

---

# 《人工神经网络》大作业开题报告

---

陶天骅

计算机科学与技术系

清华大学

@mails.tsinghua.edu.cn

杨雅儒

计算机科学与技术系

清华大学

yangyr17@mails.tsinghua.edu.cn

## 1 任务定义

本课题希望构建一个神经网络以及一些简单的界面，可以根据用户提供的一些特征的比例，自动生成一张风景图片。程序界面示意图如下。

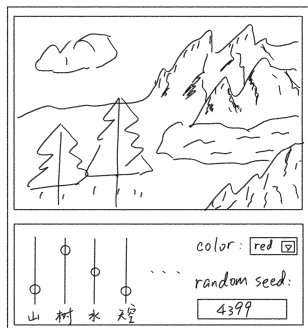


Figure 1: A Simple Sketch.

用户通过滑动滑条，确定例如树木、山、水、天空等要素在图片中的占比，并可以设定希望使用的主题颜色，程序便生成一张符合以上要求的风景图片。用户可以设置随机种子来获得不同的图片。

用数学语言形式化程序的任务即为：

令  $pic$  是一张图片， $label_{pic}$  是图片的标签，值为  $real$  或  $generated$ ，表示是真实的或者生成的。需要构建一个辨别器  $D$ ，参数是  $\theta_D$ ， $D(pic; \theta_D)$  输出一个标签，表示辨别器认为  $pic$  是否是生成的。训练目标为：

$$\arg \max_{\theta_D} (Prob(D(pic; \theta_D) = label_{pic}))$$

即  $D$  能给出正确的标签。

再构建生成器  $G$ ，参数是  $\theta_G$ ，对于一个特征向量  $x$ ， $G(x; \theta_G)$  输出一张图片。训练目标为：

$$\arg \min_{\theta_G} (Prob(D(G(x; \theta_G); \theta_D) = generated))$$

即  $G$  的输出应足够逼真，以至于  $D$  不能分辨图片是否是生成的。

同时为了让  $G$  能正确地展现  $x$  的特征，还需要一个编码器  $E$ ， $E(pic)$  输出图片对应的特征向量。 $E$  的训练就是 AutoEncoder 的 *Encoder* 的训练，而 *decoder* 的部分就是前面的  $G$ 。当  $E$  是一个理想编码器的时候， $G$  的训练目标还要包括：

$$\arg \max_{\theta_G} (Prob(E(G(x; \theta_G)) = x))$$

即对  $G$  在  $x$  上生成的图片进行编码的话，还能得到  $x$ ，这表示生成的图片确实含有  $x$  的特征。

## 2 数据集

用于训练的数据集可以从各大图片社交平台（如 Pinterest、Flicker）上下载获得的风景图片。

## 3 挑战和基线

### 3.1 挑战

- 确定使用哪些特征作为输入标签。
- 将特征和主题颜色向量化。
- 考虑到算力有限，可能无法生成分辨率较高的图片。
- 融合不同的神经网络架构，打造一个本程序专用的神经网络。

### 3.2 基线

你可以调研和你的工作相关的文献，并选择一些基线。列举的时候可以采取下述两种方式：（也可以采取其他的，这里仅作参考）

- 基线 1.
- 基线 2.

## 4 研究计划

1. 收集训练数据
2. 清洗数据
3. 构建基于 CNN、GAN、AutoEncoder 的神经网络
4. 构建 GUI

5. 训练

6. 测试

## 5 可行性