PaDiMを用いたネットワーク異常検知

劉 耀楚[1,a)]　青木 茂樹[1,b)]　宮本 貴朗[1]

概要：近年，インターネットの普及に伴ってサイバー攻撃の脅威が増大しており，対策としてネットワークの異常検知技術が注目されている．異常検知とは，正常データのパターンから逸脱したデータを異常として検知する手法であり，近年は深層学習を用いたアプローチが注目されている．特に画像処理分野では，事前学習された画像認識モデルを利用することで，学習コストを低減した異常検知モデルが多数提案されている．本研究では，事前学習モデルを用いた異常検知手法の一つである PaDiM(Patch Distribution Modeling) をサイバー攻撃検知に応用した，低コストで高精度な異常検知手法を提案する．まず，トラフィックデータ中のパケットを 1 バイト毎にバイナリデータとして読み出して画素値とし，画像に変換する．次に，正常なトラフィックデータの画像を PaDiM に入力して正常な特徴分布を学習し，分布から逸脱するデータを異常として検知する．実験では CICIDS2017 データセットを用いて，本手法の有効性を確認し，少量の正常データのみで効果的な学習が可能であることを確認した．

キーワード：ネットワークの異常検知，PaDiM

Network Anomaly Detection Using Patch Distribution Modeling

Yaochu Liu[1,a)]　Shigeki Aoki[1,b)]　Takao Miyamoto[1]

Abstract: Recently, network anomaly detection has become increasingly important due to the growing threat of cyber-attacks alongside the development of the Internet. Anomaly detection is a method for identifying data that deviate from normal patterns. Particularly in the field of image recognition, many anomaly detection methods utilize pre-trained models to reduce training costs. In this study, we apply Patch Distribution Modeling (PaDiM), an anomaly detection method besed on pre-trained models, to the detection of cyber-attacks in a cost-effective and high-performance manner. First, we extract packet data from network traffic and convert them into images. We then obtain the feature distribution of normal data using PaDiM. When test data are input into the model, data that deviate from the learned distribution are identified as anomalies. In our experiment using CICIDS2017 dataset, we confirmed the effectiveness of the proposed method and demonstrated its potential for effective learning from limited training data.

Keywords: Network Anomaly Detection, PaDiM

1. はじめに

近年，インターネットの普及に伴ってサイバー攻撃の被害が拡大しており，対策のために，異常検知技術を用いた侵入検知システム (IDS) の重要性が高まっている．IDS はネットワークを監視し，様々な異常を検知するシステムであり，近年，正常な通信パターンを基準として登録し，観測対象の通信が基準から逸脱している場合を異常として検知する，アノマリ型 IDS に関する研究が盛んに行われている．アノマリ型 IDS には未知の攻撃を検知できる利点があるが，誤検知が多いことが問題となっていた．

近年，アノマリ型 IDS の検知性能の向上のために深層学習技術を応用する研究が盛んに行われている．文献 [1] では，パケットのペイロードに対して BERT(Bidirectional

[1] 大阪公立大学情報学研究科
Graduate School of Infomatics, Osaka Metropolitan University
[a)] sp25560b@st.omu.ac.jp
[b)] aoki@omu.ac.jp

Encoder Representations for Transformers) を適用して特徴を抽出し，VAE(Variational AutoEncoder) により異常を検知する手法を提案している．実験では，各種通信プロトコルでの検知性能を評価している．文献 [2] では，単位時間ごとに分割したトラフィックデータのパケットヘッダから抽出した特徴量を，LSTM(Long Short Term Memory) で学習し，異常を検知する手法を提案している．この手法はペイロードが暗号化された通信に対しても適用可能であり，特徴抽出の処理コストが小さいという利点がある．文献 [3] では，未知の攻撃の検知と攻撃のカテゴリを分類する手法を提案している．この手法では，ノイズ除去オートエンコーダを用いて特徴を抽出した後，深層強化学習手法である DDQN(Double Deep Q Network) を用いて攻撃を分類している．文献 [4] では，学習用トラフィックデータに対して，ネットワーク固有の情報を含まないよう前処理を行い，4 層の LSTM を重ねた Stacked LSTM モデルを用いてサイバー攻撃を検知する手法を提案している．文献 [5] では，パケットヘッダをグレースケール画像に変換し，VAE で学習することにより異常を検知している．この手法では複数の種類の攻撃に対して高い検知性能を実現している．また文献 [6] では，Transformer と動的グラフ構造学習，CNN(Convolutional Neural Network) を組み合わせた MemGT を提案し，クラウドコンピューティングシステムの異常を検知している．

一方，画像処理分野では，大規模物体認識データを用いて事前学習された画像認識モデルを利用した研究が盛んに行われている．異常検知においても，事前学習モデルを用いることで学習コストを低減した手法が多数研究されており，代表的な手法として PaDiM(Patch Distribution Modeling) [7] が提案されている．PaDiM は事前学習モデルに画像を入力した際の中間層から特徴マップを抽出し，特徴ベクトルの分布を基に異常を検知する手法であり，他の一般的な異常検知手法よりも少ない学習データで高い性能を発揮できると考えられている．

事前学習モデルを用いた IDS に関する手法 [8] では，画像に変換したパケットから事前学習モデルの一つである VGG16 によって特徴ベクトルを抽出してニューラルネットワークで学習することにより，攻撃を分類している．

本研究では，IDS に PaDiM を応用することによって，低コストで高精度な異常検知手法を提案する．まず，トラフィックデータ中のパケットを画像形式のデータに変換する．次に，正常なパケット画像を PaDiM に入力して，特徴分布を求める．その後，正常データの特徴分布から逸脱したデータを異常，すなわち攻撃として検知する．実験では，CICIDS2017 データセット [9] を用いて，本手法の有効性を確認した．

以下，2 節で関連研究について述べ，3 節で提案手法を説明する．4 節では実験とその結果に対する考察を述べ，5 節でまとめる．

2. 関連研究

本研究に関連する従来研究として，深層学習技術を用いたアノマリ型 IDS に関する文献 [1–3, 5] と，事前学習された画像認識モデルを用いる手法に関する文献 [8] について述べる．

文献 [1] では，パケットのペイロードに対して BERT を用いて構造解析を行い固定長の特徴ベクトルを生成し，データ生成モデルの一種である VAE により異常を検知する手法を提案している．正常なデータから生成した特徴ベクトルのみを VAE の学習に用いると，学習後に異常なデータを入力した際の誤差関数の値が大きくなると考えられるため，誤差関数の値を異常度として異常を検知している．実験では，Modbus，BACbet，EthernetIP，https の 4 種類のプロトコルに対しての異常検知性能を確認しており，Modbus，BACbet，EthernetIP では高い精度で異常を検知できているが，ペイロードが暗号化される https では検知できないことが問題となっていた．

文献 [2] では，トラフィックデータを単位時間で分割し，分割した区間内に含まれるパケットのヘッダから，IP アドレス，ポート番号，TCP フラグなどに関する 55 次元の特徴量を抽出し，LSTM を用いて異常を検知する手法を提案している．正常データのみで学習した LSTM に対してテストデータを入力する際，テストデータの特徴ベクトルが正常データのものと類似していない場合は正しく予測ができないと考えられる．そこで，予測値と入力された特徴ベクトルとの差分を異常度として，異常度が大きくなるデータを異常データとみなして，異常を検知している．パケットのヘッダのみから特徴量を抽出することで，ペイロードが暗号化されている通信に対しても適用でき，特徴抽出の処理コストを抑えることにも成功しているが，異常検知精度に改善の余地がある．

文献 [3] では，ノイズ除去オートエンコーダを用いて抽出した特徴に，深層強化学習手法である DDQN を適用することで，未知の攻撃を検知するだけでなく，その攻撃カテゴリを分類する手法を提案している．DoS，Probe，R2L，U2R の 4 つの攻撃カテゴリに対して，各カテゴリ内の一部の攻撃を学習には用いず，それらの攻撃を正しいカテゴリに分類する実験が行われており，既存の機械学習手法と比較して高い性能を示した．

文献 [4] では，学習用トラフィックデータから，ネットワーク固有の情報を除去し，学習したモデルを異なるネットワークにも適用可能にするデータの前処理方法を提案し，Stacked LSTM モデルを用いて分類している．この手法では，前処理において，トラフィックデータに含まれる多数の IP アドレスを，プライベートアドレスとパブリックアドレスの 2 種類にエンコードしている．この処理によって

他のネットワークにも共通する普遍的な特徴が得られる．しかし，教師あり学習手法を用いているため，未知の攻撃を検知できないという問題点がある．

文献 [5] では，トラフィックデータに含まれるパケットヘッダを画像に変換し，VAE により異常を検知する手法を提案している．この手法では，トラフィックデータに含まれるパケットを宛先 IP アドレスと送信元 IP アドレスの組み合わせごとに分類し，64 パケットごとにパケットのヘッダ部分をグレースケール値として，64×64 画素の画像に変換する．この処理を学習データとテストデータに対して行う．そして学習データを VAE で学習し，学習後に入力するテストデータの再構成誤差を異常度として異常を検知している．この手法では複数種類の攻撃に対して高い検知性能を実現しているが，学習データとして必要な画像枚数が多く，学習コストが高いという課題がある．

文献 [8] では，画像化したパケットデータを，事前学習モデルの一つである VGG16 に入力し，得られた特徴ベクトルをニューラルネットワークで学習することによりサイバー攻撃を分類している．この手法では，トラフィックデータに含まれるパケットを 3 パターンの方法で分類し，128 パケットごとにパケットの先頭から 128×3 Byte のバイナリデータの数値を読み込み，RGB の 3 つの値に割り当てることで 128×128 画素のカラー画像に変換する．分類には教師あり学習手法を用いているため，教師データに含まれる複数の攻撃に対して高い検知性能を実現しているが，未知の攻撃を検知できないという課題がある．

3. 提案手法

本稿では，PaDiM を IDS に応用した，低コストな異常検知手法を提案する．図 1 に提案手法の概要を示す．まず，学習用のトラフィックデータ中のパケットのヘッダ部を画像に変換する．そして，正常パケットから変換された画像を PaDiM に入力して，得られた特徴マップの各要素の平均と共分散行列を算出し，それぞれの特徴分布を学習する．テスト時には，テスト用のパケットを PaDiM に入力して，抽出した特徴マップの各要素と，正常パケットの特徴分布のマハラノビス距離を算出し，マハラノビス距離が閾値を超えたパケット画像を異常と検知する．

3.1 パケットの画像変換

本研究では，トラフィックデータ中のパケットのヘッダ部のみを用いて異常を検知する．ヘッダ部のみを用いることで，ペイロードが暗号化されている通信に対しても，安定してデータを抽出することができる．

まずトラフィックデータ中のパケットを，宛先 IP アドレスと送信元 IP アドレスの組み合わせ毎に分類する．次に分類したパケットのヘッダ部から，オプションを除いた部分のバイナリデータを 8bit 単位で 0〜255 の数値に変換

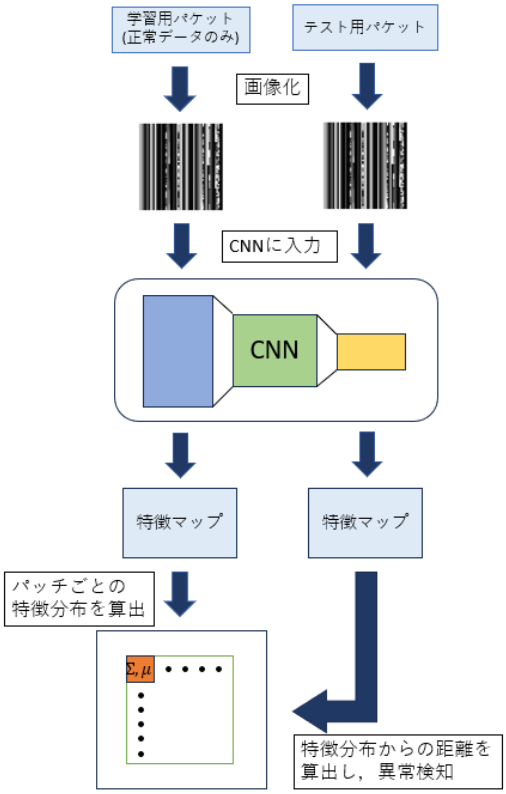

図 1 提案手法の概要
Fig. 1 Overview of the Proposed Method.

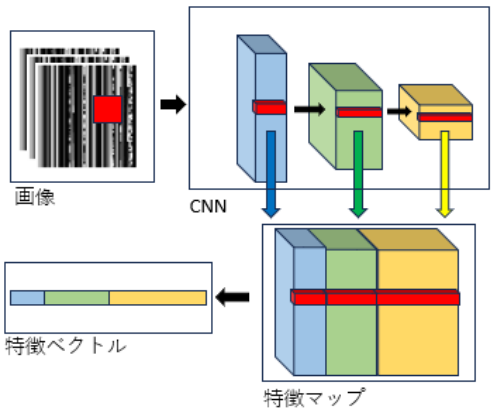

図 2 PaDiM の概要
Fig. 2 Overview of Patch Distribution Modeling (PaDiM).

する．変換した数値を 1 画素のグレースケールの値に割り当てる．その後，グレースケールの割り当てが終了した箇所から 56 バイト目まで白画素で埋める．この処理を 56 個のパケットに対して行い，56×56 画素の画像に変換する．

3.2 PaDiM による正常データの学習と異常検知

PaDiM の概要を図 2 に示す．まず，事前学習済み CNN モデルに前節で変換した画像を入力する．その後複数の中間層から特徴マップを抽出する．そして抽出した各層の特徴マップを，サイズを揃えて連結し，新たな特徴マップを生成する．生成した特徴マップ上の 1 つの要素は，元画像

表 1　CICIDS2017 データセットに含まれる攻撃
Table 1　Types of Attacks Included in the CICIDS2017 Dataset.

火曜日	FTP-Patator, SSH-Patator
水曜日	DoS slowloris, DoS Slowhttptest, DoS Hulk, DoS GoldenEye
木曜日	Brute Force, XSS
金曜日	Port Scan, DDoS LOIT

の特定の領域 (以下，パッチ) に対応している．

学習時には，学習用データを CNN モデルに入力して得られた特徴ベクトルから，平均ベクトルと共分散行列を算出する．この処理を特徴マップの全要素の特徴ベクトルに対して行うことで，正常画像の各パッチに対応する特徴ベクトルを多変量正規分布としてモデル化する．テスト時には，テスト用データから得られた特徴ベクトルと，学習時に求めた正常画像の特徴分布とのマハラノビス距離を算出する．算出したマハラノビス距離がパッチごとの異常度を表す．距離が大きいほど，正常な分布から逸脱しているため異常とする．PaDiM は事前学習されたモデルを活用した手法であるため，モデルを一から学習する必要が無く，効率的に異常を検知できる．

4. 実験

4.1 実験条件

4.1.1 実験データセット

実験には，CICIDS2017 データセットを使用した．CICIDS2017 データセットは 5 日間に渡って収集された侵入検知精度評価用のトラフィックデータセットである．学習用データとして 2017 年 7 月 3 日に収集された月曜日のデータを使用し，テストデータには 2017 年 7 月 4 日から 7 月 7 日に収集された，火曜日から金曜日のデータを用いた．月曜日の pcap 形式のデータには正常な通信データのみが含まれ，火曜日から金曜日の pcap 形式のデータには正常な通信データと各種攻撃データが含まれている．各曜日に含まれる攻撃を表 1 に示す．

本手法で用いている PaDiM では統計量 (平均および共分散行列) を用いて異常検知を行うが，これらの数値は部分集合から推定することが可能であることが知られている．したがって，学習用データの一部を用いた場合でも，全体の特徴分布を推定し，異常を検知できる可能性があると考えられる．そこで，学習データを削減した場合の異常検知性能を確認する．ここでは，全ての学習用画像を用いた場合，10%の学習用画像を用いた場合，1%の学習用画像を用いた場合の 3 パターンで実験を行った．10%，1%の学習用画像を使用する場合の実験では，データをランダムに抽出して 5 回実験し，結果を平均値で評価している．

4.1.2 PaDiM の設定

実験では，事前学習モデルとして EfficientNetV2 [10] を用いた．EfficientNetV2 は，学習時間の長さやモデルサイズの課題を克服するために設計されたモデルであり，高い精度を維持しつつ，学習コストとモデルサイズの両方を大幅に抑えることに成功している．EfficientNetV2 を用いることで，PaDiM における特徴量抽出の効率と精度向上が期待される．また用いた EfficientNetV2 の入力層のサイズが 224×224 であるため，パケット画像を同じサイズにリサイズして用いる．3.1 の手法で抽出したパケット画像は 1 辺の長さを 56 としており，224 の約数であるため，リサイズした時のデータの歪みや情報損失が抑えられると考えられる．そして，EfficientnetV2 の第 2 層，第 4 層，第 5 層を用いて特徴マップを抽出した．また，各画像のパッチごとに異常度を求め，それらの最大値を画像全体の異常スコアと設定した．

4.1.3 比較手法

提案手法の有効性を評価するために，畳み込み VAE と比較した．畳み込み VAE は学習データに類似したデータを再構成するモデルであり，異常検知によく利用されている．正常データのみを畳み込み VAE で学習すると，テスト時に異常なデータが入力された際，正確に再構成することができなくなると考えられる．そこで，テストデータを入力したときの再構成誤差を算出し，算出した値が設定した閾値を超えた場合に異常，そうでない場合に正常と判定する．本研究では，再構成誤差として入力画像と出力画像の二乗誤差を用いている．畳み込み VAE の学習の際，エポック数は 80，バッチサイズは 64，学習率は 0.001 とした．

4.1.4 評価指標

評価指標として AUC(Area Under the Curve) を用いた．AUC はモデルの分類性能を ROC(Receiver Operating Characteristic) 曲線に基づいて閾値によらず評価できる指標である．ROC 曲線は異なる閾値ごとに，モデルが異常を正しく判別した割合である真陽性率 (True Positive Rate) と正常を誤って判別した割合である偽陽性率 (False Positive Rate) をプロットした曲線である．ROC 曲線の下側の面積が AUC であり，0 から 1 の値をとり，1 に近いほど高い性能であることを示す．

4.2 実験結果

学習用画像の数を変更した実験を 5 回ずつ行った．表 2 に全ての学習用画像 (188,901 枚) を用いた場合，表 3 に 10%の学習用画像 (18,890 枚) を用いた場合，表 4 に 1%の学習用画像 (1,889 枚) を用いた場合の結果を示す．表では，各曜日の実験データに対する AUC の平均と標準偏差を示している．表中，標準偏差は括弧内に示している．

4.2.1 全ての学習用画像を用いた場合

水曜日と金曜日のデータで，PaDiM の AUC の平均値が

表 2　全ての学習用画像を用いた場合での AUC
Table 2　AUC Scores When Using All Training Images.

	VAE	PaDiM
火曜日	**0.987** (0.0051)	**0.987** (0.0007)
水曜日	0.940 (0.0270)	**0.980** (0.0005)
木曜日	**0.988** (0.0036)	0.973 (0.0007)
金曜日	0.944 (0.0275)	**0.992** (0.0030)

VAE を上回り，火曜日のデータでは同等の結果となった．

また AUC の標準偏差に注目すると，PaDiM がどの曜日のデータに対しても，VAE と比べてかなり小さい値であり，検知性能が安定していることを確認できた．

4.2.2　10%の学習用画像を用いた場合

水曜日と金曜日のデータで，PaDiM の AUC の平均値が VAE を上回り，火曜日のデータでは同等の結果となった．また VAE では全ての学習用画像を用いた場合と比べて，AUC の平均値が水曜日のデータで少し増加し，その他の曜日では変化しないか少し減少していた．PaDiM では木曜日と金曜日のデータで 0.001 のわずかな増減が見られた．

また標準偏差は VAE では水曜日と木曜日のデータ，PaDiM では金曜日以外のデータで，全画像を用いた場合より増加しているが，PaDIM の方が VAE よりも小さい値となっていた．

4.2.3　1%の学習用画像を用いた場合

水曜日と金曜日のデータで，PaDiM の AUC の平均値が VAE を上回っていた．10%の学習用画像を用いた場合と比べて，VAE では水曜日のデータを除いて，AUC の平均値が 0.002 から 0.008 の間で減少していたが，PaDiM では最大でも 0.005 程度の減少となっていた．

また標準偏差は VAE では水曜日のデータ以外で，PaDiM では全ての曜日データで増加しているが，PaDIM の方が小さく，増加量も少ない傾向にあった．

また 1%の画像を用いて学習した PaDiM は，全画像を用いて学習した VAE と比べても，水曜日と金曜日のデータで AUC の平均値が高く，また木曜日以外のデータで標準偏差が小さいことを確認できた．

4.3　考察

画像枚数を削減した場合にも，PaDiM の AUC の平均が大きく低下せず，標準偏差も大きくならなかったことから，学習データの内容に関わらず，学習した一部のデータから全データの特徴分布に近い値を学習できていることを確認できた．そのため本手法は，学習用データが少ない場合にも有効であると考えられる．しかしながら，学習用画像を 10%から 1%に削減した場合には，AUC の平均値と標準偏差の変化量が大きかったことから，学習データが少なすぎる場合には検知性能に影響を与える可能性が考えられる．今後，より多くのパターンで実験を行い，優れた検知性能を保つことのできる画像枚数を探索することが課題として挙げられる．

表 3　10%の学習用画像を用いた場合での AUC
Table 3　AUC Scores When Using 10% of Training Images.

	VAE	PaDiM
火曜日	**0.987** (0.0048)	**0.987** (0.0021)
水曜日	0.946 (0.0334)	**0.980** (0.0017)
木曜日	**0.987** (0.0079)	0.972 (0.0027)
金曜日	0.939 (0.0253)	**0.993** (0.0028)

表 4　1%の学習用画像を用いた場合での AUC
Table 4　AUC Scores When Using 1% of Training Images.

	VAE	PaDiM
火曜日	**0.985** (0.0062)	0.982 (0.0039)
水曜日	0.951 (0.0275)	**0.977** (0.0030)
木曜日	**0.979** (0.0115)	0.970 (0.0044)
金曜日	0.934 (0.0306)	**0.988** (0.0045)

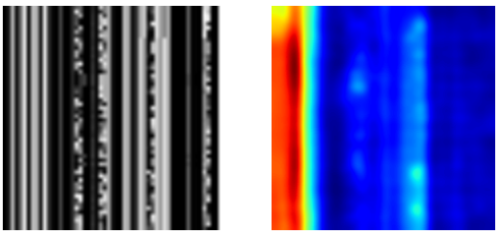

図 3　攻撃画像とその異常度の例
Fig. 3　Example of an Attack Images and Its Corresponding Anomaly Score.

VAE が，全学習用画像を用いた場合と 10%の学習用画像を用いた場合であまり結果が低下しなかったのは，元々の画像枚数が VAE に対して過剰であった可能性が考えられる．また 1%に画像枚数を削減したことで，PaDiM と比べて AUC の標準偏差が大きくなり，木曜日と金曜日のデータで平均値が大きく低下したのは，VAE では画像の特徴表現を一から学習する必要があるため，学習に用いるデータ量や，抽出したデータの内容がモデルの性能に影響を与えやすいためであると考えられる．

PaDiM と VAE で各曜日のデータごとに検知性能に差異が発生したのは，PaDiM はパッチ単位で異常度を算出し，

最大値を基に異常を検知する手法であり，部分的な差異しか持たない異常画像も検知できるためであると考えられる．水曜日と金曜日のデータには部分的な差異しか持たない異常画像が多く含まれ，他の曜日ではあまり含まれていなかった可能性が考えられる．図 3 に，PaDiM では異常度が高く算出され，VAE では異常度が低く算出された，金曜日の Portscan 攻撃の画像の一例と，この画像から PaDiM で算出された異常度を表すヒートマップを示す．ヒートマップを確認すると画像の左端の MAC アドレスを表す部分にのみ異常が現れていることが分かる．PaDiM ではこのように一部にのみ異常が現れている異常画像を検知することが出来るが，VAE は画像全体の再構成誤差を基に異常を検知するため，このような画像を検知できていない可能性がある．ただし MAC アドレスはデータセット固有の値であり，他のデータセットを用いた場合では，例のような MAC アドレスにのみ異常が現れる画像を PaDiM で検知出来ない可能性がある．従って今後は他のデータセットを用いた実験や，データセット固有の値を除いた画像化手法での実験が課題である．

実験全体を通して，VAE の AUC の標準偏差が PaDiM より大きい傾向があったのは，両モデルの学習過程に起因していると考えられる．VAE を含む深層学習モデルでは，その学習過程に重みの初期化などランダムな要素を含むため，同じ学習データを用いていてもモデルの性能が変動する．一方で，PaDiM では予め事前学習され，性能が固定されているモデルを利用しているため，これらのランダムな要素の影響を受けにくく，標準偏差が小さい値になったと考えられる．しかし，本手法で用いた EfficientNetV2 は，物体認識用のデータセットを用いて学習されており，パケット画像のようなテクスチャのみに多くの情報がある画像をあまり学習していないと考えられる．そのため，EfficientNetV2 をそのまま利用するよりも．パケット画像を用いて事前学習モデルをファインチューニングすることで，検知性能が向上する可能性が考えられる．

5. おわりに

本研究では，パケットを画像に変換し，EfficientNetV2 を用いた PaDiM に適用することで，攻撃を検知する手法を提案した．実験では，CICIDS2017 データセットを用いて，VAE との検知性能を比較し，本手法の有効性を確認した．実験の結果，大量の学習データがある場合では，本手法が他の手法と同等の検知性能と安定した標準偏差を持つことを確認できた．またデータ数を削減した場合でも安定した検知が可能であり，データ収集のコストや計算コストを抑えられることも確認できた．そのため，本手法は大規模なトラフィックデータに対しても，データを削減して低コストで異常を検知できるため，実運用中のネットワークへの適用が容易であると考えられる．今後の課題として，最適な学習用画像枚数の検討や，事前学習モデルのファインチューニング，他のデータセットや他の画像化手法を用いた実験によるさらなる有効性の確認が挙げられる．

参考文献

[1] 高橋知克, 山中友貴, 南拓也, 中嶋良彰：パケットペイロードを対象とした BERT による特徴抽出を用いた異常通信検知技術の有効性の検討，人工知能学会論文集 JSAI2023，pp.1T5GS203-1T5GS203，2023

[2] 浦川侑之介，青木茂樹，宮本貴朗：LSTM によるネットワークの異常検出，情報処理学会論文誌，vol.106，No.7，pp.1249-1254，2020

[3] 本丸真人，寺田真敏：機械学習を用いた NIDS における未知の攻撃検知手法の提案，情報処理学会論文誌，vol.62，No.12，pp.1915-1925，2021

[4] João Figueiredo, Carlos Serrão and Ana Maria de Almeida：Deep Learning Model Transposition for Network Intrusion Detection Systems, Electronics 2023, 12, 293. https://doi.org/10.3390/electronics12020293

[5] 中前諒哉，青木茂樹，宮本貴朗：VAE を用いたネットワークトラフィックの異常検出，コンピュータセキュリティシンポジウム 2021 論文集，pp.551-558，2021

[6] Huangyining Gao, Ruyue Xin, Peng Chen, Xi Li, Ning Lu and Peng You ：Memory-augment graph transformer based unsupervised detection model for identifying performance anomalies in highly-dynamic cloud environments，J Cloud Comp14, 40(2025), https://doi.org/10.1186/s13677-025-00766-5

[7] Thomas Defard, Aleksandr Setkov, Angelique Loesch and Romaric Audigier：PaDiM : a Patch Distribution Modeling Framework for Anomaly Detection and Localization, International Conference on Learning Representations (ICLR), 2020

[8] 鷺坂典雅，青木茂樹，宮本貴朗：CNN で抽出したパケットの特徴に基づくネットワークの異常検出，コンピュータセキュリティシンポジウム 2021 論文集，pp575-582，2021

[9] I.Sharafaldin, A.Habibi Lashkari and A.A.Ghorbani：Toward Genetating a New Intrusion Detection Dataset and Intrusion Traffic Characterization, 4th International Conference on Information Systems Security and Privacy (ICISSP), 2018

[10] Mingxing Tan and Quoc V. Le：EfficientNetV2 : Smaller Models and Faster Training, International Conference on Machine Learning (ICML), 2021