

Нейросетевой синтез текстур с трендами

Будалян Я. С.
Научный руководитель Грачев Е. А.

2017 г.

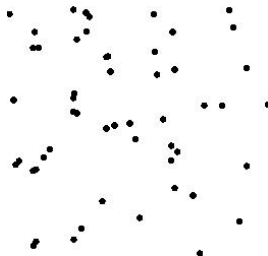
Задача состоит в построении и обучении искусственной нейронной сети, которая будет способна синтезировать текстуры с трендом. В качестве тренда в текстуре подразумевается какое-то заданное изменение статистических свойств изображения, подчиненное некоторой функциональной зависимости. В качестве таких свойств могут выступать, например, интенсивность появления частиц, пористость среды, и т.п.

Постановка задачи

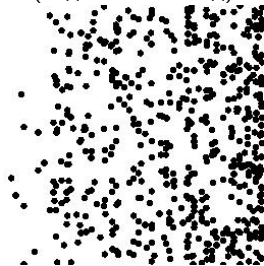
Поскольку задача достаточно общая, в работе рассматривается ее сужение:

- В качестве статистического свойства рассматривается интенсивность появления частиц λ
- Тренд направлен вдоль оси x (горизонтальный тренд)
- Тренд является линейным: $\lambda = \lambda_0 + kx$
- Обрабатываются монохромные изображения 256×256 пикселей
- На вход нейросети подается 2 изображения, с интенсивностью частиц слева и справа (при крайних значениях λ), на выходе ожидается изображение с трендом

Пример входных и выходных изображений



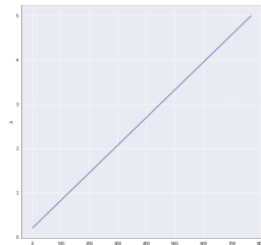
Распределение слева
(подается на вход)



Ожидаемый выход нейросети



Распределение справа
(подается на вход)



Тренд интенсивности

Математическая постановка задачи

Математически задача обучения искусственной нейросети сводится к оптимизации некоторого функционала потерь, конкретный вид которого зависит от выбранной архитектуры сети. На данный момент для решения поставленной задачи я экспериментирую с архитектурой GAN(генеративная состязательная сеть) для которой функционал выглядит следующим образом:

$$\mathcal{L}(G, D) = \mathcal{L}_{adv}(G, D) + \eta \mathbb{E}_{s_1, s_2, r \sim p_{data}(s_1, s_2, r)} (\|r - G(s_1, s_2)\|_1)$$

$$\begin{aligned} \mathcal{L}_{adv}(G, D) = & \mathbb{E}_{s_1, s_2, r \sim p_{data}(s_1, s_2, r)} [\log D(s_1, s_2, r)] + \\ & + \mathbb{E}_{s_1, s_2 \sim p_{data}(s_1, s_2)} [\log(1 - D(s_1, s_2, G(s_1, s_2)))] \end{aligned}$$

где G, D - ИНС 'генератор' и 'дискриминатор', (s_1, s_2, r) - изображения с интенсивностью слева, справа и реальное изображение с трендом соответственно, $\mathbb{E}_{s_1, s_2, r \sim p_{data}(s_1, s_2, r)}$ - математическое ожидание логарифмического правдоподобия того, что тройка изображений (s_1, s_2, r) принадлежит вероятностному распределению реальных троек $p_{data}(s_1, s_2, r)$, а $p_{data}(s_1, s_2)$ соответствует распределению реальных изображений s_1, s_2 . Оптимизация идет по весам сетей G и D соответственно.