

# ディープオートエンコーダにより生成した補間形状を用いた 形状学習データ生成

佐藤 信<sup>1</sup>

**概要:** ディープオートエンコーダ (DAE) を用いた形状学習データの生成手法を提案する。提案では始めに、与えられた曲線データを DAE の入力として、Continuous RBM を用いて曲線形状を学習する。次に、学習した DAE を用いて、学習形状を補間する曲線形状をデコードすることにより、類似の形状を生成する。そして、生成した曲線形状から形状画像データを生成し、これらの画像データを機械学習の形状学習データとする。

## Shape Learning Set Generation Using Interpolated Shape Generated with Deep Autoencoder

MAKOTO SATOH<sup>1</sup>

**Abstract:** This paper presents a method for learning set generation using a deep autoencoder (DAE). In the method, first given curve shapes are fed into an autoencoder (DAE) to be learned with continuous RBM. Then similar shapes are generated by decoding interpolating shapes of training curve shapes using trained DAE. Shape images will be generated from the generated curve shapes, and the generated shape images will be used as shape learning set.

### 1. はじめに

本稿では、形状の機械学習で用いる学習データを生成するための手法を提案する。特徴は、与えられた少数の曲線形状を基にして、それに類似な形状を生成することである。生成した形状を画像データに変換することにより、形状画像の機械学習に用いることが可能である。

提案手法では、ディープオートエンコーダを用いて、形状を表現するベクトルを、そのベクトルのもつ次元よりも低次元のベクトルに写像する。この写像ベクトルを、低次元空間で捉えた形状特徴の表現であると考えることが可能である。このような着想に基づいて、低次元空間への写像ベクトルの補間を用いて、類似な形状を生成し、それを機械学習で用いる学習データとすることが提案手法の目的である。この写像の誤差を調整すること、および、写像ベクトルを補間することにより、多様な類似性をもつ形状を生

成することが可能である。

これ以降の構成について、簡単に説明する。2 節では、関連研究について説明をおこなう、そして、提案する手法の概要と特徴について述べる。そして、3 節では、ディープオートエンコーダにより生成した補間形状を用いた学習データ生成のための手法を提案する。提案手法の実装と実験結果の検討について 4 節で説明する。そして最後に、5 節で本稿のまとめと今後について述べる。

### 2. 関連研究と提案手法の概要

#### 2.1 関連研究

提案手法では、アルゴリズムにより形状を操作することにより、機械学習で用いるための学習データを生成する。そこで、形状の類似性に注目することにより、形状を操作することを目的とする研究、そして、形状の類似性に注目した機械学習に関連する研究について述べる。

前者には、Fowler 等 [4], Welch 等 [16], Igarashi 等 [8], および、佐藤等 [11], [12] がある。これらの研究は、形状の

<sup>1</sup> 岩手大学  
Iwate University, Ueda, Iwate 020-8551, Japan

特徴を可能な範囲で維持しながら、形状を変化させるための制約条件を満たすように形状を操作するための研究である。それぞれの手法において着目する特徴量を、予め決定しておいて、その特徴量の変化を最適化するようにアルゴリズムを設計している。

後者には、形状の機械学習に用いる学習データをアルゴリズムにより生成するための研究がある。Fu 等 [5] では、少数のラベルありデータを用いて学習データを生成するための手法を提案している。画像処理アルゴリズムを用いて、与えられた簡単な線画に類似な形状を生成し、その生成した画像を学習データとしている。また、佐藤 [9] では、Bezier 曲線で表現された形状について、基準とする形状に類似な形状を生成し、その形状を学習データとして用いるための手法を提案している。

また、半教師あり学習による手描き数字画像の認識を目的とする研究には、Van Vaerenbergh, S 等 [15] がある。ここでは、少数のラベルありデータと多数のラベルなしデータを、学習データとして用いている。あるデータは、いづれかのデータを微小変化することにより生成可能であるという着想に基づいて、データの類似性を表現するために、データどうしを結合するリンクを作成している。このリンクをたどることにより、あるラベルなしデータが、どのラベルありデータの所属するクラスに分類できるのかを判定している。類似性を判断するための基準として、画像処理で用いる非類似性を表現するための基準を用いている。

機械学習を用いて与えられた形状を学習し、学習形状を補間する形状を生成するための研究には、Campbell 等 [1] がある。アルファベットの文字ごとに、複数の書体のフォントの画像データを学習することにより、学習データとして用いたフォントの形状を補間する形状を生成する手法を提案している。

## 2.2 提案手法の概要と適用分野

### 2.2.1 概要

提案手法では、形状表現の低次元空間への写像を用いた形状補間のための既提案の手法 [10] を用いて、生成した低次元ベクトルを補間することにより学習データを生成する。

そこで、既提案の手法について簡単に説明する。その手法では、形状を表現する Bezier 曲線の制御点の次元を削減することにより、補間形状を生成する。高次元データのもつ何らかの特徴を維持したままで、データの次元を削減することを目的として、多くの手法が提案されている [13]。既提案の手法 [10] では、Hinton 等 [6] において提案されたディープオートエンコーダ (deep autoencoder) を、Continuous Restricted Boltzmann machine (CRBM) を用いて訓練することにより、Bezier 曲線の次元を削減するための学習をおこなう。CRBM とは、Chen 等 [2] において提案されたものであり、実数値の入力データをエンコード

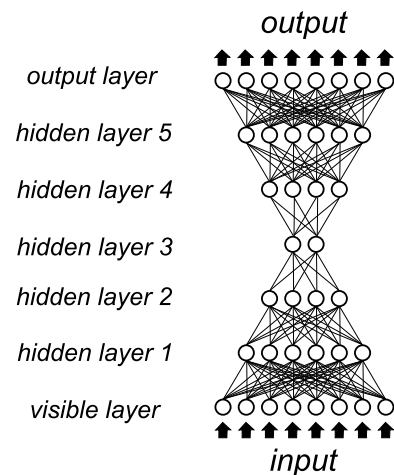


図 1 ディープオートエンコーダの例  
Fig. 1 An Example of deep autoencoder.

したものをデコードした場合の出力データの精度を考慮したものである。図 1 に、ディープオートエンコーダの例を示す。

本稿の提案手法では、[10] の手法を用いて、与えられた学習形状の表現ベクトルを、その特徴を反映するように低次元空間に写像する。そして、写像された低次元ベクトルを基準として用いて、学習形状に類似な補間形状を生成する。その類似な形状を、学習データとして用いることが可能である。

### 2.2.2 適用分野

提案手法を用いると、少数の形状データを基にして、機械学習のための学習データをアルゴリズムにより生成することが可能となる。このことから、認識する形状をカスタマイズした手描き形状の認識などへ適応することが可能であると考えられる。

例えば、類似な形状を直感的な操作で生成するためのユーザインタフェースへの適応などが考えられる。

## 3. 補間形状を用いた学習データ生成

### 3.1 既提案アルゴリズム

提案の手法では、形状の特徴を反映した低次元写像をおこなうために、[10] を用いる。既提案の手法は、以下の各段階により構成される。

Step1: 訓練データ生成 与えられた形状を基準として、それに類似な形状を生成する。類似な形状の生成には、相似変換、または、非均一相似変換 [11], [12] などの手法を用いる。

Step2: 形状学習 Step1 で生成した形状を訓練データとして、ディープオートエンコーダを用いて形状を学習する。これにより、訓練に用いた形状表現を低次元空間に写像するためのパラメータが得られる。

Step3: 補間形状生成 Step2 で学習したオートエンコー

ダのデコード部を用いて、学習データを補間する類似な形状を生成する。

Step2 の学習では、確率的勾配降下法を用いる。次式で与えられる誤差  $e$  の値が、予め与えた値より小さくなるまで計算を反復する。

$$e = \frac{1}{N} \sum_{n=1}^N \sum_i |v| (v_{n_i} - \hat{v}_{n_i})^2. \quad (1)$$

ここで、 $N$ 、ベクトル  $v$ 、そして、 $\hat{v}$  は、それぞれ、CRBM の入力データ数、入力データベクトル、そして、その推定値のベクトルである。ここでの推定値とは、CRBM によりエンコードした値をデコードして得られる値のことである。誤差の計算は、ミニバッチごとにおこない、あるエポックにおいて全てのミニバッチの誤差が条件を満たした段階で計算を終了する。

本稿の提案手法では、誤差の条件を弱めること、及び、形状を補間することにより、多様性をもつ類似な形状を生成する。

### 3.2 非均一形状補間による学習データ生成

ディープオートエンコーダを用いて生成(デコード)した形状に多様性をもたせるために、着目した点を以下に示す。学習形状 学習データとして用いる形状の組み合わせを変化させることにより、デコードされる形状に多様性をもたせる。

学習誤差 形状を学習する場合に、学習を終了するための基準として設定する学習誤差を変化させることにより、デコードされる形状に多様性をもたせる。

補間 写像点を表現するベクトルを補間することにより、デコードされる形状に多様性をもたせる。

## 4. 実装と結果の検討

3 節で提案したアルゴリズムを実装し、その有効性について検討した。

### 4.1 訓練データ

形状学習データを生成するために基準として用いた形状を、図 2 に示す。この基準の形状を相似変換することにより生成した形状を、訓練形状として用いた。それぞれの形状は、平面上の 3 次 Bezier 曲線 (4 セグメント) を用いて表現されている。その Bezier 曲線の 16 個の制御点の値 ( $x$  座標値と  $y$  座標値により構成されるので 32 個の値) を、 $[0, 1]$  の範囲にスケールしたものをディープオートエンコーダの訓練データとした。ディープオートエンコーダには、それ以外のデータは与えていない。

具体的に説明するために、Bezier 曲線の制御点の各座標値を  $(x_1, y_1), \dots, (x_{16}, y_{16})$  とする。訓練データとしては、

これを並べ替えた  $v = \{x_1, \dots, x_{16}, y_1, \dots, y_{16}\}$  を用いた。図 3 の第 1 行の左側に、訓練データの例を示す。これは、円形を近似する bezier 曲線の制御点である。ここでは、一つの訓練データベクトル  $v$  の要素を、二つづつ組にした各点  $(x_1, y_1), \dots, (x_{16}, y_{16})$  をプロットしている。実線で結ばれた大きさ、色、そして、形状が等しい印が、ひとつの訓練データベクトルである。例えば、大きい印(青色)が、ひとつの訓練データベクトルの各要素に対応している。データ相互の関連を分かりやすくするために、線の幅を変化させている。なおこの例では、接続するセグメントの端点の座標値が同一であるので、接続部分の印が重なって表示されている。

学習に用いたそれぞれのデータセットのサイズは、64 または 128 である。それぞれのデータセットを、図 2 の形状のうちのどれか 1 種類の形状または、2 種類の形状により構成した。

### 4.2 ディープオートエンコーダの構成とパラメータ

これ以降の学習においては、確率的勾配降下法を用いて学習をおこなった。そこで用いたハイパーパラメータを、表 1 に示す。パラメータの選択には、Hinton[7] を参考にした。

なお、表中の  $\sigma$  は、Chen 等 [2] の (3) 式のパラメータである。ここでは、モデリング空間をなめらかにする効果を期待して用いた。また、学習アルゴリズムにより学習したパラメータは、CRBM の各層を構成する素子間の結合を表す重み  $w$  (バイアスを含む)、活性化関数のスロープをニューロンごとに調整するための係数  $a_f, a_b$  である。

### 4.3 形状補間を用いた学習データ生成

4.2 節でのパラメータにより、4.1 節でのデータを訓練データとして、ディープオートエンコーダの学習をおこなった。そして、そのディープオートエンコーダを用いて、訓練に用いた形状に類似な形状をデコードすることにより、形状学習のための学習データを生成した。それぞれの実験で用いたディープオートエンコーダは、隠れ層が 3 層のものである。入力層と出力層を含めた各層の素子数は、32-16-8-16-32 である。素子数が 8 である層までの前半部がエンコーダであり、素子数が 8 である層からの後半部がデコーダである。

プリトレーニングにより学習したディープオートエンコーダを用いて訓練データを再構成する例を、図 3, 4 に示す。それぞれ、(1) 式の誤差が  $e = 0.0001$ 、または、0.1 の場合である。これにより、訓練データが各層の間で写像される様子が分かる。第 1 行は、左からそれぞれ、入力層、そして、第 1 隠れ層の出力ベクトルであり、第 2 行は、左からそれぞれ、第 2 隠れ層、そして、第 3 隠れ層の出力ベクトルである。第 3 行は、出力層の出力ベクトルである。

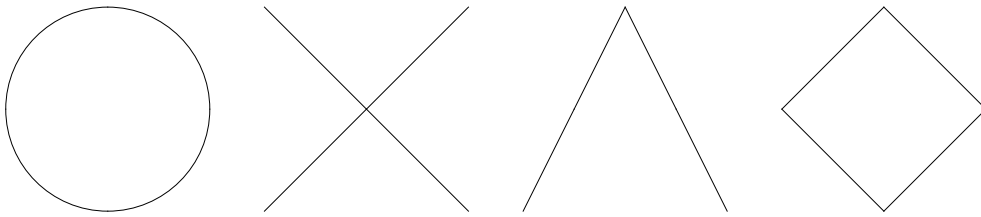


図 2 訓練データの例

Fig. 2 Examples of training shapes: From left to right: a circle, a cross, an arrow tip and a diamond shape.

表 1 ハイパーパラメータ

Table 1 Hyper-parameters

learning rate	momentum	weight cost	$\sigma$
0.001	0.5 ( <i>iter.</i> $\leq$ 5), 0.9 ( <i>iter.</i> $>$ 5)	0.00001	0.01

各層の出力ベクトルの表記法は、4.1 節と同様である。各層においてデータと、印は対応している。

学習したディープオートエンコーダを用いて生成した形状の例を、図 5 に示す。これらの形状は、訓練データを再構成した形状、または、補間形状である。(1) 式の誤差  $e$  として、 $e = 0.0001, 0.01$ , または、 $0.1$  を用いた。形状補間のための補間点は、訓練データを第 2 隠れ層に写像したベクトルを補間することにより生成した。

#### 4.4 検討

図 3, 4 から、(1) 式の誤差  $e$  の値を変化させることにより、訓練形状を再構成した形状が変化することが分かる。 $e = 0.0001$  と  $e = 0.01$  の場合を比較すると、誤差  $e$  が小さいほど、訓練形状に近い形状が再構成される。

図 5 に示す生成した形状を、訓練形状と比較すると、多様性をもつ類似な形状が生成できることが分かる。特に、Bezier 曲線セグメントの接続に着目すると、訓練形状では接続しているセグメントが離れた形状を生成することが可能であることが分かる。

これらの大部分の生成された形状は、基準とした形状に類似なことから、学習形状として適しているといえるが、図 5 の 4 行 4 列の形状のように、学習形状として適していないものもある。生成した形状データから、学習に適した形状を選択する手法については、今後の課題である。

また、提案の手法を用いると、非線形な形状変換をおこなった類似な形状を生成することが可能であることが確認できた。提案の手法により生成した形状に対して、相似変換または非均一相似変換などの線形手法による幾何学的な変換をおこなうことにより、形状に関する制約を与えて微調整をした類似な形状を生成することが可能である。

## 5. おわりに

平面上の Bezier 曲線により表現された形状を対象として、形状補間のための既提案の手法 [10] を用いて、形状の機械学習で用いる学習データを生成するための手法を提案した。この手法を用いると、与えられた少数の曲線形状を基にして、多様性を持つ類似な形状を生成することが可能である。

提案では、始めに、ディープオートエンコーダを用いて予め与えられた基準とする形状の学習をおこなう。そして、学習したディープオートエンコーダを用いて、訓練形状に類似な形状を生成する。この生成される形状に多様性をもたせるために、学習形状の組み合わせ、学習を終了するための基準として設定する学習誤差、そして、形状の補間に着目した。

簡単な形状を基準の形状として、その形状を相似変換した形状を訓練データとして、提案手法の動作の確認をおこなった。生成される形状の類似性という点では、比較的良質な形状が生成できた。

生成した形状は Bezier 曲線で表現されているので、その形状をラスタライズすることにより、画像データを生成することが可能である。そのため、形状画像の機械学習のための学習データの生成にも有効であるといえる。また、手書き形状の機械学習のための学習データには多様性をもつ類似な形状が必要であることから、提案手法により生成した形状を、学習データとして用いることが有効であるといえる。

今後の課題には、精度の改善、計算時間の減少、複雑な形状への適用、そして、機械学習の学習データとしての有効性の確認などに関する研究を挙げることができる。

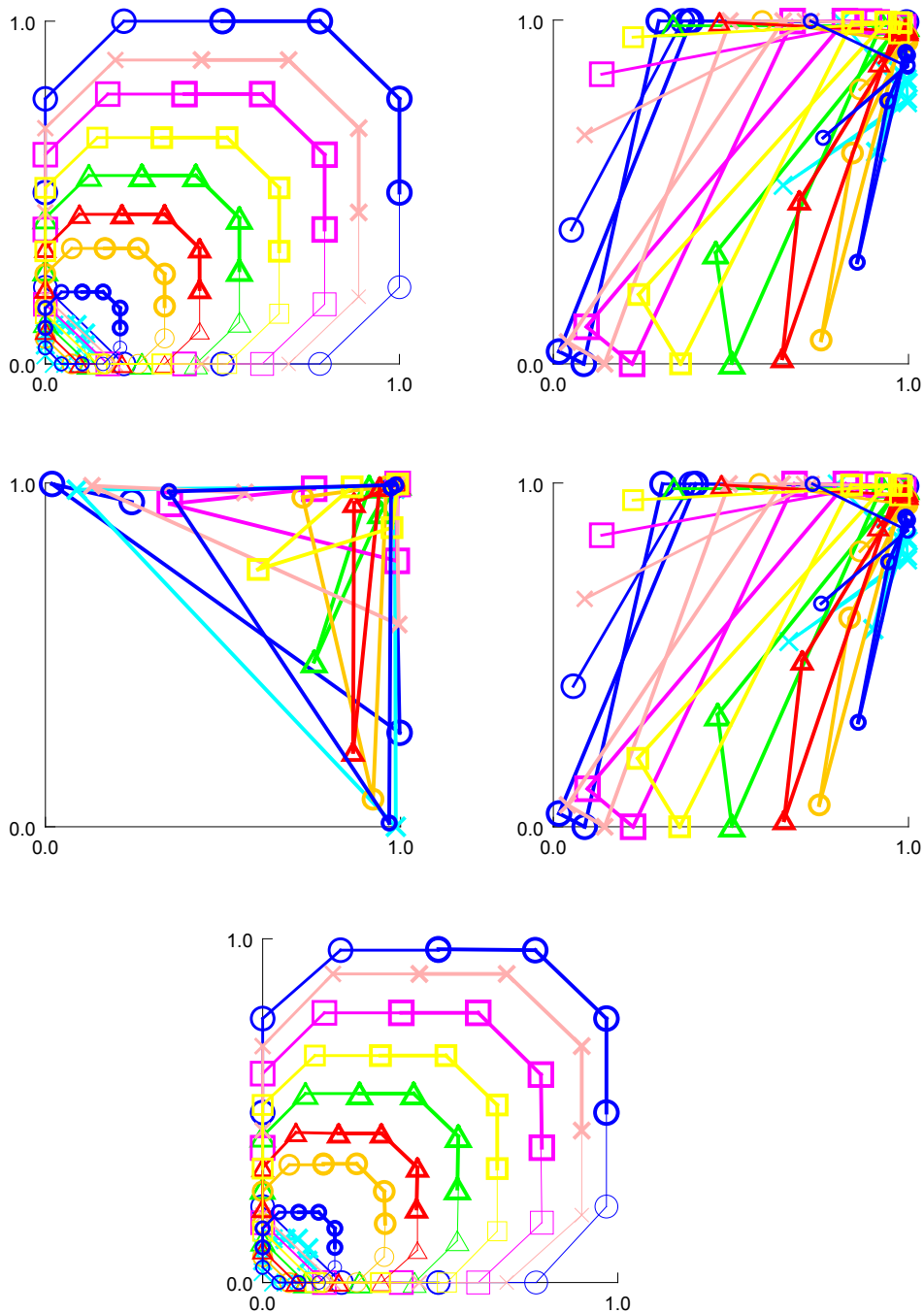


図 3 ディープオートエンコーダによる低次元空間への形状写像 ((1) 式  $e = 0.0001$ )

**Fig. 3** Shape mapping into low-dimensional spaces using deep autoencoder (eq.(1) $e = 0.0001$ ): First row: The output vectors of the input layer(left) and of the 1st hidden layer(right) are shown. Second row: The output vectors of the 2nd hidden layer(left) and of the 3rd hidden layer(right) are shown. Third row: The output vectors of the output layer are shown. The output vectors of the input layer are the same as the training data.

参考文献

[1] Campbell, N. D. F. and Kautz, J.: Learning a Manifold of Fonts, *ACM Trans. Graph.*, Vol. 33, No. 4, pp. 91:1–

91:11 (online), DOI: 10.1145/2601097.2601212 (2014).  
[2] Chen, H. and Murray, A.: Continuous restricted Boltzmann machine with an implementable training algo-

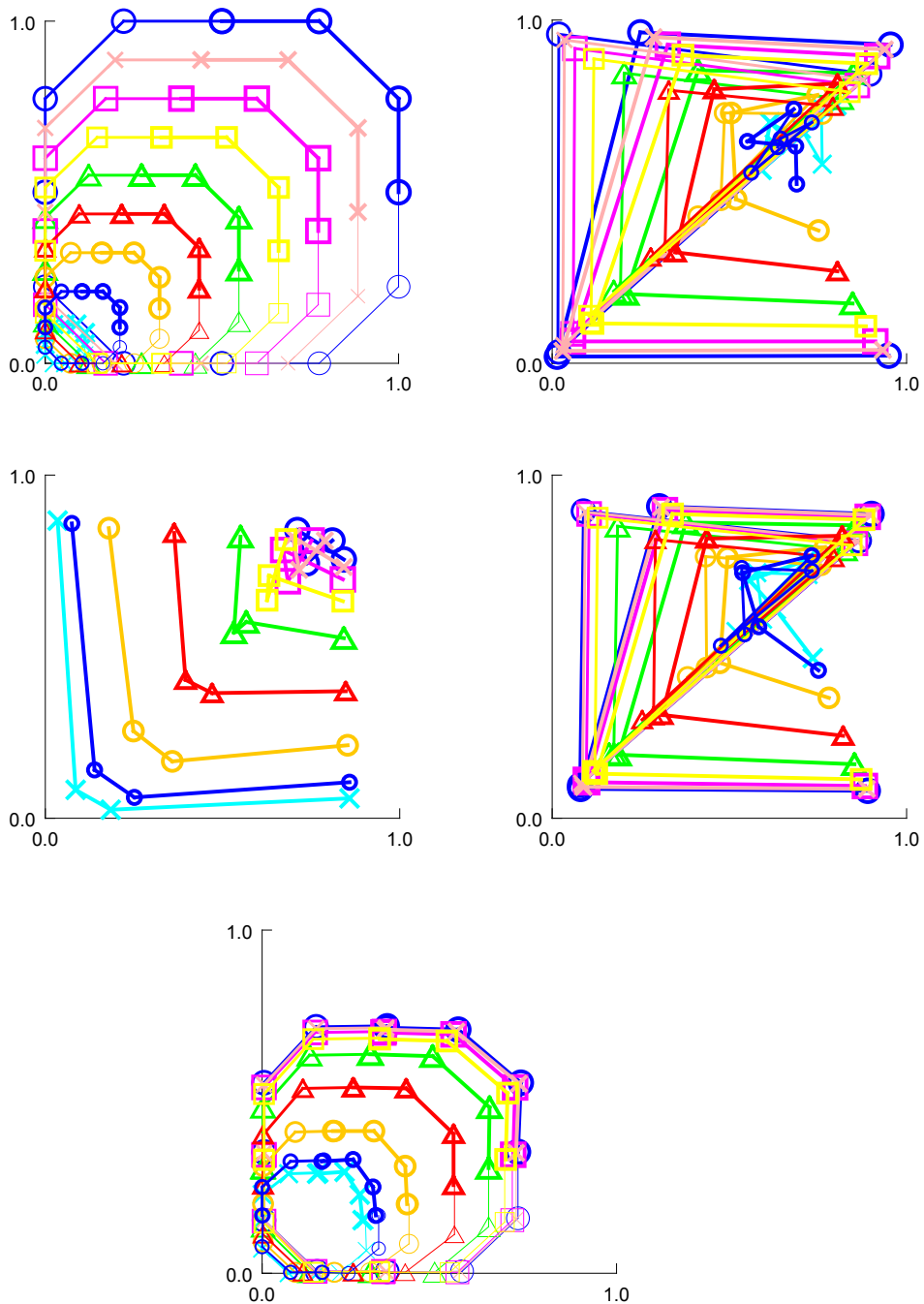


図 4 ディープオートエンコーダによる低次元空間への形状写像 ((1) 式  $e = 0.01$ )

**Fig. 4** Shape mapping into low-dimensional spaces using deep autoencoder (eq.(1)  $e = 0.01$ ): First row: The output vectors of the input layer(left) and of the 1st hidden layer(right) are shown. Second row: The output vectors of the 2nd hidden layer(left) and of the 3rd hidden layer(right) are shown. Third row: The output vectors of the output layer are shown. The output vectors of the input layer are the same as the training data.

rithm, *Vision, Image and Signal Processing, IEE Proceedings -*, Vol. 150, No. 3, pp. 153-158 (online), DOI: 10.1049/ip-vis:20030362 (2003).

[3] Forsey, D. R. and Bartels, R. H.: Hierarchical

B-spline refinement, *SIGGRAPH '88: Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, ACM, pp. 205-212 (online), DOI:

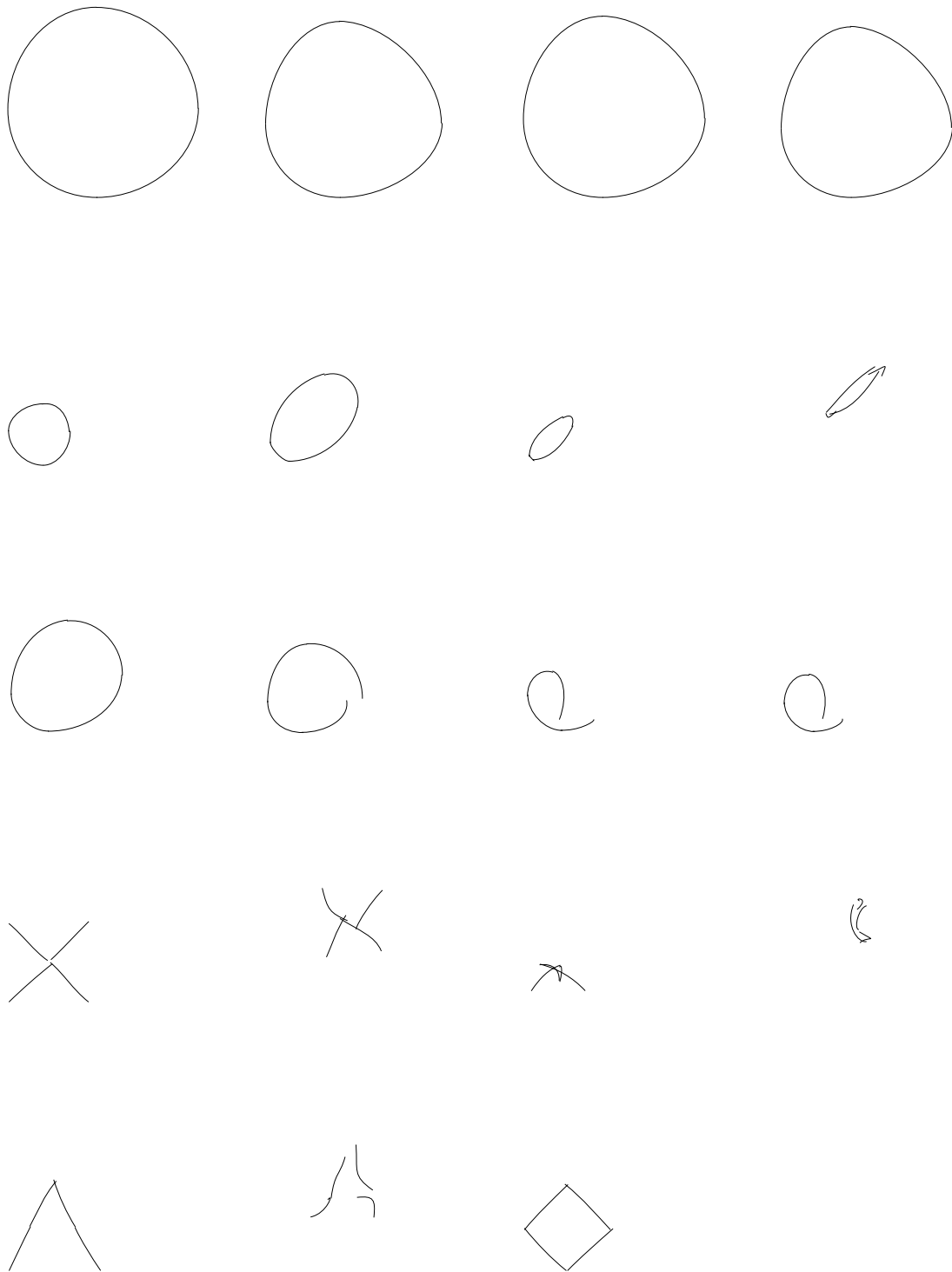


図 5 生成した形状学習データの例

Fig. 5 Examples of the generated shapes for shape learning. Almost all shapes are suitable for shape learning, while some shapes are not.

[4] <http://doi.acm.org/10.1145/54852.378512> (1988).  
Fowler, B. and Bartels, R.: Constraint-based curve manipulation, *Computer Graphics and Applications, IEEE*, Vol. 13, No. 5, pp. 43-49 (online), DOI:

10.1109/38.232098 (1993).  
[5] Fu, L. and Kara, L. B.: Neural network-based symbol recognition using a few labeled samples, *Computers & Graphics*, Vol. 35, No. 5, pp. 955 - 966 (online), DOI:

- 10.1016/j.cag.2011.07.001 (2011).
- [6] Hinton, G. E. and Salakhutdinov, R. R.: Reducing the Dimensionality of Data with Neural Networks, *Science*, Vol. 313, No. 5786, pp. 504–507 (2006).
  - [7] Hinton, G. E.: A Practical Guide to Training Restricted Boltzmann Machines, *Neural Networks: Tricks of the Trade - Second Edition*, Vol. 7700, Springer, pp. 599–619 (2012).
  - [8] Igarashi, T., Moscovich, T. and Hughes, J. F.: As-Rigid-As-Possible Shape Manipulation, *ACM SIGGRAPH 2005*, ACM, pp. 1134–1141 (2005).
  - [9] 佐藤 信：非均一相似形状補間を用いた形状学習データ生成，情報処理学会研究報告，Vol. 2013-CG-150, No. 11, pp. 1–8 (2013).
  - [10] 佐藤 信：非均一相似形状を学習したオートエンコーダを用いた形状補間，情報処理学会研究報告，Vol. 2014-CG-157, No. 3, pp. 1–8 (2014).
  - [11] 佐藤 信，三輪譲二：導関数ベクトルの非均一相似性制約に基づく曲線洗練化法，情報処理学会研究報告，Vol. 2011-CG-142, No. 12, pp. 1–6 (2011).
  - [12] 佐藤 信，三輪譲二：平面曲線形状洗練化のための導関数ベクトルの非均一相似性制約を用いた鏡映対称変換，情報処理学会研究報告，Vol. 2012-CG-146, No. 34, pp. 1–6 (2012).
  - [13] Sarveniazi, A.: An Actual Survey of Dimensionality Reduction, *American Journal of Computational Mathematics*, Vol. 4, No. 2, pp. 55–72 (2014).
  - [14] van der Maaten, L.: Learning a Parametric Embedding by Preserving Local Structure, *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (AI-STATS)*, pp. 384–391 (2009).
  - [15] Van Vaerenbergh, S., Santamaria, I. and Barbano, P.: Semi-supervised handwritten digit recognition using very few labeled data, *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pp. 2136–2139 (2011).
  - [16] Welch, W. and Witkin, A.: Variational surface modeling, *SIGGRAPH '92: Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, ACM, pp. 157–166 (online), DOI: <http://doi.acm.org/10.1145/133994.134033> (1992).