

神経力学モデルによる複雑図形の学習と生成

久保野 雄亮[†]西出 俊[‡]康 鑫[‡]任 福継[‡][†] 徳島大学 大学院創成科学研究科[‡] 徳島大学 大学院社会産業理工学研究部

1. はじめに

近年、力学情報が学習可能な深層学習モデルである Recurrent Neural Network (RNN) を用いて様々な動的対象を学習し生成させる研究が盛んに行われている。その中でも我々は、描画系列の学習と生成に注目している。Google Brain の Ha らはスケッチ RNN を用い、様々な手描きスケッチを学習・認識・生成した [1]。本研究では単一の神経力学モデル MTRNN を用いて複数の複雑図形を学習し、生成する手法を提案する。

2. 神経力学モデル：MTRNN

本研究では、神経力学モデルとして Multiple Timescale Recurrent Neural Network (MTRNN) [2] を学習モデルとして用いる。MTRNN の構造を図 1 に示す。

2.1 MTRNN の構成

MTRNN は入出力層である IO (Input-Output Unit)、コンテキスト層である C_f (Fast Context Unit)、 C_s (Slow Context Unit) の 3 種類のニューロン群からなる階層的な構造を持っており、現状態 $IO(t)$ を入力して次状態 $IO(t+1)$ を出力する時系列予測器である。各ニューロン群には時定数と呼ばれるパラメータが設定されており、時定数の大きさで発火速度が制御される。本モデルでは IO 、 C_f 、 C_s の順に発火速度が遅くなるように設定する。発火速度の違いによって MTRNN は異なるレベルの情報を階層的に学習することが可能である。各ニューロン値の計算式については紙面の都合上省略する。

2.2 MTRNN の機能

MTRNN には、学習・認識・生成の 3 つの機能がある。学習は学習させる時系列データを IO ニューロン群に入力し、結合の重みの値と C_s の初期値である $C_s(0)$ を更新する。 $C_s(0)$ は各時系列データを表現する固有の値として決定される。認識は IO ニューロン群に認識する時系列データを入力し、 $C_s(0)$ のみを更新する。求まった $C_s(0)$ は認識する時系列データを表現する値である。生成は任意の $C_s(0)$ を入力し、それが表現する時系列を逐次計算によって求めることである。

2.3 MTRNN を用いた描画モデル

前節で述べた MTRNN の機能を用い、単一の MTRNN で学習と生成を行うモデルを構築する。本研究では描画時にペンタブレットを用い、ペンの座標 (x, y) が取得できるものとする。

複雑図形を描画する際にストロークとストロークの間ではペンを持ち上げる必要があり、ペンの座標が取得できない。ペンの持ち上げに伴う座標データ欠損の補完については、データ欠損時には前時刻の出力値を現時刻の入力値とすることで対応する。また最終的にロボットがペンを上げ下げさせながら描画することを考慮すると、ペンの持ち上げ状態を推定する必要がある。ペンの状態(描画中か持ち上げ中)の推定問題については、ペンの状態によって発火する state ニューロンを付加することで対応する。state ニューロンは描画中には発火せず、ペンを持ち上げている時に発火する。提案モデルによって描画系列を生成する際は state ニューロンが未発火の時に描画し、state ニューロンが発火している時はペン持ち上げとし描画しない。モデルの学習は複雑図形を描画した時に得られる座標系列を用いて行う。MTRNN の学習後、学習機能によって得られた $C_s(0)$ を用い、生成機能によって $C_s(0)$ が表す描画系列全体を生成する。

3. 実験設定

提案手法の評価実験として、Wacom 社製の液晶タブレット Cintiq13HD(DTK-1301/KO)(図 2) に描画して取得した図形データを用いる。被験者には図形データとして、図 3 に示す 10 種類の図形(林檎・蝶・鈴・車・ひよこ・魚・花・家・桃・かたつむり)を描いてもらった。各図形とも矢印の 1 から順に描き始め、100 系列ずつ図形データを取得した。描画時に液晶タブレットの座標値 (x, y) ・state ニューロン値(描画時:1, ペン持ち上げ時:0)を 10(frames/sec) で取得した。 IO ニューロン数はペン先の座標である (x, y) であるため 2 つ、state ニューロン数は 1、 C_f 、 C_s のニューロン数は複数の構成の中から最も良いものを実験的に設定した。各ニューロン群の発火関数にはシグモイド関数を用い、 IO 、state、 C_f 、 C_s の時定数をそれぞれ、2、2、5、70 と設定した。

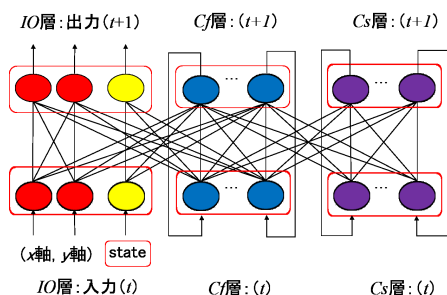


図 1: MTRNN の構造



図 2: 液晶タブレット

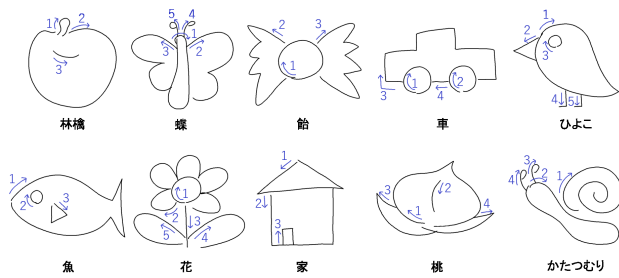


図 3: 学習用図形

4. 実験結果

実験において、各図形の学習誤差と標準偏差を表 1 に示す。図 4 では、各図形の描画系列 100 個のうち学習誤差が最小のものを生成した結果を示す。図 5 では、各図形の描画系列 100 個のうち学習誤差が最大のものを生成した結果を示す。図 4、図 5 とともにオリジナルは被験者が描画した描画系列であり、生成結果は MTRNN が生成した描画系列である。

各図形とも比較的綺麗に生成できていることを確認した。学習誤差や標準偏差が大きい飴や車の描画系列の中には、生成が上手くできずに図形が認識しづらいものがあった。

図形	学習誤差	標準偏差
林檎	0.000123	0.000058
蝶	0.000205	0.000040
飴	0.000872	0.006897
車	0.000438	0.002622
ひよこ	0.000115	0.000045
魚	0.000107	0.000060
花	0.000303	0.000845
家	0.000080	0.000022
桃	0.000105	0.000048
かたつむり	0.000297	0.000092

表 1: 各図形の学習誤差と標準偏差

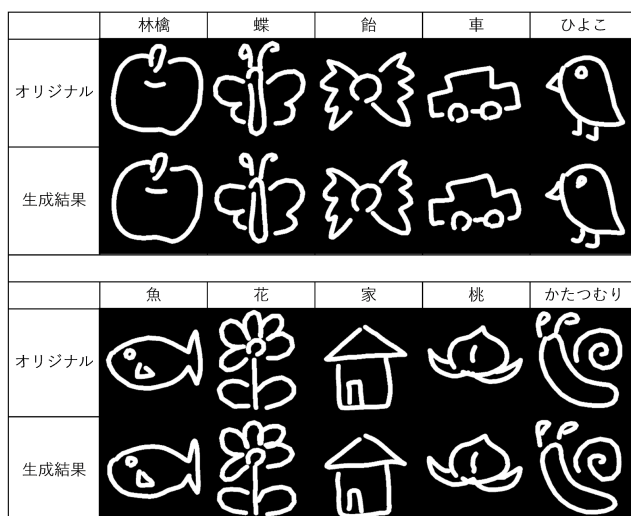


図 4: 最も誤差が小さかった生成結果

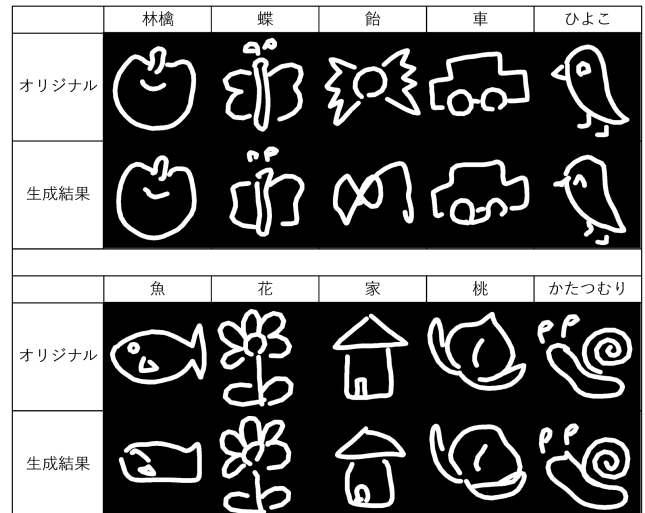


図 5: 最も誤差が大きかった生成結果

5. 考察

実験結果より、飴の学習誤差が一番大きく描画系列ごとの学習誤差のばらつきも大きいことが確認できた。この要因については、力学モデルである MTRNN の学習特性が影響していると考えられる。力学モデルを用いた時系列データの学習では急激な変化を扱うことが困難であり、飴の包装紙部分のように尖った部分の学習が難しい。車や花についても同様の原因であると考えられる。一方で他の図形については急激に変化する部分が少ないため、MTRNN による学習に適していたと考えられる。様々な描画対象を学習するためには力学モデルの学習性能を向上することが今後の課題である。

6. おわりに

本稿では単一の神経力学モデル MTRNN を用いた複雑図形の学習と生成について報告した。提案手法ではペンの持ち上げに対応したモデルを MTRNN によって構築し、学習と生成の機能を用いて複数の複雑図形を生成する。評価実験は液晶タブレットで描いた 10 種類の図形を用いた。実験により、複数の描画系列を学習し生成できることを確認した。

今後はモデルの改良を行うと共に、認識機能を用いた描画系列の予測について検証したいと考えている。

謝辞

本研究の一部は徳島大学クラスター事業 (番号: 2003002) の支援を受けたものである。

参考文献

- [1] D. Ha and D. Eck: "A neural representation of sketch drawings," arXiv preprint arXiv: 1704.03477(2017).
- [2] Y. Yamashita and J. Tani: "Emergence of Functional Hierarchy in a Multiple Timescale Recurrent Neural Network Model: a Humanoid Robot Experiment," PLoS Computational Biology, Vol.4, No.11, e1000220, 2008.