Generative Adversarial Network(GAN)

GAN模型的参考网址：https://github.com/hindupuravinash/the-gan-zoo；
里面有各式各样的GAN得模型， 到目前为止已经有接近300个不同的GAN了

Outline

- Basic Idea of GAN
- GAN as structured learning
- Can Generator learn by itself?
- Can Discriminator generate?
- A little bit theory

（接下来的4周，李宏毅老师都在讲GAN这个技术）

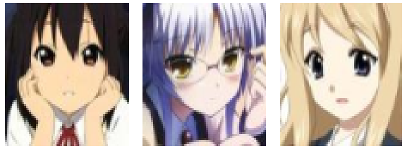
Generation

这个主要任务就是生成generator

We will control what to generate latter. → Conditional Generation

Image Generation

$$\begin{bmatrix} 0.3 \\ -0.1 \\ \vdots \\ -0.7 \end{bmatrix} \begin{bmatrix} 0.1 \\ -0.1 \\ \vdots \\ 0.7 \end{bmatrix} \begin{bmatrix} -0.3 \\ 0.1 \\ \vdots \\ 0.9 \end{bmatrix} \rightarrow \boxed{\text{NN Generator}} \rightarrow$$



In a specific range

Sentence Generation

$$\begin{bmatrix} 0.3 \\ -0.1 \\ \vdots \\ -0.7 \end{bmatrix} \begin{bmatrix} 0.1 \\ -0.1 \\ \vdots \\ 0.2 \end{bmatrix} \begin{bmatrix} -0.3 \\ 0.1 \\ \vdots \\ 0.5 \end{bmatrix} \rightarrow \boxed{\text{NN Generator}} \rightarrow$$

How are you?

Good morning.

Good afternoon.

二次元图像的生成网址：
http://mattya.github.io/chainer-DCGAN/

Basic Idea of GAN

It is a neural network (NN), or a function.

vector → Generator → image

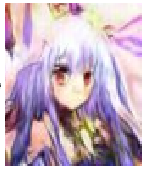
high dimensional vector

每次改变一个值

$$\begin{bmatrix} 0.1 \\ -3 \\ \vdots \\ 2.4 \\ 0.9 \end{bmatrix} \rightarrow \text{Generator} \rightarrow$$

Each dimension of input vector represents some characteristics.

$$\begin{bmatrix} \boxed{3} \\ -3 \\ \vdots \\ 2.4 \\ 0.9 \end{bmatrix} \rightarrow \text{Generator} \rightarrow$$



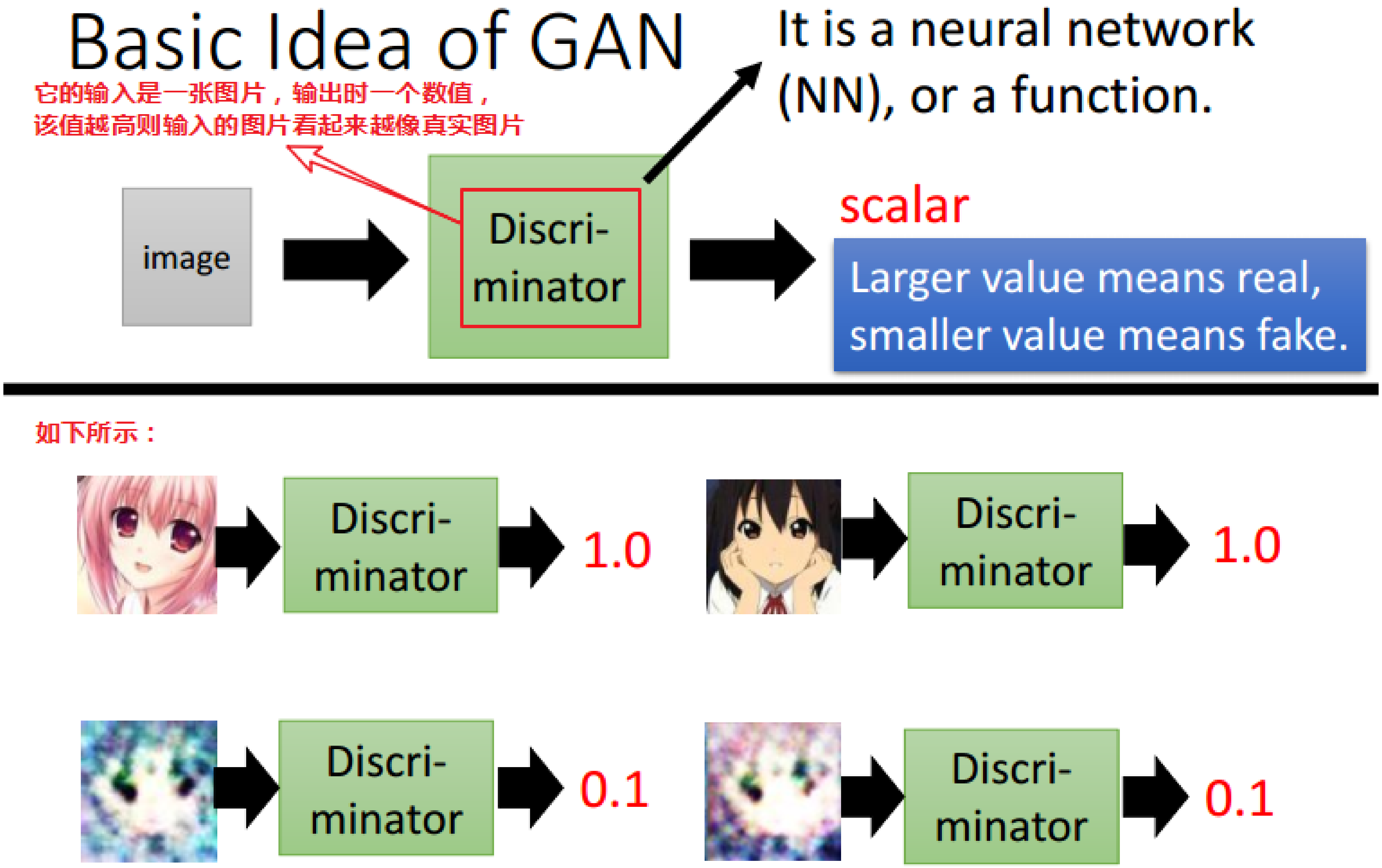
Longer hair

$$\begin{bmatrix} 0.1 \\ 2.1 \\ \vdots \\ \boxed{5.4} \\ 0.9 \end{bmatrix} \rightarrow \text{Generator} \rightarrow$$

blue hair

$$\begin{bmatrix} 0.1 \\ -3 \\ \vdots \\ 2.4 \\ \boxed{3.5} \end{bmatrix} \rightarrow \text{Generator} \rightarrow$$

Open mouth


Basic Idea of GAN
It is a neural network (NN), or a function.
它的输入是一张图片，输出时一个数值，
该值越高则输入的图片看起来越像真实图片
image
Discri-minator
scalar
Larger value means real, smaller value means fake.
如下所示：
Discri-minator
1.0
Discri-minator
1.0
Discri-minator
0.1
Discri-minator
0.1

Basic Idea of GAN

如枯叶蝶进化去防止被比比鸟吃掉（最终进化成了枯叶的这种）：

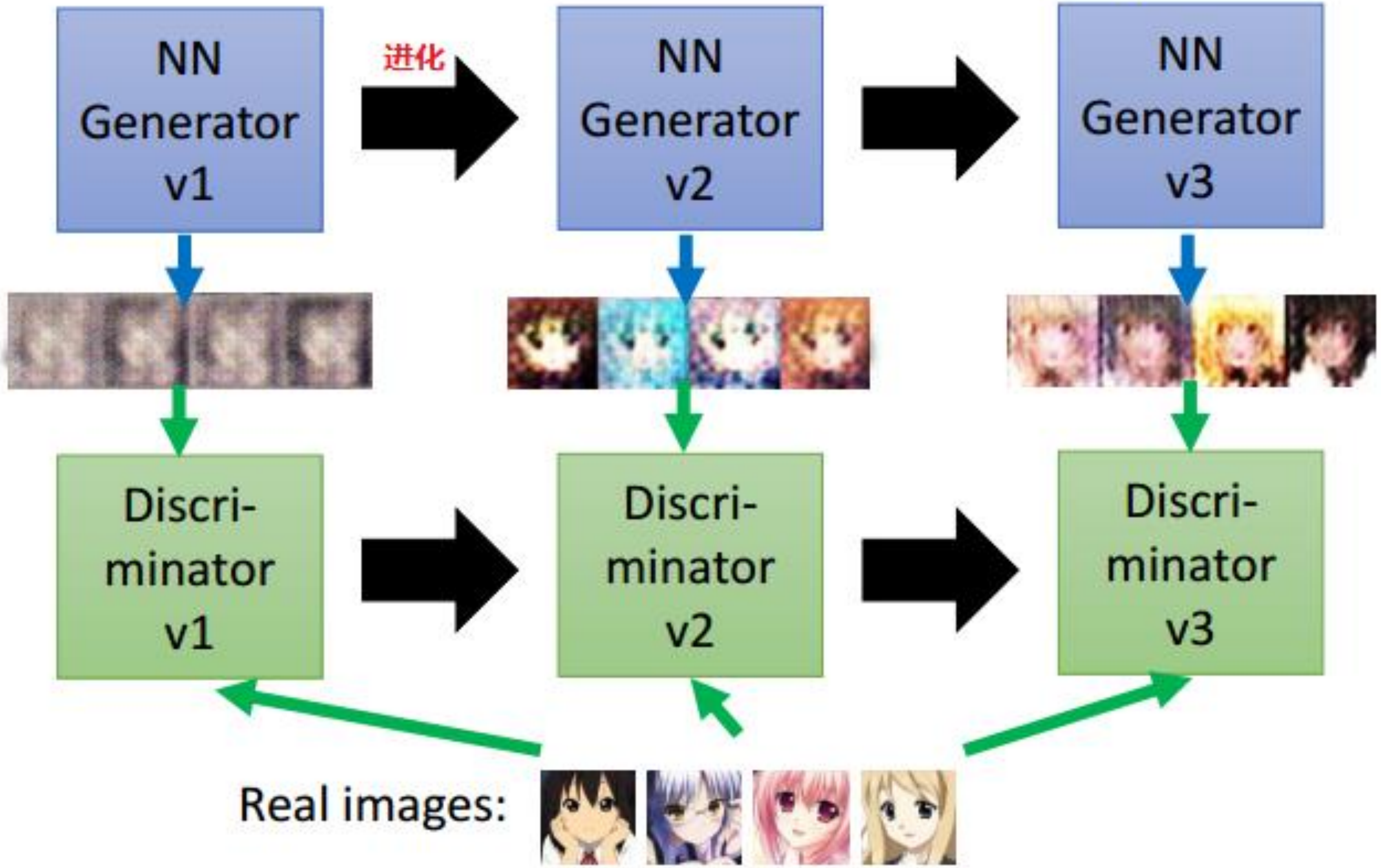
比比鸟去进化判断这个东西能不能吃：

最后两者变得越来越强。

Basic Idea of GAN

This is where the term "***adversarial***" comes from.

You can explain the process in different ways…….



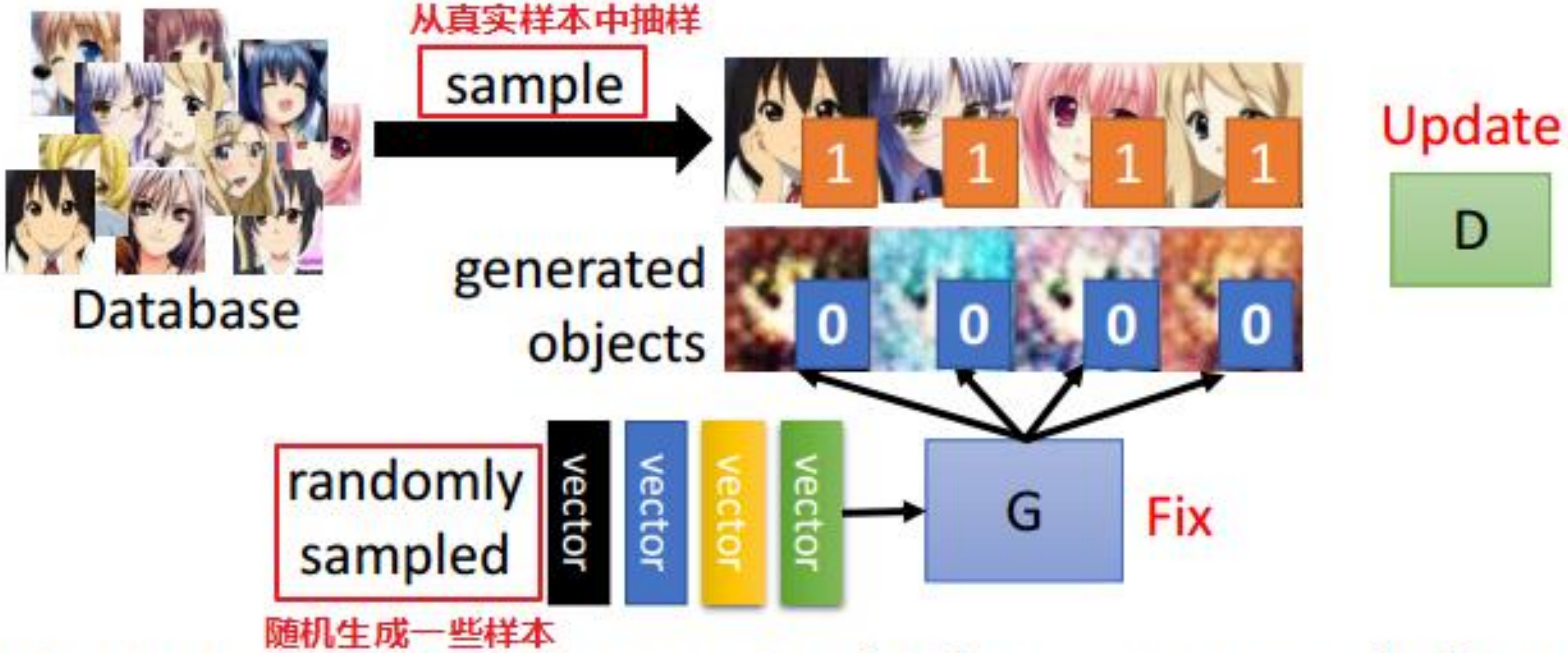
Algorithm 算法

- Initialize generator and discriminator

- In each training iteration:

step1:固定G，更新D的参数

Step 1: Fix generator G, and update discriminator D



Discriminator learns to assign high scores to real objects and low scores to generated objects.

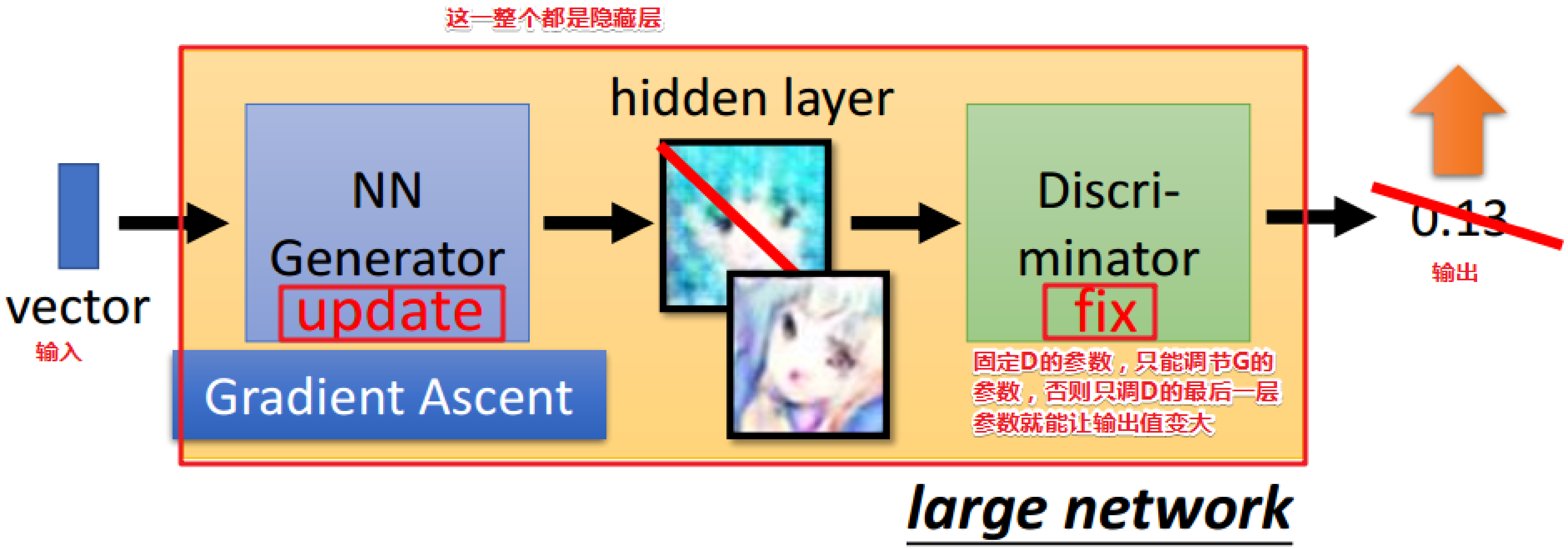
Algorithm

- Initialize generator and discriminator G D
- In each training iteration:

step2：固定D，更新G的参数

Step 2: Fix discriminator D, and update generator G

Generator learns to "fool" the discriminator



Algorithm Initialize θ_d for D and θ_g for G

- In each training iteration:

Learning D

- 1、Sample m examples $\{x^1, x^2, \ldots, x^m\}$ from database 随机选出m个样本
- 2、Sample m noise samples $\{z^1, z^2, \ldots, z^m\}$ from a distribution
- 3、Obtaining generated data $\{\tilde{x}^1, \tilde{x}^2, \ldots, \tilde{x}^m\}$, $\tilde{x}^i = G(z^i)$
- 4、Update discriminator parameters θ_d to maximize
  - $\tilde{V} = \frac{1}{m}\sum_{i=1}^{m} log D\left(\boxed{x^i}\right) + \frac{1}{m}\sum_{i=1}^{m} log\left(\boxed{1 - D\left(\tilde{x}^i\right)}\right)$
    真实图片 假的图片 这一项越大越好
  - 参数 $\boxed{\theta_d} \leftarrow \theta_d + \eta\nabla\tilde{V}(\theta_d)$

Learning G

- Sample m noise samples $\{z^1, z^2, \ldots, z^m\}$ from a distribution 不一定必须和上面的一样
- Update generator parameters θ_g to maximize
  - $\tilde{V} = \frac{1}{m}\sum_{i=1}^{m} log\left(\boxed{D\left(\boxed{G\left(z^i\right)}\right)}\right)$ 意思是希望generator得到的图片放在discriminator里面值越大越好
  - $\theta_g \leftarrow \theta_g + \eta\nabla\tilde{V}(\theta_g)$

Source of training data: https://zhuanlan.zhihu.com/p/24767059

Why Structured Learning Challenging?

为什么Structured Learing具有挑战性？

- <u>**One-shot/Zero-shot Learning**</u>: 要让机器去输出一些其从来没有见过的东西
  - In classification, each class has some examples.
  - In structured learning,
    - If you consider each possible output as a "class" ……
    - Since the output space is huge, most "classes" do not have any training data.
    - Machine has to create new stuff during testing.
    - Need more intelligence（机器需要一定的智慧）

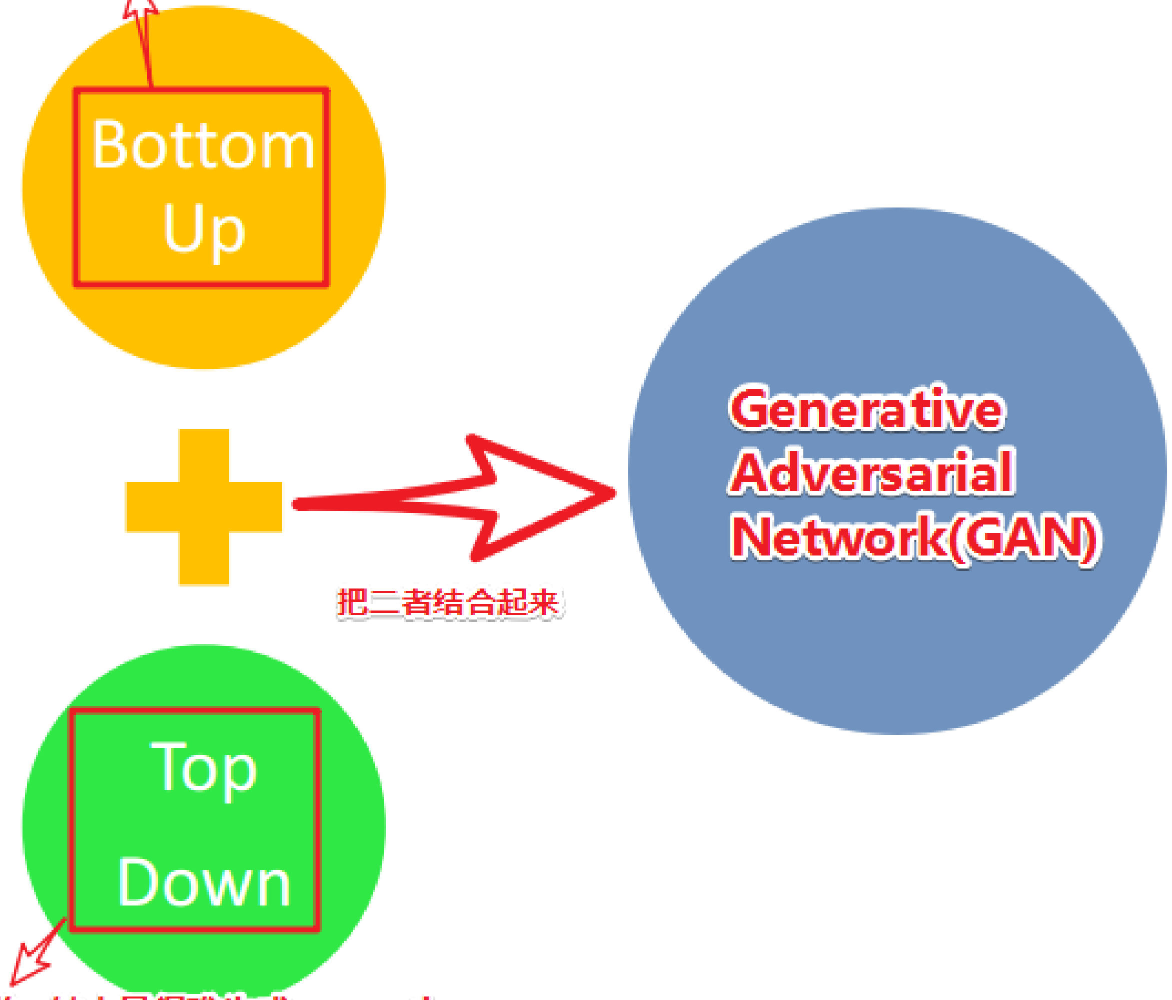
Structured Learning Approach

的两种方法：

一个component去产生，缺点是很容易失去大局观

Generator

Learn to generate the object at the component level



把二者结合起来

Discriminator

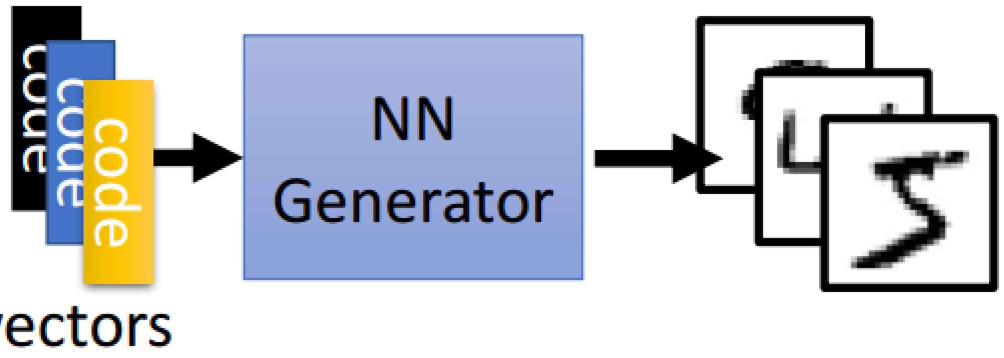
Evaluating the whole object, and find the best one

这个方法是基于大局去做，缺点是很难生成generation

1、Generator能否自己学习

Generator

如果只是用Generator，则需要更大的网络结构才能产生图像，
用GAN则相对轻松的产生图像



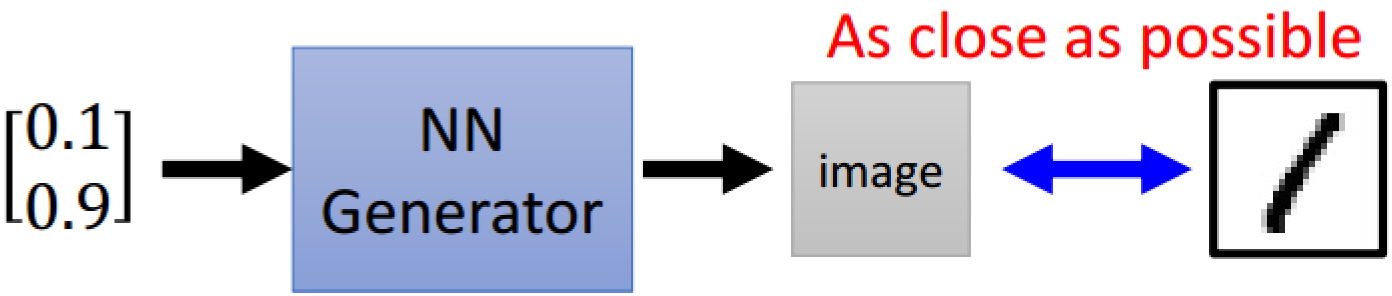
Generator的component
需要一个一个
独立的生成

code: $\begin{bmatrix} 0.1 \\ -0.5 \end{bmatrix}$ $\begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix}$ $\begin{bmatrix} 0.2 \\ -0.1 \end{bmatrix}$ $\begin{bmatrix} 0.3 \\ 0.2 \end{bmatrix}$

(where does they come from?)

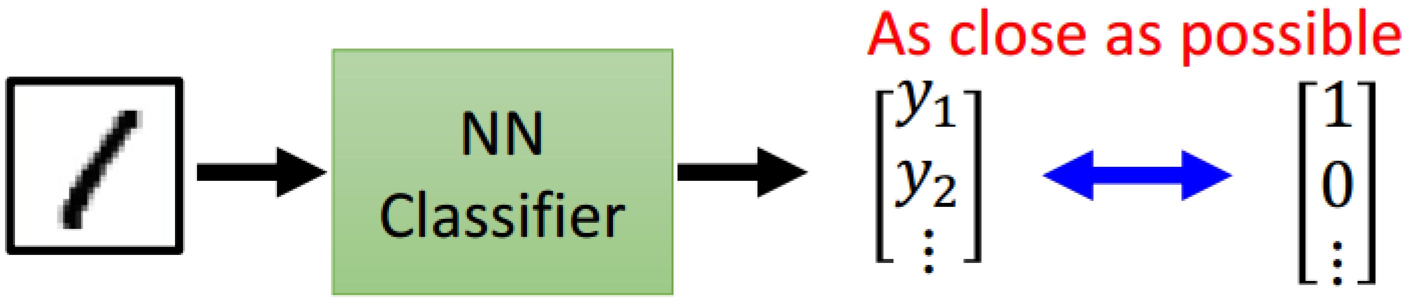
Image:

Generator：



一般的分类：

c.f.



2、能否用Discriminator去生成图片？

不能，如果只用Discrimination生成图像，它看到的只有手工绘制的正样本图像，没有负样本；
再者，它本身是用来做评价的，很难用来生成好的图像。

二者的比较

Generator v.s. Discriminator

- ***Generator***
- Pros: 优点：容易生成图片
  - Easy to generate even with deep model
- Cons: 缺点：只学到了表象，学不到大局
  - Imitate the appearance
  - Hard to learn the correlation between components

- ***Discriminator***
- Pros: 优点：可以考虑大局
  - Considering the big picture
- Cons: 缺点：很难生成图像
  - Generation is not always feasible
    - Especially when your model is deep
  - How to do negative sampling?

总结：主要讲了怎么直观地了解GAN
作业搜索关键词：MLDS hw3-1