

Model działania GAN'u

Dyskryminator

Sieć neuronowa, z funkcją aktywacji sigmoid, czyli na końcu mamy jeden neuron, który zwraca prawdopodobieństwo od 0 do 1, czy badany obraz jest prawdziwy czy nie.

Generator

Na podstawie zestawu cech(losowo generowanych liczb) generuje obrazek(banknot).

Ważne: możesz używać gotowych struktur sieci neuronowych ale nie mogą być już sieci wyuczone. Zarówno Generator jak i dyskryminator powinny być uczone równolegle

Jak ocenić GAN?

- Patrzymy po tym jak obrazki są generowane po prostu. Tą drogą się raczej idzie.
- Frechet Inception Distance
 - jest to po prostu wzór oceny na podstawie odpowiedzi modelu Inception
 - mamy rozkład danych sztucznych i prawdziwych i porównujemy te rozkłady z poziomem cech. Wykorzystuje się do tego Inceptoon v3. Daleko to źle oszukaliśmy, blisko, to dobrze.

DCGAN - 2016

- W przeciwieństwie do tradycyjnych GANów, DCGAN wykorzystuje warstwy konwolucyjne w generatorze i dyskryminatorze. W generatorze stosowane są transponowane warstwy konwolucyjne do "rozpakowywania" wektora zanurzenia w pełnowymiarowy obraz.
- DCGAN eliminuje warstwy w pełni połączone z architektury generatora i dyskryminatora, co pomaga w redukcji problemów związanych z uczeniem i stabilnością.
- Wprowadza batch normalization w obu sieciach, co pomaga w stabilizacji procesu uczenia poprzez normalizację wejść do każdej warstwy.

Jak generować konkretną cyfrę?

- Conditional GAN - przyjmowanie jakiegoś y'ka by wiedział że jest to np pies i uczył się rozróżniać

Wasserstein GAN - 2017

- dyskryminator zwraca teraz informację o tym jak bardzo źle jest obraz generowany bądź dobrze.
- ma to na celu ułatwić przyporządkowanie generowanych obrazów bliżej docelowych grup (np psa do psów)

BiGAN /AliGAN

- Teraz dodajmy jeszcze wektory cech do generatora. (Za pomocą encodera)
- lepsze zrozumienie szumu, dzięki czemu jesteśmy w stanie lepiej określić detale w wygenerowanym obrazku

CycleGAN - 2017

- Zmieniamy style w obrazkach
- jeden generator tworzy jakiś szum, który drugi generator próbuje odtworzyć
- dyskriminatorem jest Unet
- Kontrolujemy jeszcze w postaci 2 dyskriminatorów: 1 do stylu drugi do oryginału obrazka.

Pix2Pix

- pary, zamiana stylu przedmiotu na ten sam przedmiot w innym stylu, np. szkic kota na zdjęcie kota, widok z mapy satelitarnej na mapę topograficzną, przeglądową

GauGAN

- bardziej zaawansowany pix2pix, interaktywny paint

CycleGan

- mamy zdjęcia koni i chcemy przerobić na zebry, który wyglądają jak koń,
- wykorzystywane do zmiany stylu np. z obrazka zimy robimy obrazek lata