論文の内容の要旨

１　申 請 者

防衛大学校　ダオ　ヴァン　トゥアン

２　論文題目

軽量深層学習を用いた画像に基づく高精度マルウェア分類

３　論文の内容の要旨（博士:2,000 字程度）

近年，スマートデバイス，IoT(Internet of Things)デバイスなどでのマルウェア事案が深刻になっている．そこで本論では軽量深層学習を用いた高精度マルウェア分類に取り組んだ．デジタル時代において，悪意のある行為者は，脆弱性を悪用し，システムを破綻させる新たな方法を絶えず模索している．対策は講ぜられているが不十分な為，新たに出現するマルウェアの数は一向に減る気配がない．その理由として，少なくとも次の三つが考えられる．一つ目は，マルウェア分類技術の遅れである（理由 1)．二つ目は，複数かつ巨大なモデルにより高い精度が得られるが，低パワーデバイスに適用できない点である．一方で，軽量なモデルでは精度は不十分であるため，軽量でありながら高精度維持のモデルが求められている（理由 2)．三つ目は，機械学習を用いた教師学習ありのマルウェア分類が既知のものにしか対応できない点である．未知のマルウェア分類できる新たな方法が求められている（理由 3)．ここで「マルウェア分類」とは，検出したマルウェアの種類を予測する課題である．

（理由 1)に関しては，従来のマルウェアへの対応（シグニチャベース，動的解析，静的解析，ヒューリスティック・ベース）では，現在のマルウェアの開発速度にほとんど追いつくことができない．この問題に対処するために，迅速な対応が可能な機械学習，中でも深層学習が適用されつつある．深層学習を用いてマルウェアを分類するためには，まず，悪意のあるコードを学習可能な入力形式に変換する前処理が必要である．変換にはさまざまなアプローチがあり，アプローチごとに異なる前処理方法がある．代表的な前処理方法の一つに，悪意のあるコードを画像に変換するというものがある．さまざまな画像処理技術を活用することで，同じ系列に属するサンプル間の相関関係を見つけることができる．画像に基づくマルウェア分類には，マルウェアの内容や実際の動作に基づいた特徴量を人間が決定する必要がないという利点がある．さらに，学習済みの深層学習は，多くの悪意のあるコードを短時間で処理できる．

（理由 2）に関しては，限られたリソース，環境の中で実行可能なモデルが求められる，これを高精度かつ低負荷で実現する方法には二つの観点がある．一つはマルウェアの特徴抽出に対する工夫された処理方法であり，もう一つはモデルの軽量化である．深層学習を使用したマルウェア分類に関する現在の研究は，GPU（Graphics

Processing Unit）やクラウドコンピューティングといった環境の使用を前提としている．そこでは，畳み込みネットワーク(Convolutional Neural Network: CNN）のような大規模で複雑なモデルが構築される．しかし，IoT デバイスのようにリソースが限られた環境における利用を考えた場合には，より軽量でコンパクトなモデルが求められ，しかも，品質は保証されなければならない．本研究においては，限られたリソースとして，(I) 少数の GPU しか使用できない環境，及び (II) CPU (Central Processing Unit) しか使用できない環境の二つを考え，それぞれの環境において適切なモデルを提案する．

本研究では，まず上記（I）の少数の GPU しか使用できない環境のために，ローカル特徴及びグローバル特徴の双方を重視し，軽量化畳み込み層，変分オートエンコーダ（Variational AutoEncoder: VAE)，及び注意機構を組み合わせた，新たな CNN-AVAE モデルを提案した．提案手法は，画像ベースマルウェア分類タスクにおいて，他の CNN モデルと比べ高い精度を得た．

本研究では，次に上記（II）の CPU しか使用できない環境のために，CNN や注意機構を使わない，多層パーセプトロン（Multi-Layer Perceptron: MLP）に基づいたモデルである MLP-mixer を利用した．MLP-mixer は，通常の MLP に比べ大幅な精度の改善ができたものの，CNN モデルである Resnet50 には及ばなかった．そこで，本研究は MLP-mixer を軽量化し，オートエンコーダにより特徴を洗練させた．その結果，Resnet50 や，CNN を用いない既存研究より上回る性能が得られた．

（理由 3）に関しては，上記の２つの提案手法は高い分類性能を持つが，教師あり学習であるため未知のマルウェアを分類することは困難である．この問題を解決するために，本研究では，学習のためのサンプル数を N 個に抑えた N-shot 学習に着目した．これまで フューショット学習が多く研究され成果をあげているが，そこでは少なくとも一つのサンプルが必要となる．一方，近年提案されたゼロショット学習は，学習データとラベルの関係（マッピング手法）を活かすことで，サンプルがなくとも分割ができることから注目を集めている．とはいえ，一つのマッピングしか行わないため，関係性を捉えきれない部分がある．つまり，関係の多重性に関する考慮が欠けていると考えられる．そこで，本研究はマルウェア分類のための新たなマッピング手法を提案した．マルウェアの画像から軽量化 CNN により作成された特徴空間とマルウェアに関する記述から作成されたラベルとでマッピングを行う．マッピングを二段階に増やすことで，提案手法は，既存手法である巨大なモデルより高い精度を得た．

４　キーワード（５個程度）

「深層学習」，「軽量化」，「画像処理」，「マルウェア分類」，「セキュリティ」