



中国科学院大学
University of Chinese Academy of Sciences

硕士学位论文

中国科学院大学学位论文 L^AT_EX 模板 $\pi\pi\pi$

作者姓名: 胡兵兵

指导教师: 吴幼龙 助理教授 上海科技大学

学位类别: 工学硕士

学科专业: 通信与信息系统

培养单位: 中国科学院上海微系统与信息技术研究所

2020 年 6 月

L^AT_EX Thesis Template
of
The University of Chinese Academy of Sciences $\pi\pi\pi$

A thesis submitted to the
University of Chinese Academy of Sciences
in partial fulfillment of the requirement
for the degree of
Master of Engineering
in Communication and Information System

By

Hu Bingbing

Supervisor: Professor Wu Youlong

Shanghai Institute of Microsystem and Information
Technology, Chinese Academy of Sciences

June, 2020

中国科学院大学 学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。本人完全意识到本声明的法律结果由本人承担。

作者签名：

日 期：

中国科学院大学 学位论文授权使用声明

本人完全了解并同意遵守中国科学院大学有关保存和使用学位论文的规定，即中国科学院大学有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延迟期后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘 要

本文是中国科学院大学学位论文模板 ucasthesis 的使用说明文档。主要内容为介绍 L^AT_EX 文档类 ucasthesis 的用法，以及如何使用 L^AT_EX 快速高效地撰写学位论文。

关键词：中国科学院大学，学位论文，L^AT_EX 模板

Abstract

This paper is a help documentation for the \LaTeX class ucasthesis, which is a thesis template for the University of Chinese Academy of Sciences. The main content is about how to use the ucasthesis, as well as how to write thesis efficiently by using \LaTeX .

Keywords: University of Chinese Academy of Sciences (UCAS), Thesis, \LaTeX Template

目 录

第 1 章 绪论	1
1.1 研究背景	1
1.2 研究现状	1
1.3 本文贡献	1
1.4 本文结构	1
第 2 章 My Work	3
2.1 C-InfoGAN	3
2.1.1 无监督分类方法	4
2.1.2 半监督分类方法	4
2.2 InfoCatGAN	6
2.2.1 无监督分类方法	6
2.2.2 半监督分类方法	7
2.3 本章小结	8
第 3 章 模型评估	9
3.1 实现细节	10
3.1.1 C-InfoGAN	10
3.1.2 InfoCatGAN	11
3.2 实验结果	12
3.2.1 MNIST	12
3.2.2 FashionMNIST	15
3.2.3 收敛速度分析	15
附录 A 中国科学院大学学位论文撰写要求	19
A.1 论文无附录者无需附录部分	19
A.2 测试公式编号 $\Lambda, \lambda, \theta, \bar{\Lambda}, \sqrt{S_{NN}}$	19
A.3 测试生僻字	20
参考文献	21

图形列表

2.1 C-InfoGAN 模型结构示意图	5
2.2 InfoCatGAN 模型结构。图中 D 的输出为 $P(y \cdot)$ 。在训练生成器的时候，将判别器的输出 $P(y \tilde{x})$ 和隐变量 c 通过某种度量 $d(\cdot, \cdot)$ 建立联系使得条件概率的峰值与 c 的取值对应。	6
3.1 隐变量对生成图片的调控。图中每一行对应类别隐变量的一个取值，每一列对应其他某个隐变量（固定剩余隐变量）的连续变化。	10
3.2 模型在 MNIST 上得生成效果。在 InfoCatGAN 和 C-InfoGAN 的生成结果中，每一行对应隐变量 c 的一个取值，从 0 到 9。	13
3.3 模型在 FashionMNIST 上的生成效果。在 InfoCatGAN 和 C-InfoGAN 的生成结果中，每一行对应隐变量 c 的一个取值，从 0 到 9。	16
3.4 模型在 MNIST 上的收敛速度	17

表格列表

3.1 C-InfoGAN 在 MNIST 上的网络结构	11
3.2 InfoCatGAN 在 MNIST 上的网络结构	12
3.3 MNIST 分类准确率对比	14
3.4 FashionMNIST 分类准确率对比	15

符号列表

字符

Symbol	Description	Unit
R	the gas constant	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
C_v	specific heat capacity at constant volume	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
C_p	specific heat capacity at constant pressure	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
E	specific total energy	$\text{m}^2 \cdot \text{s}^{-2}$
e	specific internal energy	$\text{m}^2 \cdot \text{s}^{-2}$
h_T	specific total enthalpy	$\text{m}^2 \cdot \text{s}^{-2}$
h	specific enthalpy	$\text{m}^2 \cdot \text{s}^{-2}$
k	thermal conductivity	$\text{kg} \cdot \text{m} \cdot \text{s}^{-3} \cdot \text{K}^{-1}$
S_{ij}	deviatoric stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
τ_{ij}	viscous stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
δ_{ij}	Kronecker tensor	1
I_{ij}	identity tensor	1
DUnif(a,b)	区间 [a,b] 上的离散均匀分布	
Unif(a,b)	区间 [a,b] 上的连续均匀分布	
p	概率分布	
Pr	事件概率, 概率测度	
x	数学加粗表示矢量	

算子

Symbol	Description
Δ	difference
∇	gradient operator

缩写

CFD Computational Fluid Dynamics

第 1 章 绪论

1.1 研究背景

In the probabilistic framework, semi-supervised learning can be modeled as a missing data problem, which can be addressed by generative models such as mixture models thanks to the EM algorithm and extensions thereof [6]. Generative models apply to the joint density of patterns and class (X, Y) . They have appealing features, but they also have major drawbacks. Their estimation is much more demanding than discriminative models, since the model of $P(X, Y)$ is exhaustive, hence necessarily more complex than the model of $P(Y | X)$. More parameters are to be estimated, resulting in more uncertainty in the estimation process. The generative model being more precise, it is also more likely to be misspecified. Finally, the fitness measure is not discriminative, so that better models are not necessarily better predictors of class labels. These difficulties have lead to proposals aiming at processing unlabeled data in the framework of supervised classification [1, 5, 11]. Here, we propose an estimation principle applicable to any probabilistic classifier, aiming at making the most of unlabeled data when they are beneficial, while providing a control on their contribution to provide robustness to the learning scheme.

1.2 研究现状

1.3 本文贡献

1.4 本文结构

第 2 章 My Work

在原始的生成对抗网络中，生成器通过将噪声映射成数据，然后由对抗损失函数驱动，进而能够在训练过程中慢慢学习到真实数据的分布信息，最终生成较为逼真的虚假数据。生成对抗网络的特殊性在于它创新地结合了生成式模型和判别式模型。我们既可以认为它训练了一个判别式模型，也可以认为它训练了一个生成式模型。它的最大贡献在于提出了一个对抗训练的机制，而且本身没有过多的约束，这为后续的研究提供了巨大的可扩展性。Goodfellow 等^[1]在给出生成对抗网络的模型以及结果之后，在文章最后给出了几点扩展工作：

1. 可以为生成器和判别器添加条件信息，此时生成器可以学习到对应的条件分布；
2. 通过增加一个辅助网络来估计 $p(\mathbf{z}|\mathbf{x})$ ，可以进行进一步的统计推断；
3. 半监督学习：当拥有少量标签信息时，判别器学到特征可以用于提高分类器的性能。

鉴于生成对抗网络具有如此强的可扩展性，众多学者投身研究该模型^[2-6]。尽管如此，关于生成器如何将噪声映射成数据的细节仍有待探索。许多研究表明，通过对隐变量连续插值，会在生成的图片上得到连续平滑的变化^[3,4,7,8]。然而，大多数变化无法解释并且没有明确的意义。Chen 等^[4]提出将隐变量分解，通过在训练过程中增加隐变量和生成数据之间的互信息，达到了将特征解耦的效果。此时，InfoGAN 的隐变量可以明确对应到一个个有意义的数据特征（如 MNIST 中手写数字的角度、笔画粗细等）。这证明了互信息约束在生成对抗网络中有值得探究的作用。

2.1 C-InfoGAN

如??所述，InfoGAN 通过最大化隐变量 \mathbf{c} 和生成数据 $G(\mathbf{z}, \mathbf{c})$ 之间的互信息，可以无监督地学习到数据的解耦合特征（disentangled representation），这在一定程度上解释了隐空间的结构变化对生成图片的影响。在模型收敛之后，隐变量 \mathbf{c} 的每一维度都能绑定到数据的某个特征。比如对于 MNIST 数据集 $\mathbf{c} = (c_1, c_2, c_3)$ ，其中 $c_1 \sim \text{DUnif}(0, 9)$ ， $c_1, c_2 \sim \text{Unif}(-1, 1)$ ，离散隐变量 c_1 绑

定到数字的类别，连续隐变量 c_1, c_2 绑定到数字的倾斜角度和笔画粗细。

2.1.1 无监督分类方法

InfoGAN 通过优化隐变量与生成数据之间的互信息研究了隐空间和数据空间的联系，可以通过改变隐变量来调整生成数据的指定特征。值得注意的是，在 MNIST 数据集上，隐变量 c_1 绑定到了数字的类别特征，再加上辅助网络 Q 是对后验概率 $p(c|\mathbf{x})$ 的估计，这为分类任务提供了基础。

具体来说，本文在 InfoGAN 的目标函数上添加一个正则项 $L(c, \hat{c})$ ，其中 $\hat{c} = Q(c|\tilde{\mathbf{x}}) \in \mathbf{R}^K$ 是 Q 网络的输出。称这个分类模型为 C-InfoGAN，简称 CIG，其目标函数如下：

$$\min_{G, Q} \max_D V_{\text{CIG}}(G, D, Q, \lambda_1, \lambda_2) = V_{\text{InfoGAN}}(G, D, Q, \lambda_1) + \lambda_2 L(c, Q(c|\tilde{\mathbf{x}})), \quad (2.1)$$

其中 λ_2 是正则化系数， $L(c, \hat{c}) = L(c, Q(c|\tilde{\mathbf{x}}))$ 在实现中一般采用交叉熵，参见(??)式，模型结构见图 2.1。

2.1.2 半监督分类方法

当拥有少量标签信息时，C-InfoGAN 可以利用这些标签进一步提升分类准确率和生成效果。同时将隐变量 c 直接绑定到真实的标签，实现精准调控。不同于 Spurr 等[9] 将隐变量进一步划分为无监督版本 c_{us} 和半监督版本 c_{ss} ，同时设计两个辅助网络 Q_{us} 和 Q_{ss} 分别处理对应版本的隐变量；本文直接将标签信息加入 Q 网络，先用真实数据和标签训练，接着用生成数据和虚假标签（即隐变量 c ）来训练，这样做的目的是为了使真实标签的信息流入隐变量 c 中，或者说用真实标签指导 c 绑定到正确的类别特征。使用和??节中类似的方法，我们给出半监督 C-InfoGAN 的目标函数：

$$\min_{G, Q} \max_D V_{\text{ss-CIG}}(G, D, Q, \lambda_1, \lambda_2, \lambda_3) = V_{\text{CIG}}(G, D, Q, \lambda_1, \lambda_2) + \lambda_3 \mathbb{E}_{(\mathbf{x}^L, \mathbf{y}^L) \sim \mathcal{X}^L} [\text{CE}[\mathbf{y}^L, Q(\mathbf{y}|\mathbf{x}^L)]] . \quad (2.2)$$

模型结构参见图 2.1。

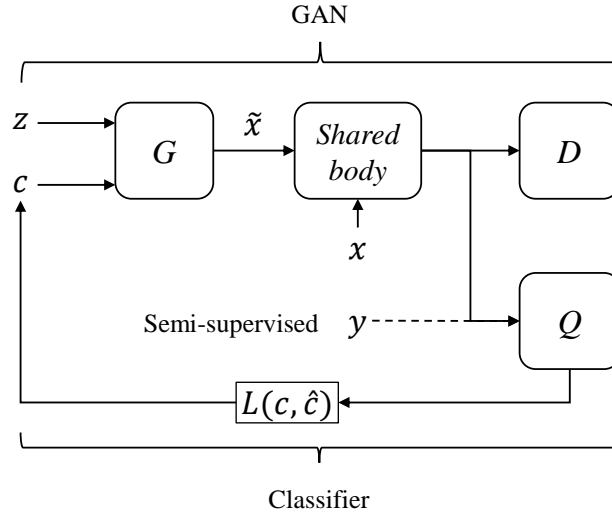


图 2.1 C-InfoGAN 模型结构。无监督情况下，生成数据 \tilde{x} 和真实数据 x 参与训练，通过和 D 共享部分结构， Q 网络可以将 GAN 模型学习到的特征加以利用，实现分类任务；在半监督情况下，一部分真实标签 y 会直接被 Q 网络利用，以得到更好的效果。优化 Q 网络的输出 \hat{c} 和隐变量 c 构成的损失函数 $L(c, \hat{c})$ 来增加 Q 的分类准确率。

Figure 2.1 The architecture of C-InfoGAN. In unsupervised case, the generated data \tilde{x} and the real data x are used for training. By sharing the body with D , Q is able to using features learned by GAN framework to perform classification. In semi-supervised case, labels y is directly fed into Q to get better performance. We optimize the some loss function $L(c, \hat{c})$ of latent code c and the output \hat{c} of Q to improve its accuracy.

2.2 InfoCatGAN

2.2.1 无监督分类方法

在训练概率分类模型的过程中，通过优化条件熵可以将分类边界调整到更自然的位置（数据分散区域）^[10]，因此 CatGAN 使用条件熵作为判别器判断真假数据的依据。但是，使用熵作为目标函数的一个缺点是没有类别指向性（ K 个类别中任意一个都可以使 $p(y|\mathbf{x})$ 呈单峰分布）。对于一个分类器，我们希望对于给定输入 \mathbf{x} ，有且仅有一个 $k \in [K]$ ，使得 $p(y = k|\mathbf{x})$ 最大，而对于任意 $k' \neq k$ ， $p(y = k'|\mathbf{x})$ 均很小。然而问题在于训练数据集没有标注，每个数据样本对应的标签无从获得。

对于上述问题本文从 InfoGAN 中获得启发，提出 InfoCatGAN 模型。InfoGAN 将输入噪声划分为 \mathbf{z} 和 \mathbf{c} ，实际上是对隐空间的结构进行了人为划分。一部分提供模型的容量，使得模型具有足够的自由度去学习数据的细节（高度耦合的特征）；一部分提供隐变量，用于在学习过程中绑定到数据的显著特征（如：MNIST 中的数字类别、笔画粗细、角度）。模型的核心思想如下：通过在隐空间构造一维隐变量 c ，在训练过程中将生成数据的类别标签与之绑定，使得可以通过 c 来控制生成数据的类别。CatGAN 对 GAN 的扩展主要在于改变了判别器的输出结构：为所有真实数据分配一个类别标签而对于虚假数据则保持一个不确定的状态。类似的，生成器应该致力于生成某个具体类别的数据而不是仅仅生成足够逼真的图片。

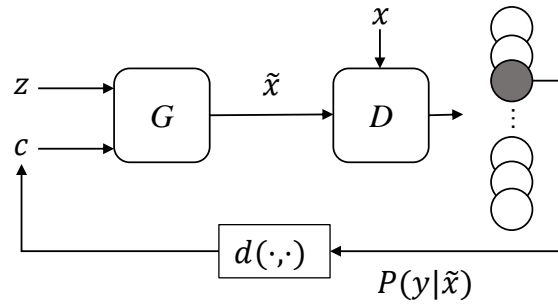


图 2.2 InfoCatGAN 模型结构。图中 D 的输出为 $P(y|\tilde{\mathbf{x}})$ 。在训练生成器的时候，将判别器的输出 $P(y|\tilde{\mathbf{x}})$ 和隐变量 c 通过某种度量 $d(\cdot, \cdot)$ 建立联系使得条件概率的峰值与 c 的取值对应。

下面给出 InfoCatGAN 的损失函数：设 $\mathbf{x} \in \mathcal{X}$ 为一个真实数据样本， $\tilde{\mathbf{x}} = G(\mathbf{z}, c)$ 为一个生成数据，其中 $\mathbf{z} \sim p_z$ 为噪声， $c \sim p_c$ 为隐变量。为了简单起见，

这里只考虑 c 为一维离散随机变量, p_c 为离散均匀分布。生成器 $G = G(\mathbf{z}, c; \theta_G)$ 和判别器 $D = D(\mathbf{x}; \theta_D)$ 均为可微深度神经网络, 其中 θ_G, θ_D 分别为生成器和判别器的参数¹。通过在 D 网络的最后一层做 Softmax 变换, 可以直接将 $D(x)$ 作为条件概率 $p(y|x)$ 的估计。注意到式(??)和(??)可以重写为:

$$L_D^{\text{cat}} = -I(X; Y) - \mathbb{E}_{\tilde{\mathbf{x}} \sim p_g} [H(p(y|\tilde{\mathbf{x}}))], \quad (2.3)$$

$$L_G^{\text{cat}} = -I(\tilde{X}; Y), \quad (2.4)$$

其中 $X \sim p_{\text{data}}, \tilde{X} \sim p_g$ 分别表示真实数据和虚假数据对应的随机变量, Y 表示未知标签对应的随机变量。从(2.3)、(2.4)式可以看出, CatGAN 在优化数据与标签之间的互信息。互信息是常用的变量间相关性的衡量标准, 所以本文用它作为生成器损失函数的正则项, 由此得到 InfoCatGAN 的损失函数如下:

$$L_D = L_D^{\text{cat}}, \quad (2.5)$$

$$L_G = L_G^{\text{cat}} - \lambda_1 I(c; \tilde{\mathbf{x}}),$$

其中 λ_1 为正则系数, 可知当 $\lambda_1 = 0$ 时, InfoCatGAN 退化为 CatGAN, 模型结构见图 2.2。参考(??)式, $I(c; \tilde{\mathbf{x}})$ 可以放缩为 $\mathbb{E}_{p(\mathbf{c}, \tilde{\mathbf{x}})} [\log p(c|\tilde{\mathbf{x}})]$, 在实现中通常使用交叉熵

$$CE[\mathbf{c}, p(c|\tilde{\mathbf{x}})] = - \sum_{i=1}^K c_i \log p(c = c_i|\tilde{\mathbf{x}}) \quad (2.6)$$

来优化此项, 这里的 $\mathbf{c} \in \mathbf{R}^K$ 是隐变量 c 经过 one-hot 编码之后的向量, $p(c|\tilde{\mathbf{x}})$ 可以用 $D(\tilde{\mathbf{x}})$ 来近似。

2.2.2 半监督分类方法

draft 作为 CatGAN 的扩展, InfoCatGAN 能够很自然地适用于半监督的情况。假设 $\mathcal{X}^L = \{\mathbf{x}_i^L\}_{i=1}^m$ 为 m 个有标签的样本, $\mathbf{y}_i^L \in \mathbf{R}^K$ 为经过 one-hot 编码之后的标签向量。对于有标签的样本, $D(\mathbf{x}^L)$ 的分布信息可以明确获得, 所以可以通过计算 y^L 和 $p(y|\mathbf{x}^L)$ 之间的交叉熵:

$$CE[\mathbf{y}^L, p(y|\mathbf{x}^L)] = - \sum_{i=1}^K y_i \log p(y = y_i|\mathbf{x}^L) \quad (2.7)$$

来辅助判别器做出更精确的判断。半监督版本的 InfoCatGAN 损失函数如下:

$$L_D^L = L_D + \lambda_2 \mathbb{E}_{(\mathbf{x}^L, \mathbf{y}^L) \sim \mathcal{X}^L} [CE[\mathbf{y}^L, p(y|\mathbf{x}^L)]], \quad (2.8)$$

其中 λ_2 为正则系数而生成器的损失函数同(2.5)式: $L_G^L = L_G$ 。

¹为了简便起见, 在无歧义的情况下通常省略网络参数。

2.3 本章小结

To be written...

第 3 章 模型评估

本章将在 MNIST^[11] 和 FashionMNIST^[12] 数据集上评估模型。MNIST 是深度学习广泛使用的评价模型的基础数据集，它包含 60000 个训练样本，10000 个测试样本，每个样本均为 28×28 的灰度手写数字图片。FashionMNIST 具有和 MNIST 相同的数据结构，同样是 60000 个训练样本和 10000 个测试样本，每个样本均为 28×28 灰度图片，也具有 10 个类别。然而相比于 MNIST，FashionMNIST 的图片构成更复杂，因此对模型具有更大的挑战。

在所有实验中，本文主要考察两个指标：分类准确率和图片生成质量。对于分类准确率，计算模型预测值并不像一般分类器那样直接。隐变量虽然可以学习到数据类别的特征，但是其取值并不一定和真实标签正确对应（例如 $c = 1$ 对应生成真实标签为 ω_2 的数据）。因此不能直接使用隐变量的取值作为模型的预测值，必须将隐变量的取值与真实标签之间做一个映射。

对于这个问题，本文采取与 Springenberg^[13] 类似的做法。首先在测试集上选取 t 个测试样本 $\mathcal{X}_t = \{\mathbf{x}_i, y_i\}_{i=1}^t$ ，计算模型在这批数据上的预测概率矩阵

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{t1} & p_{t2} & \cdots & p_{tk} \end{bmatrix},$$

其中 $p_{ij} = \Pr(y = \omega_j | \mathbf{x}_i)$ 表示给定输入样本 \mathbf{x}_i 时，模型预测 $y = \omega_j$ 的概率。然后对每一行选取最大概率的索引作为模型在这批样本上的预测值

$$\mathbf{P}^* = (\ell_1, \ell_2, \dots, \ell_t)^T, \quad \ell_i = \operatorname{argmax}_j p_{ij}.$$

显然 $\ell_i, y_i \in \Omega = \{\omega_1, \omega_2, \dots, \omega_K\}$ 。定义映射 $f: \Omega \rightarrow \Omega$,

$$f(\ell) = \operatorname{argmax}_y \left(\sum_{\mathbf{x}: \ell(\mathbf{x}) = \ell} \mathbf{1}(y(\mathbf{x}), y) \right),$$

其中 $\ell(\cdot)$ 表示模型预测值， $y(\cdot)$ 表示数据真实标签， $\mathbf{1}(a, b) = 1$ 当且仅当 $a = b$ ，否则为零。此时，给定模型预测的虚假标签 ℓ ，映射 f 都可以将它映射为真实标

签 $f(\ell)$ 。这就完成了虚假标签到真实标签的转换¹。简单来说，模型为每一个数据 \mathbf{x}_i 输出对应的概率向量 $p(y|\mathbf{x}_i)$ ，从而分配虚假标签 ℓ_i (概率向量中最大概率对应的索引)，然后将预测值和真实标签对比：将虚假标签落入最多的真实标签的取值作为该虚假标签的取值。比如在所有 10 个被分类为虚假标签 ℓ_1 的样本中，有 9 个真实标签为 ω_3 ，则将虚假标签 ℓ_1 映射到真实类别 ω_3 。

对于图片生成质量，本文采用 Fréchet Inception Distance (FID)^[14] 来进行衡量²，相较于 Inception Score^[15] 只考虑生成数据，FID 还利用了真实数据，因此更能反映生成数据和真实数据的差异。FID 越小代表生成的图片和真实图片越接近，生成质量越好。

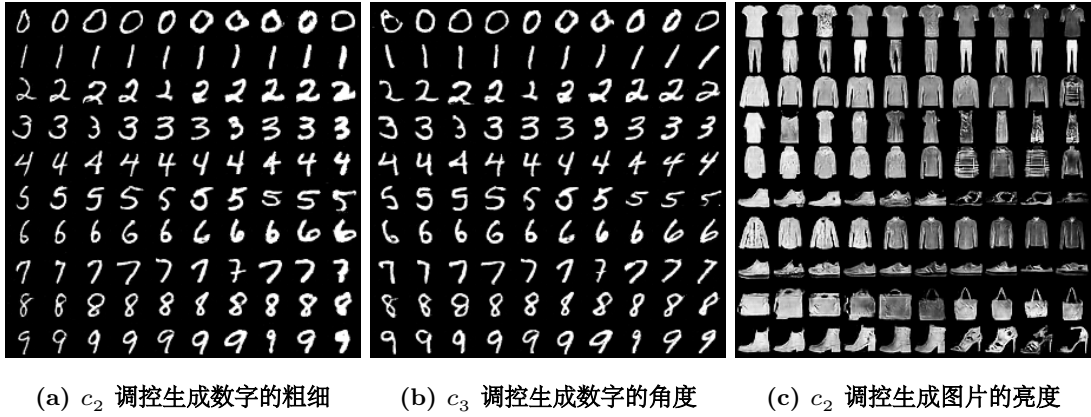


图 3.1 隐变量对生成图片的调控。图中每一行对应类别隐变量的一个取值，每一列对应其他某个隐变量（固定剩余隐变量）的连续变化。

3.1 实现细节

在实践中，模型中的所有模块（生成器、判别器、辅助网络等）均被实现为深度神经网络。具体实现可在 Github 上查看³。**网络结构在附录中给出。**

3.1.1 C-InfoGAN

对于 C-InfoGAN 模型，我们的设定和 InfoGAN 保持大致相同。辅助网络 Q 和判别器 D 共享大部分网络结构， Q 在判别器的尾部（通常是倒数第二层）

¹在实际应用中，可能会出现某一行的最大值不唯一的情况，此时选择第一个最大值索引作为输出。值得注意的是，这种情况在实验中并不多见。

²FID 一般用于彩色图片，而 MNIST 数据集是单通道的灰度图片，本文的做法是将单通道复制 3 份，形成一张 RGB 彩色图片然后再计算其 FID 值。

³实现代码地址：<https://github.com/guyueshui/baGAN>

表 3.1 C-InfoGAN 在 MNIST 上的网络结构

Table 3.1 Architectures of C-InfoGAN used for MNSIT dataset

生成器 G	判别器 D /辅助网络 Q
Input $\in \mathbf{R}^{74}$	Input 28×28 Gray image
fc 1024. bn. relu	conv 64: k4, s2, p1. lrelu
fc $7 \times 7 \times 128$. bn. relu	conv 128: k4, s2, p1. bn. lrelu
upconv 64: k4, s2, p1. bn. relu	fc 1024. bn. lrelu
upconv 1: k4, s2, p1. tanh	fc 1. sigmoid. output for D , fc 128. bn. lrelu. fc. output for Q

分离出一个全连接网络，输出后验概率的估计 $Q(\mathbf{c}|\mathbf{x})$ 。因此，C-InfoGAN 带来的计算复杂度增加较小。对于离散隐变量 c_i ，在输出之前施加一个 softmax 激活层，得到的输出代表 $Q(c_i|\mathbf{x})$ 。对于连续隐变量 c_j ，我们假设其先验分布为高斯分布，然后利用 Q 网络输出它的均值和方差，放到高斯分布中去拟合，计算差值。由于 GAN 在训练时容易坍塌，DCGAN^[3] 提出了一种训练稳定的生成对抗网络结构，所以我们采用的网络结构对此有很多参考。关于超参数的选择我们也沿用 InfoGAN 的设定，对于离散隐变量设置为 1；对于连续隐变量，设置为较小的值 0.1。

表 3.1 是给出了 C-InfoGAN 在 MNIST 上的网络结构，如前所述，判别器 D 和辅助网络 Q 共享大部分网络结构。在 MNIST 数据集上，噪声空间的构成为一个 10 维离散隐变量，两个连续隐变量和 62 维高斯噪声，最终得到的噪声维度为 74。对于 FashionMNIST，我们在实践中发现使用同样的网络结构也能达到很好的效果。因此本文对于两个数据集采用同样的网络结构。

3.1.2 InfoCatGAN

对于 InfoCatGAN，我采用的模型如下。

表 3.2 InfoCatGAN 在 MNIST 上的网络结构

Table 3.2 Architectures of InfoCatGAN used for MNSIT dataset

生成器 G	判别器 D
Input $\in \mathbf{R}^{74}$	Input 28×28 Gray image
fc 1024. bn. relu	conv 64: k4, s2, p1. lrelu
fc $7 \times 7 \times 128$. bn. relu	conv 128: k4, s2, p1. bn. lrelu
upconv 64: k4, s2, p1. bn. relu	fc 1024. bn. lrelu
upconv 1: k4, s2, p1. tanh	fc 10. softmax

3.2 实验结果

3.2.1 MNIST

图 3.2a和图 3.2b是在无监督情况下 CatGAN 和 InfoCatGAN 的生成效果, 可以看到, InfoCatGAN 的生成效果明显高于 CatGAN, 并且每一行基本是一种数字类别, 对应隐变量的不同取值。半监督情况下有类似的结果, 不同的是在少量标签信息的辅助下, InfoCatGAN 可以将隐变量 c 和真实标签正确绑定, $c = 1$ 对应生成数字 ‘1’, 见图 3.2e。CatGAN 生成的图片质量很差, 原因在于其目标函数是为了分类而设计的。生成器的作用只是为了判别器能够更加鲁棒, 如2.2.1节所述, 从(??)式中可以看到, G 的目标函数只有条件熵, 无法针对性地生成图片, 从而会降低生成图片的质量。而 InfoCatGAN 由于增加了隐变量 c , 并在训练过程中有意识地将生成数据的类别与之绑定, 所以生成的图片质量较好。

表 3.3给出了无监督和半监督情况下的分类准确率⁴。从表中看出, 在无监督的情况下 InfoCatGAN 的分类准确率相较于 CatGAN 有所提升, 而 FID 从 236.75 降低到 8.04, 图片生成质量有了极大的提升, 见图 3.2b。在半监督的情况下, CatGAN 的准确率达到 96.05%, 而 InfoCatGAN 随着正则化系数 λ_1 的变化呈现出不同的效果。当系数较小时, 分类准确率较高, 但生成图片的质量非常差; 当系数较大时, 生成的图片效果很好, 但分类准确率有所降低。通过调节参数 λ_1 , 可以实现生成效果和分类准确率之间的折中。实验使用的默认值是 $\lambda_1 = 1.1$,

⁴表中有关 CatGAN 的数据来自与本文复现的结果, 与 Springenberg^[13] 有所差距。经过多次尝试, 我们仍无法完美复现出原文中的结果。

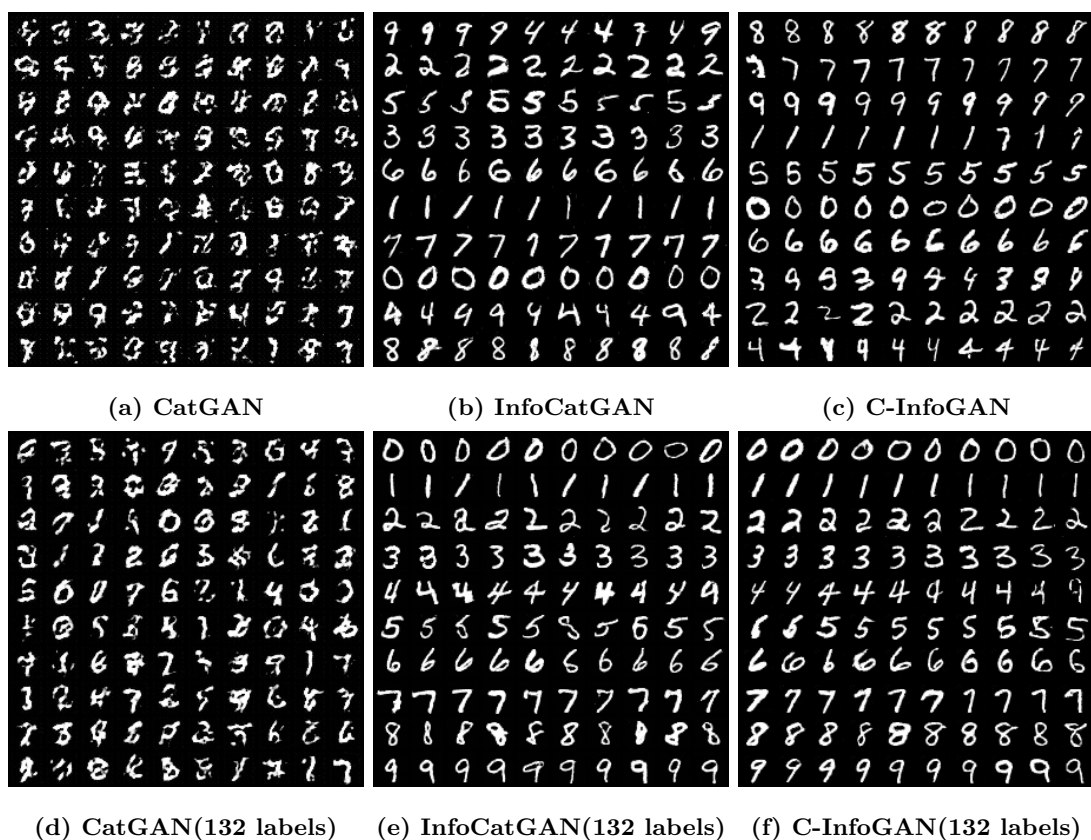


图 3.2 模型在 MNIST 上得生成效果。在 InfoCatGAN 和 C-InfoGAN 的生成结果中，每一行对应隐变量 c 的一个取值，从 0 到 9。

当 λ_1 减小时, 生成图片的质量开始下降, 同时分类准确率也会相应增加。值得注意的是, λ_1 越小并不意味着分类准确率越高。当 $\lambda_1 = 0$ 时, InfoCatGAN 退化为 CatGAN; 而从表 3.3 可以看出, 当 $\lambda_1 = 0.02$ 时, InfoCatGAN 的分类准确率高于 CatGAN, 达到 98.15%。

表 3.3 MNIST 分类准确率对比

模型	准确率 (%)	FID
CatGAN	64.41	236.75
InfoCatGAN	69.04	8.04
C-InfoGAN	77.65	7.55
CatGAN(132 labels)	96.05	127.92
InfoCatGAN(132 labels, $\lambda_1 = 0.02$)	98.15	169.72
InfoCatGAN(132 labels, $\lambda_1 = 0.03$)	80.43	6.59
C-InfoGAN(132 labels)	93.70	9.16

图 3.2c 和图 3.2f 给出了无监督和半监督情况下 C-InfoGAN 的生成结果。从图中可以看出无监督情况下, 模型已经达到了很好的生成效果, 隐变量 c 基本可以控制生成图片的类别, 但是仍有部分类别未能精确控制 (图 3.2c); 在半监督情况下, 隐变量达到了精确的绑定, 每一行对应生成一种类别的数字, 而且顺序和真实标签是对应的。另外从图 3.1 可以看出, C-InfoGAN 模型不仅可以生成指定类别的图片, 并且可以通过额外的隐变量调节图片局部特征, 如手写数字的粗细, 角度等, 这对指定特征的数据补足具有很大意义。

表 3.3 给出了 C-InfoGAN 的准确率和 FID 及其与 CatGAN 模型的性能对比。可以看出, 相较于 CatGAN 模型, C-InfoGAN 模型可以获得更高的准确率和生成质量, 而且隐变量的绑定效果也更好。而在半监督情况下, C-InfoGAN 在保证生成质量的前提下, 仍然能够达到 93.7% 的分类准确率。这是因为 InfoGAN 模型使用的是一个辅助网络 Q 来做类别绑定和分类任务, 训练过程中并没有判别器做过多约束, 所以无论如何调整分类网络或更改分类约束, 也不会对生成效果产生很大影响。这使得模型可以进一步利用生成的图片和标签扩充数据集, 以达到更进一步的性能提升。

3.2.2 FashionMNIST

FashionMNIST^[12] 是一个类似 MNIST 的数据集，二者拥有同样的结构，图像大小，同样的类别数目。但是相对于 MNIST，FashionMNIST 拥有更复杂的图像结构，以及更难获得非常高的分类准确率，所以对模型更具有检验性。

表 3.4 给出了模型在 FashionMNIST 的数值结果，其中 InfoCatGAN 的正则系数 $\lambda_1 = 0.03$ 。从表中可以看出，无论是在无监督还是半监督情况下，CatGAN 的分类准确率都是比较高的，但从生成效果来看，它却是最差的。在无监督情况下，InfoCatGAN 提高了生成图片的质量，但与此同时，牺牲了分类准确率，而 C-InfoGAN 在一定程度上二者兼顾，不仅生成质量最优，而且具有相对较高的分类准确率，此外其模型复杂度也较低。在半监督情况下，InfoCatGAN 在两个方面均体现出优势，分类准确率达到 74.21%，FID 为 16.92，生成效果见图 3.3e，这说明增加标签信息对模型的增益很大。

表 3.4 FashionMNIST 分类准确率对比

模型	准确率 (%)	FID
CatGAN	64.28	120.04
InfoCatGAN	57.66	26.43
C-InfoGAN	60.50	15.97
CatGAN(100 labels)	73.34	119.16
InfoCatGAN(100 labels)	74.21	16.92
C-InfoGAN(100 labels)	68.94	15.94

图 3.3 中给出了所有模型的生成结果。值得一提的是，加入互信息约束后的半监督版本模型从上往下每一行都对应同一个类别，并且顺序和训练数据的真实标签正确对应（图 3.3e、3.3f）。这说明了隐变量正确绑定到类别特征，并且可以精准调控生成图片的类别。

3.2.3 收敛速度分析

本文提出的两个模型在原理上都属于正则化生成对抗网络，因此相较于原先的两个模型 CatGAN 和 InfoGAN，增加的计算复杂度很小。由于 GAN 的训练方式特殊，训练的过程是生成器和判别器的对抗，因此目前没有一个统一的评

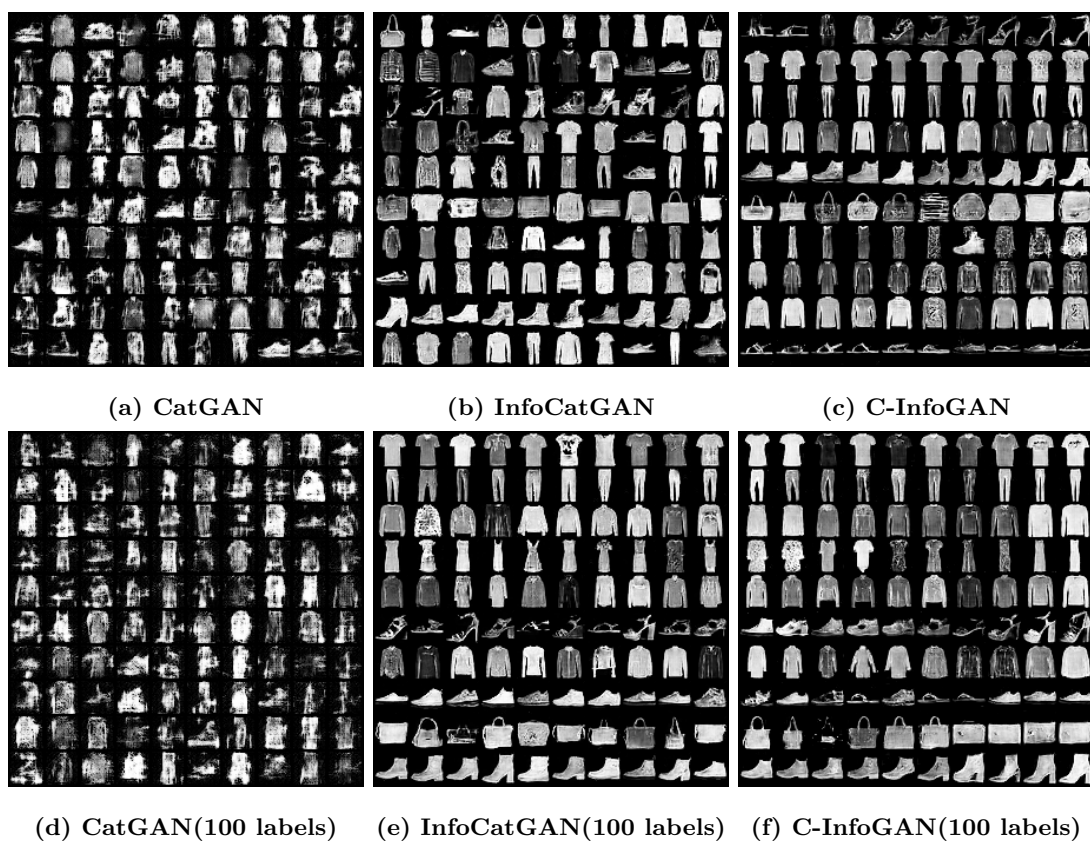


图 3.3 模型在 FashionMNIST 上的生成效果。在 InfoCatGAN 和 C-InfoGAN 的生成结果中，每一行对应隐变量 c 的一个取值，从 0 到 9。

判收敛性的标准。针对 InfoCatGAN 和 C-InfoGAN 两种模型，本文分别用条件熵损失（即判别器输出的概率分布对应的熵）以及互信息损失（实际采用交叉熵估计，详见??节）来作为模型收敛的佐证，见图 3.4。

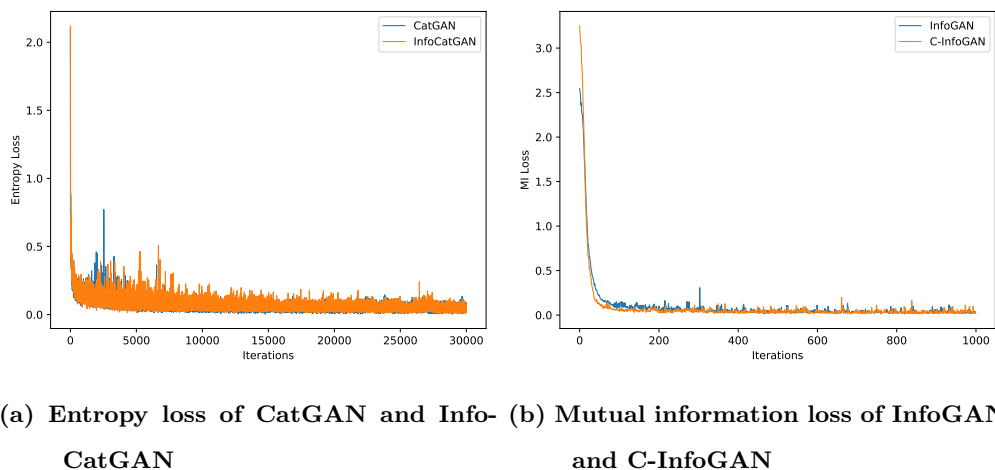


图 3.4 模型在 MNIST 上的收敛速度

Figure 3.4 Convergence speed on MNIST

附录 A 中国科学院大学学位论文撰写要求

学位论文是研究生科研工作成果的集中体现，是评判学位申请者学术水平、授予其学位的主要依据，是科研领域重要的文献资料。根据《科学技术报告、学位论文和学术论文的编写格式》(GB/T 7713-1987)、《学位论文编写规则》(GB/T 7713.1-2006) 和《文后参考文献著录规则》(GB7714—87) 等国家有关标准，结合中国科学院大学（以下简称“国科大”）的实际情况，特制订本规定。

A.1 论文无附录者无需附录部分

A.2 测试公式编号 $\Lambda, \lambda, \theta, \bar{\Lambda}, \sqrt{S_{NN}}$

$$\begin{cases} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{V}) = 0 \\ \frac{\partial(\rho \mathbf{V})}{\partial t} + \nabla \cdot (\rho \mathbf{V} \mathbf{V}) = \nabla \cdot \boldsymbol{\sigma} \\ \frac{\partial(\rho E)}{\partial t} + \nabla \cdot (\rho E \mathbf{V}) = \nabla \cdot (k \nabla T) + \nabla \cdot (\boldsymbol{\sigma} \cdot \mathbf{V}) \end{cases} \quad \dots \text{ (A.1)}$$

$$\frac{\partial}{\partial t} \int_{\Omega} u \, d\Omega + \int_S \mathbf{n} \cdot (u \mathbf{V}) \, dS = \dot{\phi} \quad \dots \text{ (A.2)}$$

$$\mathcal{L}\{f\}(s) = \int_{0^-}^{\infty} f(t) e^{-st} \, dt, \quad \mathcal{L}\{f\}(s) = \int_{0^-}^{\infty} f(t) e^{-st} \, dt$$

$$\mathcal{F}(f(x + x_0)) = \mathcal{F}(f(x)) e^{2\pi i \xi x_0}, \quad \mathcal{F}(f(x + x_0)) = \mathcal{F}(f(x)) e^{2\pi i \xi x_0}$$

mathtext: $A, F, L, 2, 3, 5, \sigma$, mathnormal: $A, F, L, 2, 3, 5, \sigma$, mathrm: $A, F, L, 2, 3, 5, \sigma$.

mathbf: $\mathbf{A}, \mathbf{F}, \mathbf{L}, \mathbf{2}, \mathbf{3}, \mathbf{5}, \mathbf{\sigma}$, mathit: $A, F, L, 2, 3, 5, \sigma$, mathsf: $A, F, L, 2, 3, 5, \sigma$.

mathtt: $A, F, L, 2, 3, 5, \sigma$, mathfrak: $\mathfrak{A}, \mathfrak{F}, \mathfrak{L}, 2, 3, 5, \sigma$, mathbb: $\mathbb{A}, \mathbb{F}, \mathbb{L}, 2, 3, 5, \sigma$.

mathcal: $\mathcal{A}, \mathcal{F}, \mathcal{L}, 2, 3, 5, \sigma$, mathscr: $\mathcal{A}, \mathcal{F}, \mathcal{L}, 2, 3, 5, \sigma$, boldsymbol: $\mathbf{A}, \mathbf{F}, \mathbf{L}, 2, 3, 5, \mathbf{\sigma}$.

vector: $\boldsymbol{\sigma}, \mathbf{T}, \mathbf{a}, \mathbf{F}, \mathbf{n}$, unitvector: $\boldsymbol{\sigma}, \mathbf{T}, \mathbf{a}, \mathbf{F}, \mathbf{n}$

matrix: $\boldsymbol{\sigma}, \mathbf{T}, \mathbf{a}, \mathbf{F}, \mathbf{n}$, unitmatrix: $\boldsymbol{\sigma}, \mathbf{T}, \mathbf{a}, \mathbf{F}, \mathbf{n}$

tensor: $\boldsymbol{\sigma}, \mathbf{T}, \mathbf{a}, \mathbf{F}, \mathbf{n}$, unittensor: $\boldsymbol{\sigma}, \mathbf{T}, \mathbf{a}, \mathbf{F}, \mathbf{n}$

20

参考文献

- [1] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Advances in neural information processing systems. 2014: 2672-2680.
- [2] MIRZA M, OSINDERO S. Conditional generative adversarial nets[J]. arXiv preprint arXiv:1411.1784, 2014.
- [3] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.
- [4] CHEN X, DUAN Y, HOUTHOOFT R, et al. Infogan: Interpretable representation learning by information maximizing generative adversarial nets[C]//Advances in neural information processing systems. 2016: 2172-2180.
- [5] CHONGXUAN L, XU T, ZHU J, et al. Triple generative adversarial nets[C]//Advances in neural information processing systems. 2017: 4088-4098.
- [6] MAO X, LI Q, XIE H, et al. Least squares generative adversarial networks[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2794-2802.
- [7] DUMOULIN V, BELGHAZI I, POOLE B, et al. Adversarially learned inference[J]. arXiv preprint arXiv:1606.00704, 2016.
- [8] MIYATO T, KOYAMA M. cgans with projection discriminator[J]. arXiv preprint arXiv:1802.05637, 2018.
- [9] SPURR A, AKSAN E, HILLIGES O. Guiding infogan with semi-supervision[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, 2017: 119-134.
- [10] GRANDVALET Y, BENGIO Y. Semi-supervised learning by entropy minimization [C]//Advances in neural information processing systems. 2005: 529-536.
- [11] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural computation, 1989, 1(4):541-551.
- [12] XIAO H, RASUL K, VOLLGRAF R. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms[EB/OL]. (2017-08-28).
- [13] SPRINGENBERG J T. Unsupervised and semi-supervised learning with categorical generative adversarial networks[J]. arXiv preprint arXiv:1511.06390, 2015.
- [14] HEUSEL M, RAMSAUER H, UNTERTHINER T, et al. Gans trained by a two

time-scale update rule converge to a local nash equilibrium[C]//Advances in neural information processing systems. 2017: 6626-6637.

- [15] SALIMANS T, GOODFELLOW I, ZAREMBA W, et al. Improved techniques for training gans[C]//Advances in neural information processing systems. 2016: 2234-2242.