

2020 年度 卒業論文

GAN を用いた手書き数字画像の生成

龍谷大学 理工学部 数理情報学科

T170068 野上 光志朗

指導教員 佐野 彰

概要

人工知能 (AI) の進歩により私たちの日常生活は少しずつ変化し、人と人工知能が共生する未来が近づいている。人工知能は人にしかできないと思われた作業をそつなくこなし、さらには人間の知能を越えようとしている。人工知能の一つの分野である機械学習は人に備わる学習能力と同じような機能をコンピュータで再現しようとする技術はあらゆる場面で応用されている。その技術の一つとして近年その中でも、GAN(敵対的生成ネットワーク)と言われる機械学習の技術が注目されている。GAN の大きな特徴は画像を変化させることである。GAN の応用例として、実在しない人物や風景の画像を作り出す画像生成や、シマウマの画像を馬の画像に変える画像変換（逆も可能）、動画内の人物の顔を別の人物の顔に入れ替えるディープフェイク、また低解像度の画像を高解像度に変換させることなどがあげられる。それらを活用してネットのファッションショッピングサイトの会社では無数の衣服を分析することにより、指定したスタイルに適合する新しいデザインを自動生成し、おすすめファッションの生成に用いる実験を行なっている。新しいアイデアを発想させる GAN の現状を知るとともに、その中でも画像生成のシステムを実装したいと考え本研究を始めるきっかけとなった。

本研究では GAN と DCGAN と CGAN の 3 つの種類の GAN を用いて Google Colaboratory のクラウドサービス上で MNIST というデータセットを使い、手書き数字画像を生成した。3 つの GAN それぞれを実装し性能を比較した。普通の GAN では 20000 回の学習で 1486 秒かかりいくつかの画像にはノイズが入り綺麗な画像は少なかった。DCGAN では他の 2 つと比べて少ない回数での学習で綺麗な画像が生成できたが生成できた画像に偏りがあった。1800 秒で普通の GAN より綺麗な画像が生成できた。CGAN では他の 2 つの GAN と比べて 320 秒の短い時間で綺麗な画像が生成される学習が行えた。CGAN と DCGAN のラベルごとの学習では DCGAN では時間がかかったが学習回数が少ないことがわかった。また CGAN は時短だが学習が多いことがわかった。3 つの GAN を実装した結果、学習のパラメータによって同じソースコードでも全く違う画像が生成されたり綺麗な画像が生成されないことが多々あった。これはモード崩壊や勾配消失などの問題が起きていたと考えられる。より深い層で学習を行うとより綺麗な画像が生成できたが、学習が複雑になる問題があった。これらのモード崩壊や勾配消失、複雑な深層学習が今後の課題として挙げられる。