階層型ニューラルネットワークの学習における ノイズ重畳に関する解析と検証

大山 輝光[†1]　三好 邦男[†1]　黒田 英夫[†2]
宮原 末治[†2]　志久 修[†3]　中村 千秋[†4]

本論文では正規ノイズや一様ノイズの付加によって生じる階層型ニューラルネットの中間素子と出力素子の出力の変化が，ニューラルネットの出力に与える影響について解析的・実験的に考察する．まずノイズ重畳に関する定性的な解析を行い，学習初期においてはノイズの影響を大きく受けて素子の出力が不確定な状態に戻される逆方向の学習と，再学習が繰り返されるような探索的な進展が見られることを明らかにする．また，学習の最終局面ではノイズ重畳による素子出力の変位が小さくなると同時に，素子出力が 0 と 1 の確定的な値となるような影響を受けることを示す．さらに，解析的な考察を実験的に理解するため，入力へのノイズ重畳と未知データに対する汎化能力との関係を調べる手書き文字データベース ETL9B の認識実験を行い，学習後の中間素子の平均的な出力とニューラルネットの構造化について検討する．実験の結果，学習初期においてはノイズの影響を受けて擬似的な学習データの増加と探索的な振舞いが生じ，学習の最終局面では素子の出力が 0 か 1 の確定的な値となるように構造化され，汎化能力の向上が期待できることが確認された．

Influence of Noise Injection on Learning of Multi-layer Neural Networks

Terumitsu Ohyama,[†1] Kunio Miyoshi,[†1] Hideo Kuroda,[†2] Sueharu Miyahara,[†2] Osamu Shiku[†3] and Chiaki Nakamura[†4]

In this paper, we discuss the structurization of the neural network and influence of noise injection into inputs for improving the generalization capability. In pattern recognition using neural networks, it is difficult for users to optimum neural network structure by using back-propagation learning. To improve the generalization, we discuss an analytic consideration about neural outputs, particularly outputs of hidden and output units with noise injection. Next, training neural networks and applies its to the pattern recognition problem by using handwritten character database ETL9B is presented. Then, we discuss an experimental consideration about outputs of hidden units with noise injection into inputs.

1. はじめに

階層型ニューラルネットの汎化能力向上に対し，学習データにノイズを重畳する方法[1)~4)]や学習によって構築された超平面の枠内で汎化能力を向上させる方法[5)]，適切な内部表現を得ることによるアプローチ[6),7)]，そして中間素子数や基底関数を実験的・理論的に決定する方法[8)~10)]などの研究が行われている．また，未学習の入力パターンに対して正しい出力を保証するような枠組みとして，ニューラルネットの隠れ素子の重み表現に対する線形の従属制約を導入し学習を行う方法も提案されている[11)]．

本論文では，手書き文字認識などの実問題にニューラルネットを適用するために，正規ノイズと一様ノイズの付加がニューラルネットの出力と汎化能力に与える影響について考察する．まず，入力素子の入出力が線形関数で規定され，中間・出力素子の入出力が微分可能な任意の関数で規定される場合について，ノイズ重畳の平均的な影響を解析的に明らかにする．また中間素子と出力素子の入出力がシグモイド関数で規定される場合，反学習と再学習が繰り返し行われるような探索的な振舞いが発生し，最終的には素子の出力が確定的な値となるような構造化が行われることを示す．

†1 和歌山信愛女子短期大学
Wakayama Shin-Ai Women's Junior College
†2 長崎大学工学部
Faculty of Engineering, Nagasaki University
†3 佐世保工業高等専門学校
Sasebo National College of Technology
†4 長崎大学教育学部
Faculty of Education, Nagasaki University

次に，手書き文字データベース ETL9B を対象に文字種数を変えながら認識実験を行い，ノイズ重畳と学習誤差・汎化誤差の関係，およびニューラルネットの構造化について検討する．実験では，学習データ数の増加にともなう計算量の増大や，高い汎化能力を獲得するために必要な学習データを準備するのが難しいなどの問題[12)~14)] を考慮し，学習データをどの程度少なくできるのかを調べるため，学習データ数の変化による認識率の変化を調べるとともに，ノイズの変化と認識率の関係を定量的に調べる．最後に，学習後のニューラルネットの結合強度を分析し，ノイズ重畳には擬似的に学習データ数を増やす効果と，中間素子の出力が 0 か 1 の確定的な値となるように構造化を進展させる効果があることを示す．

2. ノイズ重畳の解析的考察

図 1 のような入力素子 I 個，中間素子 J 個，出力素子 K 個から構成される 3 層ニューラルネットについて考える．中間素子 j に入力される入力層からの信号を ξ_j と表すと，

$$\xi_j = \sum_{i=1}^{I} a_{ji}x_i + a_{0j} \tag{1}$$

となる．ただし，a_{ji} は入力〜中間層間の結合荷重，a_{0j} は中間素子のバイアスである．同様に，出力素子 k に入力される中間層からの信号を η_k と表すと，

$$\eta_k = \sum_{j=1}^{J} b_{kj}y_j + b_{0k} \tag{2}$$

となる．ただし，b_{kj} は中間〜出力層間の結合荷重，b_{0k} は出力素子のバイアスである．文献 5) では入力層と出力層の入出力を線形関数で，中間層の入出力をロジスティック関数，$f(u) = 1/\{1 + \exp(-u)\}$ と規定して，中間素子へのノイズ付加についてのみ検討しているが，本論文では，より一般的なモデルへの応用を視野に入れてさらに議論を広げ，中間素子および出力素子へのノイズ重畳について解析する．また，素子の入出力を規定する関数についてもさらに一般化し，入力素子の入出力を線形関数，中間素子と出力素子の入出力を微分可能な任意の関数 f と g によって次のように規定する．

$$y_j = f(\xi_j) \tag{3}$$

$$z_k = g(\eta_k) \tag{4}$$

入力層の入出力が線形関数で規定されている場合，入力の変位は中間層の入力への変位として解析的に考察できる．以下，3 層ニューラルネットにおけるノイズ重畳の影響について解析するため，中間層への入力信号に変位 m_j が生じる場合と，中間層の出力信号に変位 n_j が生じる場合について，変位がない場合と比較しながら考察する．

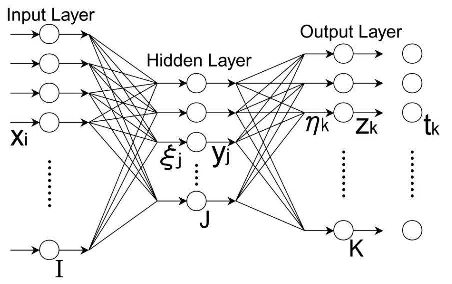

図 1 3 層ニューラルネットの構造
Fig. 1 Structure of the 3-layered neural network.

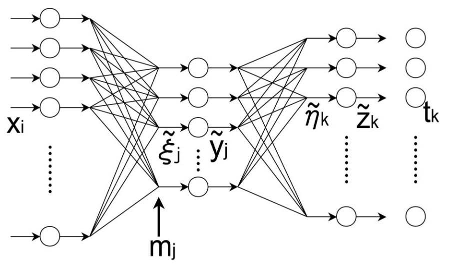

図 2 入力へのノイズ重畳
Fig. 2 Noise injection into inputs.

2.1 入力の変位に対する解析的考察

図 2 のように，式 (1) の形で現れた中間素子 j の入力に，平均が 0 の変位 m_j が生じた場合の中間素子 j の出力 $\tilde{y}_j$ は入出力関数 f によって，

$$\tilde{y}_j = f(\tilde{\xi}_j) = f(\xi_j + m_j) \tag{5}$$

と与えられる．なお，変位 m_j については平均が 0 であるということだけを規定し，以下の議論では変位 m_j が平均 0 の一様ノイズと正規ノイズの両方の場合で成り立つように解析を行う．

変位がない場合と比較するために，入出力関数 f を ξ_j の周りに Taylor 展開して 2 次の微小量までをとることにすれば，

$$\tilde{y}_j \cong y_j + \left\{ f'(\xi_j) + \frac{1}{2}f''(\xi_j)m_j \right\} m_j \tag{6}$$

となる．一方，出力層への入力 η_k は式 (2) の一般線形関数で表されるから，出力素子 k への入力は

$$\tilde{\eta}_k = \sum_{j=1}^{J} b_{kj}\tilde{y}_j + b_{0k} \tag{7}$$

となる．これを変位がない場合と比較するために，式 (6) を用いて 2 次の微小量までとれば，

$$\tilde{\eta}_k \cong \eta_k + \tilde{l}_k + \frac{1}{2}\sum_{j=1}^{J} b_{kj} f''(\xi_j) m_j^2 \tag{8}$$

となる．ここで，$\tilde{l}_k$ は 1 次の微小量

$$\tilde{l}_k = \sum_{j=1}^{J} b_{kj} f'(\xi_j) m_j \tag{9}$$

である．さらに，変位が生じた場合の出力素子 k の出力は式 (4) より，

$$\tilde{z}_k = g(\tilde{\eta}_k) \tag{10}$$

と与えられる．これを，変位がない場合の信号値で評価するために式 (8) を用いて，入出力関数 g を η_k の周りに Taylor 展開して 2 次の微小量までとれば，

$$\begin{aligned}\tilde{z}_k &\cong g(\eta_k) + g'(\eta_k)\left\{\tilde{l}_k + \frac{1}{2}\sum_{j=1}^{J} b_{kj} f''(\xi_j) m_j^2\right\} \\ &\quad + \frac{1}{2} g''(\eta_k)\left\{\tilde{l}_k + \frac{1}{2}\sum_{j=1}^{J} b_{kj} f''(\xi_j) m_j^2\right\}^2 \\ &\cong z_k + g'(\eta_k)\tilde{l}_k + \frac{1}{2} g''(\eta_k)\tilde{l}_k^2 \\ &\quad + \frac{1}{2} g'(\eta_k)\sum_{j=1}^{J} b_{kj} f''(\xi_j) m_j^2 \end{aligned} \tag{11}$$

となる．これによって，入力信号 x_i に対する教師信号 t_k と，出力信号 $\tilde{z}_k$ の 2 乗誤差

$$\tilde{\epsilon}^2 = \frac{1}{2}\sum_{k=1}^{K} (t_k - \tilde{z}_k)^2 \tag{12}$$

を，変位がない場合の 2 乗誤差 ϵ^2 と比較することができる．式 (12) の $\tilde{z}_k$ を，式 (11) で近似的に置き換えれば，

$$\begin{aligned}\tilde{\epsilon}^2 &\cong \frac{1}{2}\sum_{k=1}^{K}\Big[(t_k - z_k) - \Big\{g'(\eta_k) + \frac{1}{2} g''(\eta_k)\tilde{l}_k\Big\}\tilde{l}_k \\ &\quad - \frac{1}{2} g'(\eta_k)\sum_{j=1}^{J} b_{kj} f''(\xi_j) m_j^2\Big]^2 \end{aligned} \tag{13}$$

となるので，2 次の微小量まで評価すると，

$$\begin{aligned}\tilde{\epsilon}^2 &\cong \epsilon^2 - \sum_{k=1}^{K}(t_k - z_k) g'(\eta_k)\tilde{l}_k \\ &\quad + \frac{1}{2}\sum_{k=1}^{K}\left\{g'(\eta_k)^2 - (t_k - z_k) g''(\eta_k)\right\}\tilde{l}_k^2 \\ &\quad - \frac{1}{2}\sum_{k=1}^{K}(t_k - z_k) g'(\eta_k)\sum_{j=1}^{J} b_{kj} f''(\xi_j) m_j^2 \end{aligned}$$

となる．ここで式 (9) より，

$$\begin{aligned}&\sum_{k=1}^{K}(t_k - z_k) g'(\eta_k)\tilde{l}_k \\ &= \sum_{j=1}^{J} f'(\xi_j)\left\{\sum_{k=1}^{K}(t_k - z_k) g'(\eta_k) b_{kj}\right\} m_j \end{aligned}$$

さらに，

$$\tilde{l}_k^2 = \sum_{p=1}^{J}\sum_{q=1}^{J} b_{kp} f'(\xi_p) m_p b_{kq} f'(\xi_q) m_q$$

$$\sum_{j=1}^{J} b_{kj} f''(\xi_j) m_j^2 = \sum_{p=1}^{J}\sum_{q=1}^{J} b_{kp} f''(\xi_q)\delta_{pq} m_p m_q$$

だから，変位 m_j による 2 乗誤差への影響は

$$\tilde{\epsilon}^2 - \epsilon^2 \cong -\sum_{j=1}^{J}\tilde{S}_j m_j + \frac{1}{2}\sum_{p=1}^{J}\sum_{q=1}^{J}\tilde{T}_{pq} m_p m_q \tag{14}$$

となる．ただし，

$$S_j = \sum_{k=1}^{K}(t_k - z_k) g'(\eta_k) b_{kj}$$

$$\tilde{S}_j = f'(\xi_j) S_j$$

$$T_{pq} = \sum_{k=1}^{K}\left\{g'(\eta_k)^2 - \big(t_k - g(\eta_k)\big) g''(\eta_k)\right\} b_{kp} b_{kq}$$

$$\tilde{T}_{pq} = f'(\xi_p) f'(\xi_q) T_{pq} - f''(\xi_p) S_p \delta_{pq}$$

である．なお，δ_{pq} はクロネッカー記号である．

2.1.1　変位が正規ノイズの場合

中間素子 j の入力側で発生する変位 m_j が，次の確率密度関数で与えられる正規ノイズの場合を考える．

$$\rho(m_j) = \frac{1}{\sigma_j\sqrt{2\pi}}\exp\left(-\frac{m_j^2}{2\sigma_j^2}\right) \tag{15}$$

ノイズ m_j の期待値と分散は

$$\begin{aligned}&E[m_j] \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} m_j\rho(m_1\ldots m_J)dm_1\ldots dm_J \\ &= \int_{-\infty}^{\infty} m_j\rho(m_j)dm_j \\ &= 0 \end{aligned} \tag{16}$$

$$\begin{aligned}V[m_j] &= E\left[(m_j - E[m_j])^2\right] \\ &= \int_{-\infty}^{\infty} m_j^2\rho(m_J)dm_j \\ &= \sigma_j^2 \end{aligned} \tag{17}$$

となるので，変位 m_j が正規ノイズの場合にニューラ

ルネットに及ぼす平均的な影響は

$$E[\tilde{\epsilon}^2-\epsilon^2]\cong\frac{1}{2}\sum_{p=1}^{J}\sum_{q=1}^{J}\tilde{T}_{pq}\sigma_p^2\delta_{pq} \tag{18}$$

$$\begin{aligned}\tilde{\epsilon}^2-\epsilon^2 &\cong \frac{1}{2}\sum_{p=1}^{J}T_{pp}\sigma_p^2\\ &=\frac{1}{2}\sum_{j=1}^{J}\Bigg(\sum_{k=1}^{K}\Big[f'(\xi_j)^2\Big\{g'(\eta_k)^2\\ &\qquad -(t_k-z_k)g''(\eta_k)\Big\}b_{kj}^2\\ &\qquad -(t_k-z_k)g'(\eta_k)f''(\xi_j)b_{kj}\Big]\Bigg)\sigma_j^2\end{aligned} \tag{19}$$

となる．

2.1.2 変位が一様ノイズの場合

変位 m_j が，次の確率密度関数で与えられるような一様ノイズの場合を考える．

$$\rho(m_j)=\begin{cases}1/(d1-d2) & (d2\le m_j\le d1)\\ 0 & (m_j<d2, m_j>d1)\end{cases}$$

ただし，$d1$，$d2$ は定数 (>0) である．この場合，ノイズ m_j の期待値と分散は

$$E[m_j]=\frac{d_1+d_2}{2} \tag{20}$$

$$V[m_j]=\frac{(d1-d2)^2}{12} \tag{21}$$

となる．したがって $d2=-d1$ のように，期待値が 0 の一様ノイズがニューラルネットに及ぼす平均的な影響は

$$E[\tilde{\epsilon}^2-\epsilon^2]\cong\frac{1}{2}\sum_{p=1}^{J}\sum_{q=1}^{J}\frac{\tilde{T}_{pq}d1^2\delta_{pq}}{3} \tag{22}$$

$$\begin{aligned}\tilde{\epsilon}^2-\epsilon^2 &\cong \frac{1}{2}\sum_{j=1}^{J}\Bigg(\sum_{k=1}^{K}\Big[f'(\xi_j)^2\Big\{g'(\eta_k)^2\\ &\qquad -(t_k-z_k)g''(\eta_k)\Big\}b_{kj}^2\\ &\qquad -(t_k-z_k)g'(\eta_k)f''(\xi_j)b_{kj}\Big]\Bigg)\frac{d1^2}{3}\end{aligned} \tag{23}$$

となる．

2.1.3 入力ノイズの平均的な影響

$$\begin{aligned}e(f,g)&=f'(\xi_j)^2\Big\{g'(\eta_k)^2\\ &\qquad -(t_k-z_k)g''(\eta_k)\Big\}b_{kj}^2\\ &\qquad -(t_k-z_k)g'(\eta_k)f''(\xi_j)b_{kj}\end{aligned} \tag{24}$$

とすると，式 (19) と式 (23) はそれぞれ，

$$\tilde{\epsilon}^2-\epsilon^2\cong\frac{1}{2}\sum_{j=1}^{J}\Bigg(\sum_{k=1}^{K}e(f,g)\Bigg)\sigma_j^2 \tag{25}$$

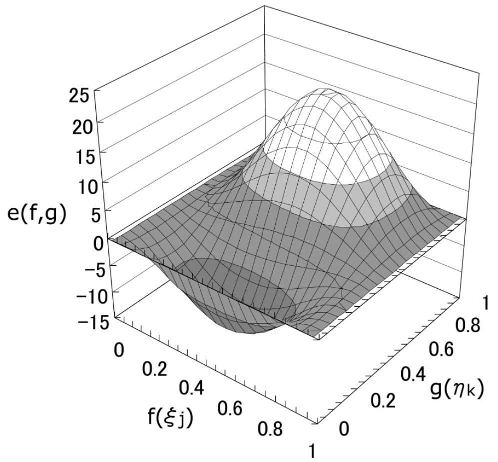

図 3 学習開始時の $e(f,g)$ の状態．$(t_k-z_k)=1$，$b_{kj}=50$ の場合

Fig. 3 The effect of $e(f,g)$ at the learning start. $((t_k-z_k)=1, b_{kj}=50)$.

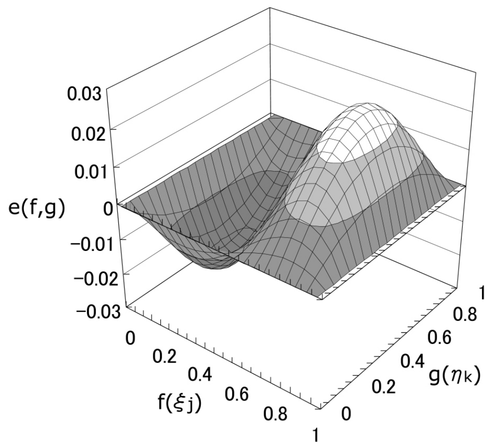

図 4 学習開始時の $e(f,g)$ の状態．$(t_k-z_k)=1$，$b_{kj}=1$ の場合

Fig. 4 The effect of $e(f,g)$ at the learning start. $((t_k-z_k)=1, b_{kj}=1)$.

$$\tilde{\epsilon}^2-\epsilon^2\cong\frac{1}{2}\sum_{j=1}^{J}\Bigg(\sum_{k=1}^{K}e(f,g)\Bigg)\frac{d1^2}{3} \tag{26}$$

となる．よってノイズ重畳による変位の平均的な影響は式 (24) の振舞いを調べることで近似的に解析できる．また，中間・出力素子の入出力を次のロジスティック関数

$$y_j=f(u)=\frac{1}{1+\exp(-u)} \tag{27}$$

で規定すると，式 (24) は $f(\xi_j)$ と $g(\eta_k)$ の関数として図によって表すことができる．学習開始時の $e(f,g)$ の様子を**図 3** と**図 4** に示す．図 3 は $(t_k-z_k)=1$，

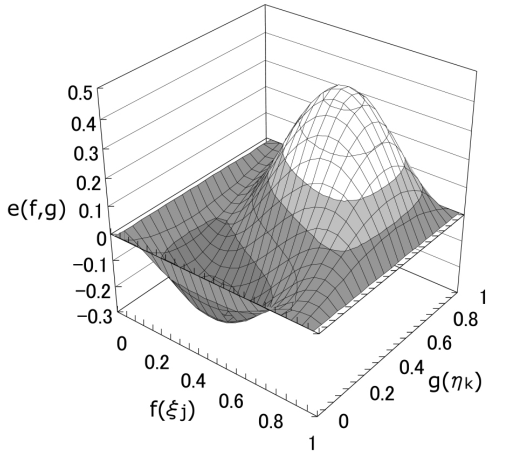

図 5　学習開始時の $e(f,g)$ の状態．$(t_k - z_k) = 1$，$b_{kj} = 7$ の場合

Fig. 5　The effect of $e(f,g)$ at the learning start. $((t_k - z_k) = 1, b_{kj} = 7)$.

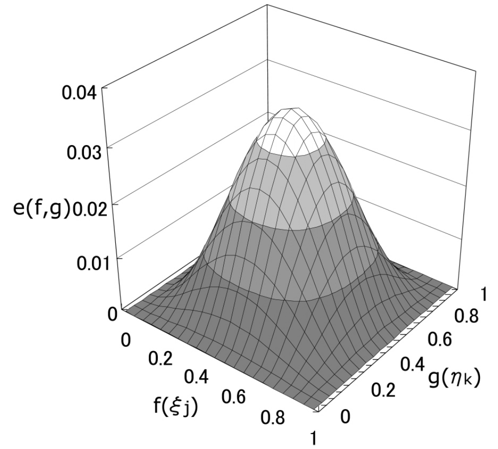

図 6　学習終了時の $e(f,g)$ の状態．$(t_k - z_k) = 0$，$b_{kj} = 3$ の場合

Fig. 6　$e(f,g)$ at the learning end. $((t_k - z_k) = 0, b_{kj} = 3)$.

$b_{kj} = 50$ の場合，図 4 は $(t_k - z_k) = 1$，$b_{kj} = 1$ の場合の $e(f,g)$ を示している．

図 3 のように $(t_k - z_k) = 1$，$b_{kj} = 50$ の場合，学習開始時のニューラルネットは，変位によって $g(\eta_k) < 0.5$ の場合には負の，$g(\eta_k) > 0.5$ の場合には正の影響を受け，出力素子の出力が 0 か 1，または 0.5 に近づくと影響が少なくなることが分かる．また，図 4 のように $(t_k - z_k) = 1$，$b_{kj} = 1$ の場合，学習開始時のニューラルネットは，変位によって $f(\xi_j) < 0.5$ の場合には負の，$f(\xi_j) > 0.5$ の場合には正の影響を受け，中間素子の出力が 0 か 1，または 0.5 に近づくと影響が少なくなることが分かる．さらに，**図 5** のように $(t_k - z_k) = 1$，$b_{kj} = 7$ の場合，負領域の中心が図 3 と図 4 の中間である $f(\xi_j) < 0.5$ と $g(\eta_k) < 0.5$ に移動する．このような正・負の領域の移動は，$(t_k - z_k)$ と b_{kj} の変化にともなって，$(f(\xi_j), g(\eta_k)) = (0.5, 0.5)$ を中心に回転するように観測される．そして，学習が進んで $(t_k - z_k)$ が小さくなると，負領域も小さくなる．

以上のことから，学習開始時において $(t_k - z_k)$ が大きい場合，変位 m_j によって素子の出力は 0 と 1 の確定的な値か，0.5 のように不確定な値となるような影響を受けることが分かる．通常の誤差逆伝播学習では，素子の出力は確定的な値になる方向のみに学習が進展するが，ノイズ重畳によって素子の出力が不確定な状態 ($f(\xi_j) = 0.5$ と $g(\eta_k) = 0.5$) に戻されるような逆方向の学習が生じ，その後，誤差逆伝播法による再学習が行われる．学習終了時の $e(f,g)$ の様子を**図 6** に示す．学習終了時のニューラルネットは，素子の出力が $g(\eta_k) = 0.5$ または $f(\xi_j) = 0.5$ 付近の場合，変位によって正の影響を受け，素子の出力が 0 か 1 に近づくような影響を受ける．そして，素子の出力が 0 または 1 のように確定的なら，変位の影響は受けない．

このように，学習終了時において $(t_k - z_k)$ が十分に小さくなると，変位 m_j の影響も小さくなる．このとき，もし中間素子と出力素子の出力が 0.5 と不確定な場合には（影響は小さいが）0 と 1 になるような修正が付加され，素子の出力がすでに確定的なら変位の影響は受けない．

以上のことより，学習初期には通常の学習に加え，変位の影響を大きく受けて素子の出力が不確定な状態に戻されるような逆方向の学習が加えられ，その後，再学習が行われるような探索的な進展が見られることが予想される．また，学習の最終局面では変位の影響は少なくなると同時に，素子の出力が 0 と 1 の確定的な値となるような影響を受けることが分かる．

2.2　出力素子の変位に対する解析的考察

図 7 のように，中間素子 j の出力に平均 0 の変位 n_j が生じた場合の出力信号を $\tilde{y}_j$ とすると

$$\tilde{y}_j = y_j + n_j$$

である．ここで，n_j は 1 次の微小量とする．中間層から出力層への信号伝搬は式 (2) の一般線形関数に従うので，出力素子 k への入力信号は

$$\begin{aligned}\tilde{\eta}_k &= \sum_{j=1}^{J} b_{kj}\tilde{y}_j + b_{0k} \\ &= \sum_{j=1}^{J} b_{kj}y_j + b_{0k} + \sum_{j=1}^{J} b_{kj}n_j \qquad (28)\end{aligned}$$

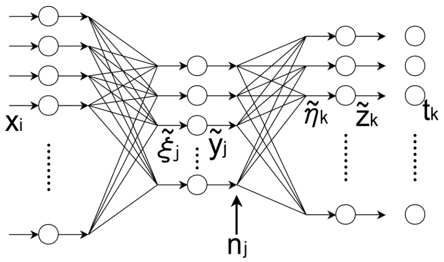

図 7 中間素子の出力へのノイズ重畳
Fig. 7 Noise injection into outputs of the hidden unit.

となる．よって，式 (2) により変位を含まない信号値で評価すれば，

$$\tilde{\eta}_k = \eta_k + l_k \tag{29}$$

ただし，l_k は n_j の 1 次の微小量

$$l_k = \sum_{j=1}^{J} b_{kj} n_j \tag{30}$$

である．出力素子の出力 z_k が微分可能な任意の関数 g によって $z_k = g(\eta_k)$ と規定されているとき，変位によって出力素子 k の出力は

$$\tilde{z}_k = g(\tilde{\eta}_k) = g(\eta_k + l_k) \tag{31}$$

となる．変位がない場合の信号値で評価するため，関数 g を η_k の周りに Taylor 展開して 2 次の微小量までとると，

$$\begin{aligned}\tilde{z}_k &\cong g(\eta_k) + g'(\eta_k) l_k + \frac{1}{2} g''(\eta_k) l_k^2 \\ &= z_k + \left\{ g'(\eta_k) + \frac{1}{2} g''(\eta_k) l_k \right\} l_k \end{aligned} \tag{32}$$

となる．これによって，入力信号 x_i に対する出力信号 $\tilde{z}_k$ と，教師信号 t_k の 2 乗誤差

$$\tilde{\epsilon}^2 = \frac{1}{2} \sum_{k=1}^{K} (t_k - \tilde{z}_k)^2 \tag{33}$$

を，変位がない場合の 2 乗誤差 ϵ^2 と 2 次の微小量まで比較することができる．式 (33) の $\tilde{z}_k$ を，近似的に式 (32) で置き換え，2 次の微小量まで評価すると，

$$\begin{aligned}\tilde{\epsilon}^2 &\cong \frac{1}{2} \sum_{k=1}^{K} \left[(t_k - z_k) - \left\{ g'(\eta_k) + \frac{1}{2} g''(\eta_k) l_k \right\} l_k \right]^2 \\ &\cong \epsilon^2 - \sum_{k=1}^{K} (t_k - z_k) g'(\eta_k) l_k \\ &\quad + \frac{1}{2} \sum_{k=1}^{K} \left\{ g'(\eta_k)^2 - (t_k - z_k) g''(\eta_k) \right\} l_k^2 \end{aligned} \tag{34}$$

となる．ここで，式 (30) を用いることで，

$$\begin{aligned}&\sum_{k=1}^{K} (t_k - z_k) g'(\eta_k) l_k \\ &= \sum_{j=1}^{J} \left\{ \sum_{k=1}^{K} (t_k - z_k) g'(\eta_k) b_{kj} \right\} n_j \end{aligned}$$

$$l_k^2 = \sum_{p=1}^{J} \sum_{q=1}^{J} b_{kp} n_p b_{kq} n_q$$

だから，

$$\begin{aligned}&\sum_{k=1}^{K} \left\{ g'(\eta_k)^2 - (t_k - z_k) g''(\eta_k) \right\} l_k^2 \\ &= \sum_{p=1}^{J} \sum_{q=1}^{J} \Bigl(\sum_{k=1}^{K} \bigl\{ g'(\eta_k)^2 \\ &\quad - (t_k - z_k) g''(\eta_k) \bigr\} b_{kp} b_{kq} \Bigr) n_p n_q \end{aligned}$$

したがって，中間素子の出力に対する変位が 2 乗誤差へ及ぼす影響は

$$\tilde{\epsilon}^2 - \epsilon^2 \cong - \sum_{j=1}^{J} S_j n_j + \frac{1}{2} \sum_{p=1}^{J} \sum_{q=1}^{J} T_{pq} n_p n_q \tag{35}$$

となる．変位 n_j が独立でありその期待値が 0 ならば，変位が正規ノイズの場合の平均的な影響は

$$E[\tilde{\epsilon}^2 - \epsilon^2] \cong \frac{1}{2} \sum_{p=1}^{J} \sum_{q=1}^{J} T_{pq} \sigma_p^2 \delta_{pq} \tag{36}$$

$$\begin{aligned}\tilde{\epsilon}^2 - \epsilon^2 &\cong \frac{1}{2} \sum_{p=1}^{J} T_{pp} \sigma_p^2 \\ &= \frac{1}{2} \sum_{j=1}^{J} \Bigl(\sum_{k=1}^{K} \bigl\{ g'(\eta_k)^2 \\ &\quad - (t_k - g(\eta_k)) g''(\eta_k) \bigr\} b_{kj}^2 \Bigr) \sigma_j^2 \end{aligned} \tag{37}$$

となる．このように，変位が一様ノイズの場合も前節と同様の形となる．前述のようにノイズ重畳によってニューラルネットが受ける平均的な影響は，中間素子と出力素子の入出力を規定する関数の式として表される．一方，中間素子の出力に対する変位の平均的影響は，中間素子の入出力を規定する関数とは無関係に，出力素子の入出力関数 g の式としてのみ表されている．この結果より，素子を規定する関数を式 (27) のロジスティック関数で規定する場合，中間素子への変位と同様に出力素子においても，学習初期には通常の

学習に加え，変位の影響を大きく受けて素子の出力が不確定な状態に戻されるような逆方向の学習が加えられ，その後，再学習が行われるような探索的な進展が見られることが分かる．また，学習の最終局面では変位の影響は少なくなると同時に，素子の出力が 0 と 1 の確定的な値となるような影響を受ける．

3. ETL9B 認識実験

理論的な考察について検証するため，ETL9B 手書き文字データベースを使用した認識実験を行う．入力画像から抽出した方向線素特徴量などを学習することで認識率を高くすることが可能であるが，今回は認識率の向上よりもノイズ重畳の影響を定量的に調べ，学習後のニューラルネットに組織化される受容野結合の構造化について明らかにするため，あえて特徴抽出を行っていない文字画像を使用する．また実験結果が，学習データの選び方や増え方に依存することを考慮して，ETL9B データの中から 100 字種，200 字種，300 字種，400 字種を取り出して文字種数の異なる 4 通りの実験データを作成する．各字種にはそれぞれ，筆跡の異なる 200 個の文字が含まれているが，それらにノイズ除去と正規化を行う前処理を施し，32×32[pixel] の 0（白画素）と 1（黒画素）からなる二値画像を作成する．そして前半の 20〜100 個を学習用に，後半の 100 個を汎化能力評価用の未知データとして使用する．したがって，100〜400 字種の 4 通りの文字セットについて，(学習データ, 未知データ）=（前半 20 個, 後半 100 個)，以下（40,100),(60,100)，(80,100)，(100,100）の 5 種類のデータセットを作成することで，合計 20 セットの実験用データが作成できる．よって，たとえば 300 字種・(20,100）のデータセットを用いた認識実験の場合，学習用の 6,000 個（300 字種×前半 20 個）の文字を用いて誤差逆伝播学習を行い，後半の 30,000 個（300 字種×後半 100 個）の未知データを提示して認識率を調べ汎化能力の基準とする．

さらに，ノイズ量の変化に対するニューラルネットの出力の変化を調べるため，ノイズ量を 0（ノイズなし）から 1.5 の範囲で 0.5 刻みに変化させながら学習時に付加し，2,000 回の学習を行う．同時に，50 回の学習ごとに未知データを提示して認識率を記録する．以上の方法で 80 種類の実験（実験用データ 20 セット×ノイズ量 4 種類）を行うが，同じ中間素子数を持つニューラルネットでも，結合係数の初期値によって内部表現が異なって構築され，汎化能力が変化することから，すべての実験で結合係数の初期値に異なる乱数値を与えることとする．

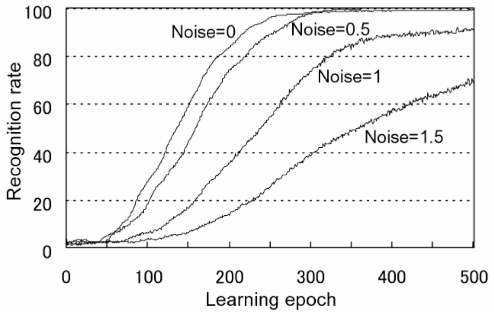

図 8　ノイズ重畳による学習過程の変化（学習初期の場合）
Fig. 8　Learning process by using noise injection into inputs at the learning start.

3.1 ニューラルネットの構成

認識部には，入力素子数 1024（32×32)，中間素子数 36，出力素子数 100〜400（認識文字種数に依存）の 3 層ニューラルネットを用い，学習方法には一般的な誤差逆伝播学習法を用いる．また，入力素子の入出力は出力範囲が 0〜1 の線形関数で，中間・出力素子の入出力はロジスティック関数 $f(u) = 1/(1 + e^{-u})$ で規定する．

誤差逆伝播学習法では，学習用データセットに対するニューラルネットの出力と目標信号とによって規定される誤差関数 E を最小化するように，結合荷重 w を $\Delta w = -\alpha(\partial E/\partial w)$ ずつ変化させる．ここで，$\alpha(> 0)$ は学習速度を決める定数であり，今回は $\alpha = 0.1$ としてすべての実験を行う．

3.2 入力へのノイズ重畳実験

平均 0 の乱数を，その絶対値の最大値を 0（ノイズなし）から 1.5 の範囲で変化させながらノイズとして付加する実験を行い，学習誤差と未知データに対する認識率を調べる．ノイズ 0，0.5，1，そして 1.5 の場合における学習の進展過程を**図 8** に示す．図は文字種数 100，(20,100) のデータに対する学習初期の様子である．

図のように，付加するノイズが大きいほど学習がゆっくり進んで収束が遅い．またノイズ 1 と 1.5 の場合，学習曲線に振動が見られる．このような学習速度の変化と振動は，2 の解析で述べたノイズによる探索的な振舞いと，反学習と再学習の繰返しを示していると考えられる．

未知データに対する認識率を**図 9** に示す．またデータセット (100,100) に対する認識率を**表 1** に示す．データセット（100,100）の場合ではノイズ 0.5 の認識率が高く，データセット（20,100）の場合ではノイズ 1 の認識率が高い．また表 1 に示すように，データセットの文字種数を 100 から 400 に増やした場合，ノイズ 0

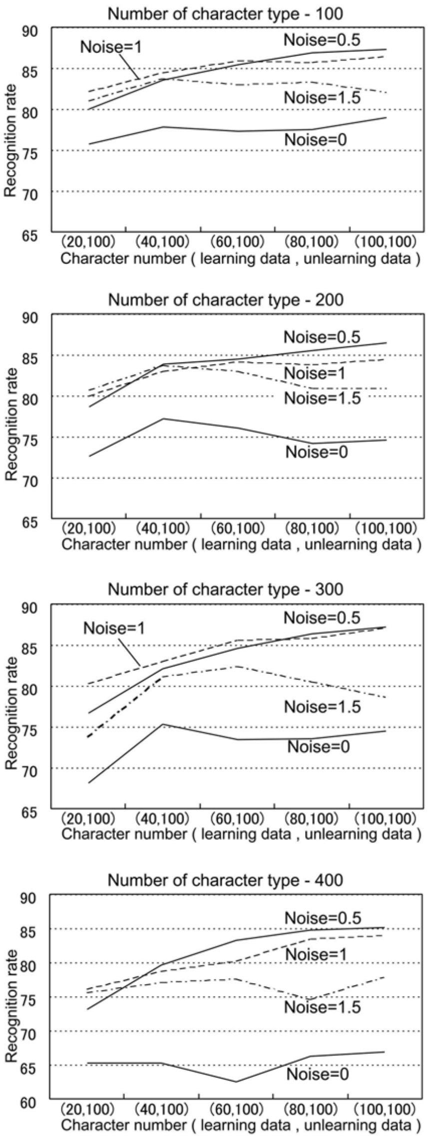

図 9 ノイズ重畳による認識率の変化

Fig. 9 The relation between the number of learning data and the recognition rate by using noise injection into inputs.

表 1 未知データセット (100,100) の認識率

Table 1 Recognition rate of unlearning data set (100,100).

ノイズ量	100 字種	200 字種	300 字種	400 字種
0（なし）	78.92	74.63	74.45	66.93
0.5	87.33	86.47	87.23	85.13
1.0	86.51	84.43	87.05	84.05
1.5	82.13	80.90	78.66	77.83

の認識率は 78.92%から 66.93%と大きく低下しているが，ノイズ 0.5 の場合 87.33%から 85.13%と 2.2%の低下に抑えられている．

一方，ノイズ 1.5 のように付加するノイズ量が大きすぎる場合，認識性能が低下している．これは付加するノイズが大きすぎると反学習・再学習の振幅と探索的振舞いが大きくなりすぎて，誤差逆伝播学習によるニューラルネットの構造化に悪影響を及ぼしているためと考えられる．ところで図 9 のノイズ 0.5 と 1 の結果を見た場合，データセット (40,100) に対する認識率が良好なら，他のデータセットについても認識率が高い．実問題において最も重要なのはデータセット (100,100) に対する認識率であるが，認識対象となる文字数やニューラルネットが大規模になると，学習と認識に要する計算量の増大が問題となる．そこで，最初にデータセット (40,100) に対して認識率が高くなるようにノイズ量を定め，そのノイズをデータセット (100,100) の学習時に付加するような認識システムへの応用が期待できる．

4. ニューラルネットの構造化に対する考察

学習後のニューラルネットがどのように構造化されているかを調べるため，次のような評価基準で中間素子の平均的な出力を算出し，ニューラルネットの構造化に対する考察を行う（以下，この評価値を構造化度と呼ぶ）．

$$W = \frac{\sum_j f\left(\gamma|\sum_i a_{ji}|\right)}{J} \tag{38}$$

$|\sum_i a_{ji}|$ は，1 つの中間素子と全入力素子との結合強度，すなわち中間素子の受容野結合の絶対値である．また，γ は微小定数で今回は $\gamma = 0.2$ とした．

実験で用いるニューラルネットは平均が 0 の乱数値で初期化されるため，学習開始時の構造化度は最小値の 0.5 である．そして，学習が進むに従って各中間素子の出力が 0 か 1 に近づくほど構造化度 W は 1 に近づく．

実験後の構造化度を**図 10** に示す．100 字種に対する結果では，ノイズが大きい方が構造化度が高くなっており，その差は特に学習データ数の少ない場合で顕著である．またノイズ 0 の場合，学習データ数が多い方が構造化度も高くなる傾向が見られる．一方，文字種数の最も多い 400 字種の結果ではノイズ 0 の構造化度が高くなり，ノイズを付加した場合の構造化度との差が小さくなっている．これらの結果より，ノイズ重畳には擬似的に学習データ数を増やして構造化度を高くする効果があり，それによって学習データ数が少ない場合でも中間素子の出力が確定的な値に近づくものと推測される．ただし構造化度が高いほど認識率が高くなるのではなく，構造化度が高すぎると反対に認

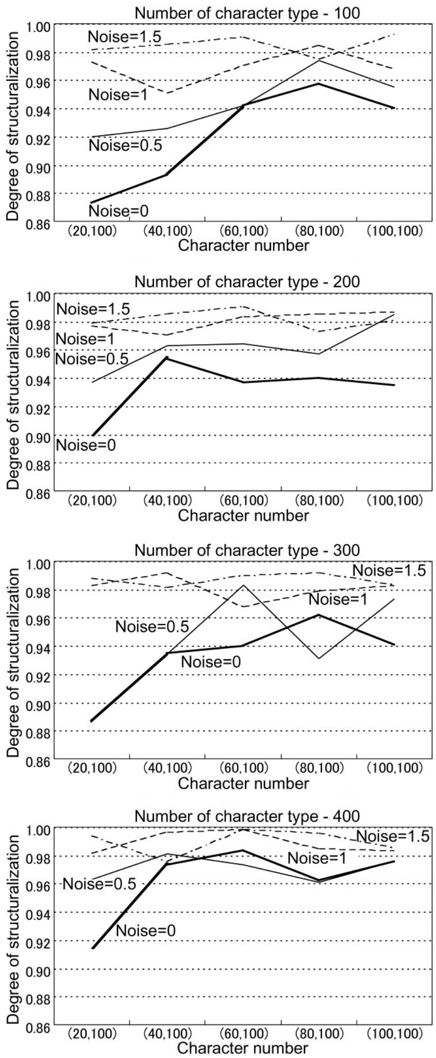

図 10　ノイズ重畳による構造化度の変化
Fig. 10　The relation between noise injection into inputs and the degree of structuralization.

識率が低くなる場合がある．その要因の 1 つとして，ノイズ量が大きすぎる場合，学習初期における探索的な振舞いと反学習・再学習の振幅の変化が大きすぎるために，誤差逆伝播学習によって構造化が進むよりも早く，ノイズの影響によって素子の出力が確定的になるように構造化されてしまうためと考えられる．

5. む す び

階層型ニューラルネットへのノイズ重畳効果を理解するために，解析的な考察と手書き文字認識実験による検証を行った．その結果，学習初期にはノイズ重畳の影響を大きく受けて素子の出力が不確定な状態に戻されるような逆方向の学習が加えられ，その後，再学習が行われるような探索的な進展が見られることを確認した．また，学習の最終局面では変位の影響が少なくなると同時に，素子の出力が 0 と 1 の確定的な値となるような影響を受けることが確認された．そして，このような学習初期の探索的な振舞いにより，ニューラルネットの汎化能力向上が期待できることを確かめた．また，学習後のニューラルネットに見られる結合強度を分析し，ノイズ重畳には擬似的に学習データ数を増やす効果があり，それによって中間素子の出力が 0 か 1 と確定的な値となるような構造化が加速されることが推測されることを示した．今後，さらに汎化能力を向上させるため，付加するノイズ量を学習の進度に応じて動的に変化させる方法や，入力素子と中間素子間の結合に，あらかじめ大きさや形状の異なる局所的な受容野結合を構築することなど，構造化度の低い学習初期での探索的な振舞いをより効果的にするような方法に対する検討を行う必要がある．また，実問題への応用に向け，より文字種の多い認識問題に対する検討や，方向特徴などの特徴ベクトルを学習し，認識率を高くすることを検討したい．

謝辞　ETL9B を提供してくださった電子技術総合研究所の皆様に深く感謝いたします．

参 考 文 献

1) Matsuoka, K.: Noise Injection into Inputs in Back-Propagation Learning, *IEEE Trans. Systems, Man and Cybernetics*, Vol.22, No.3, pp.436–440 (1992).
2) 鹿山昌宏，阿部重夫：汎化能力向上を目的としたクラスタリング用ニューラルネットワークの学習方式，信学論（D-II），Vol.J74-D-II, No.4, pp.863–872 (1993).
3) 大山輝光，黒田英夫，宮原末治，志久　修，高比良秀彰：方向線素特徴とノイズ重畳を用いたニューラルネットワークによる手書き文字認識，情報処理学会論文誌，Vol.44, No.5, pp.1368–1371 (2003).
4) Hamamoto, Y., Mitani, Y. and Tomita, S.: On the Effect of the Noise Injection in Small Training Sample Size Situations, *Proc. ICONIP*, pp.626–628 (1994).
5) 殿村正延，中山謙三：多層パーセプトロンにおける内部情報最適化アルゴリズムと汎化能力の解析，信学論（D-II），Vol.J84-D-II, No.5, pp.830–342 (2001).
6) 栗田多喜夫，麻生英樹，梅山伸二，赤穂昭太郎，細美章隆：多層パーセプトロンの学習における中間層に付加したノイズの影響とネットワークの構造化，信学論（D-II），Vol.J79-D-II, No.2, pp.257–266 (1996).

7) 渡辺栄治：パターン認識問題に対する階層型ニューラルネットワークの汎化能力改善学習法，信学論（D-II），Vol.J79-D-II, No.5, pp.917–923 (1996).
8) 小川英光：ニューラルネットと汎化能力，信学技法，NC95-8, pp.57–64 (1995).
9) 栗田多喜夫：情報量基準による 3 層ニューラルネットの隠れ層のユニット数の決定法，信学論（D-II），Vol.J73-D-II, No.11, pp.1872–1878 (1990).
10) Plaut, D.C., Nowlan, S.J. and Hinton, G.E.: *Experiments on learning by back propagation*, Carnegie-Mellon Univ., CMU-CS-86-126 (1986).
11) 石井真樹，熊沢逸夫：汎化能力改善のための階層型ネットワークの重み表現に対する線形制約の導入と文字認識への応用，信学論（D-II），Vol.J84-D-II, No.3, pp.541–548 (2001).
12) 鳥脇純一郎：X 線像のコンピュータ支援診断—研究動向と課題，信学論（D-II），Vol.J83-D-II, No.1, pp.3–26 (2000).
13) R.H. ニールセン（著），袋谷賢吉（訳）：ニューロコンピューティング，トッパン (1992).
14) 藤吉弘亘，梅崎太造，今村友彦，金 武雄：ニューラルネットワークによるナンバープレートの位置検出，信学論（D-II），Vol.J80-D-II, No.6, pp.1627–1634 (1997).

(平成 15 年 1 月 24 日受付)
(平成 16 年 1 月 6 日採録)

大山　輝光（正会員）

1990 年長崎大学工学部電気工学科卒業．1992 年同大学大学院工学研究科修士課程修了．同年和歌山信愛女子短期大学助手，1993 年講師，2004 年助教授，現在に至る．ニューラルネットワーク，パターン認識，コンピュータと教育等の研究に従事．日本神経回路学会，電子情報通信学会，IEEE 各会員．

三好　邦男

1978 年信州大学理学部物理学科卒業．1986 年大阪府立大学大学院工学研究科博士後期課程単位取得済退学．現在，和歌山信愛女子短期大学生活文化学科助教授．日本物理学会，日本応用数理学会，日本数学会各会員．

黒田　英夫（正会員）

1971 年九州工業大学大学院修士課程修了．同年日本電信電話公社電気通信研究所入社．1989 年より長崎大学工学部・大学院教授．その間 1994 年シドニー大学客員教授．画像信号高能率符号化，画像処理，ネットワーク，CG，CV 等の研究に従事．工学博士．

宮原　末治（正会員）

1971 年熊本大学大学院工学研究科修士課程修了．同年日本電信電話公社研究所入社．1997 年長崎大学工学部教授．工学博士．音声情報処理，文字認識装置の実用化，情報検索等の研究に従事．

志久　　修（正会員）

1989 年長崎大学工学部電子工学科卒業．1991 年同大学大学院工学研究科修士課程修了．同年国立佐世保工業高等専門学校助手，1994 年講師，1996 年助教授．現在に至る．博士（工学）．画像処理，パターン認識の研究に従事．電子情報通信学会会員．

中村　千秋（正会員）

1987 年佐賀大学理工学部電子工学科卒業．1989 年九州大学大学院工学研究科修士課程情報工学専攻修了．同年長崎大学工学部助手，同講師を経て，現在長崎大学教育学部助教授．コンピュータネットワークの応用，情報検索に関する研究に従事．