ラベル付きオープンデータセットを活用した2入力深層学習モデルによるネットワーク異常検知

吉村 尚人[1,a)]　白石 善明[1,b)]　森井 昌克[1,c)]

概要：近年，攻撃の多様化により，未知の攻撃も検知可能なアノマリ型侵入検知システム (Intrusion Detection System: IDS) の需要が高まっている．なかでも，機械学習を用いた手法は注目を集め，広く研究されている．しかし，これらの手法の多くは，検知に向けた特徴量の設計やパケットからの特徴抽出，攻撃の有無や種類を示すラベルの付与に多大な労力を要する．本稿では，あらかじめラベルが付与された多クラス通信によるオープンデータセットを活用し，パケットからの特徴自動抽出および異常検知を単一の深層学習モデルで行う手法を提案する．提案モデルは特徴抽出のための 1D Convolutional Neural Network (CNN) 部分と異常検知のための Autoencoder 部分で構成され，学習には異なる 3 種類の損失関数を用いる．MWS データセットを含む複数のデータセットを用いて提案手法の有効性を検証した結果，従来手法の課題であった特徴量設計や特徴抽出，ラベルの付与における労力を必要とせず，攻撃通信，特に C2 サーバとの通信に対して比較手法を上回る高い検知性能が確認された．

キーワード：侵入検知システム，深層学習，異常検知，特徴抽出，MWS データセット

Network Anomaly Detection Using Two-Input Deep Learning Model with Labeled Open Dataset

Naoto Yoshimura[1,a)]　Yoshiaki Shiraishi[1,b)]　Masakatu Morii[1,c)]

Abstract: In recent years, there has been a growing demand for anomaly-based Intrusion Detection Systems. Among them, methods based on machine learning have attracted much attention and have been widely studied. However, most of these methods require a great deal of effort in designing and extracting features, and assigning labels. In this paper, we propose a method that automatically extracts features from packets and detects anomalies with a single deep learning model by utilizing an open dataset of multi-class. The proposed model consists of a Convolutional Neural Network part for feature extraction and an Autoencoder part for anomaly detection, and uses three different loss functions for training. The effectiveness of the proposed method is verified using several datasets. The results show that the proposed method outperforms the comparative methods in detecting attacks, especially those with C2 servers.

Keywords: Intrusion Detection System，Deep Learning，Anomaly Detection，Feature Extraction，MWS Dataset

1. はじめに

コンピュータシステムに対する悪意のある活動を検知する仕組みとして侵入検知システム (Intrusion Detection System：IDS) があり，サイバー攻撃対策における主要技術の一つとして広く用いられている．この IDS は検知方法によって大きく二種類に分類され，事前に定義したルールに基づいて検知を行うシグネチャ型と，通常状態から外れる状態を異常として検知するアノマリ型が存在する．しかし，シグネチャ型 IDS に関しては，近年のインターネット

[1] 神戸大学
Kobe University
[a)] yoshimura@stu.kobe-u.ac.jp.ac.jp
[b)] zenmei@port.kobe-u.ac.jp
[c)] mmorii@kobe-u.ac.jp

普及やデジタル技術の発展に伴うサイバー攻撃の複雑化・巧妙化 [1] により，ルール作成の負担は非常に大きく，ルールにない攻撃が検知できないという問題も存在する．こうした背景から，未知の攻撃を検知可能なアノマリ型 IDS，なかでも，機械学習を用いたものが注目を集め，広く研究されている [2], [3], [4].

文献 [5], [6], [7], [8] では Autoencoder を用いた異常検知手法が提案されている．Autoencoder は Hinton ら [9] によって提案されたニューラルネットワークの一種で，入力データの再構成を行うように訓練される．Autoencoder は，学習に用いていないデータが適切に再構成されないことを利用し異常検知に広く用いられている．また，文献 [10], [11], [12] では画像認識の分野を中心に用いられている Convolutional Neural Network (CNN) を侵入検知に応用する試みがなされている．

しかし，これらの手法の多くでは，検知に向けた特徴量設計や，パケットからの特徴抽出，攻撃の有無や種類を示すラベルの付与に多大な労力を要する．そこで，本研究では，パケットからの特徴自動抽出および異常検知を単一の深層学習モデルで行う手法 DOC-IDS を提案する．DOC-IDS ではターゲットとなるネットワークの通信に加え，オープンデータセットからすでにラベルの付与された通信をモデルに入力することで，ラベルの付与を必要とせず，高精度なネットワーク異常検知を実現する．DOC-IDS は特徴抽出のための 1D CNN 部分と，異常検知のための Autoencoder 部分で構成され，学習には異なる 3 種類の損失関数を用いる．実験では，MWS データセットに含まれる BOS 2018 データセット [13] および CIC-IDS-2017 データセット [14] を用いて DOC-IDS の異常検知性能の評価を行う．異常検知性能を検討し，従来の異常検知手法と比較した結果，DOC-IDS は特徴量設計や特徴抽出に伴うラベル付けを行うことなく，より高い精度で通信の異常を検知可能であり，特に C2 サーバからの通信に対しては高い検知性能を有することを確認した．

2. One-Class 分類に向けた特徴抽出手法

異常や新規性の検出においては，データの入手困難性から通常データのみを用いてモデルの学習を行い，そこから外れるものを異常や新規性として検知するアプローチが一般的である．しかし，Perera ら [15] は，画像分野を対象に，目的とするドメインとは異なるドメインのラベル付きデータを用いることで，One-Class 分類に対して有用な特徴量を抽出する手法 Deep One-class Classification (DOC) を提案している．Perera らは異常検知や新規性検知の対象となるあらかじめ与えられていないクラスを alien クラスと呼称し，DOC では alien クラスとあらかじめ与えられた 1 クラスとの識別性を高めるために，2 種類の損失が用いられた．これらの誤差を用いて訓練される DOC のネットワークは重みを共有した対となる CNN (Reference Network と Secondary Network) で構成されており，各 CNN は特徴抽出を行うサブネットワーク g と，分類を行うサブネットワーク h_c に分割される．以下では DOC のさらなる詳細について述べる．

2.1 Reference Network

Reference Network は One-Class 分類のターゲットとは異なるドメインのラベル付き多クラスデータセット Reference Dataset からの入力をもとに，損失 descriptiveness loss (l_D) の計算を行う．l_D はクラス間距離の最大化を目的としており，Perera らは cross-entropy loss を用いている．

2.2 Secondary Network

Secondary Network はネットワークの構成は Reference Network と同一であるが，One-Class 分類のターゲットとなる 1 クラスの Target Dataset を入力とし，h_c の出力から compactness loss (l_C) の計算を行う．Target Dataset の例として，異常検知タスクにおける通常データがあげられる．l_C は Target Dataset の特徴空間における分散最小化を目的とし，式 1 によってミニバッチの分散を表すように計算される．ただし，n はミニバッチサイズ，k は h_c の出力サイズ，σ^2 は h_c の出力に対するミニバッチごとの分散を表す．

$$l_c = \frac{1}{nk} \sum_{i=1}^{n} \frac{n^2 \sigma_i^2}{(n-1)^2} \tag{1}$$

2.3 モデルの訓練

訓練の開始時，モデルは学習済み CNN モデルの重みによって初期化され，最後の 4 層を除く層の重みは固定される．また，訓練時は reference dataset と target dataset からそれぞれ Reference network と Secondary Network に入力が与えられ，得られた 2 種類の損失 l_D と l_C を組み合わせた損失 (式 2) によってモデル全体の学習を行う．この際，Perera らは compactness loss の重要性を表す係数 λ を 0.1 に設定している．

$$loss = l_D + \lambda l_C \tag{2}$$

2.4 特徴の抽出

モデルの学習が終了すると，g の出力を One-Class 分類のための特徴量として得る．g の出力は l_D によって Reference Dataset に含まれる異なるクラス間の差異を表現し，l_C によって対象となる 1 クラスのデータがコンパクトに分布すると期待される．

2.5 異常検知への利用

Perera らは DOC の異常検知への利用として，通常デー

タに対する学習済みサブネットワーク g の出力により訓練した One-Class 分類器を用いる方法を提案している．One-Class 分類器としては，one-class Support Vector Machines (SVM) や Support Vector Data Description (SVDD)，k-nearest neighbor などがあげられる．Perera らによって行われた，画像データセットに対する異常検知実験の結果，様々なケースにおいて高い検知精度が示された．

本研究の提案手法 DOC-IDS では，Perera らの手法を通信データに応用し，さらに Autoencoder をネットワークに組み込むことで，特徴抽出から異常検知までを単一の深層学習モデルで実現する．

3. 提案手法

3.1 DOC-IDS の概要

本章では，パケットからの特徴自動抽出および異常検知までを単一の深層学習モデルにより実現する DOC-IDS を提案する．DOC-IDS ではまず，フローのサンプリングを行う．ここでフローは，送信元/宛先の IP アドレスとポート番号およびトランスポート層プロトコルによる 5 タプルの値で分割される通信と定義する．

次にサンプリングした値を入力に DOC-IDS の訓練を行う．DOC-IDS は特徴抽出のための 1D CNN 部分と異常検知のための Autoencoder 部分で構成される (図 1)．ここで，特徴抽出に 1D CNN を用いるのは，各バイト間の関係性を捉えることができると考えたためである．訓練にはターゲットとなるネットワークの通常通信である Target Dataset と，オープンデータセットからのラベル付き多クラス通信 Reference Dataset が用いられる．この際，訓練には 3 種類の損失を利用し，それぞれの損失は多クラス通信のクラス間識別と，通常通信の特徴空間における分散最小化，オートエンコーダの再構成誤差最小化を目的とする．異常検知においては，学習済みモデルの Autoencoder 部における再構成誤差をもとに異常検知を行う．以降では各ステップの詳細を述べる．

3.2 フローのサンプリング

フローのサンプリングでは，Hwang ら [6] の手法を用いた．Hwang らは，送信元 IP アドレス，宛先 IP アドレス送信元ポート番号，宛先ポート番号，トランスポート層プロトコルによって定められる 5 タプルの値によって集約されるフローの中から，先頭 n パケットの各 l バイトのみを用いて異常検知を行う手法を提案している．これにより，処理すべきデータの大幅な削減に加え，早期での異常検知が可能になると考えられる．Hwang らは論文中で $n = 2$, $l = 80$ を推奨したが，本研究では TCP 通信における 3-way ハンドシェイク後のペイロードも考慮するために $n = 4$, $l = 80$ を採用した．

フローのサンプリング時には，バイト列で表されたパケットを 1 バイトずつ 0-255 の整数値に変換し，長さ l を超える部分は切り捨て，長さ l に満たないパケットにはゼロパディングを施す．また，ネットワークが送信元の識別情報に着目しないよう，学習用データに関しては IP アドレスおよび MAC アドレスをランダムな値に変換する処理を行う．

3.3 DOC-IDS の構成

本節では DOC-IDS の構成について述べる．DOC-IDS は別ドメインのラベル付き多クラスデータセットである Reference Dataset が入力される Reference Network とターゲットとなるネットワークの 1 クラス通信 Target Dataset が入力される Secondary Network によって構成される．これら 2 つのネットワークは重みを共有した同一の CNN を保有し，CNN はサブネットワーク g (表 1*) と h_c (表 1**) で構成されると考える．Secondary Network には CNN に加えて，異常検知を行うための Autoencoder (表 1***) がサブネットワーク g より接続されている．これより，ネットワークの各部について詳細を述べる．

3.3.1 Reference Network

Reference Network は異常として検知する対象と Target Dataset の識別性を高めることを目的に学習する．入力としては，Reference Dataset を用い，多クラス間の識別誤差 descriptiveness loss (l_D) を計算する．これにより，g の出力から通信の種類を識別するために有用な特徴表現を得られることが期待される．Reference Network での損失は Perera らと同様に h_c の出力から cross-entropy loss を計算する．

3.3.2 Secondary Network

Secondary Network は Reference Network と同一の CNN を保有し，サブネットワーク g から Autoencoder が接続される．Secondary Network の学習では Target Dataset を用いる．訓練時には 2 種類の損失関数が計算され，それぞれ，Target Dataset における h_c の出力分散最小化と，Autoencoder の再構成誤差最小化を目的とする．分散の最小化に対しては Perera の提案した compactness loss (l_C) (式 1) を用いる．この l_c は，学習に用いるミニバッチ単位で計算される．Autoencoder の再構成誤差の最小化に対しては損失 reconstruction loss （l_R） を用意し，mean squared error (MSE) を用いる．

3.3.3 訓練フェーズ

DOC-IDS ではモデルの訓練に，3 種類の損失 l_D, l_C, l_R を組み合わせた損失 (式 3) を用いる．このとき，λ_D, λ_C, λ_R は各損失の学習における重要度を表す係数であり，正の定数をとる．本研究では，$\lambda_D = 1$, $\lambda_C = 0.1$, $\lambda_R = 10$ を採用している．また，最適化アルゴリズムには stochastic gradient descent (SGD) を用い，学習率は 5×10^{-5}, weight decay は 0.0005 とする．学習を通して，Reference Network

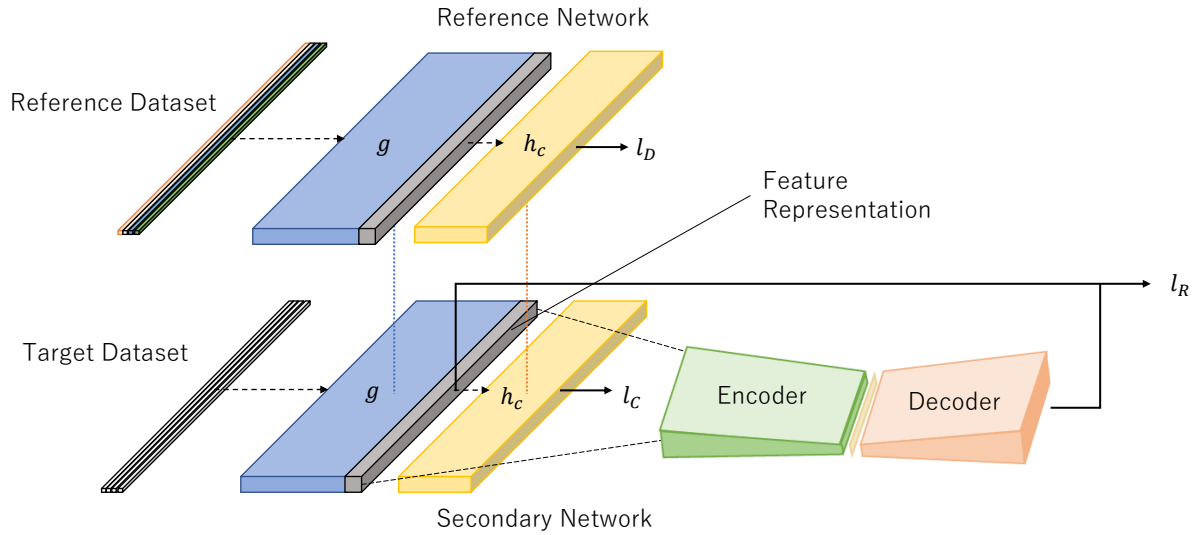


図 1　DOC-IDS の構成.
Fig. 1　Structure of DOC-IDS.

表 1　DOC-IDS のネットワーク構成.
Table 1　Structural parameters of DOC-IDS.

Layer	Type	Filters/neurons	Stride	Padding
1*	1D-Conv+Relu	32 (kernel size=6)	1	5
	+Batch Normalization			
2*	Maxpooling	kernel size=2	2	-
3*	1D-Conv+Relu	64 (kernel size=6)	1	5
	+Batch Normalization			
4*	Maxpooling	kernel size=2	2	-
5*	Dense	1024	-	-
	+Batch Normalization			
6*, ***	Dense	256	-	-
	+Batch Normalization			
7**	Dense	classes	-	-
		(the number of classes in reference dataset)		
8***	Dense	128	-	-
9***	Dense	64	-	-
10***	Dense	128	-	-
11***	Dense	256	-	-

と Secondary Network のサブネットワーク g および h_c の重みは常に共有される.

$$loss = \lambda_D l_D + \lambda_C l_C + +\lambda_R l_R \tag{3}$$

3.3.4　テストフェーズ

実際に異常検知を行う際は，学習済みのモデルから Reference Network を切り離す．そして Autoencoder による再構成誤差を異常スコアとして異常検知を行う．DOC-IDS による再構成誤差の例を図 2 に示す．このとき，Autoencoder 部分への入力となるサブネットワーク g の出力は，もとのデータと比較し，より通常通信と異常通信を分離していると考えられる (図 3) .

4.　評価実験

DOC-IDS の異常検知性能を評価するために実験を行った．実験では比較のために Autoencoder および 1D Convolutional Autoencoder による精度も調べた．以下では実

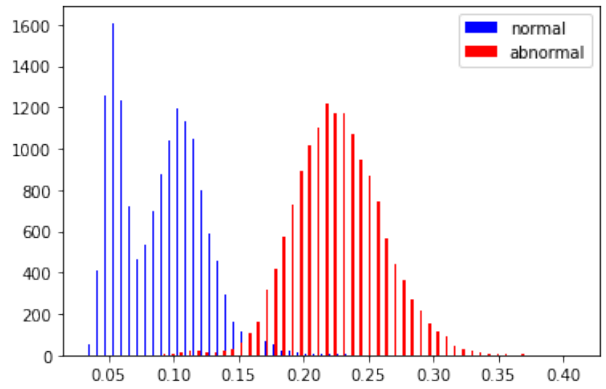


図 2　DOC-IDS による BOS 2018 データセットの再構成誤差，青のヒストグラムは通常通信に対する誤差を表し，赤のヒストグラムは C2 サーバとの通信に対する誤差を表す．

Fig. 2　DOC-IDS's Reconstruction error for BOS 2018 Dataset. The blue histogram indicates the error for benign traffic and the red one indicates the error for communication with the C2 servers.

表 2　USTC-TFC2016 データセット．

Table 2　USTC-TFC2016 Dataset.

	Class	Number		Class	Number
Normal	BitTorrent	5000	Malware	Cridex	5000
	FTP	5000		Geodo	5000
	Facetime	5000		Htbot	5000
	Gmail	5000		Miuref	5000
	MySQL	5000		Neris	5000
	Outlook	5000		Nsis-ay	5000
	SMB	5000		Shifu	5000
	Skype	5000		Tinba	5000
	Weibo	5000		Virut	5000
	WorldOfWarcraft	5000		Zeus	5000

験に用いたデータセットおよび，評価に用いる指標を示す．

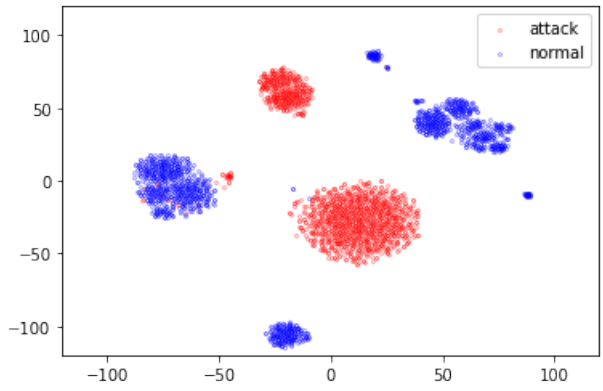
4.1　データセット

4.1.1　USTC-TFC2016

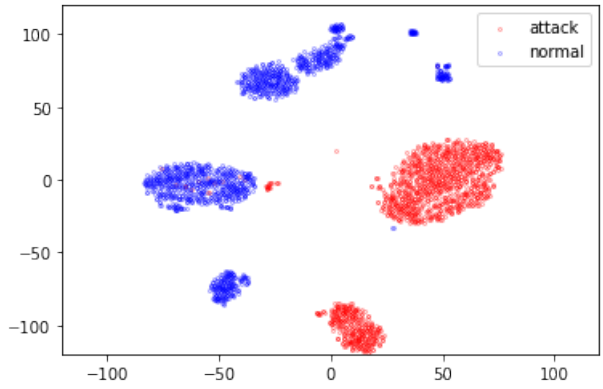
Reference Dataset として，USTC-TFC2016 [12] を用いた．USTC-TFC2016 は，Wang らによって構築され，通常通信とマルウェア通信の各 10 クラスが含まれる．本研究の対象となる生の通信データを記録した pcap ファイルに関しては，個別の通信に対して攻撃通信であるかのラベル付けが困難であることが知られている [16]．そのため，各クラスの通信が個別の pcap ファイルに分割されている USTC-TFC2016 は，ラベルの精度が担保でき，Reference Dataset として適していると考える．学習に用いたデータを表 2 に示す．

4.1.2　BOS 2018

テストデータの 1 つ目として MWS Dataset 2018 から BOS 2018[13] を利用する．BOS 2018 データセットは，標的型攻撃の観測データを記録したものであり，侵入検知の評価に広く用いられている [17], [18], [19]．今回の実験では [17], [18], [19] に倣い，C2 サーバとの通信が含まれない



(a)



(b)

図 3　BOS2018 データセットの入力ベクトルと DOC-IDS による特徴表現の可視化．DOC-IDS への入力 (320 次元) (a) とサブネットワーク g の出力 (b) を t-SNE によって可視化．青の点は通常通信，赤の点は C2 サーバとの通信を表す．

Fig. 3　Visualization of input vectors and feature representation by DOC-IDS for the BOS2018 dataset. (a) Input to DOC-IDS (320 dimensions) and (b) outputs of sub-network g (256 dimensions) are visualized by t-SNE. The blue dots indicate normal communication, and the red dots indicate communication with the C2 servers.

表 3　BOS 2018 データセット．

Table 3　BOS 2018 Dataset.

Type	Train data	Test data
通常通信	152,348	659,835
攻撃通信 (進行度 7)	-	12,051
攻撃通信 (進行度 8)	-	3,041

進行度 2 の通信ファイルを訓練に，C2 サーバとの通信が含まれる進行度 7，8 の通信ファイルをテストに用いる．この際，C2 サーバとの通信を攻撃通信とラベル付けする．実験に用いたデータを表 3 に示す．

4.1.3　CIC-IDS-2017

テストデータの 2 つ目として CIC-IDS-2017[14] を利用する．CIC-IDS-2017 データセットには，月曜日から金曜日のそれぞれに対して通信をキャプチャした pcap ファイルが用意され，月曜日以外には攻撃通信が含まれる．今回

表 4　CIC-IDS-2017 データセット.
Table 4　CIC-IDS-2017 Dataset.

Type	Attack	Train data	Test data
Benign	-	249,044	961,128
Brute Force	FTP-Patator	-	3,499
	SSH-Patator	-	2,949
DoS/DDoS	slowloris	-	3
	Slowhttptest	-	4,191
	Hulk	-	13,802
	GoldenEye	-	7,427
Heartbleed		-	1
Web Attack	Brute Force	-	791
	XSS	-	346
	SQL Injection	-	11
Infiltration		-	6
Bot		-	1,228
Port Scan		-	158,603

表 5　混同行列.
Table 5　Confusion Matrix.

		Predicted	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

の実験では月曜日の通信ファイルを訓練に用い，火曜日から金曜日のファイルで精度の検証を行った．実験に用いたデータを表 4 に示す．

4.2　評価指標

実験での評価には Receiver Operating Characteristic curve (ROC 曲線) と，Precision-Recall Curve (PR 曲線) の曲線下面積 Area Under the Curve (AUC) を用いる．ROC 曲線は横軸に false positive rate (FPR)，縦軸に true positive rate (TPR) をとり，PR 曲線は横軸に Recall, 縦軸に Precision をとる．各指標については表 5 をもとに以下の式で表される．

$$TPR(Recall) = \frac{TP}{TP + FN} \tag{4}$$

$$FPR = \frac{FP}{FP + TN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

4.3　結果と考察

BOS 2018 に対する結果を図 4 に示す．図 4(a) から DOC-IDS は他の手法と比較し，高い精度で通常通信と異常通信を識別できていることが分かる．また図 4(b) から，見逃しの少ない条件でも他の手法と比較し誤検知を抑制できている．このことから DOC-IDS では，通常通信のみの学習から C2 サーバの通信を異常として高精度で検知できると考えられる．

同じく，CIC-IDS-2017 に対する結果を図 5 に示す．こちらのデータセットも BOS 2018 データセット同様に他の手法と比較し，高い精度で異常が検知できていることが分かる．ただし，攻撃タイプごとの DOC-IDS による異常スコアを可視化したところ (図 6)，DoS/DDoS，Bot，Port Scan を除く攻撃タイプと通常通信の差異は小さく，これらを検知するためのモデル改良が今後の課題となる．

5.　まとめ

本稿では，ターゲットとなるネットワークの通信に加え，オープンデータセットからあらかじめラベルの付与された通信をモデルに入力することで，ラベルの付与を必要とせず，特徴抽出から異常検知までを可能とする手法 DOC-IDS を提案した．本稿の貢献を以下にまとめる．

- 画像分野で One-Class 分類に向けた特徴抽出法として，高い精度での異常検知に貢献した手法を通信データに応用し，有用な特徴表現の獲得方法を示した．また，これにより，生の通信データから自動での特徴抽出が可能となり，特徴量設計や特徴抽出における負担を軽減した．
- 特徴抽出と異常検知を単一の深層学習モデルによって実現した．これにより，特徴抽出のためのネットワークと異常検知のためのネットワークの同時学習を可能とした．
- 実験により，従来の異常検知手法と比較し，DOC-IDS が高い精度で通信の異常を検知可能であることを確認した．特に，C2 サーバからの通信に対しては高い検知性能を有することを示した．

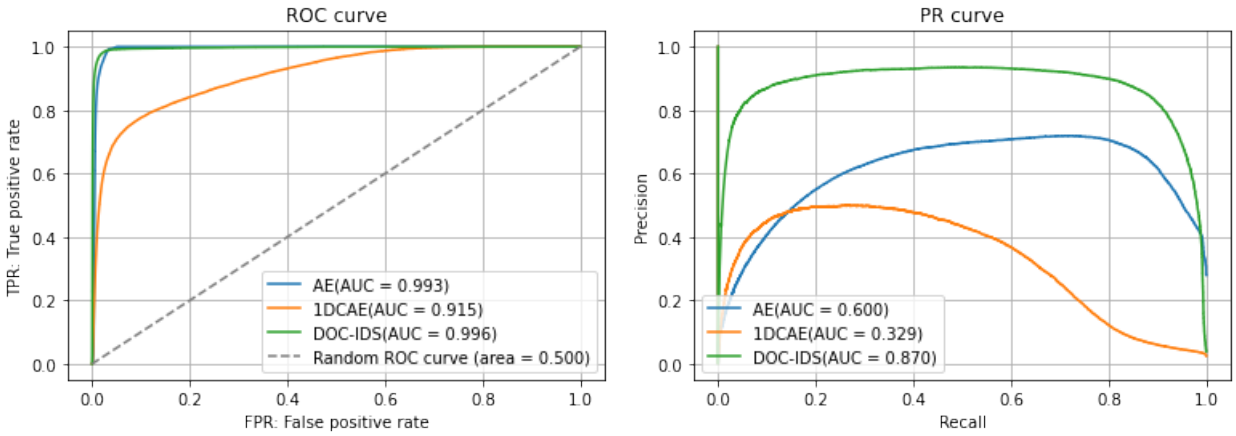
今後の課題としては，より多様な攻撃に対する検知を可能とすることがあげられる．

謝辞　研究を遂行するにあたり，協力いただいたコアマイクロシステムズ (株) 代表取締役 高橋晶三氏をはじめ，関係各位に謝意を示す．

本研究の一部は JSPS 科研費 20K11810 の助成を受けたものである．

参考文献

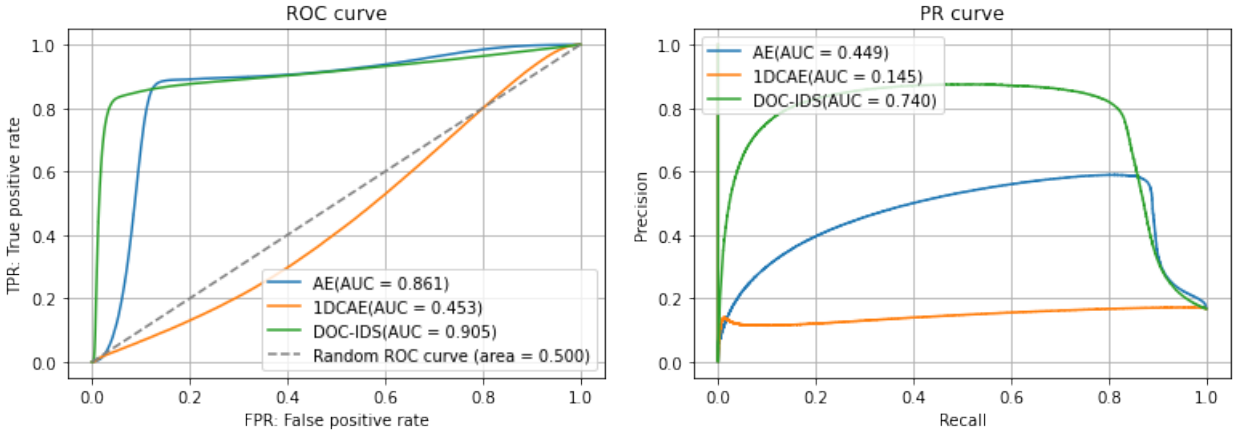
[1]　総務省, 令和 3 年版情報通信白書. 日経印刷株式会社, 2021.
[2]　R. Chalapathy and S. Chawla, “Deep learning for anomaly detection: A survey,” arXiv preprint arXiv:1901.03407, 2019.
[3]　M. Ahmed, A. Naser Mahmood, and J. Hu, “A survey of network anomaly detection techniques,” Journal of Network and Computer Applications, vol.60, pp.19–31,



(a) (b)

図 4　BOS 2018 に対する評価結果．(a)ROC 曲線．(b)PR 曲線．

Fig. 4　Evaluation results for the BOS 2018. (a)ROC curve. (b)PR curve.



(a) (b)

図 5　CIC-IDS-2017 に対する評価結果．(a)ROC 曲線．(b)PR 曲線．

Fig. 5　Evaluation results for the CIC-IDS-2017. (a)ROC curve. (b)PR curve.

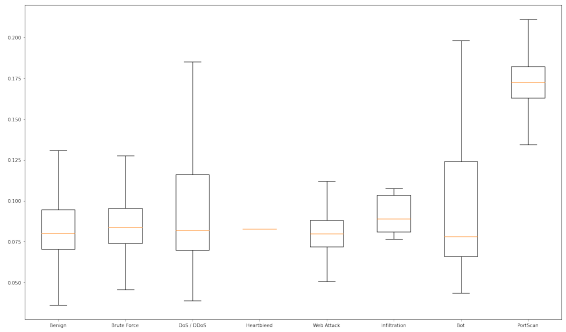


図 6　CIC-IDS-2017 に対する DOC-IDS の異常スコア．

Fig. 6　Abnormal score of DOC-IDS against CIC-IDS-2017.

2016, doi: 10.1016/j.jnca.2015.11.016.

[4] D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, “A survey of deep learning-based network anomaly detection,” Cluster Computing, vol.22, pp.949–961, 2019, doi: 10.1007/s10586-017-1117-8.

[5] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, “Kitsune: An ensemble of autoencoders for online network intrusion detection,” arXiv preprint arXiv:1802.09089, 2018, doi: 10.14722/ndss.2018.23204.

[6] R. H. Hwang, M. C. Peng, C. W. Huang, P. C. Lin, and V. L. Nguyen, “An Unsupervised Deep Learning Model for Early Network Traffic Anomaly Detection,” IEEE Access, vol.8, pp.30387–30399, 2020, doi: 10.1109/ACCESS.2020.2973023.

[7] S. Zavrak and M. Iskefiyeli, “Anomaly-Based Intrusion Detection from Network Flow Features Using Variational Autoencoder,” IEEE Access, vol.8, pp.108346–108358, 2020, doi: 10.1109/ACCESS.2020.3001350.

[8] Y. Yu, J. Long, and Z. Cai, “Network Intrusion Detection through Stacking Dilated Convolutional Autoencoders,” Security and Communication Networks, 2017, doi: 10.1155/2017/4184196.

[9] G. E. Hinton and R. R. Salakhutdinov, “Reducing the Dimensionality of Data with Neural Networks,” Science, vol.313, no.5786, pp.504–507, 2006, doi: 10.1126/science.1127647.

[10] R. Vinayakumar, K. P. Soman and P. Poornachandran, ”Applying convolutional neural network for network intrusion detection,” 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1222-1228, 2017, doi:

10.1109/ICACCI.2017.8126009.
[11] L. Yu et al., “PBCNN: Packet Bytes-based Convolutional Neural Network for Network Intrusion Detection,” Computer Networks, vol.194, pp.108117, 2021, doi: 10.1016/j.comnet.2021.108117.
[12] W. Wang, M. Zhu, X. Zeng, X. Ye, and Y. Sheng, “Malware traffic classification using convolutional neural network for representation learning,” International Conference on Information Networking, pp.712–717, 2017, doi: 10.1109/ICOIN.2017.7899588.
[13] 高田雄太，寺田真敏，松木隆宏，笠間貴弘，荒木粧子，畑田充弘, “マルウェア対策のための研究用データセット～MWS Datasets 2018～,” 研究報告コンピュータセキュリティ（CSEC），vol. 2018-CSEC-82, no.38, pp.1–8, 2018.
[14] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, “Toward generating a new intrusion detection dataset and intrusion traffic characterization,” ICISSP 2018 - Proceedings of the 4th International Conference on Information Systems Security and Privacy, pp.108–116, 2018, doi: 10.5220/0006639801080116.
[15] P. Perera and V. M. Patel, “Learning Deep Features for One-Class Classification,” IEEE Transactions on Image Processing, vol.28, no.11, pp.5450–5463, 2019, doi: 10.1109/TIP.2019.2917862.
[16] 高原尚志, “ネットワーク型侵入検知システム評価用セッション型データセット に関する一考察,” コンピュータセキュリティシンポジウム 2018 論文集, val.2018, no.2, pp.159–164, 2018.
[17] 中久木達哉, 青木茂樹, 宮本貴朗, “マルウェア感染後の HTTP 通信の特徴に基づく異常検知,” コンピュータセキュリティシンポジウム 2020 論文集, pp. 1017–1024, 2020.
[18] 二瓶凌輔, 青木茂樹, 宮本貴朗, “自己組織化マップを用いたネットワークトラフィックのエントロピーに基づく異常検出,” コンピュータセキュリティシンポジウム 2020 論文集, pp. 1025–1032, 2020.
[19] 鷺坂典雅, 青木茂樹, 宮本貴朗, “トラフィックデータのエントロピーに基づく アノマリ型 IDS の構築,” コンピュータセキュリティシンポジウム 2020 論文集, pp. 1033–1039, 2020.