

Курс Введение в нейронные сети

Практическое задание к курсу:

2. Сделайте краткий обзор какой-нибудь научной работы, посвященной тому или иному алгоритму нейронных сетей, который не рассматривался на курсе. Проведите анализ: Чем отличается выбранная вами на рассмотрение архитектура нейронной сети от других архитектур? В чем плюсы и минусы данной архитектуры? Какие могут возникнуть трудности при применении данной архитектуры на практике?

Ответ

В научной работе Large Scale GAN Training for High Fidelity Natural Image Synthesis авторы Эндрю Брок, Джефф Донахью, Карен Симонян предлагают модель BigGAN.

BigGAN является одним из видов генеративных соревновательных сетей, разработана для масштабирования генерации изображений с высоким разрешением и высокой достоверностью. Она включает ряд изменений и нововведений, а именно:

- Использование SAGAN в качестве базового решения со спектральной нормализацией для G и D, и использование TTUR.
- Использование Hinge Loss
- Использование классово-обусловленной пакетной нормализации для предоставления информации о классе в G (но с линейной проекцией не MLP)
- Использование проекционного дискриминатора для D для предоставления информации о классе в D.
- Оценка с помощью EWMA весов G, как в ProGANs.

К числу новшеств относятся:

- Увеличение batch size, что существенно влияет на Inception Score модели.
- Увеличение ширины каждого слоя приводит к дальнейшему улучшению Inception Score.
- Добавление пропускающих соединений от латентной переменной z в последующие уровни помогает производительности.
- Новый вариант ортогональной регуляризации.

Модель способна генерировать очень впечатляющие изображения с высокой точностью. При обучении на наборе данных ImageNet с разрешением 128x128, BigGAN может достичь Inception Score (IS) 166.3, что более чем на 100% лучше предыдущего результата (SotA) 52.52. Балл стартового расстояния Frechet (FID) также улучшился с 18,65 до 9,6.

Высокий Inception Score означает, что GAN может генерировать много разнообразных отличимых изображений. Однако Inception Score имеет ограничения, которые важно знать.

(1) Если генерировать что-то, не присутствующее в тренировочных данных (например, акул нет в ILSVRC 2014), то всегда можно получить низкий IS, несмотря на генерацию высококачественных изображений, поскольку это изображение не классифицируется как отдельный класс.

(2) Если вы генерируете изображения с другим набором лейблов из тренировочного набора (скажем, вы тренируете GAN для создания различных разновидностей пуделей, или только слонов и муравьев), то IS может быть низким.

(3) Если сеть не может обнаружить признаки, относящиеся к вашей концепции качества изображения (например, есть доказательства того, что CNNs сильно полагаются на локальные текстуры изображений для классификации, и грубые формы не имеют большого значения), то изображения низкого качества могут получить высокие оценки. Например, сеть может создать людей с двумя головами и не получит наказание за это.

(4) Если генератор генерирует только одно изображение из класса изображений, повторяя каждое изображение много раз, он может дать высокий IS (т.е. нет никакой меры внутриклассового разнообразия)

(5) Если генератор запомнит тренировочные данные и реплицирует их, он может получить высокий IS.