

1. はじめに

人間の脳には、役割により機能する箇所が分かれていく機能分化と呼ばれる現象がある。一方、人工ニューラルネットワークの1つである、Echo State Network (ESN) [1] においても、特定の時間または空間パターンに反応するニューロンを獲得する機能分化が確認されている [2]。本手法は、Genetic Algorithm (GA) を用いて ESN の Reservoir 層のニューロンの接続構造と重みを探索する。しかし、GA による最適な重みの探索には膨大な時間が必要になる。そこで本研究では、DARTS を用いて ESN の接続構造探索を行い、探索した接続の重みを誤差逆伝播により最適化する方法を提案する。

2. 先行研究

ESN は入力層、Reservoir 層、出力層の3層で構成され、出力層のみを学習する。入力層と Reservoir 層の重みは 0 から 1 の間でランダムに決定する。Yamaguti ら [2] は、図 1 のように Reservoir 層のニューロンを Input Neurons と Output Neurons に分けた拘束条件付き ESN を提案した。更に、この拘束条件付き ESN を対象に Genetic Algorithm (GA) を用いて構造探索を行う手法を進化 ESN として提案した。拘束条件付き ESN の計算を式 (1) と式 (2) に示す。

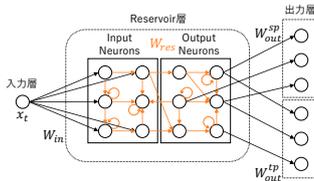


図 1. 拘束条件付き ESN のネットワーク構造

$$h_T = f(W_{in}X_T + W_{res}h_{T-1}) \quad (1)$$

$$y_T = f(W_{out}h_T) \quad (2)$$

W_{in} は入力重み、 X_T は入力であり、 h_T は時刻 T における出力である。 W_{res} は Reservoir 重みであり、前時刻の出力である h_{T-1} との積により過去の情報を考慮した Reservoir 層の出力を求める。この Reservoir 層の出力と W_{out} の積に活性化関数 f を適用し、ESN の最終出力を求める。拘束条件付き ESN を含めて一般的な ESN は出力層のみを学習対象としているため W_{out} のみが最適化される。進化 ESN はこの拘束条件付き ESN の出力層の学習後、遺伝的アルゴリズム (Genetic Algorithm) を用いて Reservoir 層の重みと接続構造を探索する。このとき、全ニューロンの構造の組み合わせは膨大なため、進化 ESN によって最適な重みと接続構造の両者を十分に探索することは困難であるという問題がある。

3. 提案手法

本研究では、DARTS を用いて ESN の接続構造の探索を行い、探索した接続に対する重みの最適化を行う DARTS-ESN を提案する。提案手法では、DARTS [3] を用いてニューロン間に評価重みを定義し、接続構造を探索する。その後、探索した接続に対する重みの値を誤差逆伝播により最適化する。この接続構造探索と重み最適化の構造探索手法を DARTS-ESN と定義する。DARTS-ESN の構造探索の流れを図 2 に示す。

3.1. DARTS による接続構造探索

ESN の接続構造の探索に DARTS を適用する。この接続構造探索時の ESN モデルを図 3 に示す。DARTS を適用した ESN は Reservoir 層の各ニューロンが接続している場合の接続評価重みと、接続していない場合の非接続評価重みの 2 つを持っており、それらの出力を Softmax 関数によってスケールする。 $\alpha_o^{(i,j)}$ はニューロン i と j の

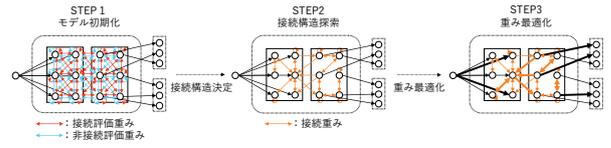


図 2. DARTS-ESN の構造探索の流れ

非接続評価重み、 $\alpha_1^{(i,j)}$ を接続評価重みを表す。式 (3) に DARTS を適用した Reservoir 層の出力計算を示す。

$$\hat{o}^{(i,j)}(X_T) = \sum_{o \in O} \frac{\exp(\alpha_o^{(i,j)})}{\sum_{o' \in O} \exp(\alpha_{o'}^{(i,j)})} o(X_T) \quad (3)$$

ここで、 O は接続を表す 2 つの候補 (非接続を 0, 接続を 1) の集合である。ニューロン間の接続 (i, j) において考え

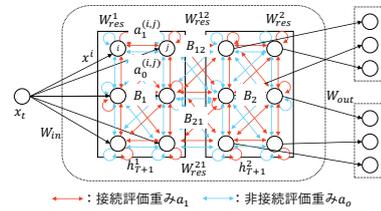


図 3. DARTS による接続構造最適化モデル

られる全ての演算候補の集合である評価重み $\alpha_o^{(i,j)}$ を誤差逆伝播により評価重みを更新する。これにより、精度に良い影響を及ぼす、接続状態 $\alpha_o^{(i,j)}$ の評価重みが増大する。誤差が収束するまで重みを更新した後、接続構造毎により大きい評価重みを持つ構造を選択することで、最適な接続を決定する。

3.2. 誤差逆伝播による重みの最適化

DARTS により接続構造を探索した後、全ての層を誤差逆伝播を行い重みを最適化する。拘束条件付き ESN は通常、入力層と Reservoir 層の学習を想定していないため、Reservoir 層において拘束条件と接続構造を維持するように計算式を変更する。重み最適化する際の拘束条件付き ESN の計算式を式 (4)、式 (5)、式 (6)、式 (7) に示す。

$$\hat{h}_T = W_{in}x_T + W_{res}^1 b_1 h_{T-1}^1 + W_{res}^{21} b_{21} h_{T-1}^2 \quad (4)$$

$$h_T^2 = f(W_{res}^{12} b_{12} \hat{h}_T^1 + W_{res}^2 b_2 h_{T-1}^2) \quad (5)$$

$$h_T^1 = f(\hat{h}_T^1) \quad (6)$$

$$y_T = f(W_{out} h_T^2) \quad (7)$$

式 (4) で Input Neurons の内部状態を式 (5) で Output Neurons の内部状態をそれぞれ計算することで、拘束条件と接続構造を維持したまま重みを最適化できる。式 (4)、式 (5) において、Input Neurons の内部状態は \hat{h}_T^1 、Output Neurons の内部状態は h_T^2 とする。Reservoir 層重みは、Input Neurons と Output Neurons の関係を考慮して W_{res}^1 , W_{res}^{12} , W_{res}^2 , W_{res}^{21} に分割する。また、学習対象とせず、誤差逆伝播後も拘束条件と接続構造を維持する b_1 , b_2 , b_{12} , b_{21} を定義する。また、活性化関数は式 (2) の h_T と式 (7) の h_T^2 の計算結果が同一になるよう式 (6) で計算する。最後に、式 (7) において、Reservoir 層出力 h_T^2 と出力重み W_{out} の積によって計算される出力 y_T と正解ラベルとの誤差を算出し誤差逆伝播によって重みを最適化する。

4. 評価実験

本研究では、進化 ESN と DARTS-ESN の構造探索手法の評価を行う。

4.1. 実験概要

学習データは空間情報と時間情報を組み合わせたパターンを用いる。このとき、縦方向の分割数を空間情報、横方向の周波数を時間情報と見立てることでマルチタスクを実現する。また、クラス分類の数を変化させ、6クラス分類、8クラス分類、10クラス分類と分ける。これにより、タスク難易度毎の初期化条件、ニューロン数と接続率の関係性を分析する。実験に使用したコサイン波の波形パターンの一例を図4に示す。

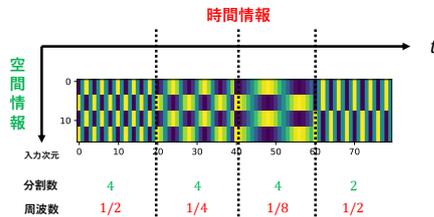


図4. 実験用データの波形パターン

4.2. ESN の構造探索

進化 ESN と DARTS-ESN の構造探索を行う。初期化手法には、He の初期化を用いる。拘束条件付き ESN を対象として進化 ESN は Reservoir 層と出力層の構造探索を行い、DARTS-ESN は全ての層の重みと Reservoir 層の接続構造の構造探索を行う。Reservoir 層のニューロン数は 32, 64, 128 とする。進化 ESN では世代数 10, 個体数 100, 生存個体数 20, Epoch 数 1 で GA を行い、DARTS-ESN は接続構造探索 100 Epoch, 重み最適化 ESN モデルでの重み最適化 100 Epoch とする。各手法の各クラス分類数における構造探索の結果から、最も良い性能を獲得したと考えられるモデルを選択した。それらモデルの空間情報精度、時間情報精度とニューロン数を比較する。各クラス分類数毎の進化 ESN の分類精度と DARTS-ESN の分類精度を表1に示す。

表1. 分類精度 [%]

		進化 ESN (ニューロン数)	DARTS-ESN (ニューロン数)
		6 クラス分類	空間情報 50.21 (32)
	時間情報 45.25 (32)	65.84 (128)	
8 クラス分類	空間情報 47.46 (64)	84.62 (64)	
	時間情報 34.89 (64)	61.25 (64)	
10 クラス分類	空間情報 51.85 (64)	85.52 (128)	
	時間情報 20.89 (64)	54.81 (128)	

表1より、進化 ESN と比較して DARTS-ESN の方が全てのクラス分類数において空間情報、時間情報の両方の分類精度が高い結果となった。DARTS-ESN は進化 ESN より高精度なモデルを獲得する構造探索ができた。また、DARTS-ESN は8クラス分類において、64ニューロン、6、10クラス分類においては128ニューロンが高い精度を示している。一方、進化 ESN はニューロン数が少ない、32もしくは64の時に高い精度を示した。これは進化 ESN の構造探索は GA で行うため、ニューロン数128において十分な構造探索できなかつたためであると考えられる。DARTS-ESN は接続構造の探索と重み最適化を分けて行うため、ニューロン数が多くても最適な構造探索が可能である。

4.3. ESN の構造分析

入力情報の空間情報と時間情報の正解ラベルと各ニューロンの出力の相互情報量の算出を行った。この算出によりニューロンが情報を特化して処理ようになる機能分化を分析する。横軸に空間情報量、縦軸に時間情報量を示す。図5(a)に8クラス分類64ニューロンの時の進化 ESN の

機能分化の分析を、図5(b)に8クラス分類64ニューロンの時の DARTS-ESN の機能分化の分析を示す。

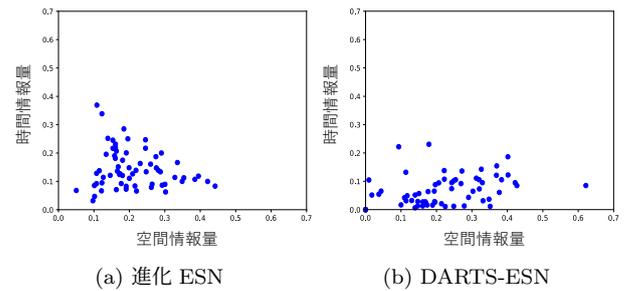


図5. 8クラス分類における機能分化の結果

図5(a)から、進化 ESN は空間情報と時間情報の両者において各情報に特化したニューロンが発生する機能分化が生じている。図5(b)から、DARTS-ESN は空間情報のみの特化したニューロンが多く発生する機能分化であることが分かる。その他の実験条件における機能分化の分析からも進化 ESN と DARTS-ESN では、機能分化は異なる傾向である事が分かった。具体的には、進化 ESN は両方の情報に対してそれぞれ特化したニューロンが発生する機能分化を誘発し、DARTS-ESN では空間情報のみの特化したニューロンが発生する機能分化を誘発することが分かった。

また、精度と機能分化の傾向についても分析する。表1の8クラス分類の進化 ESN と DARTS-ESN の空間情報の結果と図5(a)、図5(b)から、DARTS-ESN は進化 ESN と比べて、空間情報量へのニューロンの機能分化が多く、空間情報精度も高いことが分かる。一方で、表1の8クラス分類の進化 ESN と DARTS-ESN の時間情報の結果と図5(a)、図5(b)から、DARTS-ESN は進化 ESN と比べて、時間情報量へのニューロンの機能分化が少ないにもかかわらず時間情報精度が高いことが分かる。この結果から、DARTS-ESN の機能分化は進化 ESN とは違う傾向の機能分化を示す構造を獲得しているものの、問題設定に対して有効な構造を獲得していると言える。

5. おわりに

本研究では、DARTS を用いて ESN の接続構造探索を行い、誤差逆伝播を用いて重み最適化を行う手法を提案した。結果、従来の手法である進化 ESN と比較して、高精度な ESN の構造を獲得し、進化 ESN とは異なる機能分化の特徴を持つことを確認した。今後の予定としては、獲得できた ESN の構造のより詳しい分析やニューロンカットによる分類精度の確認などが挙げられる。

参考文献

- [1] H. Jaeger, *et al.*, “Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication”, Science, 2004.
- [2] Y. Yamaguti, *et al.*, “Functional differentiations in evolutionary reservoir computing networks”, Chaos 2021.
- [3] Liu, *et al.*, “Darts: Differentiable architecture search”, arXiv preprint arXiv:1806.09055, 2018.

研究業績

濱地優輝, 平川翼, 山下隆義, 藤吉弘亘, 山口裕, 津田一郎, “Hessian-Free 法を用いた Echo State Network の構造探索”, FIT2021 (第20回情報科学技術フォーラム), 2021.