

時系列信号を認識するリカレント・ニューラルネットワーク

3D-7

麻生川 稔

日本電気(株) C&C システム研究所

1 はじめに

現在、時系列構造を持つデータ(音声など)を取り扱うためのニューラルネットワークとして、時間遅れ要素を用いたTDNN[1]、DP マッチングと融合したDNN[2]、フィードバックループを持つリカレント・ニューラルネットワーク[3][4]などが提案されている。リカレント・ニューラルネットワークは内部状態を保持する機構(フィードバックループ)を持つために、階層型のニューラルネットワークと比較して、時間構造をもつ信号の認識において有利と考えられる。本実験ではJordanらによって提案されたリカレント・ニューラルネットワーク[3]を心電図波形の区分点認識に適用した。ノイズを重ねたデータによる評価を行なった。クラスタ分析を行なった結果、リカレント・ニューラルネットワークが時間構造をかなり良く抽出していることが示された。

2 リカレント・ニューラルネットワーク

構成は図1に示すような、3階層型のニューラルネットワークである。時刻Tにおいて、入力層の影響を受けた中間層の内容をコンテキスト層にコピーし時刻T+1の中間層への入力とする。その結果、中間層には過去の入力層の履歴が蓄積されることとなる。学習は、出力層の出力を教師信号に一致させるように行うので、コンテキスト層から中間層への重みは出力層の誤差が減少するように変更される。これらのことから、「ある時刻Tにおいてはコンテキスト層に時刻T-1以前に入力層の時系列信号の特徴を保持し、時刻Tでの入力の認識に用いている」と見なすことができる。

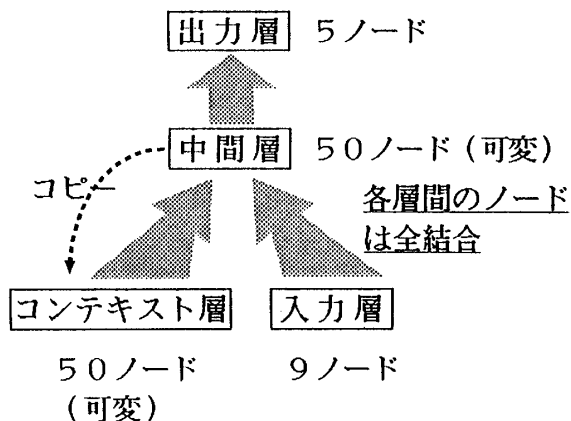


図1: リカレント・ニューラルネットワークの構成例

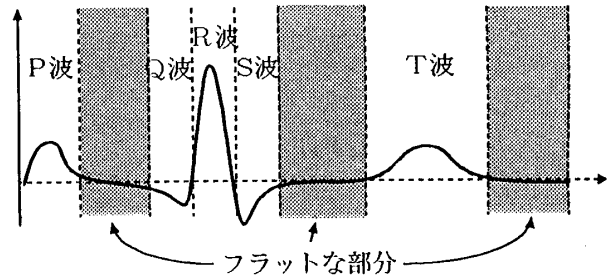


図2: 心電図波形の区分点

3 心電図波形の区分点認識

3.1 心電図波形

本実験では、時系列データとして心電図波形を用い、その区分点の認識を行なわせることを目的とした。心電図波形は一般に図2に示すように、P波、フラットな部分、Q波、R波、S波、フラットな部分、T波、フラットな部分、の時系列で構成されている(本実験ではQ波の認識を行っていない)。区分点の認識とは、各々の波とフラットな部分の始まりと終りを検出することである。

3.2 区分点認識用リカレント・ニューラルネットワーク

リカレント・ニューラルネットワークは入力層9個、中間層50個、コンテキスト層50個、出力層5個のノードによって構成されている。出力層のノードは各々、P波、R波、S波、T波、フラットな部分に対応づけた。時刻Tのリカレント・ニューラルネットワークの学習では、時刻T-4から時刻T+4の時系列信号を入力層に入力する。中間層では、入力層からコンテキスト層の出力を入力とし、入力の総和を求め、シグモイド関数を用いて出力を決定する。出力層には教師信号として時刻Tの波形区分を提示する。この時の誤差を誤差逆伝搬法(バックワード・プロパゲーション・アルゴリズム)を用いて各層間の重みの修正を行なう。この時の中間層の出力を次のコンテキスト層の内容とする。次の時刻T+1の学習では、T-3からT+5の時系列信号を入力層に入力し、出力層にはT+1の波形の区分を提示し、学習を行ない各層間の重みの修正を行なう。

認識時では重みの修正を行わない以外は前記と同じ方式とした。

4 実験結果と考察

図3に原信号にノイズを重ねさせた信号の“フラットな部分”の区分点の認識結果を示す。

比較のために、同一の学習データで学習させたコンテキスト層を持たないネットワークにノイズを重ねさせた信号を提示した時の出力を図3の上段に示した。

図3において、教師信号が1の部分(フラットな部分)では、リカレント・ニューラルネットワークの出力が低

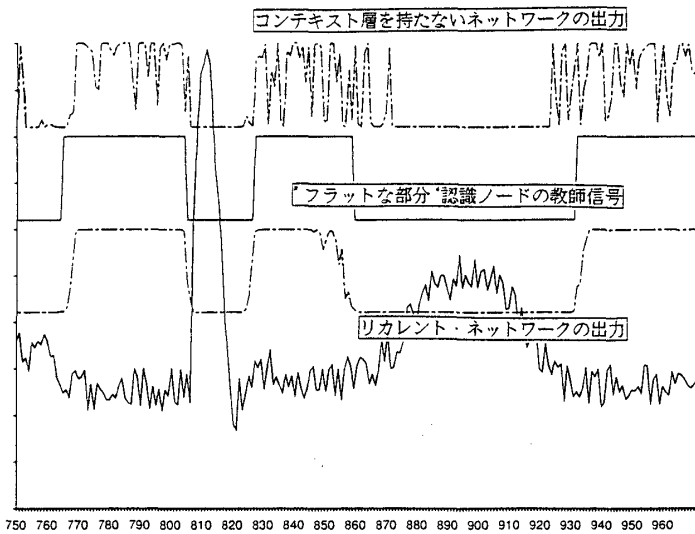


図 3: ノイズを含んだ波形の“フラットな部分”の認識結果

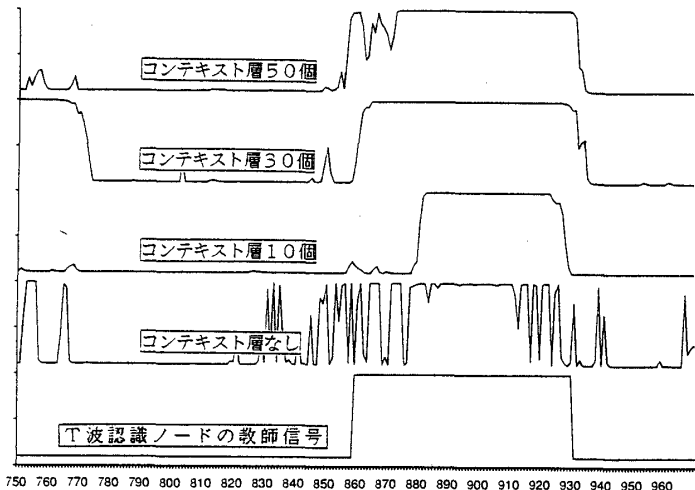


図 4: コンテキスト層のノードを増やした時の T 波認識結果

ば教師信号と一致しているのに比較して、コンテキスト層を持たないネットワークでは一致していない。また、両ネットワークとも、教師信号が 0 の部分(“フラットな部分”でない部分)では、出力は教師信号と一致している。

ここには示していないが、P 波等の区分についても同様な傾向が見られた。

図 4 にはコンテキスト層中のノードの数を変更した時の T 波の区分認識の結果を示す。コンテキスト層中のノードの数を増やすにつれて、より教師信号に一致している。

以上の 2 つの結果から、コンテキスト層に以前の時系列信号の情報が蓄えられ、それが時系列の波形を認識に、有効であることが分かった。

学習済みのリカレント・ニューラルネットワークでは、コンテキスト層に以前の時系列信号の情報がどの様に蓄えられているかを調べる為に、コンテキスト層の内容をクラスタ分析を試みた(群平均法、非類似度としてユークリッド距離)。図 5 にコンテキスト層 50 個の学習済みのリカレント・ニューラルネットワークに学習データを提示した時の、コンテキスト層の内容のクラスタ分析結果を示す。

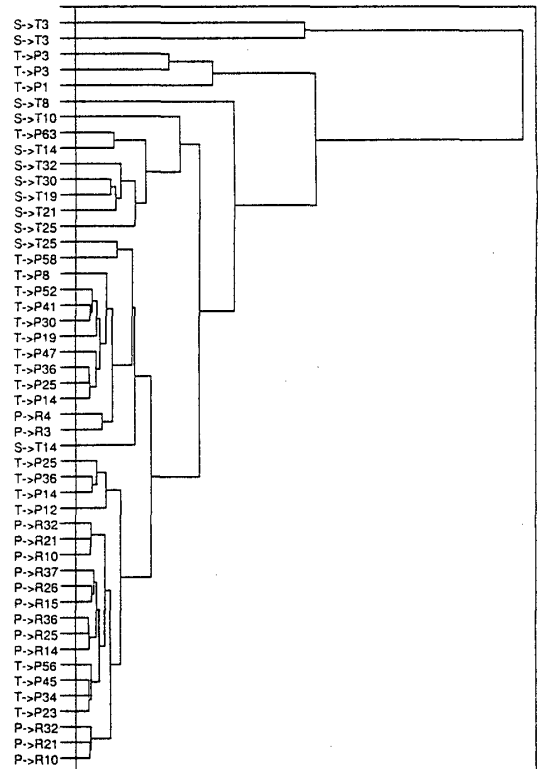


図 5: コンテキスト層の内容のクラスタ分析結果

P 波と R 波の間のフラットな部分 (P->R**), S 波と T 波の間のフラットな部分 (S->T**), T 波と P 波の間のフラットな部分 (T->P**) については、入力信号がほぼ同じにもかかわらず、各々別のクラスタを構成している。このことから、コンテキスト層には時系列信号の構造が反映されていると考えることができる。

5 まとめ

時系列の信号を認識するためのリカレント・ニューラルネットワークを提案し、実験によって時系列信号に対する有効性を確認した。また、時系列信号の構造にしたがってコンテキスト層が構成されることを確認した。

謝辞 本研究に対して有益なコメントを頂いた日本電気(株)C&C システム研究所の大野部長、久保主管研究員、中崎課長、難波主任に感謝致します。

参考文献

- [1] ワイベルほか、“時間遅れ神経回路網 (TDNN) による音韻認識”、信学技報、SP88-85, pp.19-24, (1988)
- [2] 迫江ほか、“ダイナミックニューラルネットワークの提案—神経回路網と DP マッチングに基づく新しい音声認識モデル—”、電子情報通信学会論文誌、vol. J71-D, no. 7, pp.1341-1344 (1988)
- [3] Jordan, M.I., “Attractor Dynamics and Parallelism in a Connectionist Sequential Machine”, in *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, (1986)
- [4] Servan-Schreiber, D. et.al. “Encoding Sequential Structure in Simple Recurrent Networks” Technical Report CMU-CS-88-183, Carnegie Mellon University. (1988)