动态视频具有多样性,而且不受约束,因此在训练模型时不需要关心固定对象以特定方式交互的小数据集上发生的过度拟合^[14]。

4 模型设计

4.1 前景-背景建模

GAN 网络主要应用在图像生成,由于在视频领域中复杂性较高,因此不能直接对视频进行建模,尤其是在连续的多帧图像中,否则会产生误差累积的问题。对复杂视频建模采用前景和背景分开建模的方式,将图像中稳定不变的部分归纳为背景,而发生运动的部分归纳为前景。

前景(Foreground)建模,使用 3D 转置卷积神经网络。3D 转置卷积神经网络在生成视频的时候可以保证空间和时间的一致性,有效地防止误差的累积而产生的视频内容的漂移。3D 转置卷积核第一层的大小为 $2 * 4 * 4$,其余的是 $4 *$

4×4 ,步长为 2,卷积核的选取可以更好地提升训练速度和生成视频的质量。

背景(Background)建模,使用 2D 转置卷积神经网络。根据噪声来生成背景,2D 转置卷积神经网络的网络结构可以直接借鉴深度卷积神经网络(DCGAN)网络的基本结构。

最后将背景图像和前景图像进行加权得到最终的视频输出。G 网络的网络结构如图 3 所示。

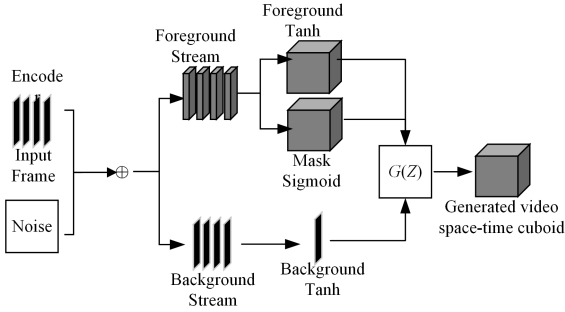


图 3 G 网络中分开建模的网络结构

Fig. 3 Network structure modeled separately in G network

在图 3 中显示的输入是高斯噪声和视频编码框架。流向 G 网络有两条独立的流向:一种是动态前景路径,另一种是静态背景路径,使用 mask 的方法融合生成视频的运动前景和静态背景。网络采用半监督模型,因此网络的输入不能为一个随机的噪声,需要在 G 网络中增加一个 Encoder 部分,然后通过 Encoder 编码成一个 Latent code 和样本的 Noise code 同时作为输入,同时,还需要添加一个损失函数,使生成视频的第一帧图片和输入的图片是相等。损失函数如式(4)所示:

$$L = \min_{W_G} \max_{W_D} E_{S \sim P_S(S)} [\log D(S, W_D)] + E_{S_0 \sim p_{S_0}(S_0)} [\log (1 - D(G(S_0; W_G); W_D))] + E_{S_0 \sim p_{S_0}(S_0)} [\lambda \| S_0 - G^0(S_0; W_G) \|_2^2] \quad (4)$$

其中, S 是从一个正态分布中取样, $S \sim P_S(S)$ 则从数据分布中取样,但是在实践中,由于我们不知道数据 $P_S(S)$ 的真实分布,因此,可以通过数据进行集中绘制,从而来估计预期值。 S_0 表示输入 $G^0(\cdot)$ 的第一个坐标系的第一帧,其中, $\lambda \in R$ 为超参数,两个网络 G 和 D 对参数 W_G 和 W_D 是可微的。

最后使用显性表达式约束网络,如式(5)所示:

$$G(z) = m(z) \otimes f(z) + (1 - m(z)) \otimes b(z) \quad (5)$$

其中, $m(z)$ 表示 Mask 函数, $f(z)$ 表示生成的前景, \otimes 符号表示基于元素的乘法, $b(z)$ 表示生成的背景。在假设参考静态背景的情况下,为了使 $b(z)$ 的维度与 $m(z)$ 和 $f(z)$ 的维度相同,应该对 $b(z)$ 沿着时间进行复制,这样可以得到最终的生成视频为 $G(z)$ 。在对数据集进行训练的时候,需要给 Mask 加了一个正则项 $\lambda |m(z)|_1$,设置权重为 0.1,目的是让 Mask 不会变得宽泛,最终得到的视频的大小为 $32 \times 64 \times 64$ 。

4.2 多重空间自适应归一化

对前景-背景分开建模后的图像输入到由多组 SPADE 层组成生成器中的 Multi-SPADE 模块中。每个 SPADE 层都获取语义分割后的图,并对中间特征图进行适当的转换,以使最终生成的输出图像看上去更加逼真,细节展示更加自然、真实。Multi-SPADE 网络结构如图 4 所示。

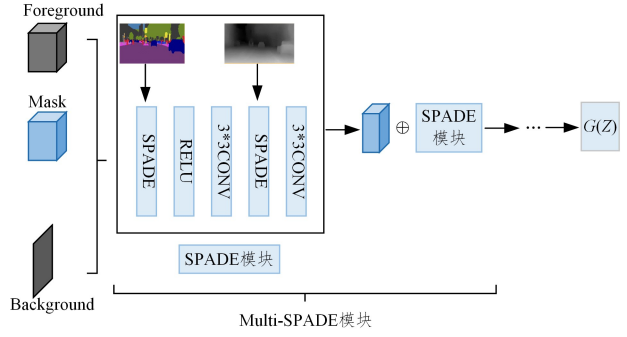


图 4 Multi-SPADE 网络结构

Fig. 4 Multi-SPADE network structure

4.3 DVD-GAN 网络

基于 BigGAN 网络模型构建的 DVD-GAN 网络,主要用于调整视频生成,使 DVD-GAN 能够在 Kinetics-600 上进行训练^[15]。

DVD-GAN 有两个判别器,分别为空间判别器和时间判别器, D_S 通过随机采样全分辨率帧然后进行单独处理,再来评估每个帧的内容以及结构; D_T 则提供一个可以生成动作的学习信号,在 DVD-GAN 网络中还有一个单独的 Transformer 模块,它可以在整合模型中传播学习信息。

D_T 向 G 提供信号来产生不被 D_S 评判的东西,因此可以不对批处理进行二次采样^[16],而是选择将 D_S 的下采样功能 $\varphi(\cdot)$ 应用于整个视频,然后将其输出到 D_T , φ 选择大小为 2×2 的平均池化函数。 D_S 与 MoCoGAN 中的每帧判别器类似^[17],但是 D_S 与 MoCoGAN 的不同在于 D_S 是 DVD-GAN 网络中高分辨率信号的唯一来源,因此 φ 不是固定函数。DVD-GAN 模型判别器网络结构如图 5 所示。

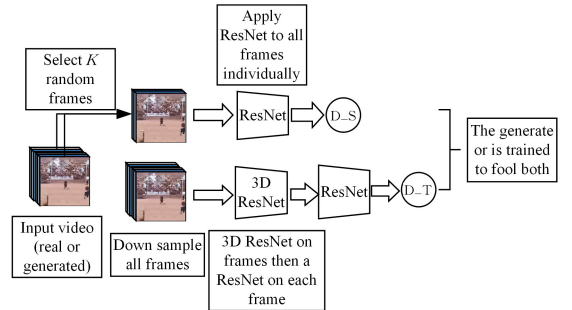


图 5 DVD-GAN 模型判别器网络结构

Fig. 5 DVD-GAN model discriminator network structure

4.4 目标函数

为使生成器生成的图像更加逼真,判别器可以更好地对生成的图像进行判别,对目标函数 \tilde{F} 的设计应如式(6)所示:

$$\tilde{F} = \min_F \max_D (L(F, D) + \lambda * L) \quad (6)$$

在对生成模型建模时为避免产生大量的冗余信息,分别在 G 和 D 网络中引用 Flow 模型,通过使用 CNN 网络进行学习来达到对气流信息的计算,计算结果也可以对下一帧图像进行预测,设置 3 个参数 W_i, H_i, M_i 分别表示当前输出的气流信息、结果和模糊程度,如式(7)~式(9)所示:

$$M_i = M(\tilde{S}_i^{-1}, S_{i-T}^{-1}) \quad (7)$$

$$W_i = W(\tilde{S}_i^{-1}, S_{i-T}^{-1}) \quad (8)$$

$$H_i = H(\tilde{S}_i^{-1}, S_{i-T}^{-1}) \quad (9)$$

其中, M_i 参数的作用是控制气流信息和生成图像的权重,以

达到对两种信息的权衡。 W_t 的作用是通过 $t-1$ 帧到 t 帧的光流,将光流与 $t-1$ 帧图像结合,以预测第 t 帧的图像。 H_t 保证了生成的图像输入与光流的输入是一致。

结合以上公式,最终的生成图像建模公式如式(10)所示:

$$\tilde{S}_t = (1 - M_t) \otimes W_{t-1} (\tilde{S}_{t-1}) + M_t \otimes H_t \quad (10)$$

5 实验细节

实验中对于G网络的每次更新,判别器D则需要被更新两次,并且我们对所有权重层使用光谱归一化^[18]和权重的正交初始化^[19]。使用G的权重的指数来移动平均值进行采样,其中在进行 2×10^4 次迭代训练步骤之后开始以衰减率 $\gamma=0.999$ 累积衰减。网络中使用的优化方程是SGD(基于梯度的方法),批量大小为256帧图像,G和D的学习率分别为 1×10^{-3} 和 2×10^{-3} 。

判别器中的空间判别器D_S对每一个均匀采样的 k 帧(默认 $k=8$)计算一个分数,则D_S输出的是每帧分数的总和;时间判别器D_T使用 2×2 平均池下采样函数 φ 对真实或生成的视频进行预处理。此外,D_T的前两个残块是3D的,其中每个卷积都替换为核大小为 $3 \times 3 \times 3$ 的3D卷积。

6 实验效果

6.1 评估分数

本文主要采用两个最常用的指标衡量生成模型的质量是否实用,即初始分数(Initial score, IS)^[20]和初始距离(Fr chet Inception Distance, FID)^[21],IS和FID在不同的视频中都具有可比性,因此这些含有标准实例化的指标将会更加适用于生成模型,可以使初始模型进行分类或者是进行特征提取,而且FID非常类似于视频距离(Fr chet Video Distance, FVD)^[22]。

6.2 Kinetics-600 数据集结果

生成视频尽管是一个具有挑战性的建模问题,但是生成视频的质量都可以用指标来体现,因此我们采用IS和FID作为指示标准。在实验中使用Kinetics-600数据集对改进的GAN网络进行评估,对合成的视频分别选取 64×64 像素和 128×128 像素进行不同像素值的比较,并且进行显示12帧和36帧的视频长度比较。不同像素值在FID与IS分数值上的对比如表1所列。

表1 不同图像值在FID与IS分数值上的对比

Table 1 FID/IS score values for different image sizes

	FID	IS
12/64 * 64	7.23	188.23
12/128 * 128	13.04	246.18
36/64 * 64	48.86	263.92
36/128 * 128	45.79	198.32

改进的模型能够在不同分辨率下生成视频。如表1所列,在低分辨率生成的图像,尽管视频较小且没有高分辨率展现的感官度好,但是也展现出高品质的纹理、物体组成和运动。然而在更高的分辨率下,生成对象会变得更加困难,需要改变更多的像素的移动位置,不过生成图像中的细节非常连贯,有良好的纹理。

为探究改进模型的效果,使用Vid2Vid、F-Vid2Vid、WC-Vid2Vid和改进模型分别在Kinetics-600数据集上测量IS值,IS值越高越好,实验结果如表2所列。

表2 不同的模型的IS值

Table 2 IS values of different models

Method	IS
Vid2Vid	16.56 ± 0.05
F-Vid2Vid	20.59 ± 0.32
WC-Vid2Vid	21.96 ± 0.58
Ours	22.65 ± 0.19

6.3 对比实验

为展示改进后的网络模型生成的图像效果是否有变化,对同一数据集Cityscape用不同方法生成的效果进行对比,结果如图6所示。对同一物体的颜色,本文模型生成的效果相比于其他两种生成方法有明显的一致性。

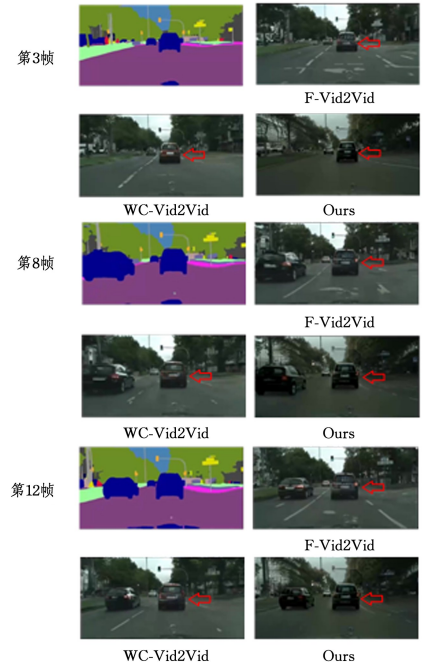


图6 不同方法生成的效果对比(电子版为彩色)

Fig. 6 Comparison of effects generated by different methods

由图6可以明显看出,在同一语义图的情况下,改进的模型要优于F-Vid2Vid和WC-Vid2Vid生成方法,尤其表现在箭头所指的目标车辆外观颜色不会随着时间的推移而发生改变。

6.4 实验结果展示

改进GAN网络后的生成视频结果如图7、图8所示。



图7 实验效果展示

Fig. 7 Experimental effect display

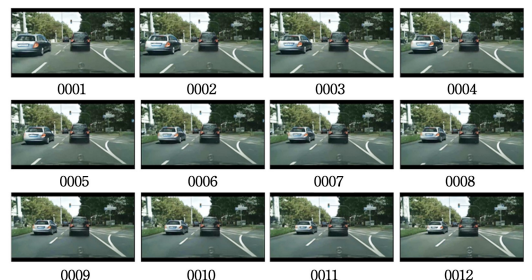


图8 实验效果展示

Fig. 8 Experimental effect display

由图 7 与图 8 所展示的效果可以看出,生成的图像在细节处所展示的效果更加逼真,如道路实线与虚线、汽车等。

6.5 实验结果分析

利用 GAN 网络生成逼真的视频是一次尝试,相比于利用 GAN 生成静态图像有很大的突破性,但是在视频领域却进展缓慢,因此基于改进 GAN 网络架构在 Kinetics-600 数据集上进行训练有着重要的意义。实验结果证明了把视频语义图进行前景和背景分开建模使用,然后利用 DVD_GAN 判别器进行分解判别,得到的视频的分辨率质量更高,对图像中的物体的约束更强,使生成的每一帧图像更加符合真实性。

结束语 本文对视频进行前景-背景分开建模,然后引入双判别器 DVD-GAN,在 Kinetics-600 数据集上合成复杂和多样化的短视频。实验结果表明,本文模型生成的视频要比 GAN 和 DVD-GAN 的效果好,尤其在图像内容中物体的颜色在前后帧也表现出了一致性。

虽然模型中引用的 Flow 算法对生成视频有很大的作用,但是会影响高质量的视频效果,因此在以后的研究中应该选择更好的模型代替 Flow 模型在 G 和 D 网络中的作用;其次应选择更好的方法去解决视频细节问题,可以尝试对标签图进行更加细致的划分或者使用 3D 渲染效果对细微处进行刻画。

参 考 文 献

- [1] GOODFELLOW I J,POUGET-ABADIE J,MIRZA M,et al. Generative adversarial nets[C]// International Conference on Neural Information Processing Systems. MIT Press,2014:2672-2680.
- [2] JIA Y F,MA L. Conditional self-attention generates adversarial networks[J]. Journal of Xidian University, 2019, 46(6): 163-170.
- [3] BROCK A,DONAHUE J,SIMONYAN K. Large scale GAN training for high fidelity natural image synthesis[C]// ICLR. 2019.
- [4] NITISH S,ELMAN M,RUSLAN S. Unsupervised learning of video representations using LSTMs[R]. In ICML,2015.
- [5] YAMAGUCHI A,CABATUAN M. Generative model based frame generation of volcanic flow video[C]// IEEE International Conference on Humanoid. IEEE,2017:1-5.
- [6] WANG T C,LIU M Y,ZHU J Y,et al. Video-to-Video Synthesis[J]. arXiv:1808.06601. 2018.
- [7] WANG T C,LIU M Y,TAO R,et al. Few-shot Video-to-Video Synthesis[J]. arXiv:1910.12713. 2019.
- [8] MALLYA A,WANG T C,SAPRA K,et al. World-Consistent Video-to-Video Synthesis[J]. arXiv:2007.08509. 2020.
- [9] CLARK A,JEFF D,KAREN S. Efficient Video Generation on Complex Datasets[J]. arXiv:1907.06571,2019.

- [10] AAYUSH B,SHUGAO M,YASER S. Recycle-GAN: Unsupervised video retargeting[C]// ECCV. 2018.
- [11] ZHOU Y P,WANG Z W,CHEN F,et al. Dance Dance Generation: Motion Transfer for Internet Videos[C]// 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE,2020.
- [12] VONDRICK C,PIRSIAVASH H,TORRALBA A. Generating videos with scene dynamics[C]// NeurIPS. 2016.
- [13] TAESUNG P,LIU M Y. Semantic Image Synthesis with Spatially-Adaptive Normalization[J]. arXiv:1903.07291.
- [14] ZHAO W Z,CHEN X,CHEN J G,et al. Sample Generation with Self-Attention Generative Adversarial Adaptation Network (SaGAAN) for Hyperspectral Image Classification[J]. Remote Sensing,2020,12(5):843.
- [15] AIDAN C,JEFF D,KAREN S,et al. Efficient Video Generation on Complex Datasets[J]. arXiv:1907.06571. 2019.
- [16] TULYAKOV S,LIU M Y,YANG X D. MoCoGAN: Decomposing motion and content for video generation[C]// CVPR. 2018.
- [17] LENA G,MOSHE B,ELI S,et al. Actions as space-time shapes [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2008,29(12):2247-2253.
- [18] ZHANG H,GOODFELLOW I J,METAXAS D,et al. Self-attention generative adversarial networks[J]. arXiv:1805.08318, 2018.
- [19] ANDREW B,JEFF D,KAREN S. Large scale GAN training for high fidelity natural image synthesis[C]// ICLR. 2019.
- [20] SALIMANS T,ZAREMBA W,CHEUNG V. Improved techniques for training GANs[C]// NeurIPS. 2016.
- [21] HEUSEL M,RAMSAUER H,UNTERTHINER T,et al. GANs trained by a twotime-scale update rule converge to a local nash equilibrium[C]// NeurIPS. 2017.
- [22] ZHENG S B,ZHOU G X,ZHANG B H,et al. A map matching algorithm based on discrete Frechet distance [J]. Journal of Hefei University of Technology (Natural Science),2017,40(1): 42-46.



YU Xiao-ming, born in 1965, associate professor. Her main research interests include intelligent information processing, graphics and image processing



HUANG Hua, born in 1995, postgraduate. His main research interests include graphics and image processing.