

5. 敵対的生成ネットワーク (GAN) による異種モダリティ画像生成

大竹 義人*¹/日朝 祐太*¹/松岡 拓未*¹/高尾 正樹*²
Jerry L. Prince*³/菅野 伸彦*²/佐藤 嘉伸*¹

*1 奈良先端科学技術大学院大学先端科学技術研究科 *2 大阪大学大学院医学系研究科
*3 Department of Electrical and Computer Engineering, Johns Hopkins University

敵対的生成ネットワーク (generative adversarial network : GAN)¹⁾ は、画像や音声、文章などのデータを生成する生成器 (generator) と呼ばれるニューラルネットワークと、生成されたデータが目的のデータ (生成したい画像や音声) に近いかどうかを判定する識別器 (discriminator) と呼ばれるニューラルネットワークを同時に学習することで、目的のデータに近い新しいデータを生成するネットワークである。文章の入力から、それが表す画像を生成する研究²⁾ や音声信号のノイズ除去³⁾ など、応用分野は非常に多岐にわたる。本稿では、特に医用画像分野での一つの応用例として、われわれが進めているMR画像とCT画像のドメイン変換⁴⁾ を紹介する。

GAN の概念・基礎技術

2014年に、Ian Goodfellowによって初めてGANが発表された論文¹⁾ の例を用いて、基本的な概念を説明する(図1)。

入力となる教師画像データセットは、生成したい画像と同じ分布を持つ(似たような)画像群である。この論文では例えば、手書き数字画像や顔写真、動物の写真などが例として用いられた。

GANは、生成器(generator)と識別器(discriminator)という2つのニューラルネットワークから構成される。生成器は、ランダムなベクトルから何らかの画像を生成する。識別器は、生成器からの出力は偽物、教師画像データセットからの出力は本物、という識別が正しくできた時に低い値を示す損失関数を最

小化する方向に学習(重みを更新)する。逆に生成器は、その損失関数を最大化する(生成器を欺く)方向に重みを更新する。生成器と識別器の重み更新を交互に行うことで、2つの敵対するネットワークの学習が同時に進み、生成器は本物と見分けのつかない画像を生成するようになる。つまり、教師画像データセットに含まれる画像の分布を学習し、その分布に含まれる(見分けがつかない)新しい画像を作り出すのがGANである。

ただし、実問題への応用においては、限られたデータセットからの学習の安定化など、まだ解決されていない問題も多くあり、ニューラルネットワークのトップ国際会議でも活発に議論が進んでいる研究途上の技術である。典型的な学習の失敗例の一つとして、mode collapseが挙げられる。例えば、生成器が、どのような入力を与えても教師画像データセットの中のどれか1枚とまったく同じ画像を常に生成するようになれば、上

記の2つの敵対するネットワークの要求条件を満たすことになり、学習は収束してしまう。このように、生成器が教師画像データセットの分布の一部(1つのmode)のみを学習してしまう状況をmode collapseと呼ぶ。生成器がより多様なデータを生成できるようにするなど、mode collapseを回避するためのさまざまな改良が提案されている。

GANによる画像のドメイン変換

コンピュータビジョンでのGANの応用分野の一つとして、画像のドメイン変換の問題が挙げられる。画像のドメイン変換は、手で描かれた絵画を写真に変換する、あるいはウマの写真をシマウマに変換するなど、性質の異なる(別のドメインの)画像同士を相互に変換する技術である。この技術は、すでにディープラーニングの台頭以前の2001年頃から、

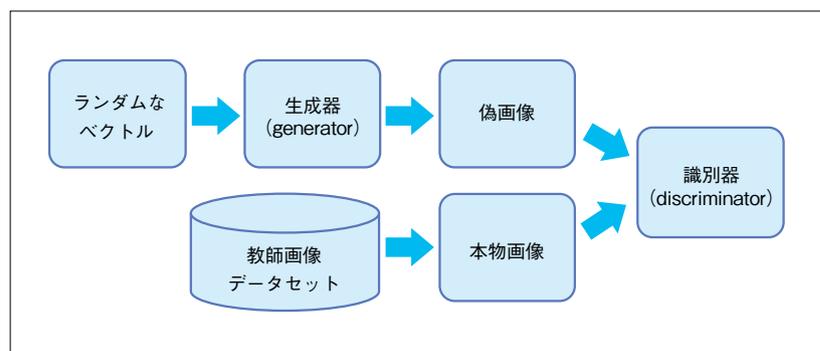


図1 GANの構造¹⁾ (Ian Goodfellowにより初めて提案されたモデル)
生成器(generator)と識別器(discriminator)という2つのニューラルネットワークから構成される。詳細は本文参照。