

少数異常データを有効活用する 部分空間法による異常検知手法

江渕 文人^{1,2,a)} 長谷川 隆徳^{2,3,b)} 村川 正宏^{1,2,c)}

受付日 2019年5月20日, 再受付日 2019年9月13日,
採録日 2019年10月2日

概要: 本稿では部分空間法において, 少数の異常データを活用して正常部分空間を生成することにより, 異常検知能力を向上させる手法を提案する. 従来の異常検知手法では, 正常データのみで学習を行うが, 現実の問題では異常データもわずかに取得できる. そこで提案手法では, 部分空間法の従来の目的関数に加えて異常データの平均射影長の最小化を導入する. このような定式化により, 異常データの分布を考慮した正常部分空間を生成できるため, 異常検知能力の向上が期待できる. また, 提案手法は平均射影長を用いて異常データの情報を与えているため, 異常データ数が極端に少ない場合においても安定した識別が期待できる. MNIST データセットを使い提案手法の有効性を評価した結果, 提案手法はわずかな異常データの利用できる条件下において既存の正常モデルや識別モデルに比べて高い異常検知性能を示し, 学習データに含まれない未知の異常パターンに対しても頑健に異常検知が行えることを明らかにした.

キーワード: 異常検知, 部分空間法, 標本特徴空間

Anomaly Detection with an Improved Subspace Method Utilizing very Little Anomalous Data

FUMITO EBUCHI^{1,2,a)} TAKANORI HASEGAWA^{2,3,b)} MASAHIRO MURAKAWA^{1,2,c)}

Received: May 20, 2019, Revised: September 13, 2019,
Accepted: October 2, 2019

Abstract: In this paper, we propose an improved subspace method utilizing very little anomalous data for anomaly detection. In the conventional anomaly detection method, the classifier is usually trained only with normal data. However, in the real world problems, we can obtain very little anomalous data. Therefore, our method introduces an objective function that minimizes the average projection length of anomaly data into the conventional objective function for the subspace method. By this formulation, a normal subspace that considers the distribution of anomaly data can be learned, resulting in improvement of anomaly detection capability. Furthermore, since the information on anomaly data is given as the average projection length, stable detection can be expected even when the number of anomaly data is extremely few. Using MNIST dataset, in order to evaluate the effectiveness of the proposed method, we show that the proposed method yield high anomaly detection performance compared with the conventional normal model or classifier model under the condition very little anomalous data can be obtained. Moreover, our proposed method could robustly detect unknown anomaly patterns that are not included in the training data.

Keywords: Anomaly detection, Subspace method, Empirical Feature Space

¹ 筑波大学大学院システム情報工学研究科
Graduate School of Systems and Information Engineering,
Tsukuba University, Tsukuba, Ibaraki 305–8577, Japan

² 産業技術総合研究所
National Institute of Advanced Industrial Science and Tech-
nology (AIST), Tsukuba, Ibaraki 305–8560, Japan

³ 早稲田大学基幹理工学研究科
School of Fundamental Science and Engineering, Waseda
University, Shinjuku, Tokyo 169–8050, Japan

a) fumito.ebuchi@gmail.com

b) takanori-hasegawa@aist.go.jp

c) m.murakawa@aist.go.jp

1. はじめに

パターン認識 [1] における識別手法の 1 つに部分空間法 (SM: Subspace Method) [2], [3] があげられる。部分空間法は汎化能力が高く、カーネル法 [4] の適用により非線形問題への拡張も容易であることから様々な分野、特に画像認識 [5] において広く用いられている。部分空間法では、あるクラスの教師データを部分空間に射影したときにその射影長が最大となるような低次元部分空間を生成する。クラスの特徴をよく表現するような部分空間をクラスごとに生成し、それらの部分空間に未知のデータを射影したときにその射影長が最大となるクラスに未知のデータを分類する。このとき部分空間は対象となるクラスの教師データのみを用いることで生成できる。このため、部分空間法はクラス追加型学習や 1 つの代表クラスしか持たないような問題にも利用できる。

一方で近年、センサデータの収集コストが低下したことにより、機械学習を用いた異常検知 [6], [7] の研究が盛んに行われており、機械の故障検知などに応用されている [8], [9]。異常検知には様々なタスクがあり、部品製造における不具合発生検出 [10] のような外れ値検出、ネットワークセキュリティにおける攻撃検知 [11] やインフラ設備点検における異常打音検出 [9] などの分布外検出があげられる。現実で得られるセンサデータからは、膨大な正常データを得ることができる一方で異常データはほとんど観測されない。そのため、機械学習を用いた多くの異常検知手法では、正常データのみを用いて正常状態を学習し、その正常状態からの逸脱度を基に異常を検知する [12], [13], [14]。この問題設定は 1 つの代表クラスしか持たないような問題設定であるため、部分空間法が利用できる [15]。ここで重要なのは、実用的な課題では、正常データだけでなくわずかな異常データも得られることである。したがって、このような貴重な少数の異常データの情報を有効活用することができれば、正常データのみで学習した部分空間モデルよりも異常検知能力を向上できると考えられる。

そこで、本稿では少数異常データを有効活用する部分空間法による異常検知手法と非線形問題への拡張方法を提案する。従来の部分空間法の目的関数は正常データの平均射影長最大化であるが、提案手法では通常の部分空間法の目的関数に加えて、異常データの平均射影長最小化を導入する。提案手法の最適化問題を解くことで、入力次元数の大きさの固有値問題に帰着し、この固有値問題を解くことで正常クラス部分空間の基底ベクトルを得る。この基底ベクトルが張る部分空間に未知のデータを射影したときの射影長に基づき正常、または、異常を判定する。このような定式化により、異常データの分布を考慮した正常クラス部分空間を生成できるため、異常検知能力の向上が期待できる。平均射影長により異常データの情報を与えているため、異

常データが極端に少ない場合においても異常データの情報を有効活用でき、安定した異常検知が期待できる。

さらに、上記手法は入力空間においては部分空間の基底ベクトルを得られるが、カーネル法への拡張を行う場合、通常の部分空間法とは異なり、そのままでは解を得ることができない。これは、カーネル法を用いた非線形問題への拡張は、一般的に RBF カーネルなどが利用されるが、これにより定義される特徴ベクトルは無次元であり、通常の部分空間法では式変形することで有限次元の問題に帰着させて解くことが可能であるが、上記手法ではそれが困難である。そこで、Xiong らにより提案されている標本特徴空間 [16] を導入することでこの問題を解決する。標本特徴空間はカーネル行列の固有値問題を解くことで得られる特徴空間であり、高次元特徴空間とカーネルの値が同じになる、すなわち、高次元特徴空間と等価な有限次元の特徴空間である。このため、標本特徴空間へ写像することにより、有限次元の特徴ベクトルを得ることができ、線形分離性の高い特徴空間でカーネル法に拡張した上記手法を解くことができる。

以下では、2 章で提案手法に関連する部分空間法と標本特徴空間について述べ、3 章で少数異常データを有効活用する部分空間法を提案する。4 章で MNIST データセットを用いた比較実験を行い提案手法の有効性を示し、5 章でまとめを述べる。

2. 関連研究

2.1 正常データのみを学習する部分空間法

部分空間法は対象のクラスの教師データの射影長の二乗和が最大となる部分空間を求める手法である。異常検知において正常クラスを表す部分空間の基底ベクトルは以下の最適化問題を解くことで得られる。

$$\begin{aligned} \max \quad & L = \frac{1}{|S_+|} \sum_{i \in S_+} (\mathbf{x}_i^\top \mathbf{v})^2 \\ \text{s.t.} \quad & \mathbf{v}^\top \mathbf{v} = 1 \end{aligned} \quad (1)$$

ここで、 S_+ は正常データの添字集合であり、 $|S_+|$ はその大きさである。また、 \mathbf{x}_i は m 次元入力データ、 \mathbf{v} は m 次元ベクトルである。式 (1) にラグランジュ定数 λ を導入すると、以下の最適化問題となる。

$$\max \quad L = \frac{1}{|S_+|} \sum_{i \in S_+} (\mathbf{x}_i^\top \mathbf{v})^2 - \lambda (\mathbf{v}^\top \mathbf{v} - 1) \quad (2)$$

式 (2) の最適化問題の \mathbf{v} に対する最適条件は次式の固有値問題となる。

$$\frac{1}{|S_+|} \sum_{i \in S_+} \mathbf{x}_i \mathbf{x}_i^\top \mathbf{v} = \lambda \mathbf{v} \quad (3)$$

次に、 m 次元入力データ \mathbf{x} を l 次元空間 ($m < l$) に写像

する写像関数 $\phi(\mathbf{x})$ を考える。ここで、 $\phi(\mathbf{x})$ はカーネル関数 $K(\mathbf{x}, \mathbf{x}') = \phi^\top(\mathbf{x})\phi(\mathbf{x}')$ を満たす関数とする [4]。式 (3) の \mathbf{x}_i を $\phi(\mathbf{x}_i)$ に置き換えると次式となる。

$$\frac{1}{|S_+|} \sum_{i \in S_+} \phi(\mathbf{x}_i)\phi^\top(\mathbf{x}_i)\mathbf{v} = \lambda\mathbf{v} \quad (4)$$

ここで、高次元特徴空間の正常データを並べた行列 $G = (\phi(\mathbf{x}_1) \dots \phi(\mathbf{x}_{|S_+|}))$ を用いると、次式のように書き換えることができる [18]。

$$\begin{aligned} \frac{1}{|S_+|} GG^\top \mathbf{v} &= \lambda\mathbf{v} \\ \iff \frac{1}{|S_+|} G^\top GG^\top \mathbf{v} &= \lambda G^\top \mathbf{v} \\ \iff \frac{1}{|S_+|} K_+ \boldsymbol{\alpha} &= \lambda \boldsymbol{\alpha} \end{aligned} \quad (5)$$

ここで、 $\boldsymbol{\alpha} = G^\top \mathbf{v}$ 、 K_+ は正常データによるカーネル行列である。式 (5) により得られた固有ベクトル $\boldsymbol{\alpha}$ より、 \mathbf{v} を求めると次式となる。

$$\mathbf{v}_i = \frac{1}{\sqrt{\lambda_i}} \sum_{j=1}^{|S_+|} \alpha_{ij} \phi(\mathbf{x}_j) \quad (\mathbf{x}_j \in S_+) \quad (6)$$

式 (5), (6) より、正常データ数と同数の固有ベクトルを得るが、冗長な表現が含まれるため、次式で定義する寄与率を用いることで、正常クラス部分空間に必要な固有ベクトルを選択する。

$$\frac{\sum_{i=1}^p \lambda_i}{\sum_{i=1}^{|S_+|} \lambda_i} > \eta \quad (7)$$

ここで、 η は累積寄与率と呼ばれ、あらかじめ人為的に定めるパラメータである。式 (7) を満たす最小の p を求め、対応する固有ベクトルが張る部分空間を生成する。未知のデータ \mathbf{x} に対して、正常クラス部分空間への射影長 $D_+(\mathbf{x})$ を次式で定義する。

$$D_+(\mathbf{x}) = \sum_{i=1}^p \left(\frac{\mathbf{v}_i^\top \phi(\mathbf{x})}{\|\mathbf{v}_i\| \|\phi(\mathbf{x})\|} \right)^2 \quad (8)$$

これを基に正常と異常を判定できる。

2.2 標本特徴空間

一般に高次元特徴空間のベクトル $\phi(\mathbf{x})$ は RBF カーネルなどを利用するため無限次元になりうる。標本特徴空間 [16] は高次元特徴空間のデータ分布を解析するために提案された有限次元の空間である。データ数を M とする。カーネル行列は半正定値対称行列であるため、 $p = \text{rank}(K)$ とすると、固有値 λ と固有ベクトル \mathbf{v} を並べた行列 $V = (\mathbf{v}_1 \dots \mathbf{v}_p \dots \mathbf{v}_M)$ を用いて次式のように表される。

$$K = V \begin{pmatrix} \lambda_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \dots & \vdots \\ 0 & \dots & \lambda_p & \dots & 0 \\ \vdots & \dots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} V^\top \quad (9)$$

ここで、固有ベクトルの定義より、 $V^\top V = VV^\top = I_{M \times M}$ である。正の固有値を対角成分に持つ行列 $\Lambda = \text{diag}(\lambda_1 \dots \lambda_p)$ と、それらの固有値に対応する固有ベクトルを並べた行列 $U = (\mathbf{v}_1 \dots \mathbf{v}_p)$ を用いると、式 (9) は以下のように書き換えられる。

$$K = U\Lambda U^\top \quad (10)$$

このとき、 $U^\top U = I_{p \times p}$ であるが、 $UU^\top \neq I_{M \times M}$ である。ここで、 $U = (\mathbf{v}_1 \dots \mathbf{v}_p) = (\mathbf{u}_1 \dots \mathbf{u}_M)^\top$ を満たす \mathbf{u}_i を考えると、教師データ \mathbf{x}_i に対するカーネルは次式となる。

$$(K(\mathbf{x}_1, \mathbf{x}_i), \dots, K(\mathbf{x}_M, \mathbf{x}_i))^\top = U\Lambda \mathbf{u}_i \quad (11)$$

このとき、 m 次元入力ベクトル \mathbf{x}_i を p 次元標本特徴空間へ写像する写像関数は次式で定義される。

$$\begin{aligned} \mathbf{h}(\mathbf{x}_i) &= \Lambda^{-\frac{1}{2}} U^\top (K(\mathbf{x}_1, \mathbf{x}_i), \dots, K(\mathbf{x}_M, \mathbf{x}_i))^\top \\ &= \Lambda^{-\frac{1}{2}} U^\top U\Lambda \mathbf{u}_i \\ &= \Lambda^{\frac{1}{2}} \mathbf{u}_i \end{aligned} \quad (12)$$

したがって、標本特徴空間で与えられるカーネル値 $K_e(\mathbf{x}_i, \mathbf{x}_j)$ 、すなわち、標本特徴空間上のベクトルの内積値は、次のようになる。

$$\begin{aligned} K_e(\mathbf{x}_i, \mathbf{x}_j) &= \mathbf{h}^\top(\mathbf{x}_i)\mathbf{h}(\mathbf{x}_j) \\ &= \mathbf{u}_i^\top \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} \mathbf{u}_j \\ &= \mathbf{u}_i^\top \Lambda \mathbf{u}_j \\ &= K(\mathbf{x}_i, \mathbf{x}_j) \end{aligned} \quad (13)$$

式 (13) より、標本特徴空間と高次元特徴空間のカーネル値が一致するため、2つの空間は等価であるといえる。この手法により得られる標本特徴空間は教師データ数と同数の次元の有限次元空間となる。

3. 少数異常データを有効活用する部分空間法

部分空間法では式 (3), (4) から分かるように正常データのみを用いて部分空間を生成しているため、異常データを考慮した正常クラス部分空間を生成することにより異常検知能力が向上しうる。そこで、目的関数に異常クラスのデータの平均射影長を最小化する条件を導入することにより実現する。式 (1) に異常クラスのデータの平均射影長を最小化する目的を加えると次式となる。

$$\begin{aligned} \max L &= \frac{1}{|S_+|} \sum_{i \in S_+} (\mathbf{x}_i^\top \mathbf{v})^2 - \frac{C}{|S_-|} \sum_{j \in S_-} (\mathbf{x}_j^\top \mathbf{v})^2 \quad (14) \\ \text{s.t. } \mathbf{v}^\top \mathbf{v} &= 1 \end{aligned}$$

ここで、 C は第一項と第二項のトレードオフを決定する人為的パラメータ、 S_- は異常データの添え字集合である。特に $C = 0$ の場合は通常の部分空間法と等価である。ラグランジュ乗数 λ を導入すると、次式の無制約最適化問題が得られる。

$$\begin{aligned} \max L &= \frac{1}{|S_+|} \sum_{i \in S_+} (\mathbf{x}_i^\top \mathbf{v})^2 - \frac{C}{|S_-|} \sum_{j \in S_-} (\mathbf{x}_j^\top \mathbf{v})^2 \\ &\quad - \lambda(\mathbf{v}^\top \mathbf{v} - 1) \quad (15) \end{aligned}$$

式 (15) の \mathbf{v} に対する最適条件は次式の固有値問題で与えられる。

$$\left(\frac{1}{|S_+|} \sum_{i \in S_+} \mathbf{x}_i \mathbf{x}_i^\top - \frac{C}{|S_-|} \sum_{j \in S_-} \mathbf{x}_j \mathbf{x}_j^\top \right) \mathbf{v} = \lambda \mathbf{v} \quad (16)$$

式 (16) の固有値問題を解くことにより、正常クラス部分空間が得られる。異常クラスのデータの情報を平均射影長により与えているため、異常データがごくわずかにしか得られていない場合でも有効に活用することができる。

本手法を高次元特徴空間に拡張する場合、通常の部分空間法と異なり、式 (5) のように式変形を行うことができないため、解が得られない。そこで、本手法に標本特徴空間を導入することでこの問題を解決する。具体的には、式 (16) の \mathbf{x} を式 (12) により得られた標本特徴空間のベクトル $\mathbf{h}(\mathbf{x})$ に置き換える。すなわち、式 (16) を以下のように書き換える。

$$\begin{aligned} (A - B)\mathbf{v} &= \lambda \mathbf{v} \quad (17) \\ A &= \frac{1}{|S_+|} \sum_{i \in S_+} \mathbf{h}(\mathbf{x}_i) \mathbf{h}^\top(\mathbf{x}_i) \\ B &= \frac{C}{|S_-|} \sum_{j \in S_-} \mathbf{h}(\mathbf{x}_j) \mathbf{h}^\top(\mathbf{x}_j) \end{aligned}$$

$\mathbf{h}(\mathbf{x})$ は有限次元の特徴ベクトルであるため、式 (17) の固有値問題の解を得ることができる。得られた固有値から式 (7) を基に正常クラス部分空間の基底ベクトルを選択する。選択した基底ベクトル \mathbf{v} を用いて、射影長は次のようにならわされる。

$$D_+(\mathbf{x}) = \sum_{i=1}^p \left(\frac{\mathbf{v}_i^\top \mathbf{h}(\mathbf{x})}{\|\mathbf{v}_i\| \|\mathbf{h}(\mathbf{x})\|} \right)^2 \quad (18)$$

未知のデータ \mathbf{x} に対して式 (18) を求め、これを基に正常と異常を判定できる。

4. 計算機実験

提案手法の有効性を検証するために MNIST データセッ

トを用いた比較実験を行った。本章では、従来手法の入力空間における部分空間法を **SM**、カーネル部分空間法を **KSM** とし、入力空間における提案手法（標本特徴空間を用いない）を **ISM**、標本特徴空間を導入した手法を **KISM** と記す。さらに、ニューラルネットワークを用いた異常検知でよく利用されるオートエンコーダの再構成誤差による異常検知モデルと比較を行う [13], [14]。オートエンコーダは正常データのみで学習される正常モデルである。また、異常データを利用した識別モデルである多層パーセプトロン、画像認識に対して有効な手法である Convolutional Neural Network (CNN) と比較を行う。本章ではオートエンコーダを **AE**、多層パーセプトロンを **MLP**、CNN を **CNN** と記す。

実験で用いた **AE** の構造は (784-500-1500-50-1500-500-784) で L2 ノルムの再構成誤差を損失関数とした。実験で用いた **MLP** の構造は (784-1000-2) で正常クラスと異常クラスの確率を出力する。**CNN** は VGG16 [19] のアーキテクチャを使用し、正常クラスと異常クラスの確率を出力する。**AE**、**MLP**、**CNN** は中間層の活性化関数に ReLU を使用した。また、バッチサイズを 128 として、Adam (学習率 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$) を用いて最適化した。**AE** では、MSE loss を損失関数とし 100 epoch 学習し、**MLP** と **CNN** では、正常データ数と異常データ数の不均衡を考慮するために、Class-balanced cross-entropy loss [20] を損失関数として 20epoch 学習した。

SM、**KSM**、**ISM**、**KISM** では、正常クラス部分空間の基底ベクトルを選択するための累積寄与率 η を $\eta = 0.99$ とした。また、**KSM** と **KISM** では、RBF カーネル ($K(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|)$) を使用し、 $\gamma = 0.005$ とした。**ISM**、**KISM** では、トレードオフパラメータとして C を設定する必要があるが、本実験では $C = 0.05$ とした。

4.1 MNIST データセットと実験条件

MNIST データセット [17] はオープンソースのデータセットの 1 つであり、教師データ数が 60,000、テストデータ数が 10,000 の 0-9 の手書き数字画像である。提案手法の有効性を評価するために、分布外検出の問題設定で次の 2 つの実験を行った。各実験において、学習に用いることができる正常データ数に対する異常データの比率を 0.1% から 10% まで変化させて学習して、全 10 クラスのテストデータに対し ROC で評価し、AUC 値を算出する。異常データはランダムサンプリングとするが、異常データのサンプリング結果により識別率が変化するため、1 つの実験条件につき 10 回試行を行いその平均 AUC により比較を行う。

実験 1：単一の正常データ分布に対する提案手法の有効性を評価する。正常データは 10 クラスのうち 1 個のクラスのデータを用い、異常データは残りの 9 個のクラスの

表 1 実験 1 の結果

Table 1 Comparison of average AUC in EXPERIMENT 1.

		0.0%	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%	
0	SM	0.973	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.996	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.937	0.974	0.984	0.989	0.992	0.992	0.995	0.995	0.995	0.996	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.999	0.998
	CNN	-	0.972	0.989	0.996	0.997	0.997	0.998	0.999	0.998	0.999	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ISM	-	0.996	0.997	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
1	SM	0.996	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.999	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.907	0.981	0.985	0.993	0.994	0.995	0.996	0.996	0.997	0.997	0.998	0.998	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.999
	CNN	-	0.992	0.99	0.994	0.998	0.997	0.998	0.997	0.999	0.999	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ISM	-	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
2	SM	0.911	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.938	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.849	0.925	0.935	0.953	0.953	0.964	0.965	0.971	0.971	0.976	0.982	0.983	0.985	0.986	0.987	0.988	0.988	0.989	0.989	
	CNN	-	0.95	0.968	0.982	0.993	0.994	0.994	0.995	0.996	0.996	0.996	0.999	0.999	0.999	0.999	1.0	1.0	1.0	1.0	1.0	1.0
	ISM	-	0.955	0.968	0.971	0.975	0.976	0.977	0.977	0.974	0.977	0.976	0.975	0.975	0.975	0.974	0.974	0.974	0.974	0.974	0.974	0.974
3	SM	0.919	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.957	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.835	0.916	0.94	0.957	0.963	0.969	0.969	0.969	0.972	0.971	0.98	0.982	0.985	0.987	0.987	0.988	0.988	0.988	0.988	
	CNN	-	0.948	0.943	0.977	0.981	0.992	0.995	0.993	0.996	0.997	0.998	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ISM	-	0.967	0.975	0.976	0.975	0.977	0.976	0.978	0.978	0.977	0.978	0.977	0.975	0.975	0.975	0.975	0.975	0.975	0.975	0.975	0.975
4	SM	0.923	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.954	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.763	0.882	0.927	0.956	0.961	0.972	0.971	0.975	0.974	0.977	0.987	0.99	0.991	0.992	0.992	0.993	0.993	0.994	0.994	
	CNN	-	0.915	0.958	0.988	0.988	0.994	0.993	0.995	0.996	0.995	0.997	0.998	0.999	0.999	0.999	0.999	1.0	1.0	1.0	1.0	
	ISM	-	0.979	0.983	0.988	0.986	0.988	0.988	0.988	0.988	0.989	0.988	0.989	0.989	0.988	0.989	0.988	0.989	0.989	0.988	0.989	
5	SM	0.923	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.965	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.707	0.826	0.879	0.894	0.924	0.933	0.94	0.954	0.945	0.952	0.975	0.982	0.981	0.986	0.987	0.988	0.989	0.99	0.99	
	CNN	-	0.949	0.966	0.986	0.992	0.994	0.996	0.995	0.997	0.998	0.998	0.999	0.999	1.0	1.0	1.0	1.0	0.999	1.0	1.0	
	ISM	-	0.958	0.971	0.976	0.979	0.981	0.983	0.982	0.982	0.982	0.983	0.983	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984	
6	SM	0.981	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.989	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.888	0.954	0.975	0.981	0.983	0.985	0.989	0.989	0.988	0.989	0.994	0.994	0.995	0.995	0.995	0.995	0.996	0.996	0.996	
	CNN	-	0.962	0.989	0.993	0.996	0.998	0.998	0.998	0.998	0.999	0.999	0.999	1.0	0.999	1.0	1.0	1.0	1.0	1.0	1.0	
	ISM	-	0.992	0.995	0.995	0.996	0.996	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	
7	SM	0.948	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.975	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.82	0.937	0.968	0.977	0.98	0.974	0.982	0.983	0.984	0.985	0.989	0.989	0.99	0.99	0.991	0.991	0.992	0.992	0.992	
	CNN	-	0.935	0.976	0.985	0.992	0.99	0.993	0.992	0.993	0.994	0.996	0.998	0.999	0.999	0.999	0.999	0.999	1.0	1.0	1.0	
	ISM	-	0.976	0.976	0.98	0.979	0.979	0.979	0.979	0.98	0.981	0.981	0.98	0.98	0.98	0.98	0.981	0.981	0.981	0.981	0.981	
8	SM	0.874	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.866	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.743	0.829	0.873	0.893	0.918	0.927	0.914	0.934	0.941	0.941	0.964	0.97	0.973	0.976	0.975	0.977	0.979	0.981	0.981	
	CNN	-	0.824	0.877	0.945	0.983	0.987	0.986	0.988	0.992	0.994	0.993	0.997	0.999	0.999	0.999	1.0	0.999	1.0	1.0	1.0	
	ISM	-	0.927	0.936	0.935	0.934	0.935	0.933	0.933	0.93	0.93	0.931	0.926	0.926	0.924	0.924	0.923	0.923	0.923	0.922	0.922	
9	SM	0.965	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.971	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	MLP	-	0.698	0.869	0.904	0.918	0.93	0.945	0.941	0.942	0.959	0.957	0.967	0.971	0.975	0.975	0.977	0.979	0.98	0.979	0.981	
	CNN	-	0.883	0.927	0.949	0.969	0.977	0.979	0.979	0.986	0.987	0.986	0.995	0.997	0.998	0.999	0.998	0.999	0.999	0.999	0.999	
	ISM	-	0.973	0.976	0.978	0.977	0.978	0.978	0.978	0.978	0.979	0.979	0.979	0.978	0.979	0.979	0.978	0.978	0.978	0.978	0.978	

データを用いてモデルを構築する。学習データのうち、正常クラスを 10 通り変化させたときの各 AUC の平均値により、SM, AE, MLP, CNN, ISM を比較する。

実験 2：正常データが単一の分布でなく、また、テストデータに未知の異常が含まれる場合の提案手法の有効性を評価する。正常クラスには、クラス ‘7’ とクラス ‘8’ をあわせたものを用い、異常クラスには、残りの 8 個のうち 1 個のクラスを用いてモデルを構築する。異常クラスを変えたときの各 AUC の平均値により、SM, AE, MLP, CNN, ISM, KISM を比較する。学習データには正常と異常データ合わせて 3 クラスのデータのみを用いるため、

残りの 7 個のクラスは学習データに含めない。本実験のテストデータにおいては、学習データに含まれる正常クラス、学習データに含まれる異常クラス、学習データに含まれない異常クラスが評価対象となるため、[実験 2-1：全テストデータに対する評価]、[実験 2-2：正常クラスと学習データに含まれる異常クラスに対する評価]、[実験 2-3：正常クラスと学習データに含まれない異常クラスの評価] の 3 つの観点からテストデータの評価を行う。

4.2 実験 1

表 1 に各手法における平均 AUC を示す。各クラスの各

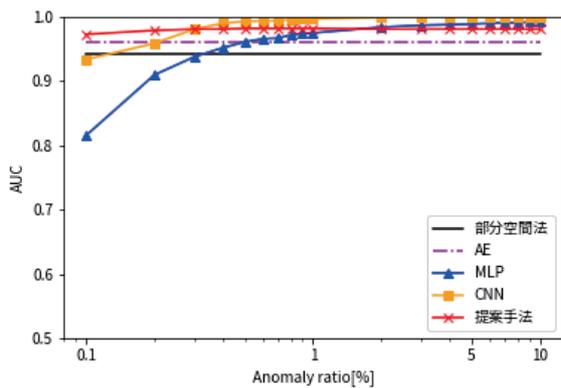


図 1 実験 1 の平均結果

Fig. 1 Result of average in EXPERIMENT 1.

表 2 異常データ率 0.1%における実験 1 の結果 요약

Table 2 Summary of average AUC for 0.1% anomaly rate in EXPERIMENT 1.

		最小値	最大値	平均値
0.1%	MLP	0.698	0.937	0.815
	CNN	0.824	0.992	0.933
	ISM	0.927	0.999	0.972

異常データ率において、平均 AUC が最も高い値を太字で記す。また、図 1 にすべての手法において、全クラスの AUC を平均した実験結果をグラフで示す。

図 1 を見て分かるように、ISM では異常データ率が小さい場合において、SM, AE, MLP, CNN よりも AUC が高く、異常データ率が高くなっても安定して異常検知ができていくことが分かる。

表 1 より、すべてのクラスにおいて ISM が SM を上回っていることが確認できる。この結果より、異常データの分布を考慮した正常クラス部分空間を生成することで、異常検知能力が向上していることが分かる。また、表 2 には、異常データ率が 0.1% の実験結果の最小値、最大値、平均値を示す。表 2 より、異常データのサンプル数が極端に少ない 0.1% の場合において、提案手法の AUC が大幅に上回っており、平均射影長により異常データの情報を与えることで少ない異常データでも有効活用できていることが確認できる。

さらに表 1 より、異常データ率が 0.1% から 0.6% の範囲においては ISM が MLP を上回っている。なお、異常データ率が 0.6% を超えると、MLP が ISM を上回るクラスが出てきて、10% では、クラス 0, 1, 6 を除いて MLP が上回っている。また、CNN と比較すると、異常データ率が 0.2% で CNN の AUC が上回るクラスが出てきて、0.8% では、すべてのクラスで CNN が上回っている。このことから、MLP, CNN ではこの問題においては異常データ数が約 50 個程度で異常クラスの情報を十分に学習できたと考えられる。

以上のことから、提案手法は単一の正常データ分布の間

題において、異常データ数が極端に少ない場合においても、それらの異常データを有効に活用して、従来手法よりも高い異常検知能力を得られることが確認できた。

4.3 実験 2

4.3.1 実験 2-1: 全テストデータに対する評価

表 3 に各手法における平均 AUC を示す。各クラスの各異常データ率において、平均 AUC が最も高い値を太字で記す。また、図 2 にすべての手法において、全クラスの AUC を平均した実験結果をグラフで示す。図 2 を見て分かるように、ISM, KISM は従来手法である SM, KSM の AUC を大きく上回っていることが確認できる。一方で、MLP, CNN と比較しても、すべての異常データ率において提案手法の AUC が大きく上回っている。これは、MLP, CNN は未知の異常データがテストデータに現れた場合に検出できないためであると考えられるため、この点については [実験 2-2], [実験 2-3] で検証する。

さらに、表 3 より、異常データ率が 1% 以下の場合において、すべてのクラスで ISM, KISM が他の手法を上回っていることが確認できる。この結果から、正常データ分布が単一でなく、また、テストデータに未知のデータが現れる場合においても提案手法は有効であることが分かる。また、ほとんどすべての場合において KISM が ISM を上回っており、カーネル法による拡張の有効性が確認できる。一方で、ISM はカーネル法を用いていないにもかかわらず、KSM, KISM と同等、または高い AUC を示す場合があり、ISM は非線形問題においても入力空間内で十分な異常検知を行える場合があることも分かる。しかしながら、表 3 より、ISM において異常データ率が上がるに従い AUC が低下していることが確認できる。これは、学習時の異常データの影響であると考えられ、パラメータ C を適切な値に設定することで影響を小さくできると考えられる。

また、表 4 には、異常データ率が 0.1% の実験結果の最小値、最大値、平均値を示す。表 4 より、異常データのサンプル数が極端に少ない 0.1% の場合において、提案手法の AUC が大幅に上回っており、平均射影長により異常データの情報を与えることで少ない異常データでも有効活用できていることが確認できる。

4.3.2 実験 2-2: 学習データに含まれる異常に対する異常検知能力

表 5 に各手法における平均 AUC を示す。各クラスの各異常データ率において、平均 AUC が最も高い値を太字で記す。また、図 3 にすべての手法において、全クラスの AUC を平均した実験結果をグラフで示す。図 3, 表 5 を見て分かるように、ISM, KISM は従来手法である SM, KSM の AUC を大きく上回っていることが確認できる。これは、提案手法は学習に含まれる異常から離れるように

表 3 実験 2-1 の結果

Table 3 Comparison of average AUC in EXPERIMENT 2-1.

		0.0%	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%	
8,7	SM	0.792	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	KSM	0.832	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	AE	0.842	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
8,7 vs 0	MLP	-	0.752	0.768	0.773	0.777	0.783	0.786	0.78	0.781	0.778	0.789	0.785	0.787	0.798	0.797	0.797	0.799	0.792	0.803	0.8	
	CNN	-	0.599	0.585	0.579	0.635	0.61	0.615	0.638	0.599	0.633	0.608	0.634	0.602	0.646	0.635	0.652	0.597	0.631	0.635	0.581	
	ISM	-	0.858	0.859	0.86	0.858	0.858	0.855	0.858	0.858	0.859	0.858	0.857	0.856	0.856	0.856	0.856	0.857	0.856	0.856	0.856	0.856
	KISM	-	0.849	0.849	0.849	0.849	0.849	0.848	0.849	0.849	0.849	0.849	0.85	0.849	0.849	0.849	0.849	0.849	0.849	0.849	0.849	0.849
	KISM	-	0.849	0.849	0.849	0.849	0.849	0.848	0.849	0.849	0.849	0.849	0.849	0.85	0.849	0.849	0.849	0.849	0.849	0.849	0.849	0.849
8,7 vs 1	MLP	-	0.52	0.514	0.525	0.549	0.539	0.552	0.563	0.567	0.538	0.568	0.579	0.594	0.599	0.597	0.615	0.621	0.615	0.625	0.627	
	CNN	-	0.644	0.667	0.636	0.641	0.684	0.665	0.7	0.657	0.718	0.705	0.741	0.74	0.772	0.775	0.793	0.771	0.8	0.769	0.828	
	ISM	-	0.839	0.843	0.843	0.845	0.843	0.844	0.845	0.844	0.843	0.844	0.845	0.846	0.845	0.845	0.845	0.845	0.845	0.845	0.845	0.845
	KISM	-	0.92	0.922	0.923	0.923	0.923	0.923	0.923	0.923	0.923	0.923	0.924	0.923	0.924	0.923	0.923	0.923	0.923	0.923	0.924	0.923
8,7 vs 2	MLP	-	0.743	0.762	0.752	0.779	0.771	0.773	0.763	0.77	0.766	0.755	0.777	0.772	0.775	0.779	0.783	0.779	0.782	0.787	0.785	
	CNN	-	0.734	0.768	0.769	0.83	0.795	0.818	0.82	0.864	0.845	0.851	0.87	0.868	0.899	0.886	0.889	0.872	0.882	0.894	0.901	
	ISM	-	0.869	0.865	0.864	0.861	0.861	0.862	0.861	0.858	0.858	0.859	0.856	0.857	0.856	0.855	0.855	0.855	0.854	0.854	0.855	
	KISM	-	0.864	0.866	0.864	0.864	0.864	0.866	0.866	0.866	0.866	0.867	0.865	0.867	0.866	0.866	0.867	0.866	0.867	0.866	0.866	
8,7 vs 3	MLP	-	0.622	0.646	0.654	0.644	0.676	0.688	0.666	0.707	0.701	0.706	0.721	0.731	0.758	0.767	0.764	0.788	0.78	0.783	0.793	
	CNN	-	0.632	0.733	0.733	0.735	0.72	0.746	0.74	0.751	0.783	0.782	0.798	0.794	0.785	0.816	0.818	0.803	0.818	0.81	0.806	
	ISM	-	0.879	0.878	0.877	0.874	0.875	0.874	0.875	0.874	0.873	0.876	0.868	0.869	0.87	0.87	0.868	0.869	0.868	0.867	0.868	
	KISM	-	0.869	0.869	0.869	0.869	0.869	0.87	0.87	0.87	0.871	0.871	0.87	0.87	0.871	0.871	0.871	0.871	0.871	0.871	0.871	
8,7 vs 4	MLP	-	0.593	0.612	0.621	0.628	0.633	0.641	0.632	0.651	0.651	0.655	0.66	0.654	0.673	0.668	0.67	0.671	0.674	0.675	0.68	
	CNN	-	0.632	0.691	0.669	0.723	0.703	0.737	0.694	0.728	0.732	0.733	0.756	0.776	0.75	0.759	0.777	0.774	0.787	0.774	0.75	
	ISM	-	0.888	0.896	0.894	0.892	0.889	0.89	0.888	0.886	0.887	0.886	0.884	0.883	0.882	0.882	0.882	0.882	0.882	0.881	0.881	
	KISM	-	0.884	0.888	0.89	0.888	0.888	0.889	0.89	0.89	0.89	0.889	0.889	0.89	0.89	0.889	0.89	0.89	0.89	0.89	0.89	
8,7 vs 5	MLP	-	0.677	0.734	0.738	0.752	0.763	0.77	0.77	0.766	0.78	0.775	0.798	0.806	0.809	0.812	0.817	0.818	0.827	0.83	0.829	
	CNN	-	0.754	0.789	0.77	0.803	0.84	0.854	0.837	0.832	0.863	0.858	0.881	0.881	0.894	0.895	0.9	0.901	0.917	0.9	0.907	
	ISM	-	0.865	0.871	0.862	0.862	0.862	0.859	0.859	0.857	0.861	0.857	0.854	0.851	0.852	0.85	0.849	0.849	0.85	0.849	0.85	
	KISM	-	0.866	0.871	0.87	0.871	0.87	0.869	0.871	0.87	0.87	0.871	0.871	0.871	0.871	0.871	0.871	0.871	0.871	0.871	0.871	
8,7 vs 6	MLP	-	0.773	0.784	0.805	0.815	0.805	0.813	0.817	0.813	0.821	0.817	0.83	0.825	0.826	0.831	0.831	0.832	0.832	0.832	0.835	
	CNN	-	0.79	0.818	0.828	0.863	0.852	0.844	0.868	0.881	0.835	0.858	0.888	0.891	0.89	0.892	0.906	0.904	0.914	0.915	0.909	
	ISM	-	0.873	0.874	0.871	0.872	0.873	0.872	0.871	0.872	0.872	0.873	0.871	0.871	0.871	0.87	0.871	0.871	0.87	0.871	0.871	
	KISM	-	0.869	0.871	0.871	0.869	0.871	0.872	0.87	0.871	0.872	0.872	0.872	0.872	0.872	0.871	0.872	0.872	0.872	0.872	0.872	
8,7 vs 9	MLP	-	0.524	0.508	0.529	0.55	0.562	0.55	0.586	0.593	0.576	0.578	0.618	0.63	0.653	0.662	0.662	0.71	0.689	0.691	0.724	
	CNN	-	0.601	0.606	0.689	0.734	0.644	0.738	0.711	0.788	0.798	0.79	0.844	0.861	0.878	0.9	0.891	0.916	0.899	0.92	0.898	
	ISM	-	0.862	0.867	0.863	0.863	0.864	0.861	0.86	0.858	0.858	0.859	0.854	0.851	0.85	0.852	0.85	0.851	0.85	0.85	0.85	
	KISM	-	0.882	0.883	0.885	0.884	0.886	0.886	0.887	0.887	0.887	0.887	0.888	0.888	0.888	0.888	0.889	0.888	0.889	0.889	0.888	

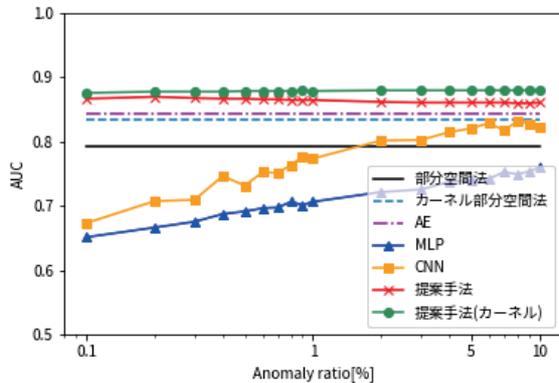


図 2 実験 2-1 の平均結果

Fig. 2 Result of average in EXPERIMENT 2-1.

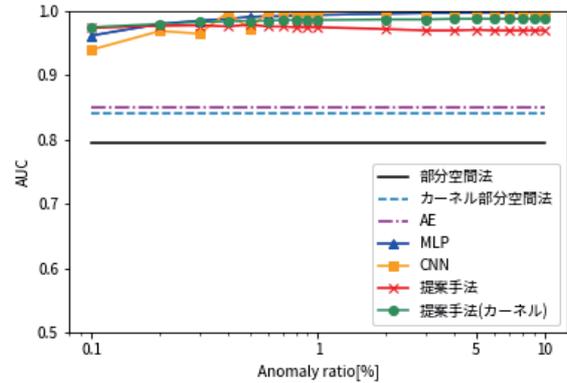


図 3 実験 2-2 の平均結果

Fig. 3 Result of average in EXPERIMENT 2-2.

表 4 異常データ率 0.1%における実験 2-1 の結果要約

Table 4 Summary of average AUC for 0.1% anomaly rate in EXPERIMENT 2-1.

		最小値	最大値	平均値
0.1%	MLP	0.520	0.773	0.651
	CNN	0.599	0.790	0.673
	KISM	0.849	0.92	0.875

正常クラス部分空間を生成するため、学習データに含まれるクラスの異常との分離度が向上したことが要因であると考えられる。さらに、KISMとMLP、CNNを比較すると、ほとんど同程度のAUCであることが確認できる。

4.3.3 実験 2-3: 学習データに含まれない異常に対する異常検知能力

表 6 に各手法における平均 AUC を示す。各クラスの各異常データ率において、平均 AUC が最も高い値を太字

表 5 実験 2-2 の結果
Table 5 Comparison of average AUC in EXPERIMENT 2-2.

		0.0%	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
8,7	SM	0.84	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.986	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.966	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.994	0.997	0.998	0.999	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
	CNN	-	0.988	0.996	0.991	0.996	0.999	0.998	1.0	1.0	0.999	0.997	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ISM	-	0.993	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994
0	KISM	-	0.995	0.996	0.996	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997
	SM	0.662	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.489	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	AE	0.461	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MLP	-	0.992	0.996	0.997	0.998	0.997	0.997	0.998	0.998	0.998	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
	CNN	-	0.91	0.919	0.966	0.994	0.904	0.997	0.999	0.975	0.999	0.998	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1	ISM	-	0.981	0.993	0.993	0.993	0.994	0.994	0.995	0.996	0.995	0.995	0.995	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996
	KISM	-	0.983	0.996	0.997	0.998	0.998	0.998	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
	SM	0.931	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8,7	KSM	0.881	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.969	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MLP	-	0.963	0.98	0.983	0.985	0.985	0.987	0.988	0.989	0.989	0.99	0.992	0.994	0.994	0.995	0.995	0.995	0.995	0.995	0.996
vs	CNN	-	0.891	0.988	0.991	0.996	0.995	0.997	0.996	0.999	0.999	0.999	0.999	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ISM	-	0.986	0.986	0.986	0.987	0.988	0.988	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.988	0.988	0.988	0.988	0.989
	KISM	-	0.985	0.986	0.986	0.987	0.988	0.988	0.989	0.989	0.989	0.989	0.989	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
8,7	SM	0.795	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.896	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.922	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.94	0.96	0.968	0.974	0.979	0.982	0.982	0.985	0.985	0.986	0.989	0.992	0.993	0.994	0.993	0.995	0.995	0.995	0.995
	CNN	-	0.895	0.995	0.977	0.996	0.999	0.999	1.0	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ISM	-	0.974	0.975	0.976	0.977	0.977	0.976	0.976	0.978	0.976	0.977	0.977	0.977	0.977	0.977	0.976	0.976	0.977	0.977	0.976
3	KISM	-	0.977	0.977	0.979	0.981	0.982	0.982	0.982	0.984	0.983	0.983	0.984	0.984	0.985	0.985	0.985	0.985	0.985	0.985	0.985
	SM	0.798	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.871	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	AE	0.876	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MLP	-	0.974	0.985	0.99	0.995	0.996	0.996	0.997	0.997	0.998	0.998	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
	CNN	-	0.935	0.975	0.949	0.999	0.955	0.999	0.998	0.999	0.999	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
4	ISM	-	0.982	0.987	0.988	0.988	0.987	0.987	0.987	0.988	0.988	0.987	0.987	0.987	0.987	0.987	0.987	0.987	0.987	0.987	0.987
	KISM	-	0.972	0.979	0.983	0.982	0.983	0.983	0.984	0.984	0.986	0.985	0.986	0.986	0.987	0.986	0.987	0.987	0.987	0.987	0.987
	SM	0.824	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8,7	KSM	0.894	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.929	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MLP	-	0.911	0.973	0.977	0.984	0.988	0.99	0.992	0.992	0.994	0.994	0.996	0.997	0.998	0.998	0.998	0.998	0.999	0.999	0.999
vs	CNN	-	0.984	0.997	0.994	0.999	0.997	0.999	0.999	1.0	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ISM	-	0.969	0.977	0.978	0.978	0.979	0.981	0.977	0.978	0.979	0.977	0.976	0.976	0.976	0.976	0.976	0.976	0.976	0.975	0.976
	KISM	-	0.973	0.982	0.985	0.985	0.986	0.987	0.988	0.988	0.987	0.987	0.988	0.989	0.99	0.99	0.99	0.99	0.99	0.99	0.99
8,7	SM	0.971	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.964	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.976	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.989	0.995	0.996	0.998	0.998	0.998	0.999	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
	CNN	-	0.995	0.997	0.996	0.999	0.999	0.997	0.999	0.999	0.999	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ISM	-	0.998	0.999	0.999	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
6	KISM	-	0.997	0.997	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998
	SM	0.542	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.738	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	AE	0.701	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MLP	-	0.928	0.948	0.963	0.965	0.974	0.975	0.977	0.978	0.977	0.979	0.985	0.988	0.99	0.991	0.991	0.994	0.994	0.994	0.995
	CNN	-	0.911	0.877	0.849	0.994	0.917	0.995	0.992	0.997	0.998	0.987	0.999	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0
9	ISM	-	0.901	0.894	0.9	0.888	0.903	0.881	0.886	0.874	0.874	0.876	0.854	0.838	0.836	0.846	0.837	0.839	0.835	0.838	0.838
	KISM	-	0.908	0.916	0.929	0.933	0.942	0.94	0.944	0.941	0.943	0.943	0.946	0.947	0.949	0.949	0.949	0.949	0.95	0.95	0.949

で記す。また、図 4 にすべての手法において、全クラスの AUC を平均した実験結果をグラフで示す。図 4 を見て分かるように、ISM, KISM は学習データに含まれない異常データに対しても、従来手法と比較して異常検知能力が高いことが分かる。これは、表 6 から分かるように、すべての実験において、ISM, KISM の AUC が従来手法の SM, KSM の AUC よりも向上している。実験から

得られた傾向であり、理論的な保証はないが、高次元の特徴空間において、データがないような領域が存在し、そのような方向に正常クラスの部分空間が生成されたことで、未知の異常クラスのデータの射影長が小さくなり、従来手法よりも異常検知能力が向上したと考えられる。なお、すべてのクラスの ISM, KISM において、異常データ率が 0.1% から 10% までの AUC がほとんど変化していない。こ

表 6 実験 2-3 の結果
Table 6 Comparison of average AUC in EXPERIMENT 2-3.

		0.0%	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
8,7	SM	0.785	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.811	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.826	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.718	0.736	0.742	0.746	0.753	0.756	0.749	0.751	0.748	0.76	0.756	0.757	0.77	0.769	0.768	0.771	0.763	0.775	0.772
	CNN	-	0.737	0.748	0.733	0.734	0.771	0.773	0.786	0.794	0.79	0.781	0.818	0.82	0.842	0.853	0.838	0.843	0.843	0.847	0.848
0	ISM	-	0.839	0.84	0.842	0.839	0.839	0.836	0.839	0.839	0.84	0.838	0.838	0.837	0.837	0.837	0.837	0.837	0.837	0.837	0.837
	KISM	-	0.828	0.828	0.829	0.828	0.829	0.828	0.829	0.828	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829
8,7	SM	0.813	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.889	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.904	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.494	0.5	0.489	0.487	0.492	0.486	0.499	0.489	0.518	0.508	0.528	0.533	0.546	0.548	0.558	0.556	0.567	0.567	0.577
	CNN	-	0.6	0.625	0.582	0.583	0.648	0.61	0.651	0.605	0.671	0.657	0.698	0.698	0.734	0.738	0.759	0.733	0.767	0.731	0.799
1	ISM	-	0.815	0.818	0.819	0.82	0.819	0.819	0.82	0.819	0.818	0.819	0.82	0.821	0.82	0.82	0.82	0.82	0.82	0.82	0.82
	KISM	-	0.909	0.91	0.91	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911
8,7	SM	0.772	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.825	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.823	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.677	0.679	0.683	0.706	0.69	0.678	0.676	0.697	0.696	0.723	0.704	0.726	0.734	0.733	0.742	0.746	0.749	0.751	0.746
	CNN	-	0.71	0.735	0.737	0.806	0.765	0.792	0.794	0.843	0.822	0.829	0.85	0.849	0.884	0.87	0.872	0.853	0.864	0.879	0.886
2	ISM	-	0.851	0.847	0.846	0.842	0.842	0.844	0.842	0.839	0.839	0.839	0.836	0.837	0.837	0.836	0.835	0.835	0.834	0.834	0.836
	KISM	-	0.846	0.848	0.845	0.846	0.846	0.848	0.848	0.847	0.849	0.846	0.849	0.848	0.848	0.848	0.848	0.848	0.847	0.848	0.848
8,7	SM	0.792	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.823	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.831	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.574	0.608	0.591	0.603	0.611	0.621	0.647	0.633	0.666	0.666	0.716	0.731	0.766	0.774	0.765	0.764	0.779	0.779	0.782
	CNN	-	0.594	0.695	0.698	0.697	0.68	0.709	0.703	0.715	0.752	0.751	0.769	0.764	0.754	0.79	0.792	0.775	0.791	0.782	0.778
3	ISM	-	0.865	0.864	0.863	0.859	0.86	0.859	0.86	0.859	0.858	0.862	0.852	0.854	0.855	0.855	0.852	0.853	0.853	0.852	0.853
	KISM	-	0.854	0.853	0.853	0.853	0.853	0.854	0.854	0.854	0.855	0.854	0.853	0.854	0.855	0.854	0.855	0.854	0.855	0.854	0.855
8,7	SM	0.791	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.827	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.836	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.562	0.592	0.606	0.587	0.582	0.616	0.612	0.58	0.611	0.595	0.627	0.658	0.669	0.678	0.676	0.686	0.685	0.681	0.706
	CNN	-	0.59	0.651	0.63	0.685	0.668	0.7	0.651	0.69	0.695	0.696	0.721	0.744	0.715	0.726	0.746	0.742	0.757	0.742	0.714
4	ISM	-	0.875	0.883	0.88	0.878	0.875	0.877	0.874	0.872	0.873	0.872	0.869	0.868	0.867	0.868	0.867	0.867	0.867	0.867	0.866
	KISM	-	0.871	0.876	0.877	0.874	0.874	0.876	0.876	0.876	0.876	0.876	0.875	0.876	0.876	0.876	0.876	0.876	0.876	0.876	0.876
8,7	SM	0.788	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.825	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.831	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.647	0.704	0.708	0.723	0.734	0.742	0.742	0.737	0.754	0.748	0.773	0.782	0.786	0.789	0.794	0.795	0.806	0.809	0.807
	CNN	-	0.725	0.763	0.742	0.778	0.82	0.836	0.817	0.812	0.846	0.84	0.866	0.867	0.881	0.882	0.887	0.889	0.906	0.887	0.896
5	ISM	-	0.851	0.858	0.848	0.848	0.847	0.844	0.844	0.842	0.847	0.842	0.838	0.836	0.836	0.834	0.833	0.833	0.834	0.834	0.834
	KISM	-	0.853	0.857	0.856	0.856	0.855	0.854	0.856	0.855	0.855	0.856	0.856	0.856	0.855	0.856	0.855	0.855	0.855	0.855	0.855
8,7	SM	0.768	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.815	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.823	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.744	0.756	0.779	0.79	0.779	0.788	0.792	0.787	0.796	0.793	0.806	0.801	0.803	0.808	0.808	0.809	0.809	0.809	0.812
	CNN	-	0.762	0.793	0.805	0.845	0.832	0.823	0.85	0.865	0.813	0.839	0.873	0.876	0.875	0.877	0.894	0.89	0.902	0.904	0.897
6	ISM	-	0.856	0.857	0.854	0.855	0.856	0.855	0.853	0.855	0.855	0.856	0.853	0.853	0.854	0.853	0.853	0.854	0.853	0.853	0.854
	KISM	-	0.851	0.854	0.853	0.852	0.854	0.854	0.852	0.854	0.855	0.855	0.855	0.855	0.855	0.854	0.855	0.855	0.855	0.855	0.855
8,7	SM	0.828	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	KSM	0.846	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AE	0.862	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
vs	MLP	-	0.466	0.444	0.466	0.49	0.503	0.488	0.529	0.538	0.518	0.521	0.565	0.578	0.605	0.614	0.614	0.669	0.645	0.648	0.685
	CNN	-	0.557	0.566	0.666	0.697	0.605	0.701	0.671	0.758	0.769	0.761	0.821	0.841	0.86	0.886	0.875	0.904	0.884	0.909	0.884
9	ISM	-	0.856	0.863	0.858	0.859	0.858	0.858	0.857	0.856	0.855	0.857	0.854	0.853	0.852	0.853	0.852	0.853	0.852	0.852	0.852
	KISM	-	0.878	0.878	0.878	0.877	0.878	0.878	0.879	0.879	0.879	0.879	0.879	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88

これは、提案手法において異常データの情報を平均射影長により与えているため、同じクラスの異常データしか与えられなければほとんど同じ正常クラス部分空間が生成されるためである。

表 5, 表 6 より, ISM, KISM どちらにおいても [実験 2-2] の AUC の値が [実験 2-3] の AUC を上回っていることから, 提案手法においては学習データ中の異常デー

タ数よりも, 含まれる異常の種類が重要であることが示唆される. たとえば, 実験 2-1 において, 学習データに含まれる異常データの種類を増やした場合, 表 5, 表 6 の混合割合により, 全体の AUC が決定されると考えられる. すなわち, 学習に含まれる異常データの種類が増えれば増えるほど, 図 3 のような結果に近づいていき, 種類が少なくても図 2 のような高い AUC が得られると推測される.

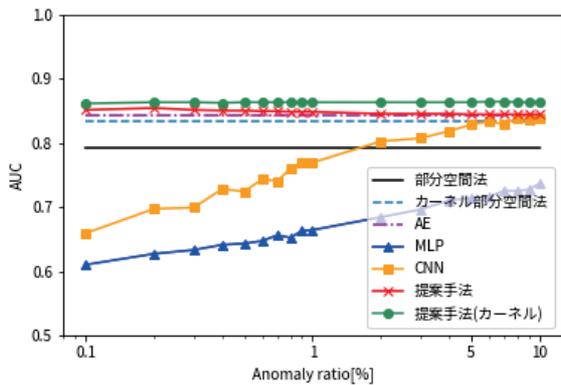


図 4 実験 2-3 の平均結果

Fig. 4 Result of average in EXPERIMENT 2-3.

以上のことから、提案手法は既知の異常データを有効活用して安定的に正常クラス部分空間を生成できているため、既知の異常データに対する異常検知能力が大幅に向上し、また、未知の異常データに対しても従来手法と比較して実験的に異常検知能力が向上することが確認できた。また、モデル構築に用いることのできる異常データ数が極端に限られた場合においても、それらの異常を有効に活用して異常検知できていることが確認できた。

5. まとめ

本稿では、少数異常データを有効活用する部分空間法と非線形問題への拡張を提案した。提案手法では正常クラス部分空間を生成する目的関数に、異常データの平均射影長最小化の目的関数を導入することにより、異常データの分布を考慮した正常クラス部分空間を生成する。また、提案手法はカーネル法への拡張が式変形により得られないため、標本特徴空間を導入することでこの問題を解決した。提案手法の有効性を検証するために、MNIST データセットを用いた 2 つの実験を行った。実験結果から、提案手法は正常データのみで学習したモデルよりも異常検知能力が高く、また、異常データ数が非常に少ない場合においてもそれらを有効活用して安定した異常検知が可能であることが確認できた。さらに、提案手法は正常クラス部分空間を生成しているため、テストデータに未知の種類異常が含まれている場合においても、それらの異常を検出することができた。これらのことから、提案手法は通常の部分空間法や AE, MLP, CNN と比較して、少数の異常データを有効活用して安定して異常検知能力を向上できる手法であり、実用上の応用範囲が広い。

参考文献

[1] Bishop, C.M.: *Pattern Recognition and Machine Learning*, Springer, New York, NY (2006).
 [2] Watanabe, S. and Pakvasa, N.: Subspace Methods of Pattern Recognition, *Proc. 1st International Joint Conference on Pattern Recognition*, pp.283-328 (1973).

[3] Oja, E.: *Subspace Methods of Pattern Recognition*, Research Studies Press, Letch-worch, UK (1983).
 [4] Schölkopf, B.: The kernel trick for distances, *Proc. Neural Information Processing Systems 13 (NIPS2000)*, pp.301-307 (2000).
 [5] Belhumeur, P.N., Hespanha, J.P. and Kriegman, D.J.: Eigenfaces vs. fisherfaces, Recognition using class specific linear projection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp.711-720 (1997).
 [6] Denning, D.E.: An intrusion-detection model, *IEEE Trans. Software Engineering*, Vol.2 (1987).
 [7] Laskov, P. et al.: Learning intrusion detection: Supervised or unsupervised?, *International Conference on Image Analysis and Processing*, Springer, Berlin, Heidelberg (2005).
 [8] Hasegawa, T., Ogata, J., Murakawa, M. and Ogawa, T.: Tandem Connectionist Anomaly Detection: Use of Faulty Vibration Signals in Feature Representation Learning, *2018 IEEE International Conference on Prognostics and Health Management (ICPHM)* (2018).
 [9] Ye, J. et al.: Statistical impact-echo analysis based on grassmann manifold learning: Its preliminary results for concrete condition assessment, *EWSHM-7th European Workshop on Structural Health Monitoring* (2014).
 [10] Moyne, J. and Iskandar, J.: Big data analytics for smart manufacturing: Case studies in semiconductor manufacturing, *Processes*, Vol.5, No.3 (2017).
 [11] Barford, P. et al.: A signal analysis of network traffic anomalies, *Proc. 2nd ACM SIGCOMM Workshop on Internet Measurement*, ACM (2002).
 [12] Wang, Y., Wong, J. and Miner, A.: Anomaly intrusion detection using one class SVM, *Proc. 5th Annual IEEE SMC Information Assurance Workshop* (2004).
 [13] An, J. and Cho, S.: Variational autoencoder based anomaly detection using reconstruction probability, *Special Lecture on IE 2* (2015).
 [14] Zhou, C. and Paffenroth, R.C.: Anomaly detection with robust deep autoencoders, *Proc. 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM (2017).
 [15] Shyu, M.-L. et al.: A novel anomaly detection scheme based on principal component classifier, *Proc. ICDM Foundation and New Direction of Data Mining Workshop* (2003).
 [16] Xiong, H., Swamy, M.N.S. and Ahmad, M.O.: Optimizing the kernel in the empirical feature space, *IEEE Trans. Neural Networks*, Vol.16, No.2, pp.460-474 (2005).
 [17] Lecun, Y., Bottou, L., Bengio, Y. and Haffner, P.: Gradient-based learning applied to document recognition, *Proc. IEEE*, Vol.86, No.11, pp.2278-2324 (1998).
 [18] Maeda, E. and Murase, H.: Multi-category classification by kernel based nonlinear subspace method, *Proc. 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol.2, IEEE (1999).
 [19] Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* (2014).
 [20] Xie, S. and Tu, Z.: Holistically-nested edge detection, *Proc. IEEE International Conference on Computer Vision* (2015).



江渕 文人

2018年富山高等専門学校エコデザイン工学専攻修了。現在、筑波大学システム情報工学研究科修士課程在学中。機械学習を用いたパターン認識、異常検知の研究に従事。



長谷川 隆徳

2018年早稲田大学基幹理工学研究科修士課程修了。現在、早稲田大学基幹理工学研究科博士課程在学中。機械学習に基づく異常検知の研究に従事。



村川 正宏

1999年東京大学大学院工学系研究科博士課程修了。博士（工学）。同年電子技術総合研究所（現、産業技術総合研究所）入所，2011年情報技術研究部門スマートシステム研究グループ長，2015年人工知能研究センター人工知能応用研究チーム長，現在，情報技術研究部門副研究部門長。筑波大学連携大学院教授，東邦大学連携大学院教授兼任。機械学習，適応信号処理の研究に従事。平成29年度産総研理事長賞（研究）受賞。