GASF 이미지 인코딩 CNN 기반 데이터 효율적 봇넷 탐지: MLP·FT-Transformer와의 동일 조건 비교

김도현, 하상우, 안균승, 조성제

KCC 2026

2026.06.26

컴퓨터보안 및 OS 연구실

INDEX

01 서론

서론

❖ 네트워크 보안 환경에서는 정상 트래픽으로 위장한 악성 행위를 조기 탐지하는 기술 중요

- 주요 탐지 대상: 봇넷(Botnet)

❖ 봇넷(Botnet)

- **감염 호스트를 원격으로 제어** ⇒ 분산 서비스 거부(DDoS), 정보 유출, 스팸 전송 등 다양한 공격에 활용

❖ 실제 환경에서 봇넷 탐지가 어려운 이유

- **플로우 통계 피처간 복잡한 비선형 관계**

  ⇒ 단순 규칙 기반 탐지, 얕은 분류 모델만으로는 안정적 탐지와 레이블 확보가 어려움

서론

데이터

Tabular 네트워크 플로우 데이터

입력 표현

1차원 수치 벡터

- 입력 해석이 직관적
- 피처 간 복잡한 비선형 관계를 충분히 반영 X

2차원 GASF 이미지

- 지역 패턴과 상호관계를 잘 포착
- 변환 과정에서 표현 손실 가능

토큰 임베딩

- 구조가 복잡하고 연산 비용이 큼
- 어텐션으로 피처 간 상호작용을 직접 학습

모델

MLP | **CNN** | **FT-Transformer**

데이터셋/출처

- **NSL-KDD**
- 김태희, 강승호, "실시간 탐지를 위한 인공 신경망 기반의 네트워크 침입탐지 시스템," 2017.

- **NSL-KDD**
- M. A. Faiaz, D. Mitra, and R. D. Prangon, "Intrusion Detection Using CNN: A Color Mapping Approach on NSL-KDD Dataset," NSysS '24, ACM, 2024.

- **CIC-IoMT2024**
- "Enhanced Intrusion Detection System for IoMT Devices Using Improved Human Evolutionary Optimization Algorithm and Tabular Transformers," IEEE, 2025.

한계점

기존 연구들은 각 방식을 서로 다른 실험 환경에서 개별 평가

입력 표현 방식 자체가 성능에 미치는 영향을 공정하게 비교 불가

본 연구

동일 조건에서 1차원 수치 벡터(MLP), 2차원 GASF 이미지(CNN), 토큰 기반 어텐션(FT-Transformer) 의 3가지 표현 방식을 대조 평가하기 위한 연구

동일 조건 비교 필요 - 피처 구성, 전처리 방식, 모델 규모 등
모델 파라미터 규모를 유사하게 통제 ⇒ 입력 표현 방식의 영향만 분리 분석
학습 데이터 비율 단계적 축소를 통해 **GASF 기반 CNN의 학습 데이터 효율성(샘플 효율성)을 정량적 평가**

Attacker	Victim	Attack Name	Date	Attack Start Time	Attack Finish Time
18.219.211.138	18.217.218.111- 172.31.69.23 18.222.10.237- 172.31.69.17 18.222.86.193- 172.31.69.14 18.222.62.221- 172.31.69.12 13.59.9.106- 172.31.69.10 18.222.102.2- 172.31.69.8 18.219.212.0- 172.31.69.6 18.216.105.13- 172.31.69.26 18.219.163.126- 172.31.69.29 18.216.164.12- 172.31.69.30	Bot	Friday-02-03-2018	10:11	11:34
18.219.211.138	18.217.218.111- 172.31.69.23 18.222.10.237- 172.31.69.17 18.222.86.193- 172.31.69.14 18.222.62.221- 172.31.69.12 13.59.9.106- 172.31.69.10 18.222.102.2- 172.31.69.8 18.219.212.0- 172.31.69.6 18.216.105.13- 172.31.69.26 18.219.163.126- 172.31.69.29 18.216.164.12- 172.31.69.30	Bot	Friday-02-03-2018	14:24	15:55

[그림 1] CSE-CIC-IDS2018 봇넷 공격 정보 (Friday-02-03-2018)

출처: Canadian Institute for Cybersecurity (UNB), CSE-CIC-IDS2018

02 제안 방법

데이터셋 및 전처리

❖ 데이터셋

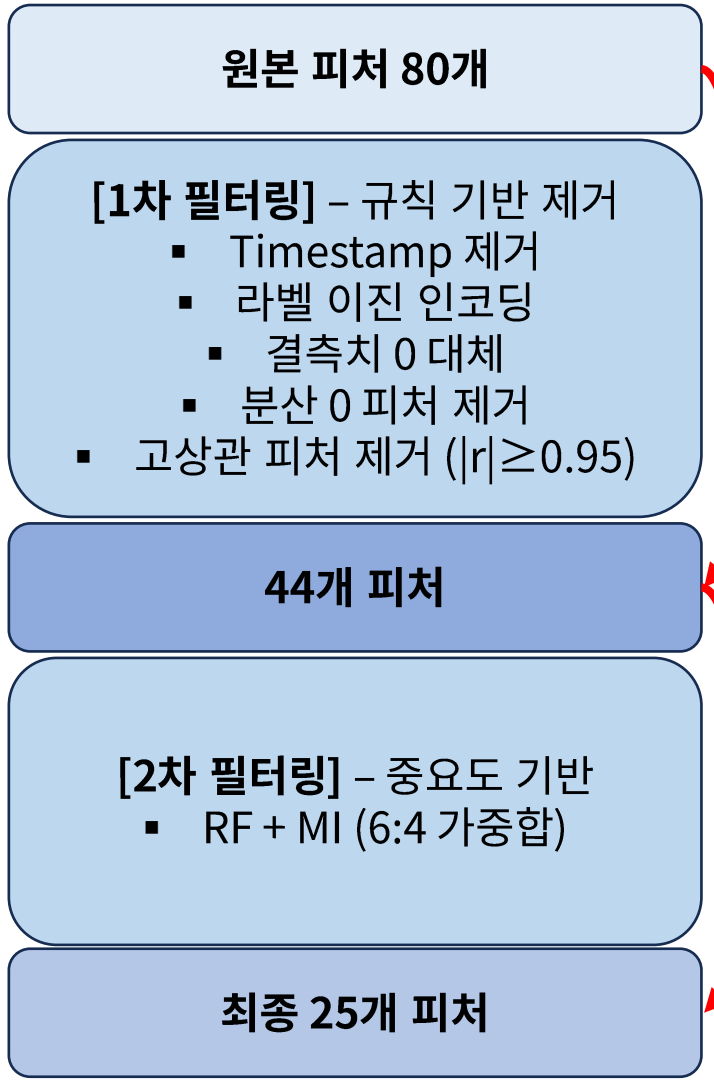
- **CSE-CIC-IDS2018 中 Friday-02-03-2018**
  - ***봇넷 공격의 플로우 메타데이터가 포함된 유일한 파일***
  - **구성:** 1,048,575개 플로우 레코드와 80개의 CICFlowMeter 기반 피처

❖ 피처 선택(feature selection)

- **1차 필터링:** Timestamp 제거, 라벨 이진 인코딩, 결측치 0 대체, 분산 0 피처 제거, 고상관 피처($|r| \geq 0.95$) 제거 ⇒ **44개 피처**
- **2차 필터링:** Random Forest 피처 중요도(feature importance) + Mutual Information 점수

✓ **데이터 누수 방지** ⇒ 학습 분할의 200,000행에서만 점수 산출

✓ 동점 시 봇넷 C&C 주기 통신 특성을 반영한 도메인 우선순위 적용 ⇒ **최종 25개 피처**



데이터 분할 및 학습 설정

❖ 데이터 분할

- 최종 선정된 25개 피처로 구성된 1,048,575개 샘플

⇒ **6:2:2(학습/검증/테스트),** 계층적(stratified) 분할로 Bot:Benign 비율(27:73) 유지

- **학습(Training) 세트: Bot:Benign = 1:1 언더샘플링**

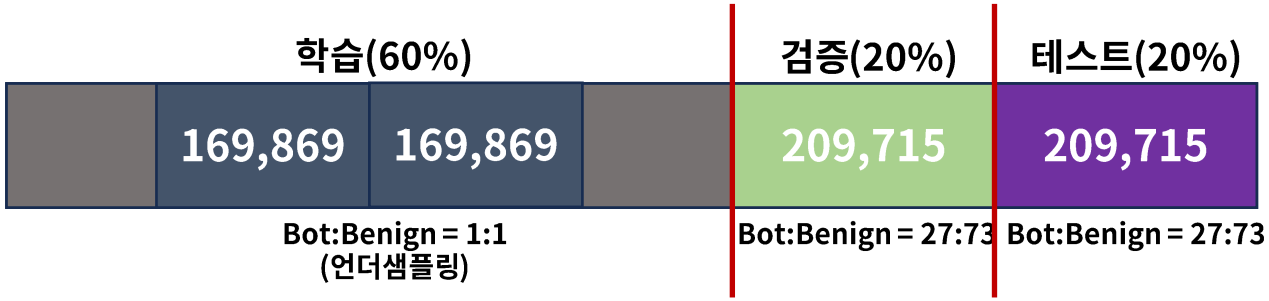
⇒ 다수 클래스(정상)로의 편향 방지

- **검증·테스트 세트: 원본 분포(27:73) 유지**

⇒ 실제 운영 환경에서 평가, 성능 과대평가 방지

- 전 모델이 동일 인덱스 공유

⇒ 입력 표현만 다르고 데이터는 동일 (공정 비교 보장)



❖ 학습 설정

- **검증 F1 기준 Early Stopping(patience=5)**

⇒ 불균형 환경에 강건한 지표로 멈춤 시점 결정

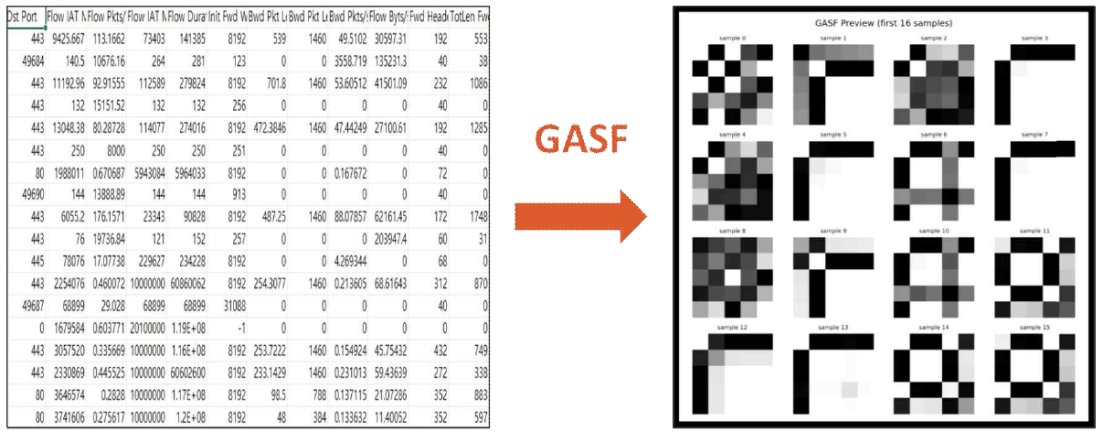
- 손실 함수: BCEWithLogitsLoss
- 옵티마이저: MLP·CNN = Adam, FT-Transformer = AdamW (세부 하이퍼파라미터는 Appendix)

GASF 이미지 변환

❖ GASF (Gramian Angular Summation Field)

- **1차원 수치 벡터 ⇒ 2차원 이미지로 인코딩하는 기법**
- **피처 중요도 순으로 정렬된 25개 피처에 적용**

⇒ 피처 쌍 간 결합 관계를 2차원 공간 패턴으로 표현



[그림 1]. GASF 이미지 변환 과정

❖ 변환결과 및 학습

- 25개 피처 ⇒ **25 × 25 × 1 이미지** : (i, j) 위치에 피처 i·j의 결합 관계가 인코딩
- **이 이미지를 2D CNN으로 학습 ⇒** 1차원에서 드러나지 않던 피처 쌍 관계를 공간적으로 포착
- 정규화 통계량은 학습 분할에서만 산출

⇒ 데이터 누수 방지

비교 모델 구조

모델	입력 표현(shape)	핵심 연산	데이터를 보는 방식	파라미터 수
MLP	1D 수치 벡터(25,)	4층 FC (BatchNorm·Dropout)	모든 피처를 동등하게 처리	1,173,505
CNN	25×25×1 GASF 이미지	합성곱·풀링 2단계 -> FC	이웃 값들의 공간 패턴에 주목	1,198,913
FT-Transformer	피처 토큰+ CLS 토큰	Transformer Self-Attention	피처 간 관계를 직접 학습	1,229,697

❖ **공통 ⇒ 25개 피처 입력 · 이진 분류 · 파라미터 ≈120만으로 통제**

03 실험 및 결과

모델 성능 비교

❖ 평가설정

- 고정 Test set · **서로 다른 랜덤 시드 3개로 반복 학습** ⇒ 평균

❖ 성능

- **세 모델 모두 정확도 99.96%, F1-score 99.93% 대로 포화 수준**
- 단순 성능 지표로는 모델 간 우열을 가리기 어려움 **⇒ 학습 데이터 양에 따른 분석으로 확장**

[표 1]. 테스트 세트 성능 비교

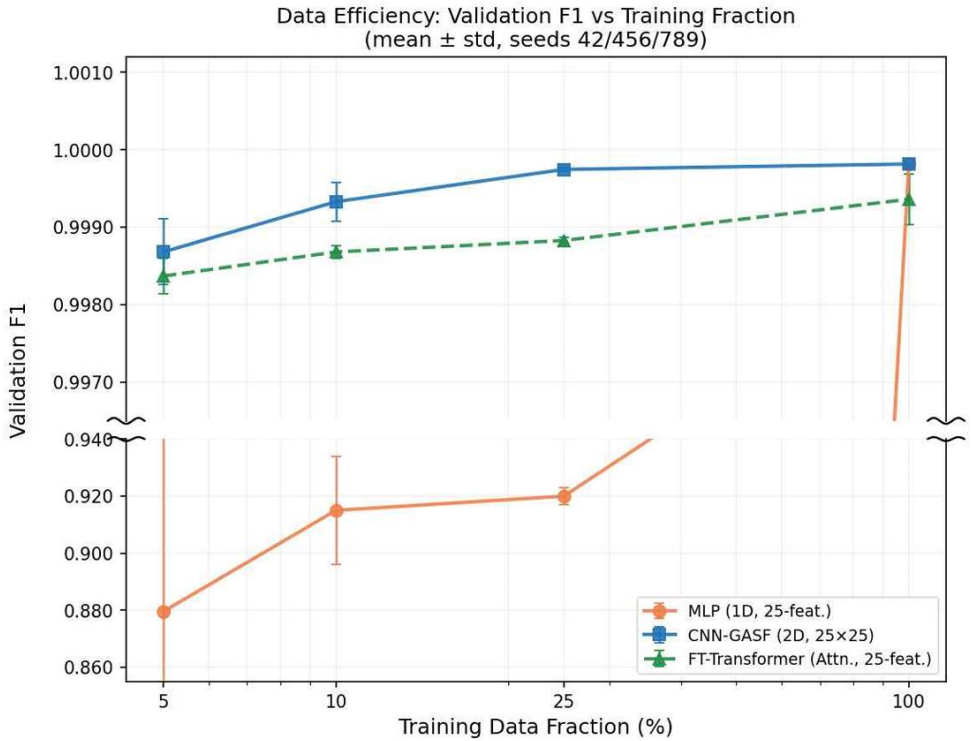
Metric \ Model	MLP	CNN-GASF	FT-Transformer
Accuracy (%)	99.990	99.990	99.965
Precision (%)	99.988	99.985	99.962
Recall (%)	99.974	99.978	99.911
F1-score (%)	99.981	99.982	99.936
오분류(건)	21	21	73

데이터 경량화 실험

- **실험: 학습 데이터를 5% ⇒ 100%로 줄이며 검증 F1 측정**
- **결과 - 5% 조건**에서 격차가 드러남: MLP **87.96% 급락** ↔ CNN-GASF **99.87% 유지**
  - 25% 데이터의 CNN ≈ 100% 데이터의 MLP (1/4 데이터로 동등)
- **Trade-off:** GASF는 1D 대비 약 **16.9배** 용량
- **학습 시간** (3-seed 평균): MLP **138초** < CNN **412초** < FT-Transformer **898초**
- **분석:** 데이터가 적을수록 GASF·토큰 표현이 1D 벡터(MLP)보다 강건 → 데이터 경량화에 GASF-CNN이 가장 효율적
  단, 자원은 완벽한 승자 없음 (저장 용량=1D · 학습 속도=MLP · 데이터 효율=GASF-CNN 우위)

[표 2]. 학습 데이터 비율별 검증 F1 및 학습 용량

	Training Data Ratio			
Category	5%	10%	25%	100%
1D 용량 (MB)	2.9	5.8	14.5	57.9
2D 용량 (MB)	46.6	93.2	233.0	931.6
MLP F1 (%)	87.961	91.503	91.999	99.981
CNN-GASF F1 (%)	99.868	99.933	99.975	99.982
FT-Transformer F1 (%)	99.837	99.868	99.883	99.936



04 결론 및 향후연구

결론 및 향후 연구

Conclusion

- 레이블 확보가 어려운 네트워크 보안 환경에서 **GASF 이미지 인코딩이 데이터 효율적 탐지 수단이 될 수 있음을 반복 실험을 통해 검증**
- **전체 학습 데이터의 5%**라는 동일 조건에서 **CNN-GASF가 F1 99.868%를 달성하여, MLP(87.961%) 대비 이미지 기반 인코딩의 우수한 샘플 효율성을 확인**
- GASF 이미지가 1차원 수치 데이터 대비 약 16.9배의 저장 용량을 차지하나, 정교한 피처 선택(25개 미만)을 통해 **이미지 크기와 전체 저장 용량을 추가로 절감할 수 있는 가능성을 제시**
- **저장 용량·학습시간 등 자원 측면에서는 표현 방식별 trade-off가 존재** — 저장 용량은 1D(MLP·FT-T), 학습 속도는 MLP, 데이터 효율성은 CNN-GASF가 우위로, 단일 최적 모델은 없음을 정량 분석

Future Work

- 단일 데이터셋·봇넷 단일 공격 유형이라는 한계 ⇒ **다양한 데이터셋·모델과 추가 비교 실험**
- GASF 외 MTF·RP등 **다양한 이미징 기법 비교, 이미지 크기(피처 수)에 따른 성능-용량 최적화**
- 이진분류(Bot/Benign)를 넘어 공격 유형별 **다중분류(multi-class)로 확장**하여, DDoS·침투 공격 등 **신규 환경에서 데이터 효율적 탐지 검증**

Acknowledgement

본 연구는 2026년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업 지원을 받아 수행되었음(2024-0-00035).

또한 본 연구는 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학ICT연구센터(ITRC)의 지원을 받아 수행되었음(IITP-2026-RS-2023-00259967).

Thank you

Q&A

Appendix.

피처 랭킹 Top 20

순위	피처	score	RF	MI
1	**Dst Port**	0.943	0.100	0.256
2	**Init Fwd Win Byts**	0.833	0.072	0.299
3	**TotLen Fwd Pkts**	0.725	0.076	0.202
4	**Fwd Pkt Len Max**	0.647	0.064	0.199
5	**Fwd Pkt Len Mean**	0.570	0.051	0.199
6	**Init Bwd Win Byts**	0.539	0.062	0.125
7	**Flow IAT Mean**	0.473	0.047	0.146
8	**Flow Duration**	0.443	0.041	0.146
9	**Flow IAT Max**	0.437	0.037	0.163
10	**Pkt Len Mean**	0.419	0.029	0.184

순위	피처	score	RF	MI
11	**Fwd Pkts/s**	0.408	0.033	0.159
12	**Flow Pkts/s**	0.406	0.036	0.141
13	**Fwd Seg Size Min**	0.400	0.040	0.119
14	**Bwd Pkt Len Mean**	0.379	0.021	0.192
15	**Bwd Pkt Len Max**	0.368	0.019	0.189
16	**Pkt Len Var**	0.366	0.020	0.183
17	**Flow Byts/s**	0.304	0.028	0.104
18	**Bwd Pkts/s**	0.295	0.019	0.137
19	**Tot Fwd Pkts**	0.260	0.027	0.074
20	**Flow IAT Std**	0.229	0.019	0.086

학습설정 – 세부 하이퍼파라미터, 모델 구조

항목	설정
손실 함수	BCEWithLogitsLoss
Early Stopping	검증 F1 기준, patience=5
Batch Size	학습 512, 평가 2048
반복 시드	42, 456, 789 (3회)
학습 epoch(최대)	50

모델	옵티마이저	학습률(LR)	Weight decay
MLP	Adam	5e-4 (0.0005)	미적용, 기본값 0
CNN	Adam	1e-3 (0.001)	미적용, 기본값 0
FT-Transformer	AdamW	**3e-4** (0.0003)	**1e-4** (0.0001)

모델	구조
MLP	은닉층 2048 → 512 → 128, Dropout 0.3, BatchNorm
CNN	Conv 2블록(32·64 필터, 3×3, padding=1), MaxPool(2), Dropout2d 0.25, FC 128
FT-Transformer	d_model=128, heads=8, layers=6, FFN=256, Dropout 0.1, Pre-LN(norm_first), CLS 토큰