

# CycleGAN よる音楽の自動変換

陶山大輝

所属：教養学部後期課程学際科学科総合情報学コース

学籍番号：08-192021

指導教員：金子知適准教授

2020 年 12 月 26 日

## 概要

# 目次

第 1 章	はじめに	2
第 2 章	背景	3
2.1	ニューラルネットワークとは . . . . .	3
2.2	ディープラーニングとは . . . . .	3
2.2.1	GAN とは . . . . .	3
2.2.2	pix2pix とは . . . . .	4
第 3 章	実験の詳細	5
3.1	データセット . . . . .	5
3.2	評価 . . . . .	5
第 4 章	提案手法	6
第 5 章	まとめ	7

## 第1章 はじめに

## 第2章 背景

### 2.1 ニューラルネットワークとは

ニューラルネットワークとはニューロンとニューロン間のシナプスによる結合で形成される脳のネットワークを模した数理モデルのことである。入力層と出力層を持ち、シナプスの結合強度を変化させることで問題に最適なネットワークを構成することを目標とする。

また、入力層と出力層の間に隠れ層を加えて多層にし層間に活性化関数を用いて非線形分離を行うことで、複雑なネットワークを作ることができる。さらに、任意の活性化関数が微分可能であれば、誤差逆伝播法により損失関数を高速に求めることができる。

### 2.2 ディープラーニングとは

ディープラーニングとは層をより深くしたニューラルネットワークを用いた機械学習の手法のことである。多大な計算資源を必要とするが、GPUを含む計算機の性能の向上により実用的な手法となった。

また、層を増やすと勾配の減衰による勾配消失や訓練データへの最適化による過学習などの問題が発生するが、前者の場合は活性化関数に ReLU 関数を用い後者の場合は汎化性能を測定することで避けることができる。

#### 2.2.1 GAN とは

GAN(敵対的生成モデル) は生成モデルと識別モデルが競合して学習を行うディープラーニングのネットワークモデルである。生成モデルは訓練データの分布を捉えようとし、識別モデルは訓練データの分布から出てきたものである確率を推定する。

具体的には、訓練データの分布を  $p_{data}$ , 生成モデルの入力となるノイズの分布を  $p_z$ , ノイズを元にデータを生成する関数を  $G$ , 生成モデルのデータではなく訓練データである確率を返す関数を  $D$  とした時に、下記の式を  $G$  は最大化し  $D$  は最大化をすることを目標として学習を行う [1]。

$$\mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

### 2.2.2 pix2pix とは

## 第3章 実験の詳細

### 3.1 データセット

### 3.2 評価

## 第4章 提案手法



## 第5章 まとめ

## 謝辭

## 関連図書

- [1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.