Adaptive Memory Networks with Self-supervised Learning for Unsupervised Anomaly Detection

한양대학교 산업공학과
IDSL 석사과정 고경준

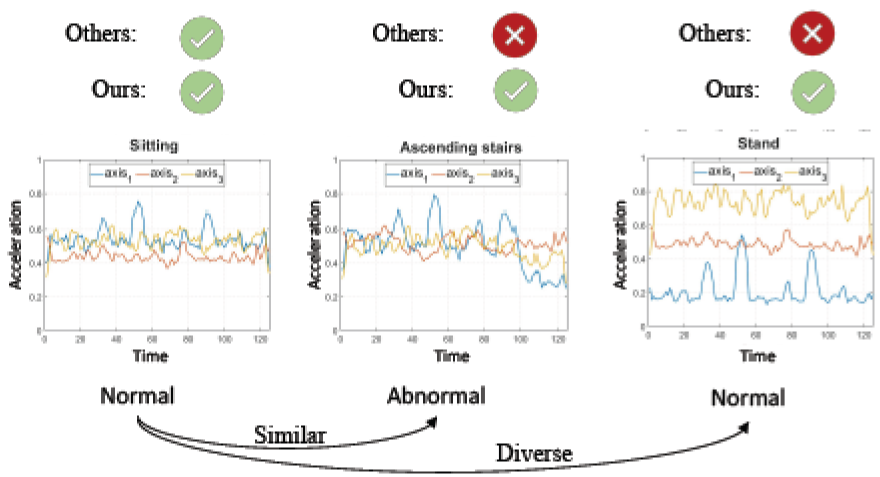
HANYANG UNIVERSITY

Contents

1. Introduction

Introduction

- 일반적인 Time-series Anomaly Detection에서는 abnormal data의 cost가 커서 normal data만을 이용한 unsupervised learning models가 사용됨.
- 하지만 대부분 limited normal data와 limited feature representations의 한계를 가짐.
(limited normal data는 normal data diverse with trained normal data를 탐지 X,
 limited feature representations는 abnormal data similar with normal data를 탐지X)



2. Related Work

2.1 Reconstruction-based methods

- Reconstruction methods (PCA, Kernel PCA)
- Clustering methods (GMM, K-Means)
- One-class learning methods (OCSVM, SVDD)

2.2 Prediction-based methods

2.3 Self-supervised learning

2.4 The vanilla memory network

3. Proposed Approach

3.1 Problem statement

- Definition 1 (Multivariate time series)

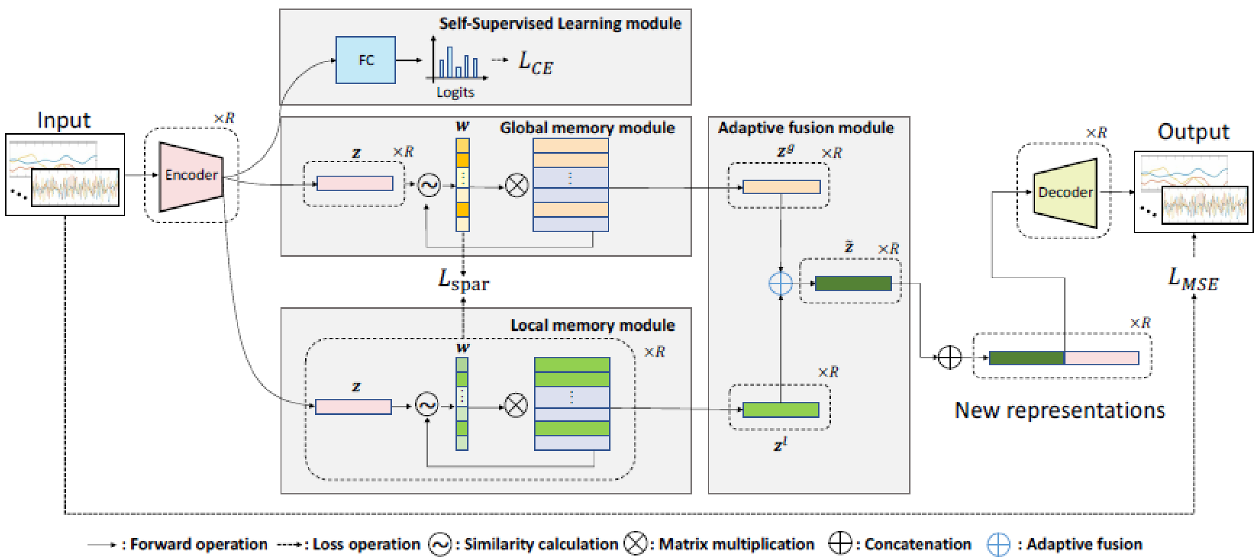
$$X = (x_1, x_2, \cdots, x_N)^\top \in \mathbb{R}^{N \times V}, \quad y \in \mathcal{Y}, \quad \mathcal{Y} = \{1, \cdots, K\}$$

(N variable and Length T. K normal states.)

- Definition 2 (Anomaly)

$$(x_a, y_a),\ y_a \notin \mathcal{Y}$$

3.2 Overview



3. Proposed Approach

3.3 Self-supervised learning

- **Data augmentation**

  Noise : add noise, **Reverse :** reverse along the temporal dimension,

  Permute : randomly perturbs (slicing and shuffle..), **Scale :** change magnitude,

  Negate : multiply -1, **Smooth :** smoothing (apply Savitzky-Golay methods)

- **Distinguish transformation types**

$$L_{CE} = -\sum_{i=1}^{R} y_i \log(p_i), \qquad (3)$$

(R denotes the number of transformations +1 (original))

3. Proposed Approach

3.4 Adaptive Memory Fusion Module

- **3.4.2 Adaptive fusion module**

  global memory (common representation)와 local memories (augmentation-specific representation)를 build

$$z_i^g = f_g(f_e(x_i;\theta_e);\theta_g), i \in [R], \quad (7)$$

$$z_i^l = f_l(f_e(x_i;\theta_e);\theta_l^i), i \in [R], \quad (8)$$

$$\tilde{z}_i = [\alpha_i^g \quad \alpha_i^l] \begin{bmatrix} z_i^g \\ \\ z_i^l \end{bmatrix}, i \in [R], \quad (9)$$

(f_g와 f_l는 각각 global, local memory module, local module은 R개.)

$$L_{MSE} = \sum_{i=1}^{R} \| f_d(\mathrm{concat}(\tilde{z}_i, z_i);\theta_d^i) - x_i \|_2^2. \quad (10) \qquad L_{spar} = \sum_{i=1}^{C} -w_i \cdot log(w_i). \quad (11)$$

reconstruction error (10)와 weight sparsity에 penalty를 주는 sparsity error (11)

3. Proposed Approach

3.5 Training and Inference

- **3.5.1 Training**

$$J(\theta) = L_{MSE} + \lambda_1 L_{CE} + \lambda_2 L_{spar}. \qquad (12)$$

- **3.5.2 Inference**

  threshold μ = 99th percentile of $Err(X_i)$

  ($Err(X_i) = \sum_{j=1}^{R} L(X'_{ij}, X_{ij})$)

  일반적인 방식..

4. Experiment Evaluation

4.1 Dataset

- WMT’14 [English to French dataset], total 850M words
  (Europarl 61M, news commentary 5.5M, UN 421M, two crawled 90M & 272.5M words)
- It reduced to 348M words
- Concatenated validation set (news-test-2012 & 2013), test set (news-test-2014)

4.2 Comparison Methods

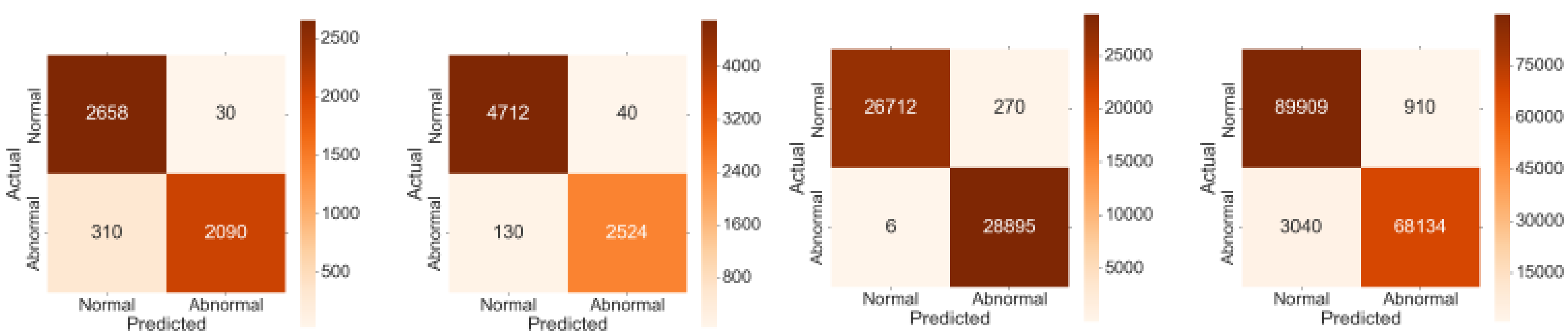
- KPCA, ABOD, OCSVM, HMM
- CNN-LSTM, LSTM-AE, MSCRED, ConvLSTM-composite, BeatGAN, MNAD, GDN, UODA.

(Detailed information of models are in paper…)

4. Experiment Evaluation

4.3 Results and Analysis

Method	DSADS dataset				PAMAP2 dataset				WESAD dataset				CAP dataset			
	mPre	mRec	mF1	Acc	mPre	mRec	mF1	Acc	mPre	mRec	mF1	Acc	mPre	mRec	mF1	Acc
Kernel PCA [11]	0.6184	0.6182	0.6183	0.6186	0.7236	0.6579	0.6892	0.5645	0.5496	0.5495	0.5495	0.5486	0.7603	0.5847	0.6611	0.5892
ABOD [50]	0.6880	0.6510	0.6690	0.6554	0.8653	0.9022	0.8834	0.8985	0.8782	0.8786	0.8784	0.8783	0.7867	0.6365	0.7037	0.6326
OCSVM [15]	0.7608	0.7277	0.7439	0.7312	0.7600	0.7204	0.7397	0.7679	0.6092	0.5631	0.5852	0.5518	<u>0.9267</u>	<u>0.9259</u>	<u>0.9263</u>	<u>0.9257</u>
HMM [51]	0.6959	0.6917	0.6937	0.6901	0.6950	0.6553	0.6745	0.5725	0.6123	0.6060	0.6097	0.6018	0.8238	0.8078	0.8157	0.8090
CNN-LSTM [52]	0.6845	0.6270	0.6545	0.6425	0.6680	0.5392	0.5968	0.6131	0.5883	0.5440	0.5653	0.5318	0.6159	0.5217	0.5649	0.5762
LSTM-AE [8]	0.8471	0.7729	0.8083	0.7852	0.8619	0.7997	0.8296	0.8426	0.2353	0.4762	0.3150	0.4599	0.7147	0.6253	0.6671	0.6286
MSCRED [9]	0.7540	0.5602	0.6428	0.6192	0.6997	0.7301	0.7146	0.7517	0.8850	0.8124	0.8471	0.8420	0.6410	0.5784	0.6081	0.5819
ConvLSTM-AE [16]	0.8164	0.6951	0.7509	0.7121	0.7359	0.7361	0.7360	0.7341	<u>0.9733</u>	<u>0.9698</u>	<u>0.9716</u>	<u>0.9709</u>	0.8150	0.8194	0.8172	0.8165
ConvLSTM-COMP [16]	0.8229	0.7379	0.7781	0.7518	<u>0.8844</u>	<u>0.8842</u>	<u>0.8843</u>	<u>0.8844</u>	0.9626	0.9629	0.9627	0.9619	0.8367	0.8377	0.8372	0.8394
BeatGAN [53]	0.9517	0.5663	0.7100	0.7818	0.7981	0.7420	0.7691	0.8369	0.7586	0.5000	0.6027	0.5172	0.5251	0.5002	0.5123	0.8437
MNAD-P [43]	0.5816	0.5783	0.5799	0.5721	0.8198	0.8176	0.8186	0.8135	0.7600	0.6938	0.7254	0.6849	0.7742	0.7489	0.7613	0.6960
MNAD-R [43]	0.8337	0.7694	0.8003	0.7811	0.8350	0.8355	0.8353	0.8334	0.7426	0.6677	0.7031	0.6579	0.8189	0.8235	0.8212	0.7871
GDN [54]	0.8706	0.8151	0.8419	0.8251	0.8129	0.8104	0.8116	0.8123	0.7520	0.5434	0.6309	0.5590	0.6831	0.6237	0.6520	0.6569
UODA [23]	<u>0.8679</u>	<u>0.8281</u>	<u>0.8475</u>	<u>0.8365</u>	0.8957	0.8513	0.8730	0.8823	0.7623	0.6314	0.6907	0.6191	0.7557	0.5124	0.6107	0.5173
AMSL (Ours)	**0.9407**	**0.9298**	**0.9352**	**0.9332**	**0.9788**	**0.9713**	**0.9750**	**0.9770**	**0.9953**	**0.9949**	**0.9951**	**0.9951**	**0.9771**	**0.9736**	**0.9753**	**0.9756**
Improvement	7.28%	10.17%	8.77%	9.67%	9.44%	8.71%	9.07%	9.26%	2.20%	2.51%	2.35%	2.42%	5.04%	4.77%	4.90%	4.99%



(a) DSADS (b) PAMAP2 (c) WESAD (d) CAP

4. Experiment Evaluation

4.4 Ablation Study

Variant	mPre	mRec	mF1	Acc
CAE	0.8728	0.8711	0.8719	0.8676
CAE + Mem	0.8808	0.8819	0.8813	0.8809
CAE + SSL	0.9393	0.9393	0.9393	0.9364
CAE + SSL + Mem	0.9616	0.9420	0.9517	0.9554
CAE + SSL + Ada Mem	**0.9788**	**0.9713**	**0.9750**	**0.9770**
CAE + Mem w/o Spar	0.8713	0.8723	0.8719	0.8718
CAE + SSL w/o SSL loss	0.8231	0.8175	0.8203	0.8205

- First 4 line에서 memory fusion module, self-supervision, adaptive fusion module 확인.
- Last 2 line에서 sparsity loss, self-supervision loss 확인.

(= Adaptive fusion module with self-supervision and sparsity loss가 best다.

이후 Detailed Analysis, Robustness to Noisy Data, Percentage of Anomaly, Case Study, Parameter sensitivity analysis, Convergence, Time and Space Complexity 생략.)

감사합니다.