

1 E-6 複数の異種ベクトルを入力とするニューラルネットの学習・統合化に関する考察

加藤誠巳 水谷明弘 炭 親良
(上智大学理工学部)

1. まえがき

誤差逆伝搬法(BP法)を用いて学習した多層型ニューラルネットワークにおいては入力とする特徴ベクトルとしてどのようなものを選ぶかが重要であり、複数の異種ベクトルを入力すると有利であることが予想される。本稿ではこれら複数の異種ベクトルを入力とするニューラルネットにおいて、異種入力ベクトルが互いに認識能力を補うような学習・統合化の手法について検討を行った結果について述べる。

2. 統合ネットワークの構築

2.1 ネットワークの冗長性と多様性

学習を終了していないネットワークは認識能力が未熟であり、冗長なネットワークといえる。しかし、複数の異種特徴ベクトルを用意し、各特徴ベクトルを個別に冗長的に学習したネットワークを結合させた後に再学習を行う(統合する)ことにより、互いに劣る認識能力を補完することが期待される。この冗長性を有効に活かすためには、各ネットワークが多様であることが要求される。ここで言う多様性とは、各特徴ベクトルを学習したネットワークの誤認特性が異なることを意味する。

特徴ベクトルを一通り学習させる度(これを学習回数一回に数える)に、クローズ認識率および下記の条件を満たす比の値 PCR(Pass Condition Ratio)を求めることで、ネットワークの冗長度を調べる。

$$[\text{条件}] \quad \text{認識すべきカテゴリを表すセルが最大に発火し、且つ、} \\ (\text{最大発火値}) - (\text{二番目に大きく発火したセルの発火値}) > \text{TH}$$

クローズ認識率が1.0付近となるまで学習してもPCRが低いときのネットワークは、冗長であるとみなせ、クローズPCR(このとき、閾値THは0.7)が1.0となったとき、学習を終了させる。

2.2 統合ネットワークの構造

各特徴ベクトルを3層ネットワークで個別に学習させた後(図1(a))、各ネットワークの相応する出力セルを結合し、更に学習させる(図1(b))。但し結合させる際、中間層と出力層を結合するネットワーク数で割り、各出力セルの閾値は、各ネットワークの各セルの閾値を平均したものとする。

3. 手書き数字認識への応用

上述の統合ネットワークを手書き数字認識に応用した。今回用いた数字データは、タブレットのペンより入力された点列情報、およびペンアップ/ダウン情報を基に点列間を均一の太さを有する線分で結んだものである。また、大きさおよび位置に依存しないようにするため枠取り・正規化処理を行う。即ち外接長方形を用いて枠取り処理を行い、長辺の長さを一定の値(256ドット)に正規化し、更にこの長方形を外接正方形の中央になるように配置する。手書き数字データとしては、12人が0~9までの数字をそれぞれ8個ずつ作成した計960個を用いた。そのうち、240個をクローズデータ、720個をオープンデータとした。

3.1 特徴ベクトル

特徴ベクトルとして以下の3種類の情報を用意した(文献[1])。

- (1) 原イメージ情報 (イメージデータ)
- (2) 2次元FFTパワー情報 (FFTデータ)
- (3) 直交軸方向面積投影情報 (投影データ)

イメージデータは16×16のメッシュに相当する256次元ベクトルデータである。FFTデータは16×16の原イメージに対する2次元FFTを求め、その256個のパワーベクトルから直流成分および対称成分を除いた129次元ベクトルである。投影データは原イメージをX軸、Y軸方向へそれぞれ面積投影した32次元ベクトルである。

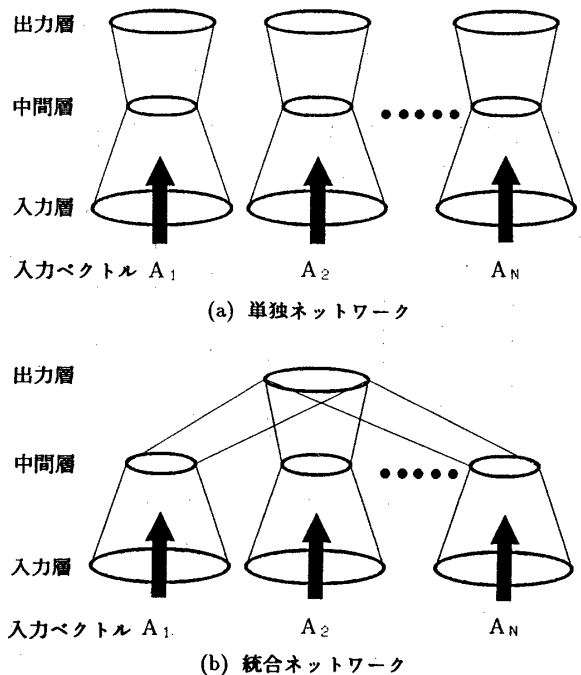


図1 ネットワークの統合

3.2 学習の方法

- (1) 各特徴ベクトルをそれぞれ単独に学習させ、クローズ PCR の値が 0.1 の整数倍付近のネットワークの状態をセーブする。(単独ネットワーク)
- (2) 各クローズ PCR のネットワークの統合学習を行う。(統合ネットワーク)
- (3) 本手法の有効性を確認するために、統合ネットワークと同じ構造のネットワークに重み・閾値を一様な乱数で与え、学習を行う。(結合ネットワーク)

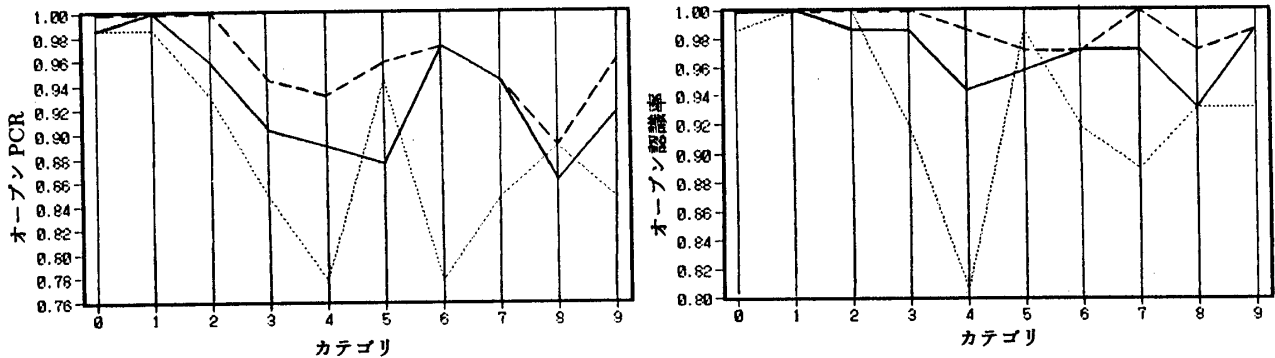
3.3 学習および認識結果

各ネットワークの認識結果を表 1 に示す。尚、統合学習は、イメージおよび FFT で夫々単独学習させた二つの単独ネットワークならびに、三種類の情報で夫々単独学習させた三つの単独ネットワークを統合する二つの場合について行った。クローズ PCR が 0.1 から 0.4 付近の単独ネットワークを統合したときオープン認識率が最高になる傾向があり、表 1 の統合ネットワークの認識結果は、いずれも最高オープン認識率が得られたクローズ PCR が 0.1 の単独ネットワーク統合時を示した。統合する単独ネットワークのクローズ PCR が大きくなるにつれ、統合学習回数は減少しオープン PCR は向上するのに対し、オープン認識率は低下する傾向があった。更に、図 2 にイメージおよび FFT で夫々学習した単独ネットワークとそれらを統合したネットワークの各カテゴリに対する認識能力 ((a) オープン PCR, (b) オープン認識率) を示す。

表 1 各ネットワークの学習認識・結果

	入力ベクトル			各層のノード数			学習回数	オープン	
	イメージ	FFT	投影	入力	中間	出力		PCR	認識率
単独ネット	○			256	10	10	25	0.583	0.974
		○		129	20	10	261	0.624	0.936
			○	32	12	10	195	0.681	0.947
結合ネット	○	○		385	30	10	57	0.675	0.981
	○	○	○	417	42	10	62	0.771	0.985
統合ネット	○	○		385	30	10	53	0.708	0.989
	○	○	○	417	42	10	50	0.738	0.992

(注) このオープン PCR の閾値 TH は 0.7 である。



(a) オープン PCR (注) このオープン PCR の閾値 TH は 0.3 である。

(b) オープン認識率

図 2 各カテゴリに対する認識能力 (——— イメージ単独 FFT単独 - - - - イメージ, FFT 統合)

4. 考察

表 1、図 2 から、各ネットワークの認識能力がうかがえ、ネットワークの統合により、互いに劣る認識能力を補完していることがわかる。また、単独ネットワークに比べ統合ネットワークの方が、そして統合する数が多い方が全体的にオープン PCR およびオープン認識率が高くなっており、統合学習により冗長度は低下し、認識能力は向上することが確かめられた。

更に、クローズ PCR が低い冗長な単独ネットワークを統合した方がオープン認識率が高くなることがわかった。一方、クローズ PCR が高く冗長度の低い単独ネットワークを統合するとオープン PCR は向上して冗長度は低下するが、オープン認識率も低下してしまい、統合後のネットワークでは冗長性と認識能力は一種の競合関係にあると考えられる。

5. むすび

手書き数字の学習・認識実験を通して、ここで提案したネットワークの学習・統合手法の有効性を確認した。更には、ネットワークの冗長性と統合力との興味深い関係が明らかになった。

終わりに、御討論いただいた本学マルチメディア・ラボの諸氏に謝意を表す。

参考文献

[1] 加藤、高木: " 複数の特徴ベクトルを利用したニューラルネットによる手書き数字認識 ". 信学技報, PRU88-151(平 01).