

# 構造適応型 Deep Belief Network 事前学習を 考慮した知識獲得の検討

## Knowledge Acquisition in Consideration of Pre-training for Adaptive Structural Learning of Deep Belief Network

市村 匠

Takumi Ichimura

県立広島大学 高度人工知能プロジェクト研究センター長  
経営情報学部 経営情報学科

Email: ichimura @ pu-hiroshima.ac.jp

鎌田 真

Shin Kamada

広島市立大学大学院情報科学研究科  
日本学術振興会特別研究員 (DC2)

Email: da65002 @ e.hiroshima-cu.ac.jp

**Abstract**—Deep Learning has a hierarchical network architecture to represent the complicated feature of input patterns. We have developed the adaptive structure learning method of Deep Belief Network (DBN) that can discover an optimal number of hidden neurons for given input data in a Restricted Boltzmann Machine (RBM) by neuron generation-annihilation algorithm, and hidden layers in DBN. The knowledge extraction method from the developed DBN and the rectification method of the signal flow on the wrong path have been developed. The fine-tuning method can reach an incredible high accuracy of classification (the best record). In Deep Learning, the layer-wise unsupervised pre-training can construct abstract and concrete modes of information processing. In this paper, we improve the knowledge acquisition method to adopt a distinction between abstract and concrete. The empirical study was executed on the ChestX-ray8 database.

### I. はじめに

Deep Learning とは大規模な階層構造を持ったニューラルネットワークにおける学習法 [1], [2] であり, 従来の階層型ニューラルネットワークとの最大の違いは事前学習にあると言われている. 最も知られた Deep Learning 学習法は, Convolutional Neural Network(CNN)[3] であろう. CNN には, Convolutional 層と Pooling 層を組み合わせながら, 複数のフィルタと呼ばれる検出器のパラメタを自動で学習する仕組みがあり, 畳み込み, 位置不変性, 合成性の3個の特徴を実現している. これらの層をうまく組み合わせながら, 独自の階層構造を構築するためには相当のスキルが必要であるため, 精度の高いネットワークとして知られている, AlexNet [4], GoogLeNet [5], VGG16 [6], ResNet [7] など, 既存の学習済みのネットワークを用いて, 再学習を行う転移学習と呼ばれる手法により, Deep Learning を応用している. しかしながら, 高精度なシステムを構築する

ためには, 転移学習よりも該当するデータのみを用いて学習する必要があるが, ネットワーク構造やパラメタ設定などは試行錯誤的に求めなければならなかった.

我々はこのような問題を解決するために, 入力データに応じて適切なネットワーク構造を探索する構造適応型学習法を, 尤度の概念を取り入れた Deep Learning 手法である Restricted Boltzmann Machine(RBM)[8], 及び事前学習した RBM を階層化し学習する Deep Belief Network(DBN)[9] において開発した [10]. 構造適応型学習法では, 入力データ空間に応じて適切な RBM の隠れニューロン数, 及び DBN の隠れ層数が学習中に自動で探索され, 画像ベンチマークデータセット CIFAR-10 及び CIFAR-100[11] に対し, 従来手法よりも高い分類精度が得られている [12], [13], [14].

構造適応型 Deep Learning 学習アルゴリズムを様々なデータに適用しているが, 今のところどのデータに対しても, 最も高い分類精度をもつシステムを構築している. しかしながら, 一般的に, 学習後のネットワーク構造はブラックボックスであるため, 明確な知識を抽出することができないと言われている. 我々は, RBM の入出力値がバイナリであることを利用して, 与えられた入力に対し, どのニューロンが発火するかを層ごとに探索し, 出力層に至るまで, ニューロン間のパスを特定した. パスの特定は, 可能な入出力パターンに対し, すべての組に対する割合を層ごとに計算し, 出現率が高いパスとした. 特定したパスを IF-THEN 形式で表現することで, 多段的な推論知識として表現することに成功した. テスト事例において, 誤判定を生じるケースについては, パスの一部分の重みを修正することで, 正しい計算結果を得る手法を開発した [15], [16], [17]. ところが, 学習するデータによっては, 事前学習時, 層ごとに形成される知識が異なり, 低い層(入力に近い層)では推論のための知識が形成されていない

ことが分かった。つまり、人が判断できる程度の具体的な認識には、出力層に近い層で判定可能な知識が形成されている一方、抽象的な (abstract な) 学習が行われるには、入力層に近い複数の層での学習が必要であると考えられる。そこで、本論文では、低い層の学習で学習誤差の値が減少していない場合は知識抽出を行わず、検出器として全結合ネットワークをそのまま利用し、誤差が小さくなる上位層からのみ知識を抽出する改良を行った。本論文では、胸部レントゲン写真のデータベース CXR8[18] を用いて検証をしたので、ここに報告する。

## II. 構造適応型 DEEP BELIEF NETWORK 学習

### A. Restricted Boltzmann Machine

RBM[8] は図 1 のように可視層と隠れ層の 2 層から構成されるネットワーク構造を持ち、確率分布に基づいて入力データに含まれる特徴を隠れニューロン上で学習することができる教師なし学習モデルである。

RBM では与えられた入力データに対して式 (1) のエネルギー関数を最小にするパラメタ  $\theta = \{\mathbf{b}, \mathbf{c}, \mathbf{W}\}$  が最尤推定により求められる。ここで、 $\mathbf{b} = \{b_1, \dots, b_i, \dots, b_I\}$ ,  $\mathbf{c} = \{c_1, \dots, c_j, \dots, c_J\}$ ,  $\mathbf{W} = \{W_{11}, \dots, W_{ij}, \dots, W_{IJ}\}$  である。

$$E(\mathbf{v}, \mathbf{h}) = -\sum_i b_i v_i - \sum_j c_j h_j - \sum_i \sum_j v_i W_{ij} h_j, \quad (1)$$

$$p(\mathbf{v}, \mathbf{h}) = \frac{1}{Z} \exp(-E(\mathbf{v}, \mathbf{h})), \quad (2)$$

$$Z = \sum_{\mathbf{v}} \sum_{\mathbf{h}} \exp(-E(\mathbf{v}, \mathbf{h})) \quad (3)$$

ここで、 $v_i$  は  $i$  番目の可視ニューロンの状態変数、 $h_j$  は  $j$  番目の隠れニューロンの状態変数である。 $b_i$  は  $v_i$  に対するパラメタ、 $c_j$  は  $h_j$  に対するパラメタ、 $W_{ij}$  は  $v_i$  と  $h_j$  間の重みである。 $\mathbf{v} \in \{0, 1\}^I$  と  $\mathbf{h} \in \{0, 1\}^J$  はそれぞれ可視層と隠れ層に含まれるニューロンの状態変数のベクトルであり、 $I$  と  $J$  は可視ニューロン数、隠れニューロン数である。式 (2) は入力  $\mathbf{v}$  と  $\mathbf{h}$  の確率分布である。式 (3) の  $Z$  は分配関数 (partition function) であり、 $\mathbf{v}$  と  $\mathbf{h}$  の全ての 2 値のパタンの組み合わせに対するエネルギーを示す。最適なパラメタ  $\theta$  は与えられた入力データの尤度  $p(\mathbf{v}|\theta)$  を最大化することで求められるが、モデルの尤度に含まれる分配関数  $Z$  の計算は現実的に困難である。そのため、この値を推定するために Contrastive Divergence(CD) 法 [19] のサンプリングによって近似が行われる。このサンプリングの過程で条件付き確率である連続値が 2 値に変換されてしまい、機械学習が収束するための条件である連続性を満たさない状態を生じることが考えられる。この問題に対して文献 [20], [21] では、RBM の各パラメタの変分がリップシツツ連続で、その上限がシャッテンノルムにより与えられ、これを満たすことで学習が収束することが示されている。

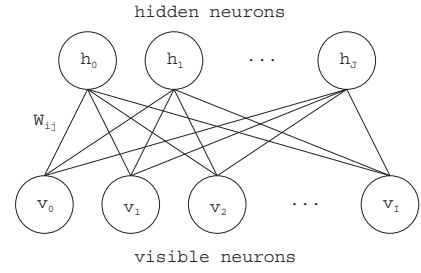


図 1. RBM ネットワーク構造

### B. Deep Belief Network

Deep Belief Network(DBN) は、事前学習を行った RBM を複数積み重ねることで入力データの特徴を階層的に学習することができる Deep Learning 手法である [9]。図 2 は 4 つの RBM を持つ DBN のネットワーク構造を示している。入力データを  $\mathbf{h}^0 = \mathbf{v}$  とすると、 $l(> 0)$  層目の RBM の  $j$  番目の隠れニューロンの発火確率は式 (4) のように定義される。

$$p(h_j^l = 1 | \mathbf{h}^{l-1}) = \text{sigm}(c_j^l + \sum_i W_{ij}^l h_i^{l-1}) \quad (4)$$

ここで、 $c_j^l$  と  $W_{ij}^l$  はそれぞれ  $l$  層目の RBM の  $j$  番目の隠れニューロンのパラメタと  $i, j$  番目の重みである。 $\text{sigm}()$  はシグモイド関数で、文献 [9] と同じ関数を使用している。DBN 自体は RBM を積み重ねたモデルであるため教師なし学習モデルであるが、最上位層に出力層を加え、ソフトマックス法等により計算された出力値と教師信号との誤差 (損失) を最小化することで、教師あり学習として適用できる (Fine-Tuning)。ソフトマックス法では、与えられた入力データがカテゴリ  $k$  に属する確率  $y_k$  が式 (5) により求められる。

$$y_k = \frac{\exp(z_k)}{\sum_{j=1}^M \exp(z_j)} \quad (5)$$

ここで、 $z_j$  は出力層における  $j$  番目のニューロン出力活性化値である。 $M$  は出力層のニューロン数である。

### C. ニューロン生成/消去アルゴリズム

階層型ニューラルネットワークでは、入力データに対して十分な隠れニューロン数があれば学習が進むにつれて重みの変分は小さくなり、ある一定の値に収束すると考えられる。一方、この条件を満たさず重みが収束しない場合は、学習データに対するネットワークの表現能力が不足しており、データの 패턴を隠れニューロン上で表現できていないことが原因だと考えられる。このような場合、関連する位置に十分な数の隠れニューロンを挿入することでデータの 패턴を表現できると考えられる。文献 [22] では学習中における重みの変分ベクトルを Walking Distance(WD) と定義している。WD は式 (6) により、 $\tau - 1$  回目までの学習によるパラメタの変分と  $\tau$  回目の学習のパラメタの変分の和によ

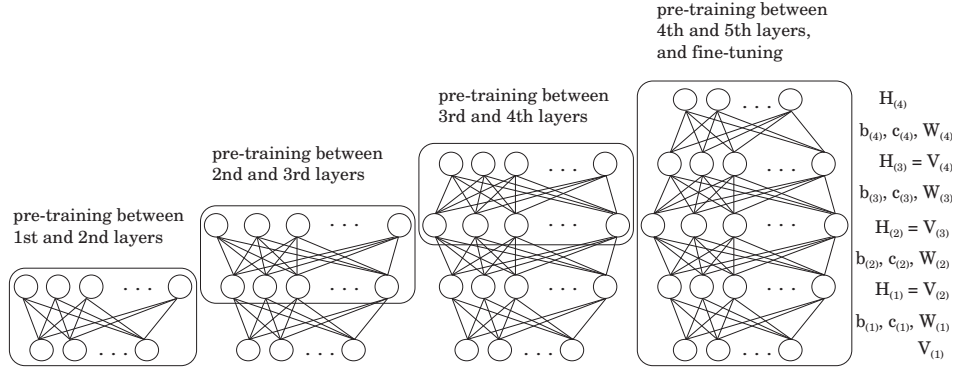


図 2. DBN ネットワーク構造

り計算される.

$$WD_j[\tau] = \gamma_W WD_j[\tau - 1] + (1 - \gamma_W) Met(\mathbf{W}_j[\tau], \mathbf{W}_j[\tau - 1]) \quad (6)$$

ここで,  $\mathbf{W}_j[\tau]$  は  $\tau$  回目の学習における隠れニューロン  $j$  に関連する重みである.  $Met$  はベクトル間の距離を計算する関数で, 本論文ではユークリッド距離  $Met(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^N \sqrt{(a_i - b_i)^2}$  を用いる.  $\gamma_W$  は  $\tau - 1$  回目までの変分と  $\tau$  回目の変分の影響度を調整するパラメタで  $0 < \gamma_W < 1$  である. 文献 [22] ではある一定回数の訓練後, 隠れニューロン  $j$  の  $WD_j[\tau]$  の値が予め定められた閾値よりも大きい場合, その隠れニューロンの重みを 2 分割し, 新しく隠れニューロンを追加する手法が提案されている.

階層型ニューラルネットワークと異なり, RBM には 3 種類の学習パラメタ  $\theta = \{\mathbf{b}, \mathbf{c}, \mathbf{W}\}$  がある. 我々の調査では, RBM 学習における 3 つのパラメタは, 隠れニューロンに関するパラメタ  $\mathbf{c}$  と重み  $\mathbf{W}$  が隠れニューロンの学習状況によって変化し, 入力に関するパラメタ  $\mathbf{b}$  が入力データの分布に応じて変化することが分かった [23]. 各パラメタの WD は式 (6) の  $WD_j[\tau]$  の計算方法により求めた. すなわち,  $\tau - 1$  回目までの学習による変分と,  $\tau$  回目の学習の変分の和に基づいて, 各パラメタの変分が計算される. 本論文では,  $\mathbf{b}$  は入力に関するパラメタであるため考慮せず,  $\mathbf{c}$  と  $\mathbf{W}$  は隠れニューロンの学習状況に依存するパラメタであるため, この 2 つの変分の積空間を用いた指標を考え, ニューロン生成の条件を式 (7) に示すように定義した.

$$WD_{c_j} \cdot WD_{W_j} > \theta_G, \quad (7)$$

$$WD_{c_j} = \gamma_c WD_{c_j} + (1 - \gamma_c)(|c_j[\tau] - c_j[\tau - 1]|), \quad (8)$$

$$WD_{W_j} = \gamma_W WD_{W_j} + (1 - \gamma_W) Met(\mathbf{W}_j[\tau], \mathbf{W}_j[\tau - 1]) \quad (9)$$

ここで,  $WD_{c_j}$ ,  $WD_{W_j}$  はそれぞれ  $j$  番目の隠れニューロンに関するパラメタと重みに関する学習の変分であり, 0 以上の値をとる. これらは, 式 (6) と同様の計

算方法で求められ, 隠れニューロン  $j$  に関連する  $c_j$  と  $\mathbf{W}_j$  の変分が観察される.  $\theta_G$  はニューロン生成に関する閾値であり  $\theta_G > 0$  の値をとる.  $\theta_G$  の値が小さくなるに従って, ニューロン生成条件は満たされやすくなる. 学習中に式 (7) を満たす隠れニューロンがある場合, 該当する隠れニューロンの属性を継承したニューロンが式 (10) のように生成され, 図 3(a) のように隣接する位置に挿入される [10].

$$c_j^{new} = c_j + N(0, \sigma^2), W_{ij}^{new} = W_{ij} + N(0, \sigma^2) \quad (10)$$

一方, 入力データに対して十分な数の隠れニューロンが生成された後, そのうち出力に寄与していない, すなわち冗長な値を出力する隠れニューロンが存在する場合がある. 構造適応型 RBM では, 一定の学習の後, 式 (11) を満たす隠れニューロンが存在すれば, 該当する隠れニューロンを図 3(b) のように消去する. 式 (11) ではすべての入力データに対する隠れニューロンの出力値の平均を観察し, その値が一定の閾値以下のニューロンを冗長なニューロンとみなし, 削除する.

$$\frac{1}{N} \sum_{n=1}^N p(h_j = 1 | \mathbf{v}_n) < \theta_A \quad (11)$$

ここで,  $p(h_j = 1 | \mathbf{v}_n)$  は  $n$  番目の入力データ  $\mathbf{v}_n$  を与えた際の隠れニューロン  $h_j$  の発火確率を示している.  $\theta_A$  はニューロン消去に関わる閾値であり,  $0 < \theta_A < 1$  の値をとる.

#### D. 層の生成条件

事前学習において学習が適切に行われれば, 上位層ほど与えられた入力データに対するエネルギーやパラメタの変化が小さくなることを用いて, DBN における層の生成条件を提案した.

構造適応型 RBM におけるニューロン生成アルゴリズムでは学習中の WD (パラメタ  $\mathbf{c}$  と  $\mathbf{W}$  の変分) を観察したが, 構造適応型 DBN [12] では個々の RBM ではなくネットワーク全体の WD とエネルギーを式 (12) と式 (13) を用いて観察した.

$$\sum_{l=1}^k WD^l > \theta_{L1}, \quad (12)$$



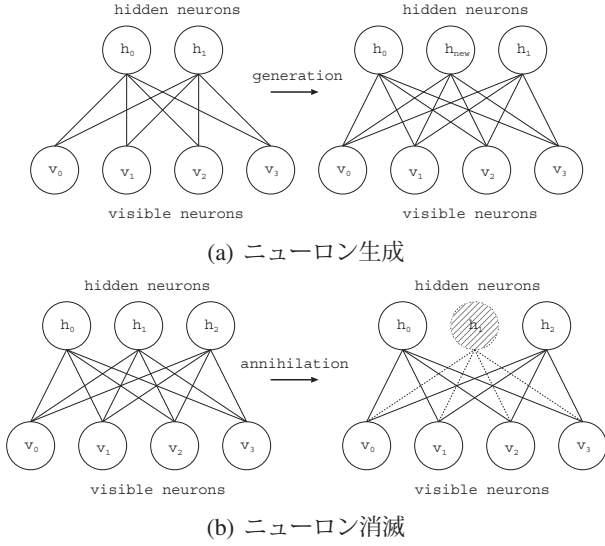


図 3. 構造適応型 RBM 学習法

### Algorithm 1 従来の Fine Tuning アルゴリズム

- 1:  $L(1 \leq l \leq L)$  層の構造を持つ学習済み DBN ネットワークが与えられる.
- 2: 1 層から  $L$  層にかけて順に, **Algorithm 2** により  $l$  層に対する Fine Tuning を実行する.

$$\sum_{l=1}^k E^l > \theta_{L2} \quad (13)$$

ここで,  $WD^l = \sum_{j=1}^J (WD_{c_j}^l \cdot WD_{W_j}^l)$  であり,  $WD_{c_j}^l$  と  $WD_{W_j}^l$  はそれぞれ  $l$  層目の RBM における  $c_j$  と  $W_j$  に関する WD を示している.  $E^l$  はエネルギー関数である.  $\theta_{L1}$ ,  $\theta_{L2}$  は予め与えられる閾値である. ある  $k$  層目の RBM の学習中に式 (12) 及び式 (13) を同時に満たす場合, 入力データに対するネットワーク全体の表現能力が不足していると考え,  $k$  層目の学習が終了した時点で,  $k+1$  層目の RBM を生成し学習を行う. ここでも RBM のニューロン生成アルゴリズムと同様に,  $k+1$  層目の RBM のパラメタ  $b$ ,  $c$ ,  $W$  の初期値が  $k$  層目の RBM から継承される.

## III. 知識獲得手法の改良

### A. 従前の手法

本研究では, 学習済みの構造適応型 DBN のネットワークに実行可能な 2 値な入力パターンを与え, 各層において得られるニューロンの入出力パターンを分析することで, 知識を獲得する手法を開発している. 文献 [15], [16], [17] では, テスト事例において, 誤判定を生じるケースに対する特徴的なパスを抽出し, その一部分の重みを修正することで, 正しい計算結果を得る Fine Tuning 手法を開発した. **Algorithm 1** は, 従来の Fine Tuning 法のアルゴリズムを示している. 学習済みの DBN ネットワークの下位層から上位層に向けて, 各層ごとに **Algorithm 2** で示される手順を実行し, 該当す

### Algorithm 2 $l$ 層の重みを修正する Fine Tuning アルゴリズム

- 1: 入力データを  $X(x_1, \dots, x_p, \dots, x_N)$  とする.  $N$  は入力データ数である.  $X$  には教師ラベル  $Y(y_1, \dots, y_p, \dots, y_N)$  が与えられている. 学習済みの DBN ネットワークを,  $L(1 \leq l \leq L)$  層の構造を持つネットワークとし,  $l$  を Fine Tuning の対象となる層とする.
- 2: 学習済みの DBN ネットワークに入力データ  $X$  を与え, 入力層から出力層に至るまでの前向き計算を行う. ここで, 各入力データ  $x_p$  ごとに, 各層において発火したニューロンのインデックス情報を保存する. また, 教師ラベル  $Y$  に対して正答したパターンを  $X^T$ , 誤答したパターンを  $X^F$  とする.
- 3:  $l$  層において, 式 (14) を満たすニューロン  $j$  を求め, このニューロンに接続されている重みの値を  $w^{correct}$  に修正する. ここで, ニューロン  $j$  は,  $l$  層 (隠れ層) のニューロンを示す. 式 (14) は,  $X$  の中で  $X^T$  に対しニューロン  $j$  のみが発火する頻度を示している. また,  $w^{correct}$  は定数であり, ここでは,  $w^{correct} = 1$  とした.

$$\frac{|Act_j^T|}{|X^T| + |X^F|} \geq \theta^T \quad (14)$$

ここで,  $\theta^T (0 \leq \theta^T \leq 100)$  は閾値である.  $|Act_j^T|$  は,  $X^T$  のみに対しニューロン  $j$  が発火する数を示している.

- 4:  $l$  層において, 式 (15) を満たすニューロン  $j$  を求め, このニューロンに接続されている重みの値を  $w^{wrong}$  に修正する. 式 (15) は,  $X$  の中で  $X^F$  に対しニューロン  $j$  が発火する頻度を示している. また,  $w^{wrong}$  は定数であり, ここでは,  $w^{wrong} = 0$  とした.

$$\frac{|Act_j^F|}{|X^T| + |X^F|} \geq \theta^F \quad (15)$$

ここで,  $\theta^F (0 \leq \theta^F \leq 100)$  は閾値である.  $|Act_j^F|$  は,  $X^F$  のみに対しニューロン  $j$  が発火する数を示している.

る重みを修正する. **Algorithm 2** では, 与えられた入力パターンに対して, 発火するニューロンの頻出度に基づいて, 重みを修正している. 文献 [15], [16], [17] では, 実験結果から, Fine Tuning を行うことで分類精度が数%向上することが示されていた.

### B. 改良方法

本論文では, 知識形成過程において明確な知識を抽出することで, 分類精度の改善を図ることを考え, 従来の Fine Tuning 法を改良する. 従来の Fine Tuning 法では, **Algorithm 1** に示すように, 下位層から上位層の全ての重みを修正していたが, 学習するデータによっては, 事前学習時において層ごとに形成される知識が

### Algorithm 3 改良した Fine Tuning アルゴリズム

- 1:  $L(1 \leq l \leq L)$  層の構造を持つ学習済み DBN ネットワークが与えられているとする. ここで,  $L$  層のうち,  $L_A$  より下位層を抽象的な特徴が学習されている層とし,  $L_A$  以降の上位層を具体的な特徴が学習されている層をとす ( $1 \leq L_A \leq L$ ). なお,  $L_A$  は, **Algorithm 1** により定まる.
- 2:  $L$  層から  $L_A$  層にかけて順に, **Algorithm 2** により  $l$  層に対する Fine Tuning を実行する.  $L_A$  より下位層に対する Fine Tuning は実行しない.

異なり, 低い層(入力に近い層)では推論のための知識が形成されていないことが分かった. これまでは, IoT 機器により収集された数値データの分析や比較的分類しやすい画像のデータを用いていたが, 胸部レントゲンや CT 画像に見られるような医療画像では, 疾患の局所的な部位と抽出した画像の特徴が一致しないこともあり, 抽出した知識をそのまま推論に利用できない課題が生じた. つまり, 人が判断できる程度の具体的な認識には, 出力層に近い層で判定可能な知識が形成されている一方, 抽象的な (abstract な) 学習が行われるには, 入力層に近い複数の層での学習が必要であると考えられる. そこで, 本論文では, 低い層の学習で学習誤差の値が減少していない場合は知識抽出を行わず, 検出器として全結合ネットワークをそのまま利用し, 誤差が小さくなる上位層からのみ知識を抽出する改良を行った. 具体的には, **Algorithm 1** の Step2) において, エネルギー関数の値に対する閾値を設定し, その値より小さくなった, すなわち精度の高い学習が行われたとき, **Algorithm 2** が動作するようになった. **Algorithm 3** は, 改良した Fine Tuning 法のアルゴリズムである. 従来とは異なり, 全ての層を修正するのではなく, 推論のための知識が形成されている上位層のみを修正している.

## IV. 実験

### A. 構造適応型 DBN による CXR 胸部 X ベンチマーク画像の分類

「ChestX-ray8 (CXR8)」は, 文献 [18] によって公開されている胸部 X 線画像のベンチマークテストである. 30,805 人の患者から収集した 112,120 枚の胸部 X 線画像が含まれており, 表 I に示すように, 各画像は, 正常ラベル及び 8 種類の異常ラベルに分類される. これらの 8 種類の異常ラベルの中には, 癌に関連する症状として, “Mass” と “Nodule” が定義されている. なお, 全 112,120 個のデータは, 訓練データ 86,524 個とテストデータ 25,596 個に予め分割されている.

提案手法の構造適応型 DBN を用いて CXR8 の学習を行った. 提案手法の学習パラメタとして, 勾配の学習手法を Stochastic Gradient Descent (SGD), 学習係数を 0.01, バッチサイズは 100, 学習の最大反復回数を 500 回, 初期隠れニューロン数を 300 とした. 構造適

表 I  
CXR8 のカテゴリ

分類		データ数	
正常	正常: No Finding	60,361	
異常	癌	形状大: Mass	5,782
		形状小: Nodule	6,331
	その他	肺拡張不全: Atelectasis	11,559
		心肥大: Cardiomegaly	2,776
		胸水: Effusion	13,317
		肺浸潤: Infiltration	19,894
		肺炎: Pneumonia	1,431
		気胸: Pneumothora	5,302
合計		112,120	

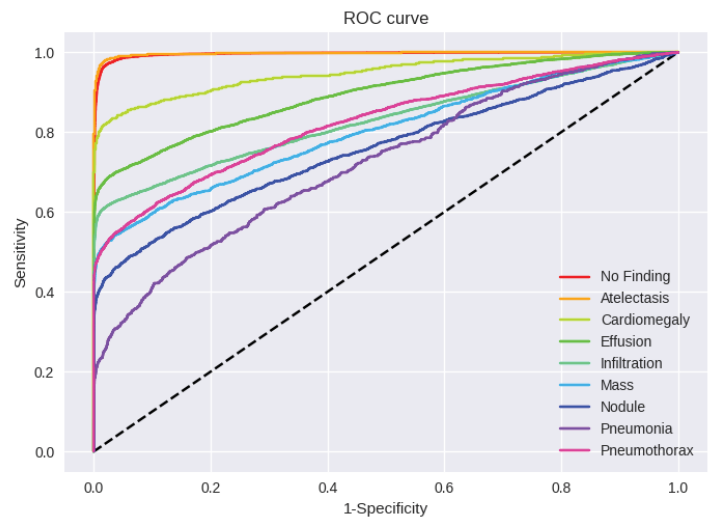


図 4. ROC 曲線

応型 RBM のパラメタとして,  $\theta_G = 0.050, \theta_A = 0.100$  とし, 構造適応型 DBN のパラメタとして,  $\theta_{L1} = 0.1, \theta_{L2} = 0.1$  とした. これらのパラメタの値は, 予備実験の結果による最良の組み合わせである.

表 II は, テストデータの各症状に対する分類精度を示している. 表 II の “ResNet-50” は ResNet[7] と呼ばれる深層学習モデルを用いた精度である. この数値は文献 [18] に記載されている CNN 手法のなかで最も良い精度を示したものである. ResNet と比較して, 提案手法は全ての症状ラベルに対して高い分類精度を示すことができた.

図 4 は, 提案手法の各症状ラベルの分類結果に対する ROC (Receiver Operating Characteristic) 曲線を示している. ROC 曲線とは, 縦軸に, 陽性を陽性と正しく予測する率である Sensitivity, 横軸に, 陰性を誤って陽性と予測する率である 1-Specificity をプロットした曲線である. ROC 曲線と縦軸と横軸で囲まれた面積ができるだけ大きいものほど良いモデルである. 文献 [18] に記載されている ResNet の ROC 曲線と比較して, 図 4 の ROC 曲線の方が, 各症状に対して良い精度を示している.

表 II  
CXR8 の分類精度

分類		Accuracy				
		ResNet-50	DBN	従来 FT 法 DBN	改良 FT 法 DBN	
正常	正常: No Finding	-	90.0%	94.4%	97.7%	
異常	癌	形状大: Mass	56.0%	96.3%	96.9%	98.6%
		形状小: Nodule	71.6%	97.2%	97.5%	98.1%
	その他	肺拡張不全: Atelectasis	70.6%	94.5%	96.0%	98.5%
		心肥大: Cardiomegaly	81.4%	98.1%	98.5%	99.2%
		胸水: Effusion	73.6%	97.2%	97.2%	98.7%
		肺浸潤: Infiltration	61.2%	96.0%	96.1%	99.4%
		肺炎: Pneumonia	63.3%	99.9%	99.9%	99.9%
		気胸: Pneumothora	78.9%	98.1%	98.1%	99.3%

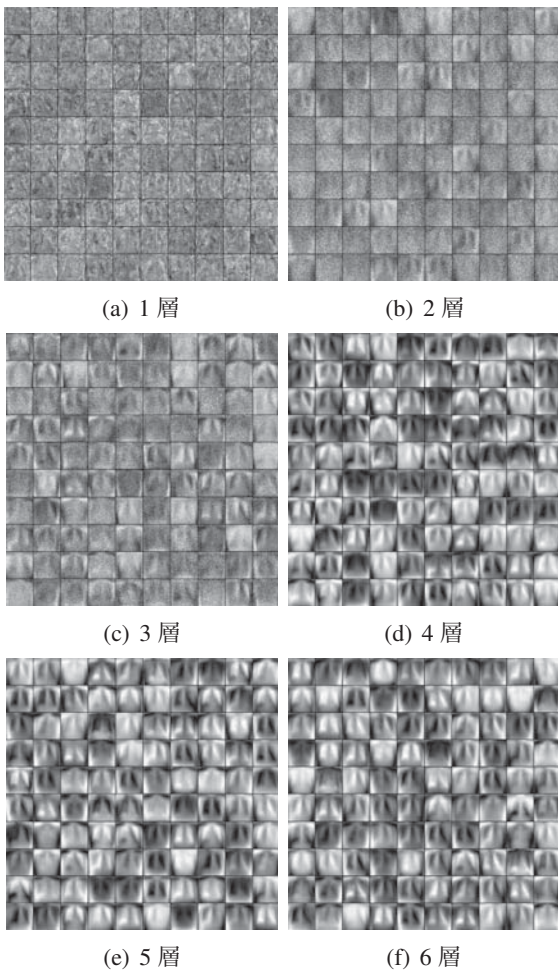


図 5. 各層において抽出された知識

## B. 知識獲得

IV-A 節の構造適応型 DBN の学習により、6 層のネットワークが形成された。ここで得られた学習済みのネットワークに対して、改良した Fine Tuning 法を適用し、分類精度の改善を試みた。図 5 は、与えられた入力パターンに対して、各層において得られた出力パターンを描画した図を示している。この結果から、下位層から上位層に向かうにつれて、徐々に具体的な特徴が現れ、特に、4 層以降において、具体的な特徴が顕著に現れた。

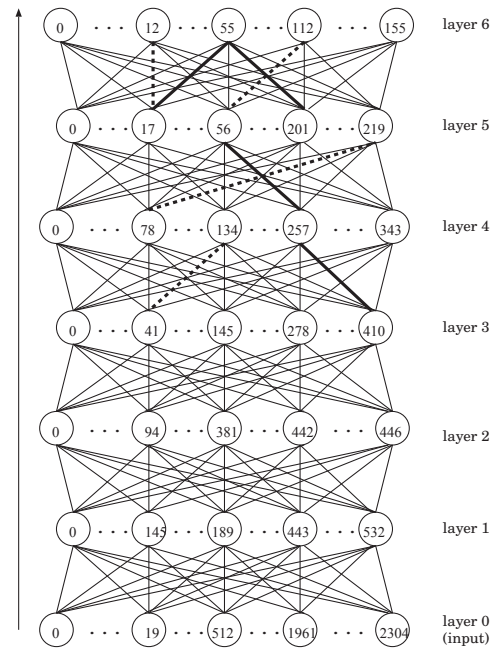


図 6. Fine Tuning により修正されたネットワーク構造

この結果に基づき、本論文では、改良した Fine Tuning 法において、テストデータに対し誤った判定を行った入出力信号の流れに基づいて、4 層以降のネットワークの重みを修正した。

図 6 は、改良した Fine Tuning により修正された構造適応型 DBN のネットワーク構造を示している。図 6 における円はニューロンを示しており、円の中の数字は各ニューロンを識別するインデックスを示している。ニューロン間を接続する線は重みを示し、Fine Tuning により、重みの値が  $w^{correct}$  により増加した接続は太線で、 $w^{wrong}$  により減少した接続は点線で表現している。例えば、5、6 層間の重みでは、5 層の 201 番目のニューロンから 6 層の 55 番目のニューロンの重みの値が  $w^{correct}$  に修正され、一方、5 層の 56 番目のニューロンから 6 層の 112 番目のニューロンの重みの値が  $w^{wrong}$  に修正され、該当するパスを選択しないようになった。これは、変更されたパスを選択しても重みの値が小さいため、積和の計算値への寄与が少な



くなり、結果として、ネットワークの修正が部分的行われていることになる。全体的に3層や4層に比べて、出力層に近い5層や6層の重みが多く修正された。

表IIの“従来FT法DBN”、“改良FT法DBN”は、それぞれ従来と改良したFine Tuning法を適用したDBNによる分類精度を示している。従来では、正常(No Finding)と異常(その他)の肺拡張不全(Atelectasis)に対する精度は向上したものの、それ以外の症状に対しては大きな精度の改善は見られなかった。一方、改良したFine Tuning法では、全てに症状に対する分類精度が従来よりも大きく向上し、全ての症状ラベルに対して97%以上の分類精度を示した。

## V. おわりに

本論文では、胸部X線画像のベンチマークテストCXR8を用いて構造適応型DBN学習を行った。さらに、学習した深層学習モデルを用いて、知識獲得を行った。CXR8で学習した深層学習モデルは、Fine Tuning前では90%以上、Fine Tuning後では97%以上の高い確率で癌の有無を分類できている。ベンチマークテストを用いたシステムによる評価の結果、判定は概ね一致していた。また、提案手法により、事前学習を生かした知識獲得が行っていた。今後は、抽象的な特徴を構築できる層について、学習エネルギーの違いから閾値を設定するための実験を行う予定である。

## 謝辞

本研究開発はJSPS科研費(課題番号: JP17J11178)の助成を受けたものです。

## 参考文献

- [1] Y.Bengio: *Learning Deep Architectures for AI*, Foundations and Trends in Machine Learning archive, vol.2, no.1, pp.1-127 (2009)
- [2] V.Le.Quoc, R.Marc's Aurelio, et.al.: *Building high-level features using large scale unsupervised learning*, Proc. of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp.8595-8598 (2013)
- [3] Y.LeCun, et.al.: *Backpropagation applied to handwritten zip code recognition*, Neural Computation, vol.1, no.4, 541/551 (1989)
- [4] A.Krizhevsky, I.Sutskever, G.E.Hinton, *ImageNet Classification with Deep Convolutional Neural Networks*, Proc. of Advances in Neural Information Processing Systems 25 (NIPS 2012) (2012)
- [5] C.Szegedy, W. Liu, Y.Jia, P.Sermanet, S.Reed, D.Anguelov, D.Erhan, V.Vanhoucke, A.Rabinovich, *Going Deeper with Convolutions*, Proc. of CVPR2015 (2015)
- [6] K.Simonyan, A.Zisserman, *Very deep convolutional networks for large-scale image recognition*, Proc. of International Conference on Learning Representations (ICLR 2015) (2015)
- [7] K.He, X.Zhang, S.Ren, J.Sun, J, *Deep residual learning for image recognition*, Proc. of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.770-778 (2016)
- [8] G.E.Hinton: *A Practical Guide to Training Restricted Boltzmann Machines*, Neural Networks, Tricks of the Trade, Lecture Notes in Computer Science (LNCS, vol.7700), pp.599-619 (2012)
- [9] G.E.Hinton, S.Osindero and Y.Teh: *A fast learning algorithm for deep belief nets*, Neural Computation, vol.18, no.7, pp.1527-1554 (2006)

- [10] S.Kamada and T.Ichimura: *An Adaptive Learning Method of Restricted Boltzmann Machine by Neuron Generation and Annihilation Algorithm*, Proc. of 2016 IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2016), pp.1273-1278 (2016)
- [11] A.Krizhevsky: *Learning Multiple Layers of Features from Tiny Images*, Master of thesis, University of Toronto (2009)
- [12] 鎌田真, 市村匠: *忘却機能を持った Deep Belief Network による構造適応型学習*, 計測自動制御学会第9回コンピュータショナル・インテリジェンス研究会講演論文集, pp.92-97 (2016)
- [13] S.Kamada and T.Ichimura: *A Structural Learning Method of Restricted Boltzmann Machine by Neuron Generation and Annihilation Algorithm*, Neural Information Processing, Lecture Notes in Computer Science (LNCS, vol.9950), pp.372-380 (2016)
- [14] S.Kamada and T.Ichimura: *An Adaptive Learning Method of Deep Belief Network by Layer Generation Algorithm*, Proc. of 2016 IEEE Region 10 Conference (TENCON), pp.2971-2974 (2016)
- [15] 鎌田真, 市村匠, *構造適応型 Deep Belief Network からの知識獲得に関する一考察*, 2016 IEEE SMC Hiroshima Chapter Young Researchers WorkShop, pp.61-66(2016)
- [16] 鎌田真, 市村匠, *リカレント構造適応型 Deep Belief Network 学習における事前知識の分析*, 第44回知能システムシンポジウム講演論文集, SY004/17/B2-2 (2017)
- [17] S. Kamada and T.Ichimura, *Fine Tuning of Adaptive Learning of Deep Belief Network for Misclassification and its Knowledge Acquisition*, International Journal Computational Intelligence Studies, Vol.6, No.4, pp.333-348(2017)
- [18] X.Wang, Y.Peng, L.Lu, Z.Lu, M.Bagheri, R.M.Summers, *ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases*, Proc. of IEEE Computer Vision and Pattern Recognition, pp.3462/3471 (2017)
- [19] G.E.Hinton: *Training products of experts by minimizing contrastive divergence*, Neural Computation, vol.14, no.8, 1771/1800 (2002)
- [20] D.Carlson, V.Cevher and L.Carin: *Stochastic Spectral Descent for Restricted Boltzmann Machines*, Proc. of the Eighteenth International Conference on Artificial Intelligence and Statistics, 111/119 (2015)
- [21] 鎌田真, 藤井哲崇, 市村匠: *Lipschitz 連続による Restricted Boltzmann Machine のエネルギー安定性に関する一考察*, 2015 IEEE SMC Hiroshima Chapter Young Researchers WorkShop, pp.53-56 (2015)
- [22] T.Ichimura, S.Oeda, Ma.Suka, and K.Yoshida: *A learning method of immune multi-agent neural networks*, Neural Computing and Applications, vol.14, no.2, 132/148 (2005)
- [23] 鎌田真, 市村匠, 原章: *ニューロン生成/消滅アルゴリズムによる構造適応型 Restricted Boltzmann Machine*, 計測自動制御学会第8回コンピュータショナル・インテリジェンス研究会講演論文集, pp.90-96 (2015)

問い合わせ先

〒734-8558

広島県広島市南区宇品東一丁目1番71号

県立広島大学

高度人工知能プロジェクト研究センター長

市村 匠