

Unsupervised Discovery of Interpretable Directions in the GAN Latent Space

1. Опишите подход к поиску интерпретируемых направлений в скрытом пространстве, предложенный в работе Unsupervised Discovery of Interpretable Directions in the GAN Latent Space. Какие дополнительные модули потребовались обучить для реализации подхода?
2. Как авторы Unsupervised Discovery of Interpretable Directions in the GAN Latent Space предлагают использовать собственные результаты в задаче Saliency Detection?
3. Какие две метрики авторы Unsupervised Discovery of Interpretable Directions in the GAN Latent Space использовали для оценки качества найденных направлений? Объясните как и почему они менялись в эксперименте, где модель не училась предсказывать длину смещения вдоль направления.

The Hessian Penalty: A Weak Prior for Unsupervised Disentanglement

1. Какой эффект на порождающую модель может оказывать регуляризация с помощью гессиана? Какой неформальный аргумент приводят авторы для объяснения?
2. Какую оценку авторы предлагают для суммы не лежащих на диагонали элементов гессиана? Какое минимальное число вызовов генератора необходимо для вычисления оценки?
3. Как гессиан генератора может быть использован для выделения независимых направлений в скрытом пространстве генератора?

Parallel WaveNet: Fast High-Fidelity Speech Synthesis

1. За что отвечает параметр расширения (dilation) в свертке? Как и для чего он применяется в архитектуре WaveNet?
2. Как устроена генерация в моделях WaveNet и Parallel WaveNet?
3. Для обучения нормализующих потоков необходимо уметь находить z_t , порождающий обучающий объект x_t . Блок Inverse Autoregressive Flow для входа $z = (z_1, \dots, z_T)$ возвращает вектор $x = (x_1, \dots, x_T)$ согласно правилу $x_t = z_t * s(z_{1:t-1}) + m(z_{1:t-1})$, где s и m задаются некоторой авторегрессионной архитектурой. Как по данному вектору x построить исходный вектор z ?
4. Опишите процедуру Probability Density Distillation, предложенную для обучения модели Parallel WaveNet.

NVAE: A Deep Hierarchical Variational Autoencoder

1. Опишите архитектуру NVAE: как кодировщик вычисляет скрытое представление изображения? Каким образом скрытые представления подаются на вход кодировщику?
2. Какие решения авторы NVAE предлагают для стабилизации на этапе обучения слагаемого с дивергенцией Кульбака-Лейблера?
3. На этапе обучения каждый слой нейронной сети хранит массивы чисел, необходимые для вычисления градиентов на обратном проходе. Для экономии памяти авторы NVAE использовали решение, которое позволяет для последовательной пары блоков Batch Norm + Swish Activation не хранить в памяти результат вычисления Batch Norm. В чем оно заключается? Какие у него недостатки?

GAN Dissection: Visualizing and Understanding Generative Adversarial Networks

1. Назовем блоком (он же unit) некоторый канал на промежуточном этапе вычисления генератора. Как авторы GAN Dissection предлагают находить блоки, коррелированные с генерацией заданных наперед классов на изображении?
2. Как авторы GAN Dissection предлагают выделять блоки, изменение которых приводит к появлению или исчезновению объектов заданного класса на изображении?
3. Всегда ли в работе GAN Dissection удастся избавиться от объектов данного класса на сгенерированном изображении? Как авторы поясняют такую особенность предложенного метода?