新手指南综述 | GAN模型太多，不知道选哪儿个？

原创 bryant8 机器学习与生成对抗网络 2月23日

点击上方蓝字，关注~

相关阅读：

GAN整整6年了！是时候要来捋捋了！

数百篇GAN论文已下载好！搭配一份生成对抗网络最新综述！

有点夸张、有点扭曲！速览这些GAN如何夸张漫画化人脸！

天降斯雨，于我却无！GAN用于去雨如何？

脸部转正！GAN能否让侧颜杀手、小猪佩奇真容无处遁形？

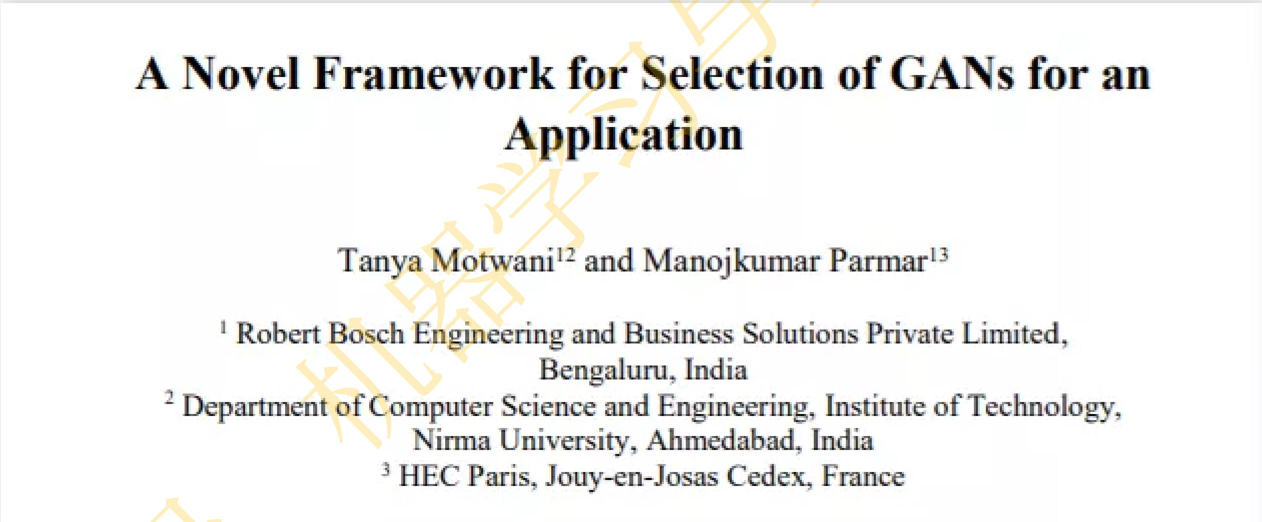
容颜渐失！GAN来预测？

弱水三千，只取你标！AL（主动学习）结合GAN如何？

异常检测，GAN如何gan？

【1】GAN在医学图像上的生成，今如何？

01-GAN公式简明原理之铁甲小宝篇



A Novel Framework for Selection of GANs for an Application

Tanya Motwani[12] and Manojkumar Parmar[13]

[1] Robert Bosch Engineering and Business Solutions Private Limited, Bengaluru, India
[2] Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, India
[3] HEC Paris, Jouy-en-Josas Cedex, France

今天看到这么一个论文题目“**A Novel Framework for Selection of GANs for an Application**”，这名字有、6啊，好久没有出厉害的GAN的变体了吧？新颖的GAN框架？决定下载下来看！引入眼帘的是**摘要**：

生成对抗网络（GAN）是当前的研究焦点。但由于其知识体系零散，导致在为给定应用场景选择合适GAN模型时可能尝试多次不合适的算法模型。本文从GAN的诞生到发展至今的变体做较全面的总结，包括如何解决模式崩溃，梯度消失，不稳定的训练和不收敛等问题。从应用的角度，对其表现和实现细节方面提供了比较。提出了一个新的框架，在特定场景下，用于从网络架构、损失、正则化手

段和散度衡量等方面去辨别备选GAN。通过一个简单示例的讨论，证明可以显著减少GAN的变体搜索空间。这种方法可以降低AI开发成本。

额，这不就是一个综述吗，2333 行吧，都来了，就做个记录吧，新手学习GAN也可概览一波GAN这个领域。【后面的几个表格总结，可以仔细看看~

其实之前也写了两篇综述导读，均有配套相应论文下载：

GAN整整6年了！是时候要来捋捋了！

数百篇GAN论文已下载好！搭配一份生成对抗网络最新综述！

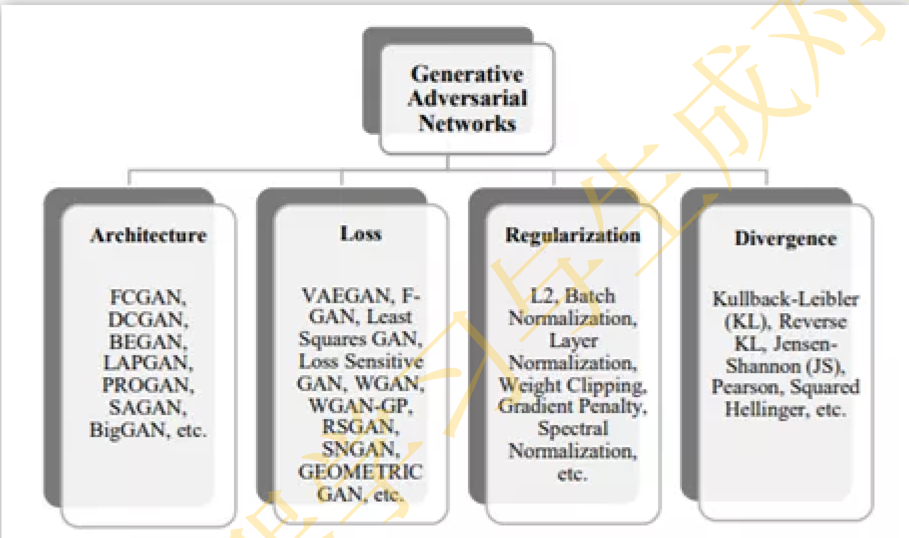


Fig. 1 Proposed Decision Parameters for GANs

简介

生成对抗网络（GAN）是基于博弈论构建的一类生成模型。这种模型的典型架构包括两个神经网络-鉴别器和生成器：发生器将输入噪声转换为潜在的高维数据；判别器评估所生成的数据是否源自原始分布。根据结果，生成器学习生成与原始分布相似的样本。

在计算机视觉[2] [3] [4] [5]，时序合成[6] [7] [8] [9]，自然语言处理[10] [11] [12] [13]等一些领域，GAN是一类可生成多样好、逼真数据的有效手段。它们属于隐式模型[14]。这些模型从学习到的分布中采样生成图像，并不表示出数据样本的潜在分布。与其他显式生成模型相比，GAN具有诸如并行生成，通用逼近，更好的质量，清晰的密度估计以及对样本结构层次的理解等优点。这些特性有助于GAN在深度学习社区中的广泛普及，尤其是在计算机视觉领域。

尽管取得了成功，但GAN仍然难以训练。每次更新网络的任何参数（判别器或生成器）时，都会导致不稳定。当前的研究致力于为各种应用（例如图像和视频生成[15] [16] [17]，领域自适应[3] [18] [19] [20]，语音合成[21] [22] [23]，语义图像编辑[2] [24]等）寻找【稳定的架构，损失和超参数组合】。尽管这些模型在特定应用中获得了有趣的结果，但尚无透彻的共识或研究可用来参考、了解哪种GAN的性能更优。在本文中旨在缓解上述问题，并通过技术框架缩小备选GAN的范围。

本文的组织如下：

第2节定义整个框架：最常用的网络结构、损失函数、正则化和散度方案。

第3节重点介绍了训练GAN时出现的问题，

第4节概述了GAN流行的损失函数。

第5节基于应用、表现和实现等方面进行GAN之间的对比，

第6节通过示例说明了框架的使用。

第7节强调了未来的研究范围，总结。

第6节

以使用CIFAR-10数据集生成图像为应用场景，来说明框架😅的使用。考虑到该应用要求生成的质量良好且具有多样性。如果没有本文理清的逻辑框架，就必须搜索爆炸性的组合空间。例如，完成这一任务可能有近5000种潜在的GAN组合：基于网络架构，损失，差异等。而在该框架的帮助下，可以缩小到5-6种候选的GAN。这相当于搜索空间减少了1000倍。为了减少组合搜索空间，提出以下四个问题，答案根据表1-7得出。

1、鉴别器和生成器使用什么网络结构?

根据表1，可选的方案包括全连接、卷积、反卷积网络，或DCGAN的修改。

2、哪些损失函数合适?

在表2-7的应用、实现和表现的详细比较中，有关于在该数据集上的研究细节。由于该应用任务需要多样性和质量良好，所以表格建议使用WGAN-GP，最小二乘LSGAN，RSGAN和SNGAN模型。

LSGAN与RSGAN结合可以生成更高质量的图像。正则化模型例如Loss Sensitive GAN和SNGAN展示了更好的泛化能力。

3、GAN是否需要正则化？如果需要，那么哪个有效？

本文研究表明梯度惩罚可以提高图像质量但不能稳定训练。谱归一化与梯度惩罚相比，计算效率更高。【47】表明生成器中的BN可提高模型质量而在鉴别器中使用的话则表现糟糕。

4、GAN是否需要与KL散度不同的衡量方式？如果是，哪儿个最合适？

[48]介绍并尝试了各种方式，包括GAN，Kullback-Leibler和Squared-Hellinger，都可以生成同样逼真的效果。

附上表1-7如下：

Table 1 Summary of loss variants in GANs

YEAR	NAME	EXPERIMENTATION		DATASET	ISSUES ADDRESSED	PERFORMANCE METRIC
2014	SGAN [1]	LOSS	Minimax	MNIST [28], TFD [29], CIFAR-10 [30]		Log-likelihood under distribution (kernel density) of Gaussian Parzen window [31]
		ARCHITECTURE	D: FC, CON G: FC, DECON			
		REGULARIZATION	NO			
2015	VAEGAN [32]	LOSS	Combines loss of Variational Autoencoder and GAN, while sharing parameters between Dec and G	CelebA [32], CIFAR-10, STL-10 [33], LFW [34]	Mode Collapse	Visual perception (Qualitative Assessment)
		ARCHITECTURE	Enc, D: CON Dec: DECON O: RMSProp [35] N: BN [36]			
		REGULARIZATION	NO			

2016	F-GAN [38]	**LOSS**	Introduces variational divergence estimation framework with generalization to f-divergences [37]	MNIST, LSUN		Same as SGAN
		ARCHITECTURE	**D, G:** Based on DCGAN **O:** Adam [38] **N:** BN			
		REGULARIZATION	NO			
2016	LEAST SQUARES GAN	**LOSS**	Adopt least squares function in SGAN's objective	HWDB1.0 [39], LSUN, Gaussian Mixtures [40]	Stable Training, Vanishing Gradient, Mode Collapse	Same as VAEGAN
		ARCHITECTURE	**D, G:** Based on DCGAN **O:** Adam, RMSProp **N:** BN			
		REGULARIZATION	NO			
2017	Loss Sensitive GAN [44]	**LOSS**	Objective function is based on minimization of data dependent margin (the loss of a real sample is less than that of generated sample). Formulated CLSGAN (fully and semi supervised), GLSGAN	SVHN [41], CelebA, CIFAR-10, MNIST, tiny ImageNet	Vanishing Gradient, Generalizability across different data distributions	Visual perception, image classification for quantitative evaluation, MRE for generalizability
		ARCHITECTURE	**D, G:** Based on DCGAN **O:** Adam **N:** BN			
		REGULARIZATION	YES (Enforces Lipschitz regularity using loss margin)			
2017	WGAN [46]	**LOSS**	Objective based on EM Distance or Wasserstein-1	LSUN	Stable Training, Vanishing Gradient, Mode Collapse	Same as VAEGAN
		ARCHITECTURE	**D:** DCGAN **G:** DCGAN, MLP **O:** Adam, RMSProp **N:** BN			
		REGULARIZATION	YES (Enforces Lipschitz regularity using weight clipping)			

2017	GEOMETRIC GAN [47]	**LOSS**	Geometric GAN with SVM hyperplane, Geometric interpretations of SGAN, FGAN, WGAN, EBGAN []	MNIST, LSUN, CelebA, Gaussian mixture	Vanishing Gradient, Imbalance between D and G	Same as VAEGAN
		ARCHITECTURE	**D, G:** MLP, DCGAN **O:** Adam, RMSProp **N:** BN			
		REGULARIZATION	YES (Weight clipping, weight decay)			
2017	WGAN-GP [48]	**LOSS**	Wasserstein objective along with gradient penalty (GP)	LSUN, CIFAR-10, Billion Word [42]	Stable Training, Vanishing Gradient	Inception Score [43]
		ARCHITECTURE	**D, G:** Based on DCGAN, ResNet [44] **G:** MLP **O:** Adam **N:** BN, LN			
		REGULARIZATION	YES (GP with comparison to weight clipping)			
2018	SN-GAN [52]	**LOSS**	Optimized the objective of SGAN with normalization of weight matrices. Comparisons with Minimax, Hinge [45], WGAN-GP	CIFAR-10, STL-10, ImageNet	Stable Training	Inception Score, Frechet Inception Distance (FID) [46]
		ARCHITECTURE	**D, G:** CON, ResNet **O:** Adam **N:** BN, LN			
		REGULARIZATION	YES (Spectral Norm (SN) and comparison with other normalization techniques (GP, Weight normalization, and Orthonormal))			
2018	RGAN [55]	**LOSS**	Relativistic SGAN such that G pushes D towards 0.5, rather than 1. Introduced Relativistic average loss	CIFAR-10, CAT []	Stable Training, Vanishing Gradient	Frechet Inception Distance
		ARCHITECTURE	**D, G:** DCGAN, CON (same as SNGAN) **O:** Adam **N:** BN			
		REGULARIZATION	YES (Gradient Penalty, Spectral Norm)			

[a] CLSGAN, Conditional Loss Sensitive GAN
[b] EM, Earth Mover
[c] GLSGAN, Generalized Loss Sensitive GAN
[d] HWDB 1.0, Chinese Handwritten Database
[e] LFW, Labeled Faces in the Wild
[f] MRE, Minimum Reconstruction Error
[g] STL-10, Shop The Look
[h] SVHN, Street View House Numbers
[i] SVM, Support Vector Machine
[j] TFD, Toronto Face Dataset

Comparison of various GANs

Table 2 Comparison of GANs with FCGAN

GAN	APPLICATION	BEHAVIOUR	IMPLEMENTATION
		FCGAN	
VAEGAN	• VAEGAN had lower entropy of each single image generated showing high quality output on synthetic MNIST distribution.	• VAEGAN does not suffer from mode collapse. It preserves the functionality to map a single image to its latent variables. • VAEGAN implicitly encourages similarity between synthetic and training data unlike FCGAN.	• VAEGAN has the same loss function for the discriminator as FCGAN.
Least Squares GAN (LSGAN)	• LSGAN observes successful learning on $BN_{Generator}$ with Adam optimizer unlike FCGAN on LSUN dataset.	• LSGAN exhibits less mode seeking behavior and training instability compared to FCGAN. • LSGAN penalizes samples which are on the correct side of the decision boundary but far away from the real data as log loss doesn't care about distance but only the sign. • LSGAN achieves good convergence without BN. • Least squares function (LSGAN) is flat at only one point unlike sigmoid cross entropy, not causing values to saturate when x is too large.	• LSGAN removed log from D in FCGAN and instead used L2 loss []. • RMSProp performs more stable than Adam optimizer for both the GANs.
WGAN	• WGAN was able to create better quality of samples than FCGAN but lower than that of DCGAN for LSUN dataset.	• WGANs are more robust than FCGAN when one varies the architectural choices for the generator without any evidence of mode collapse. • Gradient with respect to input in WGAN is better behaved, making optimization of G better. • Training of WGAN with weight clipping is slower than that of original GAN. • EM distance guarantees continuity and differentiability, which KL divergence and JS divergence lack. • No balance between D and G, or careful network architecture is required in WGAN. • WGAN is trained optimally which makes it impossible to collapse modes and not limit to imperfect gradients. • WGAN value function correlates with sample quality unlike FCGAN.	• The last layer of sigmoid in D of FCGAN is removed in WGAN, as WGAN performs regression, not binary classification. • WGAN recommends RMSProp rather than a momentum-based optimizer like Adam, as it causes instability in model training.
GEOMETRIC GAN	• Geometric GAN demonstrated less mode collapse with Lipschitz continuity regularization on Gaussian mixture dataset.	• Geometric GAN differs in the definition of normal vector of the separating plane. • Geometric GAN successfully converges to Nash Equilibrium between D and G. • Geometric GAN has a linear hyperplane consistent approach compared to non-linear separating hyperplane of FCGAN.	

WGAN-GP	• WGAN-GP achieves comparable sample quality to FCGAN objective for equivalent architectures on CIFAR-10 and LSUN datasets but has increased stability which is used to explore a range of architectures.	• WGAN-GP doesn't always converge to the equilibrium point with a finite number of updates of D per G update, unlike FCGAN which focuses training with consensus optimization, zero-centered gradient penalties or instance noise and therefore, converges when provided with enough capacity.	• No BN in WGAN-GP because it changes D's function mapping from one input to output in whole batches while WGAN-GP penalizes the norm independently.
RSGAN	• RSGAN performed only slightly better on CIFAR-10. Claimed that the dataset was too easy to realize the stabilizing effects of RSGAN. • In CAT dataset with high resolution pictures, relativism showed more improvement than spectral norm or gradient penalty.	• RSGAN fixes the generator's objective in FCGAN such that it not only increases the probability of fake data being real but also decreases the probability of real data being real. • RSGAN works very well in conjunction with gradient penalty, even when using only one D update per G update. • Relativism significantly improves data quality and stability of GANs at no computational cost. • FCGAN becomes stuck early in training as when the D reaches optimality, the gradient completely ignores real data. As RSGAN estimates the probability of real data being more realistic than a randomly sampled fake data, both real and fake data will always be incorporated in the gradient of the D's loss function.	

Table 3 Comparison of GANs with WGAN

GAN	**APPLICATION**	**BEHAVIOUR**	**IMPLEMENTATION**
		WGAN	

WGAN-GP	• WGAN-GP significantly outperforms weight clipping by higher inception scores and convergence rates on CIFAR-10. • Training loss of WGAN-GP on MNIST gradually increases even when the validation loss has dropped unlike WGAN.	• WGAN fails to capture higher moments of the data distribution. They end up learning simple functions and models very simple approximations to the optimal functions. • Detection of overfitting in WGAN-GP is faster compared to WGAN when given enough capacity and too little training capacity. • The loss quality of both correlates with sample quality and converges toward a minimum.	• WGAN-GP recommends LN as a replacement for BN. • Adam performs better than RMSProp as an optimizer with its objective in WGAN-GP.
Least Squares GAN	No comparative study available with respect to application.	• Though having a similar setup compared to WGAN, Least Squares GAN minimizes Pearson $\chi2$ divergence and learns a L_2 loss function instead of critic function.	• Least Squares GAN also uses regression, and therefore sigmoid layer is removed compared to FCGAN.
Loss Sensitive GAN (LS-GAN)	• GLS-GAN attained a smaller MRE on tiny ImageNet dataset compared to WGAN.	• Introduced generalized LS-GAN (GLS-GAN). GLS-GAN contains a large family of regularized GANs which contain both LS-GAN and Wasserstein as its special cases. • WGAN seeks to maximize the mean under the densities of real and generated samples, and clips the network weights on a bounded box to prevent the loss function from becoming unbounded. LS-GAN treats real and generated samples in pairs and maximizes the difference in their losses up to a data-dependent margin, which not only prevents their losses from being decomposed into two separate first order moments but also enforces them to coordinate with each other to learn the optimal loss function.	• Both can produce unclasped natural samples without using BN and address vanishing gradient while training G. • Both use weight regularization as a means of ensuring the model function has a bounded Lipschitz constant. WGANs use weight clipping, LS-GANs use weight decay.

VAEGAN	• WGAN generated images that did not belong to any of the 10 classes in MNIST dataset. This indicates that an application that cannot handle out-of-dataset images cannot utilize WGAN.	• WGAN varies smoothly even when two distributions overlap but the generated samples are not realistic. VAEGAN implicitly encourages similarity between training and synthetic data. • Both the models successfully address mode collapse. • Entropy of each generated image is higher in WGAN because of its low quality (suspected that the FC network is not powerful enough).	
SNGAN	• SNGAN is relatively robust and produces diverse and complex images with aggressive learning and momentum rates compared to weight clipping, evaluated on CIFAR-10 and STL-10.	• In WGAN, the number of features is diminished by weight clipping leading to a random model that matches the target distribution at select features. SNGAN augments the cost function with a sample data dependent regularization term. • Weight matrices of layers trained with weight clipping are rank deficit which proves to be fatal in lower layers (unnecessary restricts the search space of the discriminator) unlike those of SNGAN which are broadly distributed.	• In the absence of regularization techniques, SN provides better sample quality compared to weight normalization and gradient penalty.
GEOMETRIC GAN	No comparative study available with respect to application.	• WGAN follows a mean-difference driven approach and leads to generation of mean of arbitrary number of modes in true distributions. • Geometric GAN follows a linear separating hyperplane, and shows robust convergence behavior.	

Table 4 Comparison of GANs with WGAN-GP

GAN	APPLICATION	BEHAVIOUR	IMPLEMENTATION
WGAN-GP			
Least Squares	• LSGANs perform better than WGAN-GP on datasets of LSUN, CAT and CIFAR-10 datasets while	• WGAN-GP is more computational intensive than LSGAN-GP and requires multiple updates for discriminator.	

	performed poorly on ImageNet. • LSGAN and WGAN-GP achieve similar FID on LSUN but LSGAN much less time to reach the optimal FID. • Both LSGAN-GP and WGAN-GP succeed in training difficult architectures and generate higher quality images with 101-layer ResNet.		
RSGAN	• WGAN-GP produces high quality images (a low FID score) on stable setup in CIFAR-10 compared to RSGAN.	• As WGAN-GP is an integral probability metric GAN, both the real and fake data equally contribute to the gradient of D's loss function. They implicitly assume that some of the samples are fake, similar to the function of relativism. • In unstable setups, WGAN-GP performed very poorly because of a single discriminator update per generator update. Relativism provides a greater improvement in difficult settings compared to gradient penalty.	
SNGAN	• WGAN-GP fails to train GANs at high momentum and learning rates on both STL-10 and CIFAR-10. • The combination of WGAN-GP and parameterization with spectral normalization achieves better quality images than WGAN-GP.	• WGAN-GP heavily depends on the support of current generative distribution. • As they change over the course of training, they destabilize the effect of GP. • SN can be used with GP (local regularizers), because it provides global regularization on the D.	• Requires less computational cost compared to WGAN-GP.

Table 5 Comparison of F-divergence GAN with Geometric GAN

GAN	APPLICATION	BEHAVIOUR	IMPLEMENTATION
		Geometric GAN	
F-diverge	No comparative study available with respect to application.	• In F-GAN, as the scaling factors that reflect geometric space are asymmetric, it is	

		difficult to control the balance between D and G updates.	

Table 6 Comparison of RSGAN with Least Squares GAN

GAN	APPLICATION	BEHAVIOUR	IMPLEMENTATION
	LEAST SQUARES GAN		
RSGAN	• The combination of LSGAN with relativism (RaLSGAN) performed better than simple LSGAN in both unstable and stable setups, evaluated using FID on CIFAR-10. • LSGAN produced high quality 64*64 resolution CAT images (low FID score) but produced them in a very unstable manner.	• LSGAN is unable to converge in high resolution (256*256 or more) image dataset while RSGAN can generate images in all resolutions.	

Table 7 Comparison of Loss Sensitive GAN with DCGAN

GAN	APPLICATION	BEHAVIOUR	IMPLEMENTATION
	DCGAN		
Loss Sensitive GAN (LSGAN)	• LS-GAN outperformed DCGAN on classification of CIFAR-10 and SVHN datasets by a higher accuracy and lower error rate respectively. • Regularized models such as LS-GAN have better generalization performances and more stable training while achieving a low MRE on CIFAR-10 dataset.	• LS-GAN's loss comprises of linear constraints and objective, contrary to log loss which causes vanishing gradient. The linear gradient, rather than being saturated, provides sufficient gradient to continuously update the generator. • LS-GAN is not affected by over-trained loss function, unlike DCGAN.	• LS-GAN does not use a sigmoid layer as the output of the loss function. • BN is known to prevent mode collapse in DCGAN; without BN, DCGAN cannot produce any images and would collapse. LS-GAN proves to be more resilient with different structure changes and performs very well even if BN layers are removed.

▲ 长按关注我~