

スパイキングニューラルネットワークの 時空間データ処理への応用

日大生産工(院) ○神田 宗一 日大生産工 松田 聖

1 まえがき

近年の研究で、生体では情報をニューロンの発火頻度を表す従来の発火率表現ではなく、発火の時刻自体や前後関係などが情報を表しているという報告がなされている。そこで、発火の時刻自体を直接用いるタイミング表現に基づいているスパイキングニューラルネットワークは脳により近いニューラルネットワークとして注目されている。特に、時空間データ処理への適性に期待が寄せられている。

本研究では、従来のパーセプトロン型のニューラルネットワークとスパイキングニューラルネットワークをNetTalkを用いて比較実験を行う。

2 NETTalk

文字で書かれた英語の文章を読む際に、例外的な読み方が多くあることは、よく知られている。DECTalkと呼ばれる商用の文章一発音変換システムでは、例外的な処理をするために大量の単語と発音の対応表が用意されている。NetTalkとは、このDECTalkと同機能を持つように訓練されたニューラルネットワークのことである。

図1にNETTalkの構造を示す。NETTalkのニューラルネットワークの入力層は7つのグループに分けられている。各グループはアルファベット26文字と空白、句読点（“、”、“または”、“.”）からなる29の入力ユニットから構成されている。この入力ユニットのどれか1つが1の値をとることにより提示された文字を表している。この7つのグループは、各時点に入力される7つの文字列に対応しており、中央の文字に対応する発音情報がNETTalkから出力される。また両側の3文字は中央の文字の発音を決定するための文脈情報として用いられ、各時点毎に一文字ずつ右

に移動することにより、英文章が音読されることになる。

NETTalkの出力である発音情報は調音点、有声無声、母音の高さなどに分類された21の調音特徴とアクセントや音節境界を表す5つの特徴の組み合わせで表現されている。なお、発音されない文字や文字素の一部になっている文字に関しては、“—”で表現される。

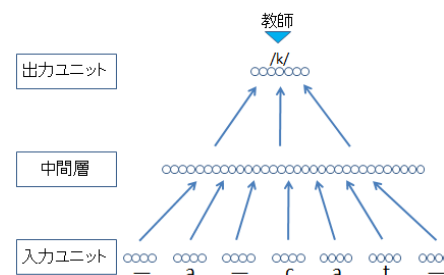


Fig. 1 NETTalkの構造の図式表現

NETTalkの学習には発話データと辞書データが用いられた。発話データは、単語間に区切りのデータを入れた文章を1文字ずつずらして入力した。同様に辞書データもランダムに1つの単語を選択し、1文字ずつずらして入力した。これらのデータをNETTalkに入力し、出力データと教師信号との誤差を最小にするようにバックプロパケーション法を適用する。ただし、結合荷重の修正量は、1単語ごとに修正される量を平均化して用いる。

3 スパイキングニューラルネットワーク

従来のパーセプトロン型ニューラルネットワークで時空間データの処理を行うには、NETtalkのように前後の3文字を与えるような特殊な方法を用いなければ処理することが

できなかった。一方、スパイキングニューロンではメモリ機能があり、Fig.2のように1つのニューラルネットワークに、時間とともに次々と入力を与えられ、一定の条件が成り立つたびにスパイクが出力を行うことで処理することができる。図2ではスパイキングニューラルネットワークの構造を図で示した。出力されたスパイクはスパイクトレインと呼ばれ $F = \{t^1, t^2, \dots, t^p\}$ あるいは

$$s(t) = \sum_{t^f \in F} \delta(t - t^f) \quad (1)$$

として表される。

t^f はスパイクの発火時刻、 $\delta(x)$ はディラックのデルタ関数で、 $x=0$ のときは $\delta(x)=1$ 、それ以外ならば $\delta=0$ である。

また後シナプスニューロン i の膜電位 $u_i(t)$ は次のように表される：

$$u_i(t) = \sum_j w_{ij} x_j(t) = \sum_j \sum_{t^f \in F_j, t^f < t} w_{ij} \varepsilon_{ij}(t - t^f) \quad (2)$$

ここで w_{ij} は前シナプスニューロン j から後シナプスニューロン i へのシナプスの結合荷重、 ε_{ij} は j から i への後シナプス電位、 t^f 及び $F = \{t^1, t^2, \dots, t^p\}$ は後シナプスニューロン i のスパイクの発火時刻及びスパイクトレインである。(2)の式から、膜電位は入力を受け取るたびに加算されているため、メモリの機能を持ち、時系列データ処理が期待できる。また、膜電位が閾値を超えると後シナプスニューロンはスパイクを出力し、同時に膜電位は静止電位となる。

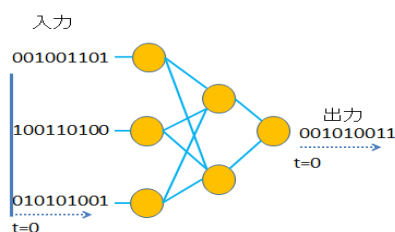


Fig. 2 スパイキングニューラルネットワークの構造

そこで、自然な時空間データ処理機能が期待できるスパイキングニューラルネットワークでNETtalkと同一の機能をもつシステムを自然に実現することができる。

4 実験方法

辞書データの英単語約500語をNETTalkと同一の機能を持つシステムに従来のニューラルネットワークとスパイキングニューラルネットワークの両方に学習させた後に、テストデータで発音記号の正答率を従来の手法や学

習アルゴリズムごとに比較する。しかし、本論で用いるスパイキングニューラルネットワークの学習法は中間層には適用できないため、本研究では入力層から中間層までの荷重結合の値をランダムとして実験を行い、線形分離性の克服を行う。

本研究で用いる学習アルゴリズムは中間層から出力層にかけて適応する。従来のニューラルネットワークにはバックプロパケーション法を用いる。スパイキングニューラルネットワークの学習アルゴリズムは数多く提案されているが、本研究では下記の2つを用いて実験を行う。^{1) 2) 3)}

1つはスパイクを指数関数カーネル化し、2つのスパイクトレイン間での対応や、ずれ、欠損などを区別せずにスパイクトレイン間での相違を誤差関数として、従来のニューラルネットワークで用いられてきたデルタルールを直接適応する。

2つ目はニューロンの閾値に対する学習である。手法は従来のニューラルネットワークで行われた閾値学習と同様に行う。閾値学習だけでは、スパイクの出力数にだけ適応され、ずれの補正目的ではないが、閾値が変動するため、結果としてスパイクの出力時間が変動し、誤差を修正できると考えられている。

この2つの学習アルゴリズムを用いて、正答率などを比較し、考察していく。

「参考文献」

- 1) 松田聖, スパイキングニューラルネットワークの新学習アルゴリズム, 電子情報通信学会技術研究報告 NLP 非線形問題, No111, (2011) p.71-76
- 2) Terrence J. Sejnowski and Chael R. Rosenberg, Parallel Networks that Learn to Pronounce English Text, Complex System, 1987, p.145-168.
- 3) A. Mohemmed, S. Schliebs, S. Matsuda, and N. Kasabov, SPAN: Spike pattern association neuron for learning spatio-temporal spike patterns, International Journal of Neural Systems, vol.22, no.4, pp.1250012-1-17, 2012.
- 4) A. Mohemmed, S. Schliebs, S. Matsuda, and N. Kasabov, Teaching a precisely timed input-output behavior to a spiking neuron, Neurocomputing, 2012.12 掲載予定