

文字画像データの位相的データ解析による筆跡鑑定への応用

岩頭 由樹¹ 佐藤 好久²

概要: 現在収集できるようになったデータの量や種類は、情報技術の発展により、大きく増加している。それに伴い、たくさんのデータを分別するデータ解析技術の発展も求められている。例えば、画像に書かれた文字をデータ解析によって認識して、その画像に何が書かれているか抽出することである。本研究の目的は、ある文字の書かれた画像に TDA (Topological Data Analysis) 技術を使い、その文字の特徴を抽出し、誰が書いた文字かを正しく分類できるか確かめるものである。データ解析の対象として、手書き文字を用いる。文字が TDA とパターン認識によって、筆跡がどう解析されるか示していく。

キーワード: 情報数学, 計算幾何学, 機械学習

Application of topological data analysis of character image data to handwriting analysis

Abstract: The amount and type of data that can now be collected has increased significantly due to developments in information technology. This has also required the development of data analysis technology to sort through the large amount of data. An example is the recognition of characters written on an image by data analysis to extract what is written on the image. The aim of this research is to see if it is possible to extract the features of the characters in an image of a certain character using TDA (Topological Data Analysis) techniques, and to correctly classify who wrote the characters. Handwritten characters are used as the object of data analysis. The paper shows how handwriting can be analyzed by TDA and pattern recognition.

Keywords: Information Mathematics, Computational Geometry, Machine Learning

1. はじめに

現在収集できるようになったデータの数や種類は、情報技術の発展により、大きく増加している。それに伴い、たくさんのデータを分別するデータ解析技術の発展も求められている。例えば、画像に書かれた文字をデータ解析によって認識して、画像に何が書かれているか抽出することである。TDA (Topological Data Analysis) 技術とは、近年データ解析技術に対する新しい手法として注目されていて、あるデータが持つ「形」を認識し、その形がもつ情報を位相的な観点から解析する手法である。従来の方法では、

データを分析する際、統計的モデルという「形状」を仮定して分析するため、統計的モデルに当てはめることができないデータは解析がうまくいかないことがあるが、TDA は統計的モデルという「形状」に依存しないため、様々なデータの「形状」を解析できるとされている。本研究の目的は、ある文字の書かれた画像に TDA (Topological Data Analysis) 技術を使い、その文字の特徴を抽出し、誰が書いた文字かを正しく分類できるか確かめるものである。データ解析の対象として、手書き文字を用いる。文字が TDA とパターン認識によって、筆跡がどう解析されるか示していく。

筆跡鑑定の先行研究 [5] では、手書き文字の濃度ヒストグラム、画素値、文字の占有率 (筆跡の太さ、大きさ) の 3 つの特徴値を検出した後、機械学習 (ニューラルネットワーク) により研究を行っている。また、TDA を利用した文字認識の先行研究 [4] では、画像データから得られた

¹ 九州工業大学大学院情報工学府情報創成工学専攻
Department of Creative Informatics,
Graduate school of Computer Science and Systems Engineering,
Kyushu Institute of Technology

² 九州工業大学大学院情報工学研究科 知能情報工学研究系
Faculty of Computer Science and Systems Engineering,
Kyushu Institute of Technology

パーシステント図をベクトル化し、機械学習(サポートベクターマシンと平均最近傍法)を用いて文字画像の識別を行っている。本研究では、画像データから TDA によって得られたパーシステント図をベクトル化し、機械学習(サポートベクターマシン, 平均最近傍法, ニューラルネットワーク)を用いて筆跡鑑定を行う。

本研究の目的は、TDA 解析により筆跡鑑定を行い、文字画像は誰が書いたものかをどれだけ正しく識別されるかを示すことで、筆跡鑑定における TDA の働きを確かめることにある。また、文字画像データの筆跡鑑定の精度をより良いものにする可能性を検討するために本研究を行った。

2. パーシステントホモロジー群

文字画像データの解析のために、位相的データ解析のパーシステントホモロジー群と、その結果を図にしたパーシステント図を利用する。

2.1 準備

ここからは、パーシステントホモロジー群のための事前知識を示していく。

有限個の点 $P = \{x_i \in \mathbb{R}^N | i = 1, \dots, m\}$ からなるデータを考える。この P に対し、抽象的単体複体を考える。有限集合 V と V の部分集合の有限個の族 \mathcal{K} が、(1) $v \in V \implies \{v\} \in \mathcal{K}$, (2) $\sigma \in \mathcal{K}, \tau \subset \sigma \implies \tau \in \mathcal{K}$ を満たすときの (V, \mathcal{K}) を、**抽象的単体複体**, または簡単に**抽象的複体**という。位相的データ解析では、有限個のデータ(点)を「図形」として幾何的に表現するため、抽象的単体複体を利用する。その代表的なものに、**チェック複体**, **ヴィートリス・リップス複体**があり、それぞれ $\mathcal{C}(P, r)$, $\mathcal{R}(P, r)$ と書く。その抽象的単体複体は、それぞれ次のような性質をもつ。 $0 \leq \forall s, t \leq k$ とする。

$$\{x_{i_0}, \dots, x_{i_k}\} \in \mathcal{C}(P, r) \Leftrightarrow \bigcap_{j=0}^k B_r(x_{i_j}) \neq \emptyset$$

$$\{i_0, \dots, i_k\} \in \mathcal{R}(P, r) \Leftrightarrow B_r(x_{i_s}) \cap B_r(x_{i_t}) \neq \emptyset$$

ここで、 P に対して、各点 x_i を中心とした半径 $r (> 0)$ の球 $B_r(x_i) = \{x \in \mathbb{R}^N | \|x - x_i\| \leq r\}$ を配置する。ここで $\|x\|$ はユークリッドノルムを表す。この2つの抽象的単体複体は、半径 r が $0 < r < r'$ のとき、 $\mathcal{C}(P, r) \subset \mathcal{C}(P, r')$, また $\mathcal{R}(P, r) \subset \mathcal{R}(P, r')$ が成立する。これより、正の数からなる増大列 $r_0 < r_1 < \dots < r_T$ が与えられたとき、 $\mathcal{C}(P, r_0) \subset \mathcal{C}(P, r_1) \subset \dots \subset \mathcal{C}(P, r_T)$, また $\mathcal{R}(P, r_0) \subset \mathcal{R}(P, r_1) \subset \dots \subset \mathcal{R}(P, r_T)$ となる。この増大列を**フィルトレーション**と呼ぶ。

また、抽象的単体複体の「時系列」 K^t ($t = 0, 1, 2, \dots$) が $\mathbb{K} : K^0 \subset K^1 \subset K^2 \subset \dots \subset K^t \subset \dots$ を満たすフィルトレーション $\mathbb{K} = \{K^t | t = 0, 1, 2, \dots\}$ が、 $K^j = K^\Theta$ を満たすとき、フィルトレーション \mathbb{K} は**有限型**であるという。

Θ の最小値を**飽和時刻**という。

2.2 パーシステントホモロジー群

有限個の単体からなる抽象的単体複体の多面体と同相な位相空間 X の環 \mathbb{Z}_2 係数 q 次元ホモロジー群 $H_q(X)$ は、有限生成 \mathbb{Z}_2 加群として与えられる。直感的に、 $H_q(X)$ の生成元は X の q 次元の穴を表し、その個数は X 内の本質的な q 次元の穴の個数を表す。例えば、 $q = 0, 1$ はそれぞれ連結成分, 輪を表す。

加群の構造定理により、係数環 \mathbb{Z}_2 による K^t のホモロジー群 $H_q(K^t)$ は、いくつかの因子 \mathbb{Z}_2 による直和分解

$$H_q(K^t) \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \dots \oplus \mathbb{Z}_2 \quad (*)$$

を持つ。

一方、 $t < s$ より、 $K^t \subset K^s$ である。包含写像 $i^{ts} : K^t \hookrightarrow K^s$ が誘導する準同型写像 $\varphi : H_q(K^t) \rightarrow H_q(K^s)$ を通じて、直和分解 (*) の各生成元の時間変化を見ることができ。例えば、直和分解 (*) の第 i 番目の因子の時間変化 $t_1 < t_2 < \dots < t_n < \dots$ に対応する部分を取り出すと、

$$\dots \rightarrow 0 \rightarrow \dots \rightarrow 0 \rightarrow \mathbb{Z}_2 \xrightarrow{t_1} \mathbb{Z}_2 \xrightarrow{t_2} \dots \rightarrow \mathbb{Z}_2 \xrightarrow{t_n} 0 \rightarrow \dots$$

のようになる。このとき、対応する生成元は時刻 $t = t_1$ で発生し、時刻 $t = t_n$ で消滅したと考えられる。

このようなホモロジー群の時系列を多項式環 $\mathbb{Z}_2[x]$ による加群としてとらえたものがフィルトレーション \mathbb{K} の q 次元のパーシステントホモロジー群 $PH_q(\mathbb{K})$ である。このとき、 $PH_q(\mathbb{K})$ は、次数付き $\mathbb{Z}_2[x]$ 加群として、次のような一意的な分解をもつ。

$$PH_q(\mathbb{K}) \simeq \bigoplus_{i=1}^p I(b_i, d_i)$$

$b_i, d_i \in \mathbb{R}$ かつ $b_i < d_i$ である。ただし、

$$I(b_i, d_i) = (x^{b_i}) / (x^{d_i})$$

である。ここで、 (x^k) は x^k で生成される $\mathbb{Z}_2[x]$ の単項イデアルである。

各 $I(b_i, d_i)$ の生成元は、 $t = b_i$ で発生し、 $t < d_i$ まで持続し、 $t = d_i$ で消滅する q 次元の穴を表現している。 $b_i, d_i, d_i - b_i$ をそれぞれ**発生時刻**, **消滅時刻**, **持続時間**という。このとき、各生成元の発生消滅時刻対 (b_i, d_i) からなる多重集合

$$D_q(\mathbb{K}) = \{(b_i, d_i) | i = 1, \dots, p\}$$

を \mathbb{K} の q 次元パーシステント図という。持続時間の長い発生時刻と消滅時刻のペア $(b_i, d_i) \in D_q(\mathbb{K})$ は \mathbb{K} の信頼できるトポロジー構造として評価し、反対に、持続時間が短いものはノイズになりやすい。

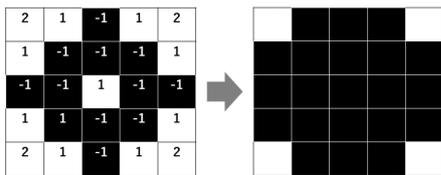


図 1 白黒二値画像についてのフィルトレーション例

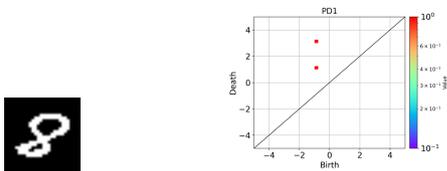


図 2 白黒二値画像の数字の 8 図 3 数字の 8 に対する結果

ここで、白黒二値画像のパーシステント図について記す。マンハッタン距離を導入し、その増減に伴う形の変化を調べていく。マンハッタン距離は、あるピクセルから別のピクセルまでに到達するために横切る辺の数の最小値で与えられる。

二値画像は黒の領域と白の領域に分けられ、計算の際はどちらの領域の表面までの距離を示すかを選択する。選択した領域の外側に太らせる操作なら横切る辺の数は距離として正の数で与えられ、内側に細らせる操作なら、距離は負の値で表される。

図 1 の各ピクセルには、黒の領域までの距離が記載してある。図 1 のフィルトレーションは、単体複体によるものではないが、図 1 の左から黒の領域を 1 ピクセル太らせる操作を行うと、図 1 の右ようになる。この操作は、マンハッタン距離が 1 の白ピクセルを黒ピクセルにすることである。図 1 においては、白の領域を囲む黒の領域のわかさを太らせて、図 1 左の中心の白の領域 (穴) を消滅させている。

図 1 の左を $t = 0$ 、右を $t = 1$ とすると、 $t = 0$ 時点の中心にある黒い部分のわかさの存続期間は 1 で、1 次元のパーシステント図は $\{(0, 1)\}$ で与えられる。

また、白黒二値画像の文字データとその 1 次元のパーシステント図を示す。白黒二値画像は、[2] の MNIST データを PNG 画像化したもののうちの 1 枚を用いた。

図 2 に示すような文字の画像データに対し、その白い部分に注目して、その文字に存在する穴がいつ発生し、いつ消失するかを確認する。その結果として図 3 が得られる。この図 3 で確認できる点は、画像の文字が持つ穴の数に対応している。この場合、赤色の点が 2 個存在し、図 2 に書かれた数字の 8 の 2 個の穴は、図 2 時点よりも白の領域が膨張してから消滅することがわかる。

3. 機械学習

ここからは、筆跡鑑定に用いるパターン認識のための機

械学習の手法について示す。

3.1 サポートベクターマシン

サポートベクターマシンとは、主に二値分類に用いられる。多クラス分類への応用も可能である。[3], [7] を参考にした。

データは特徴ベクトル $\mathbf{x} = {}^t(x_1, x_2, \dots, x_p)$ と目的変数 y からなる (\mathbf{x}, y) からなり、その学習データ $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m)$ が得られているとする。目的変数 y は、本研究においては手書き数字の認識で分類のアルゴリズムを使うのため、クラスのラベルとして使われる。

分類のアルゴリズムは、ラベルの種類の数に応じて、2 値分類、多クラス分類に分けられる。前者は 2 つのクラスに分類する判別のことを指し、後者は 3 つ以上のクラスタリングに分類する判別のことを指す。多クラス分類では、1 つのクラスと残りのクラスに分類することでクラスを決めていく 1 対他方式、全体から 2 つのペアを作って分類し、すべての組み合わせに対して、多数決を行う 1 対 1 方式がある。これらは 2 値分類を複数回行うことで得られる。

学習データから、その多くに対して $h(\mathbf{x}_i) = y_i$ となるような判別器 $h: \mathbb{R}^p \rightarrow \{1, -1\}$ を学習し、新たな入力 \mathbf{x} のラベルを $h(\mathbf{x})$ で予測することになる。

判別器について書く。符号関数 sign を

$$\text{sign}(z) := \begin{cases} 1 (z \geq 0) \\ -1 (z < 0) \end{cases}$$

と定める。このとき、関数 $f: \mathbb{R}^p \rightarrow \mathbb{R}$ で、多くのデータに対して $\text{sign}(f(\mathbf{x}_i)) = y_i$ を満たすものを見つけられたとき関数 f を判別関数といい、判別式 h を $h(\mathbf{x}) = \text{sign}(f(\mathbf{x}))$ と定められる。

p 次元空間 \mathbb{R}^p における 1 つの超平面 $w_1x_1 + w_2x_2 + \dots + w_px_p + b = {}^t\mathbf{w}\mathbf{x} + b = 0$ (${}^t\mathbf{w} = {}^t(w_1, w_2, \dots, w_p)$) を境界面にするにより、学習データ $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m)$ が 2 種類のラベルをもつグループに完全に分離される状況を考える。この場合、判別関数 f を $f(\mathbf{x}) := w_1x_1 + w_2x_2 + \dots + w_px_p + b = {}^t\mathbf{w}\mathbf{x} + b$ と定義する。

サポートベクターマシンでは、この判別関数を与える超平面を「マージンが最大」となるようにする。このような超平面をマージン最大化超平面という。マージンとは、各学習データと分離超平面との距離の最小値のことである。

マージン最大化超平面は、マージン最大化問題 $y_i({}^t\mathbf{w}\mathbf{x}_i + b) > 0$ ($i = 1, 2, \dots, m$) となるような

$$\max_{\mathbf{w}, b} \min_{i=1, 2, \dots, m} \frac{|{}^t\mathbf{w}\mathbf{x}_i + b|}{\|\mathbf{w}\|}$$

の最適解として得られる。すなわち、 $y_i({}^t\mathbf{w}\mathbf{x}_i + b) > 0$ ($i = 1, 2, \dots, m$) となるような

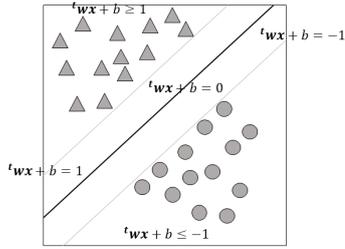


図 4 サポートベクターマシンによる識別境界の例

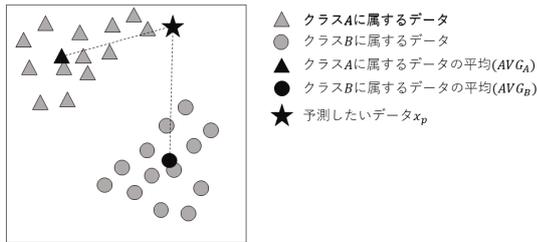


図 5 2次元特徴ベクトルによる平均最近傍法の例

$$\max_{w \in \mathbb{R}^p, b \in \mathbb{R}} \min_{i=1,2,\dots,m} \frac{1}{2} \|w\|^2$$

の最適解が求めたいマージン最大超平面を与える。これをサポートベクターマシン (SVM) という。

この SVM の定義により、図 4 が表せる。

3.2 平均最近傍法

$$\text{多次元の特徴ベクトル } \mathbf{x} = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}$$

と目的変数 $\mathbf{y} = (y_1, \dots, y_l)$ が与えられ、学習データ $(x_{11}, \dots, x_{1n}, y_1), \dots, (x_{m1}, \dots, x_{mn}, y_l)$ が得られているとする。このとき、各種類の特徴ベクトルの個数を $N_{y_1}, N_{y_2}, \dots, N_l$ としたとき、各種類の目的変数ごとの特徴ベクトルの平均 AVG_{y_i} を

$$\begin{aligned} AVG_{y_i} &:= (\overline{x_{i1}}, \dots, \overline{x_{in}}, y_i) \\ &= \left(\frac{1}{N_{y_i}} \sum x_{i1}, \dots, \frac{1}{N_{y_i}} \sum x_{in}, y_i \right) \end{aligned}$$

と定義できる。 X_1, \dots, X_n のデータを加算するのは、目的変数 y_i に分類されたものだけである。

このとき、分類したいデータ $\mathbf{x}_p = (x_{p1}, \dots, x_{pn})$ に対して、 $\min_{i=1,\dots,l} \{(\overline{x_{i1}} - x_{p1})^2 + \dots + (\overline{x_{in}} - x_{pn})^2\}$ となるような i の目的変数 y_i を x_p に与える。これを平均最近傍法という。

例として、特徴ベクトルが2次元の場合のときのイメージ図として図 5 を示す。図 5 の場合、予測したいデータ x_p はクラス A に属するデータの平均値に近いので、目的変数として A を与える。

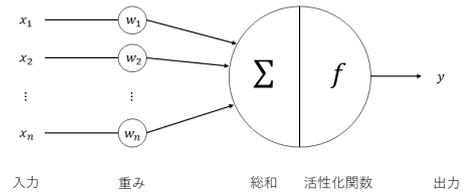


図 6 ニューロンのモデル

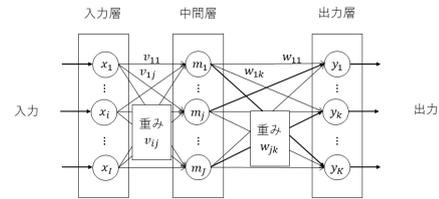


図 7 ニューラルネットワークのモデル

3.3 ニューラルネットワーク

ニューラルネットワーク (NN) とは、参考文献 [3], [8] より、モデル化したニューロンを複数つなげたネットワーク構造において、信号伝達を繰り返して情報処理を行うアルゴリズムである。ニューラルネットワークには学習機能が備わっており判断規則を構築する。

ニューロンのモデルは、図 6 のように表せる。 x_1, x_2, \dots, x_n は入力、それぞれに対応した w_1, w_2, \dots, w_n は重み、ニューロンは入力の重み付き総和 \sum と活性化関数 f で発火する。活性化関数は非線形関数である。本研究においては、活性化関数は **ReLU 関数** を使用した。ReLU 関数は、以下のように定義できる。

$$f(x) = \begin{cases} 0 & (x < 0) \\ x & (x \geq 0) \end{cases}$$

他のニューロンから伝播される入力情報の総和は、これにより、値の変換をされた後に出力される。入力の総和から閾値を引いた値が正なら 1、負なら 0 を出力するなど、二値の値を出力する。閾値を a とすると、出力値 y は $y = f\left(\sum_{i=1}^n x_i w_i - a\right)$ になる。この式において、 $-a = x_0 w_0$ とし、 x_0 は常に 1 の入力があるユニットとする。これは、ニューロンの中のバイアスとして利用する。 $-a = b$ として、 b をバイアスとすると、 $y = f\left(\sum_{i=1}^n x_i w_i + b\right)$ になる。

図 7 のように、ニューラルネットワークは表すことができる。 x_i は入力層の i 番目の要素、 m_j は中間層の j 番目のニューロンの出力、 y_k は出力層の k 番目のニューロンの出力、 v_{ij} は入力層の i 番目から中間層の j 番目のニューロンとの間の結合重み、 w_{jk} は中間層の j 番目から出力層の k 番目のニューロンとの間の結合重みを表す。また、入力層から出力層へ向かう演算は順方向演算といい、出力層

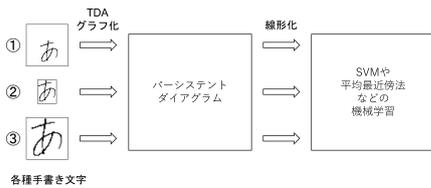


図 8 TDA を使った筆跡鑑定の概略図

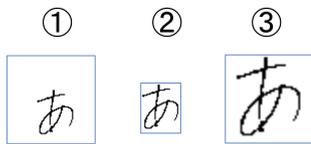


図 9 3種類の文字の切り出し方

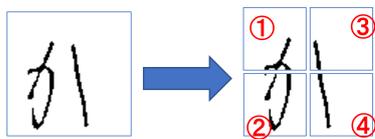


図 10 切り出した画像の4等分処理

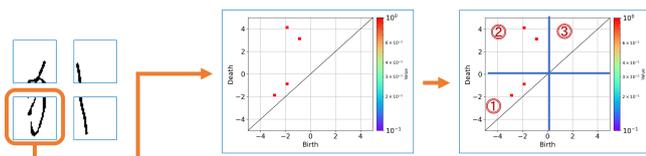


図 11 4等分した画像のパーシステント図算出とベクトル化

から入力層へ向かう演算は逆方向演算という。

4. 実験手順

本研究における実験では、TDA 解析結果のパーシステント図をベクトル化して、3種類の機械学習の手法 (SVM, 平均最近傍法, NN) で筆跡鑑定を行った。また、TDA を用いずに、NN を用いた筆跡鑑定も行った。

本実験の概略図を図 8 に示す。手書き文字のデータは、独自に収集した、4人分のものを用いる。

「あいうえおかきくけこさしすせそ月火水木金土日」を、1文字につき7回書かせた画像を白黒二値画像にし、1文字ずつ切り出したものをを入力する。切り出し方は、図 9 より、左から、①のように文字を書く枠に合わせて画像データを1文字ごとに切り出すもの、②のように文字の大きさに合わせて画像データを1文字ごとに切り出すもの、③のように、切り出した②から処理を行い、その画像の長辺に長さを合わせた正方形にして、その大きさを①に合わせたものの3つの方法を使う。

パーシステントホモロジー群で解析を行う際には、図 10 のように、文字画像を4等分してから行う。この1文字の画像を4等分したものを、各部分ごとにパーシステントホモロジー群、パーシステント図を算出する。パーシステント図を一定数に分割することによって、ベクトル化を行う。

表 1 TDA と SVM の筆跡鑑定による識別率 (約~%)

	①	②	③
あ	67	75	50
い	75	42	75
う	83	42	25
え	75	42	50
お	67	83	50
か	83	83	42
き	67	67	42
く	67	67	25
け	75	58	67
こ	92	100	42
さ	58	75	25
し	83	58	50
す	83	75	42
せ	83	58	42
そ	83	100	67
月	83	75	58
火	83	58	67
水	83	75	42
木	58	67	58
金	83	67	75
土	67	67	58
日	58	83	58

表 2 TDA と平均最近傍法の筆跡鑑定による識別率 (約~%)

	①	②	③
あ	67	75	42
い	75	42	83
う	100	67	42
え	75	42	50
お	50	83	33
か	83	75	42
き	58	67	50
く	67	75	25
け	50	50	67
こ	83	100	50
さ	67	50	33
し	83	75	58
す	75	67	50
せ	92	50	42
そ	75	100	75
月	83	75	42
火	83	75	58
水	83	83	50
木	58	50	67
金	75	67	75
土	67	67	67
日	67	67	33

図 11 ではデータ処理の流れを表している。左から、4等分に分割した文字画像、その4等分した画像のうちの1つの画像における0次元のパーシステント図、0次元のパーシステント図を格子状に分け、対角線を通る部分から上の部分をベクトルの要素としたものである。今回は3次元のベクトルへと線形化を行っている。各部分でベクトル化した4つのデータを、1文字分に統合し、合計12次元のベクトルデータにする。

5. 実験結果

以下の実験結果において、機械学習によるパターン認識では、4人分の、7回書いてもらったうちの4回分を学習データ、残りの3回分をテストデータとして行う。つまり、各文字で16枚の文字画像データを学習データ、12枚の文字画像データをテストデータとする。

5.1 TDA とサポートベクターマシンによる筆跡鑑定

機械学習として、サポートベクターマシン (SVM) を選択した場合、図 8 の機械学習の欄をサポートベクターマシンに指定した概略図の流れで筆跡鑑定を行う。

その実行結果は、表 1 の通りになる。①、②、③は、図 9 の画像データの文字の切り出し方に対応している。

表 1 は、「あ」から「日」までの文字画像のうち、ほとんどの場合、図 9 の①のように切り出した画像を TDA で解

表 3 TDA と中間層のニューロン数 200 の NN の
筆跡鑑定による識別率 (約~%)

	1000			5000			10000		
	①	②	③	①	②	③	①	②	③
あ	67	67	67	42	42	42	25	25	33
い	58	58	58	42	42	50	42	42	42
う	75	75	75	75	83	83	67	58	75
え	92	92	92	50	50	50	50	50	50
お	25	25	25	83	83	83	50	58	42
か	75	83	92	75	75	83	42	42	42
き	75	83	75	50	50	50	58	50	50
く	83	83	83	67	58	67	42	58	50
け	75	75	67	33	42	42	75	75	75
こ	83	83	83	75	75	75	50	42	42
さ	50	67	58	75	75	75	25	17	25
し	75	67	67	83	83	83	42	42	42
す	75	67	75	67	67	67	58	58	58
せ	92	92	92	50	50	50	50	42	50
そ	75	75	75	92	92	92	83	83	83
月	67	67	67	67	67	67	58	50	50
火	75	92	92	50	50	50	67	67	67
水	83	83	83	67	67	67	33	33	33
木	75	67	75	67	58	67	58	50	50
金	83	83	83	75	67	67	75	58	67
土	58	50	58	67	67	67	50	50	50
日	58	58	58	75	83	75	58	58	58

表 4 TDA と中間層のニューロン数 520 の NN の
筆跡鑑定による識別率 (約~%)

	1000			5000			10000		
	①	②	③	①	②	③	①	②	③
あ	67	67	67	42	42	50	42	33	25
い	58	58	58	42	50	42	42	42	42
う	75	75	75	75	83	75	67	58	58
え	92	92	92	50	50	50	50	50	50
お	25	25	25	83	83	83	50	50	50
か	83	83	83	75	75	75	33	25	33
き	75	75	75	50	50	50	50	50	50
く	83	83	83	58	58	58	58	58	50
け	75	67	75	42	42	42	83	75	83
こ	83	83	83	75	75	75	50	42	42
さ	58	58	58	83	75	75	17	25	25
し	67	75	67	75	83	83	50	50	50
す	67	67	75	67	67	67	58	50	50
せ	92	92	92	50	50	50	42	42	50
そ	75	75	75	92	92	92	83	83	83
月	67	67	67	67	67	67	58	50	58
火	83	83	83	50	50	50	75	75	67
水	83	83	83	67	67	67	33	33	33
木	67	67	67	58	67	67	50	50	50
金	83	83	83	67	67	67	67	75	75
土	58	67	58	67	67	67	42	58	50
日	58	58	58	83	75	83	58	58	50

析してベクトル化したものが、最も筆者を正しく識別できていることが多かったことが分かる。また、②の切り出し方において、「こ」と「そ」の識別率が100%とわかる。

5.2 TDA と平均最近傍法による筆跡鑑定

機械学習として、平均最近傍法を選択した場合、図8の機械学習の欄を平均最近傍法に指定した概略図の流れで筆跡鑑定を行う。

その実行結果は、表2の通りになる。①、②、③は、図9の画像データの文字の切り出し方に対応している。

表2は、表1のときと同様に、「あ」から「日」までの文字画像のうち、ほとんどの場合、図9の①のように切り出した画像をTDAで解析してベクトル化したものが、最も筆者を正しく識別できていることが多かったことが分かる。また、②の切り出し方において、「こ」と「そ」の識別率が100%とわかる。

また、筆跡鑑定の際に使う機械学習の比較としては、表1の結果は、表2よりも識別率が高くなったものもあれば、低くなったものもあることが分かる。

5.3 TDA とニューラルネットワークによる筆跡鑑定

機械学習として、ニューラルネットワーク(NN)を選択した場合、図8の機械学習の欄をニューラルネットワークに指定した概略図の流れで筆跡鑑定を行う。

その実行結果は、表3と表4の通りになる。①、②、③は、図9の画像データの文字の切り出し方に対応している。中間層のニューロンの数は、200と520の2通りに指定して行った。表3と表4にある1000、5000、10000は、学習の繰り返し回数である。学習の繰り返し回数は、1000回、5000回、10000回の3通りで行った。

実行結果より、表3と表4で中間層のニューロン数を変化させても識別率に大きな変化はなく、また学習の繰り返し回数を増加させても識別率に大きな改善が見られることはないことが分かる。

表3と表4も、表2と表1のときと同様に、「あ」から「日」までの文字画像のうち、ほとんどの場合、図9の①のように切り出した画像をTDAで解析してベクトル化したものが、最も筆者を正しく識別できていることが多かったことが分かる。

また、筆跡鑑定の際に使う機械学習の比較としては、表3と表4の結果は、表1と表2よりも識別率が高くなったものもあれば、低くなったものもあることが分かる。また、表3と表4より、NNでの筆跡鑑定では識別率が100%だったものはないことも分かる。

5.4 TDA を用いずに筆跡鑑定

また、参考として、TDAを使わずに筆跡鑑定を行った結果を示す。画像の画素値を検出したデータを、ニューラ

表 5 画像値だけ検出したデータで行った NN の筆跡鑑定 (約~%)

中間層の ニューロン数	200			520		
	1000	5000	10000	1000	5000	10000
あ	58	58	33	25	33	25
い	33	8	8	33	33	33
う	8	33	33	33	25	33
え	66	33	33	100	33	33
お	50	42	33	42	42	33
か	33	8	33	33	25	8
き	33	8	42	33	42	8
く	58	25	25	8	33	33
け	100	33	33	33	8	8
こ	25	42	25	25	58	8
さ	100	33	33	67	33	8
し	33	25	42	8	50	17
す	33	8	33	8	8	33
せ	8	25	25	33	33	25
そ	100	33	25	25	25	58
月	58	25	8	100	8	8
火	33	100	25	33	8	8
水	25	33	8	25	33	50
木	100	58	100	33	33	33
金	100	100	100	33	33	67
土	100	25	100	33	33	33
日	8	100	33	25	8	58

表 6 濃度ヒストグラム, 画像値, 文字の占有率を
検出したデータで行った NN の筆跡鑑定 (約~%)

中間層の ニューロン数	200			520		
	1000	5000	10000	1000	5000	10000
あ	75	8	33	67	67	92
い	33	67	75	33	67	42
う	42	33	67	67	33	8
え	42	0	67	25	8	25
お	8	33	58	33	92	92
か	25	25	33	92	92	67
き	58	92	42	67	33	75
く	33	33	67	92	67	42
け	33	92	8	58	67	67
こ	8	92	8	42	42	8
さ	8	75	8	33	33	33
し	8	50	35	25	25	33
す	50	67	8	8	58	33
せ	100	92	33	92	33	33
そ	50	25	75	58	92	25
月	50	33	58	67	83	58
火	33	8	33	58	58	8
水	67	8	58	8	58	42
木	25	8	50	58	75	67
金	67	33	58	42	25	100
土	75	17	33	17	92	83
日	33	33	25	8	8	67

ルネットワークを用いて筆跡鑑定を行う。なお、ここで筆跡鑑定に使用した画像データに用いた切り出し方は、図 9 における①のみである。

その実行結果は、表 5 の通りになる。中間層のニューロンの数は、200 と 520 の 2 通りに指定して行った。表 5 にある 1000, 5000, 10000 は、学習の繰り返し回数である。学習の繰り返し回数は、1000 回, 5000 回, 10000 回の 3 通りで行った。

表 5 より、文字画像の画素値を検出したデータで筆跡鑑定を行うと、識別率が 50% を超えた文字は少ないことが分かる。また、表 5 は、表 3 と表 4 の結果と比較すると、ほとんどの文字で識別率が下がっていることが分かる。中間層のニューロン数 200 のときの「さ」のように、識別率繰り返し回数 1000 回では 100% なのに回数を重ねると識別率が 33% のようにとても低くなってしまふものもあることが分かる。

また、筆跡鑑定の先行研究 [5] のように、画像の c(筆跡の太さ, 大きさ) を検出したデータを、ニューラルネットワークを用いて筆跡鑑定を行う。なお、ここで筆跡鑑定に使用した画像データに用いた切り出し方は、図 9 における①のみである。

その実行結果は、表 6 の通りになる。中間層のニューロンの数は、200 と 520 の 2 通りに指定して行った。表 5 にある 1000, 5000, 10000 は、学習の繰り返し回数である。

学習の繰り返し回数は、1000 回, 5000 回, 10000 回の 3 通りで行った。

表 6 より、文字画像の画素値, 濃度ヒストグラム, 文字の占有率検出したデータで筆跡鑑定を行うと、識別率が 50% を超えた文字は表 5 のときよりは多くなっていることが分かる。また、表 6 は、表 3 と表 4 の結果と比較すると、ほとんどの文字で識別率が下がっていることが分かる。表 6 は、中間層のニューロン数 200 のときの「あ」のように、識別率繰り返し回数 1000 回では 75% なのに回数を重ねると識別率が 8% のようにとても低くなってしまふものもあることが分かる。

6. 結論

文字画像データの画素値を検出したデータでの筆跡鑑定, 文字画像の画素値, 濃度ヒストグラム, 文字の占有率検出したデータで筆跡鑑定のよう、TDA を使用せずに行った筆跡鑑定よりも、TDA を用いた筆跡鑑定のほうが、その文字を書いた人物が誰かの正しく識別する精度が高くなったことが分かる。TDA を使用したほうが、繰り返し回数で識別率が急激に変化することがないことが分かった。

TDA を利用した筆跡鑑定において、NN と平均最近傍法と SVM は、比較をすると識別率が高くなった文字と低くなった文字の両方が同じくらい存在したため、どれが最

も良い結果が得られているかは断定できないことが分かった。また、TDA を用いた筆跡鑑定に使う文字画像データは、ほとんどの場合で図9の①のような、文字を書く枠に合わせて画像データを1文字ごとに切り出すと識別率が高くなることが分かった。これは、文字を書く枠に合わせて画像データを1文字ごとに切り出した場合、文字を書く位置も文字画像データの情報として保存され、そのうえで4等分にしてTDA を用いるからと考えられる。

今後の課題としては、表3における繰り返し回数1000回のときの「お」の識別率が低く、「せ」の識別率が高いのに、5000回になると「お」の識別率が高く、「せ」の識別率が低くなる原因を探るために、ボトルネック距離によってそれぞれのパーシステント図同士が似ているかどうか算出することと、ベクトル化したデータの分散を算出することが挙げられる。また、今回は0次元のパーシステント図を用いているため、1次元のパーシステント図を用いて筆跡鑑定を行った場合、どのような結果になるか調べることも行う予定である。他にも、ボトルネック距離でゴシック体の文字と手書き文字のパーシステント図の類似度を測った後で平均最近傍法で筆跡鑑定を行った場合の識別率がどうなるかの確認、筆跡鑑定における学習データを増やすことによる識別率の改善、[5]で行われているような複数の文字列での筆跡鑑定のTDAを用いて行うことによる識別率の変化とTDAの働きの確認が挙げられる。

謝辞 本研究を進めるにあたり、研究に使用したデータを提供していただいた友人に感謝いたします。

参考文献

- [1] 平岡裕章. タンパク質構造とトポロジー. 初版, 共立出版, 2013, 131p.
- [2] Yann LeCun, Corinna Cortes, Christopher JC Burges. "MNIST handwritten digit database, Yann LeCun, Corinna Cortes and Chris Burges". 入手先 (<http://yann.lecun.com/exdb/mnist/>) (参照 2021-11-10).
- [3] 林澤央. 位相的データ解析を用いたパターン認識の研究～手書き数字による実証～. 九州工業大学, 修士論文, 2020.
- [4] 岩頭由樹. 位相的データ解析を用いた手書き数字の認識. 九州工業大学, 卒業論文, 2021.
- [5] 高橋 真奈茄, 小出 洋. 機械学習を用いたパターン認識による筆者識別. 第57回プログラミング・シンポジウム, 2016.1.8-10.
- [6] トポロジカルデータ解析コミュニティ. "HomCloud (基礎編)", 東北大学. 2021-05-17. 入手先 ([https://www.wpi-aimr.tohoku.ac.jp/TDA/members/files/HomCloud\(基礎編\).pdf](https://www.wpi-aimr.tohoku.ac.jp/TDA/members/files/HomCloud(基礎編).pdf)), (参照 2022-01-15).
- [7] 金子真也. "コンピュータ先端ガイド2巻3章 勉強会 (SVM)", slideshare. 2017-05-23. 入手先 (<https://www.slideshare.net/MasayaKaneko/svm-76257267>) (参照 2022-01-15), p.10-11.
- [8] 小林一郎. 人工知能の基礎. 初版, 株式会社サイエンス社, 2016, 209p.
- [9] A.Zomorodian and G.Carlsson. Computing Persistent Homology. Discrete Comput. Geom. Vol.33, 2005. p249-274.
- [10] Ippei Obayashi, Yasuaki Hiraoka. Persistence Diagrams with Linear Machine Learning Models, arXiv. 2017, arXiv:1706.10082v2 [math.AT] 6 Jul 2017. 入手先 (<https://arxiv.org/pdf/1706.10082.pdf>), (参照 2022-02-03).
- [11] Chazal, F., de Silva, V., and Oudot, S. Persistence stability for geometric complexes. Geometriae Dedicata, 173(1), 2014a, p193-214.
- [12] Chazal, F., Glisse, M., Labruere, C., and Michel, B. Convergence rates for persistence diagram estimation in topological data analysis. In Proceedings of the 31st International Conference on Machine Learning (ICML14), 2014b. pp. 163-171.
- [13] Cohen-Steiner, D., Edelsbrunner, H., and Harer, J. Stability of persistence diagrams. Discrete Comput. Geom. 37(1): 2007. p103-120.