TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data

Shreshth Tuli, Giuliao Casale, Nicholas R. Jennings

2022

임동주

목차

Introduction

- Challenges
  - Data Modality의 증가
  - IoT 장비 증가로 인한 센서 데이터 급증
  - 데이터의

  ➔ 데이터의 이상을 “탐지” 하는 것 외에도 그 “원인”을 찾아내는 것도 중요
- Existing Solutions
  - SAND(Statistical Anomaly Detection)
  - openGauss / LSTM-NDT
  - MTAD-GAT / GDN

  Etc.

Related Work

- Classical Methods
  - k-means
  - SVM
  - PCA
  - ARIMA
  - MERLIN

- DL-based Methods
  - LSTM-NDT
  - DAGMM → Autoencoding
  - LSTM-VAE
  - ConvLSTM
  - MAD-GAN
  - MTAD-GAN
  - CAEM
  - GDN 등등

Method

- Data Preprocessing

$$x_t \leftarrow \frac{x_t - \min(\mathcal{T})}{\max(\mathcal{T}) - \min(\mathcal{T}) + \epsilon'},$$

Method

- Model Architecture

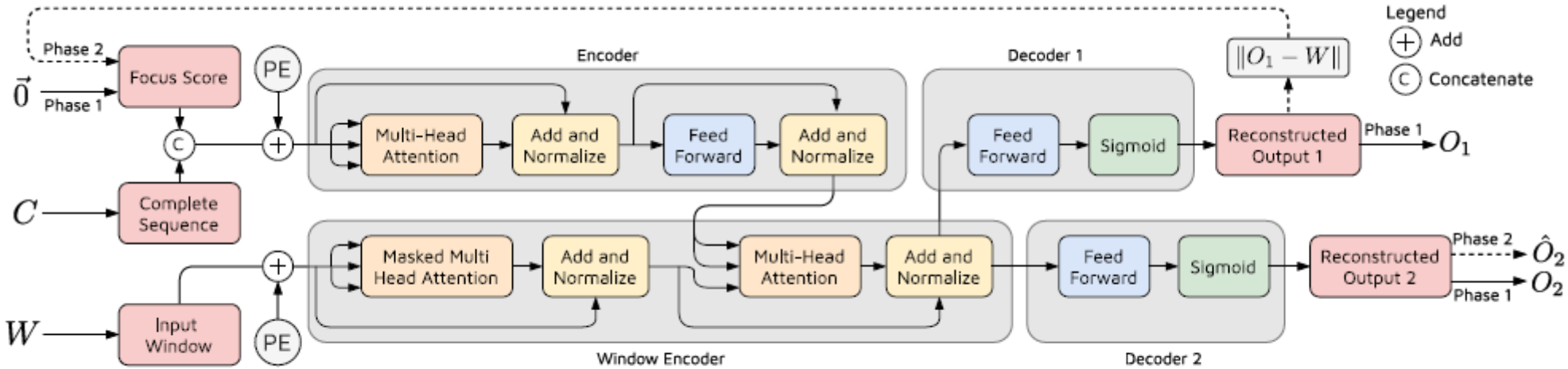


Figure 1: The TranAD Model.

Method

- Loss Function(Adversarial Training)

$$L_1 = \epsilon^{-n}\|O_1 - W\|_2 + (1 - \epsilon^{-n})\|\hat{O}_2 - W\|_2,$$

$$L_2 = \epsilon^{-n}\|O_2 - W\|_2 - (1 - \epsilon^{-n})\|\hat{O}_2 - W\|_2,$$

Experiments

- Hyperparameters

- Window size = 10.
- Number of layers in transformer encoders = 1
- Number of layers in feed-forward unit of encoders = 2
- Hidden units in encoder layers = 64
- Dropout in encoders = 0.1

Experiments

- Datasets
  - Numenta Anomaly Benchmark(NAB)
  - HexagonML(UCR)
  - MIT-BIH supraventricular Arrhythmia Database(MBA)
  - Soil Moisture Active Passive(SMAP)
  - Mars Science Laboratory(MSL)
  - Secure Water Treatment(SWaT)
  - Water Distribution(WADI)
  - Server Machine Dataset(SMD)
  - Multi-Source Distributed System(MSDS)

- Metrics
  - Anomaly Detection
    > precision, recall, ROC/AUC
  - Anomaly Diagnosis
    > HitRate@P%

Experiments

- Result(Detection)

Method	NAB				UCR				MBA			
	P	R	AUC	F1	P	R	AUC	F1	P	R	AUC	F1
MERLIN	0.8013	0.7262	0.8414	0.7619	0.7542	0.8018	0.8984	0.7773	0.9846	0.4913	0.7828	0.6555
LSTM-NDT	0.6400	0.6667	0.8322	0.6531	0.5231	0.8294	0.9781	0.6416	0.9207	0.9718	0.9780	0.9456
DAGMM	0.7622	0.7292	0.8572	0.7453	0.5337	0.9718	0.9916	0.6890	0.9475	0.9900	0.9858	0.9683
OmniAnomaly	0.8421	0.6667	0.8330	0.7442	0.8346	0.9999	0.9981	0.9098	0.8561	1.0000	0.9570	0.9225
MSCRED	0.8522	0.6700	0.8401	0.7502	0.5441	0.9718	0.9920	0.6976	0.9272	1.0000	0.9799	0.9623
MAD-GAN	0.8666	0.7012	0.8478	0.7752	0.8538	0.9891	0.9984	0.9165	0.9396	1.0000	0.9836	0.9689
USAD	0.8421	0.6667	0.8330	0.7442	0.8952	1.0000	0.9989	0.9447	0.8953	0.9989	0.9701	0.9443
MTAD-GAT	0.8421	0.7272	0.8221	0.7804	0.7812	0.9972	0.9978	0.8761	0.9018	1.0000	0.9721	0.9484
CAE-M	0.7918	0.8019	0.8019	0.7968	0.6981	1.0000	0.9957	0.8222	0.8442	0.9997	0.9661	0.9154
GDN	0.8129	0.7872	0.8542	0.7998	0.6894	0.9988	0.9959	0.8158	0.8832	0.9892	0.9528	0.9332
TranAD	0.8889	0.9892	**0.9541**	**0.9364**	0.9407	1.0000	**0.9994**	**0.9694**	0.9569	1.0000	**0.9885**	**0.9780**

Method	SMAP				MSL				SWaT			
	P	R	AUC	F1	P	R	AUC	F1	P	R	AUC	F1
MERLIN	0.1577	0.9999	0.7426	0.2725	0.2613	0.4645	0.6281	0.3345	0.6560	0.2547	0.6175	0.3669
LSTM-NDT	0.8523	0.7326	0.8602	0.7879	0.6288	1.0000	0.9532	0.7721	0.7778	0.5109	0.7140	0.6167
DAGMM	0.8069	0.9891	0.9885	0.8888	0.7363	1.0000	0.9716	0.8482	0.9933	0.6879	0.8436	0.8128
OmniAnomaly	0.8130	0.9419	0.9889	0.8728	0.7848	0.9924	0.9782	0.8765	0.9782	0.6957	0.8467	0.8131
MSCRED	0.8175	0.9216	0.9821	0.8664	0.8912	0.9862	0.9807	0.9363	0.9992	0.6770	0.8433	0.8072
MAD-GAN	0.8157	0.9216	0.9891	0.8654	0.8516	0.9930	0.9862	0.9169	0.9593	0.6957	0.8463	0.8065
USAD	0.7480	0.9627	0.9890	0.8419	0.7949	0.9912	0.9795	0.8822	0.9977	0.6879	0.8460	0.8143
MTAD-GAT	0.7991	0.9991	0.9844	0.8880	0.7917	0.9824	0.9899	0.8768	0.9718	0.6957	0.8464	0.8109
CAE-M	0.8193	0.9567	0.9901	0.8827	0.7751	1.0000	0.9903	0.8733	0.9697	0.6957	0.8464	0.8101
GDN	0.7480	0.9891	0.9864	0.8518	0.9308	0.9892	0.9814	**0.9591**	0.9697	0.6957	0.8462	0.8101
TranAD	0.8043	0.9999	**0.9921**	**0.8915**	0.9038	0.9999	**0.9916**	0.9494	0.9760	0.6997	**0.8491**	**0.8151**

Method	WADI				SMD				MSDS			
	P	R	AUC	F1	P	R	AUC	F1	P	R	AUC	F1
MERLIN	0.0636	0.7669	0.5912	0.1174	0.2871	0.5804	0.7158	0.3842	0.7254	0.3110	0.5022	0.4353
LSTM-NDT	0.0138	0.7823	0.6721	0.0271	0.9736	0.8440	0.9671	0.9042	0.9999	0.8012	0.8013	0.8896
DAGMM	0.0760	0.9981	0.8563	0.1412	0.9103	0.9914	0.9954	0.9491	0.9891	0.8026	0.9013	0.8861
OmniAnomaly	0.3158	0.6541	0.8198	0.4260	0.8881	0.9985	0.9946	0.9401	1.0000	0.7964	0.8982	0.8867
MSCRED	0.2513	0.7319	0.8412	0.3741	0.7276	0.9974	0.9921	0.8414	1.0000	0.7983	0.8943	0.8878
MAD-GAN	0.2233	0.9124	0.8026	0.3588	0.9991	0.8440	0.9933	0.9150	0.9982	0.6107	0.8054	0.7578
USAD	0.1873	0.8296	0.8723	0.3056	0.9060	0.9974	0.9933	0.9495	0.9912	0.7959	0.8979	0.8829
MTAD-GAT	0.2818	0.8012	0.8821	0.4169	0.8210	0.9215	0.9921	0.8683	0.9919	0.7964	0.8982	0.8835
CAE-M	0.2782	0.7918	0.8728	0.4117	0.9082	0.9671	0.9783	0.9367	0.9908	0.8439	0.9013	0.9115
GDN	0.2912	0.7931	0.8777	0.4260	0.7170	0.9974	0.9924	0.8342	0.9989	0.8026	**0.9105**	0.8900
TranAD	0.3529	0.8296	**0.8968**	**0.4951**	0.9262	0.9974	**0.9974**	**0.9605**	0.9999	0.8626	0.9013	**0.9262**

Experiments

- Result(Diagnosis)

Table 4: Diagnosis Performance.

Method	SMD				MSDS			
	H@100%	H@150%	N@100%	N@150%	H@100%	H@150%	N@100%	N@150%
MERLIN	**0.5907**	0.6177	0.4150	0.4912	0.3816	0.5626	0.3010	0.3947
LSTM-NDT	0.3808	0.5225	0.3603	0.4451	0.1504	0.2959	0.1124	0.1993
DAGMM	0.4927	0.6091	0.5169	0.5845	0.2617	0.4333	0.3153	0.4154
OmniAnomaly	0.4567	0.5652	0.4545	0.5125	0.2839	0.4365	0.3338	0.4231
MSCRED	0.4272	0.5180	0.4609	0.5164	0.2322	0.3469	0.2297	0.2962
MAD-GAN	0.4630	0.5785	0.4681	0.5522	0.3856	0.5589	0.4277	0.5292
USAD	0.4925	0.6055	0.5179	0.5781	0.3095	0.4769	0.3534	0.4515
MTAD-GAT	0.3493	0.4777	0.3759	0.4530	**0.5812**	0.5885	0.5926	0.6522
CAE-M	0.4707	0.5878	**0.5474**	**0.6178**	0.2530	0.4171	0.2047	0.3010
GDN	0.3143	0.4386	0.2980	0.3724	0.2276	0.3382	0.2921	0.3570
TranAD	0.4981	**0.6401**	0.4941	**0.6178**	0.4630	**0.7533**	**0.5981**	**0.6963**

Experiments

- Result(Training Time)

Table 5: Comparison of training times in seconds per epoch.

Method	NAB	UCR	MBA	SMAP	MSL	SWaT	WADI	SMD	MSDS
MERLIN	3.28	4.09	20.19	6.89	**5.12**	10.12	132.69	72.32	42.22
LSTM-NDT	10.64	8.71	27.80	27.62	26.24	26.43	297.12	373.14	361.12
DAGMM	25.38	20.78	74.62	19.05	16.41	18.51	178.17	204.36	187.54
OmniAnomaly	38.27	27.96	109.86	27.05	21.31	28.39	212.99	276.97	277.10
MSCRED	258.86	262.45	592.13	16.13	33.47	183.67	1349.05	237.66	109.63
MAD-GAN	39.80	25.71	160.29	29.49	26.27	27.79	293.60	314.82	285.25
USAD	31.21	21.10	120.86	23.63	21.22	22.72	242.86	250.97	232.82
MTAD-GAT	145.00	97.12	233.08	1015.03	1287.42	103.92	9812.13	6564.11	1304.09
CAE-M	22.48	19.42	67.44	187.35	575.96	41.25	5525.62	3102.12	552.83
GDN	83.84	58.78	159.01	62.33	96.71	59.40	4063.05	809.94	585.34
TranAD	**1.25**	**0.84**	**4.08**	**3.55**	5.27	**0.87**	**115.91**	**43.56**	**17.15**

Experiments

- Ablation Studies

Table 6: Ablation Study - F1 and F1* scores for TranAD and its ablated versions.

Method	NAB		UCR		MBA	
	F1	F1*	F1	F1*	F1	F1*
TranAD	**0.9364**	**0.8421**	**0.9694**	**0.9399**	**0.9780**	**0.9617**
w/o transformer	0.8850	0.8019	0.8466	0.5495	0.9749	0.9584
w/o self-condition	0.8887	0.8102	0.9191	0.9028	0.9770	**0.9617**
w/o adversarial training	0.9012	0.8102	0.9634	0.9289	0.9752	0.9592
w/o MAML	0.9068	0.8210	0.9689	0.9304	0.9756	**0.9617**
Method	SMAP		MSL		SWaT	
	F1	F1*	F1	F1*	F1	F1*
TranAD	**0.8915**	**0.8889**	**0.9494**	**0.9172**	**0.8151**	**0.8094**
w/o transformer	0.8643	0.8147	0.9137	0.9037	0.8143	0.6360
w/o self-condition	0.8894	0.8153	0.9175	0.8913	0.7953	0.8094
w/o adversarial training	0.8906	0.8476	0.9455	**0.9172**	0.8028	0.7832
w/o MAML	**0.8915**	**0.8899**	0.9466	0.6732	0.8143	0.8079
Method	WADI		SMD		MSDS	
	F1	F1*	F1	F1*	F1	F1*
TranAD	**0.4951**	**0.0649**	**0.9605**	**0.9478**	**0.9262**	**0.8391**
w/o transformer	0.2181	0.0037	0.9071	0.9032	0.8867	0.8389
w/o self-condition	0.3620	0.0631	0.9502	0.8847	0.8748	0.8214
w/o adversarial training	0.3820	0.0621	0.9177	0.8667	0.9181	0.8389
w/o MAML	0.4815	0.0553	0.9433	0.8164	0.8870	0.8389

Experiments

- Ablation Studies

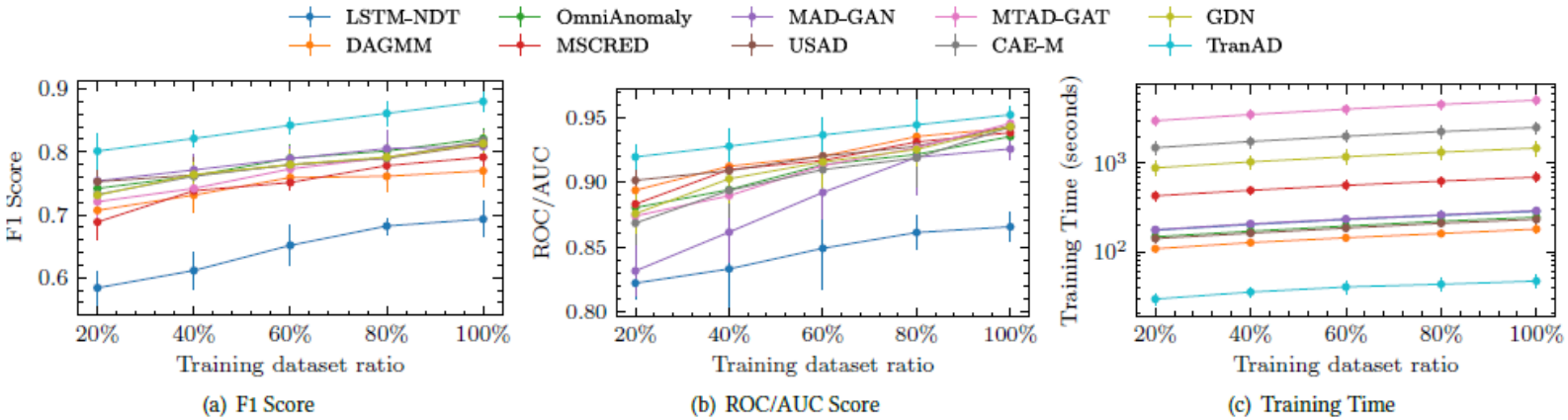


Figure 6: F1 score, ROC/AUC score and training times with dataset size.

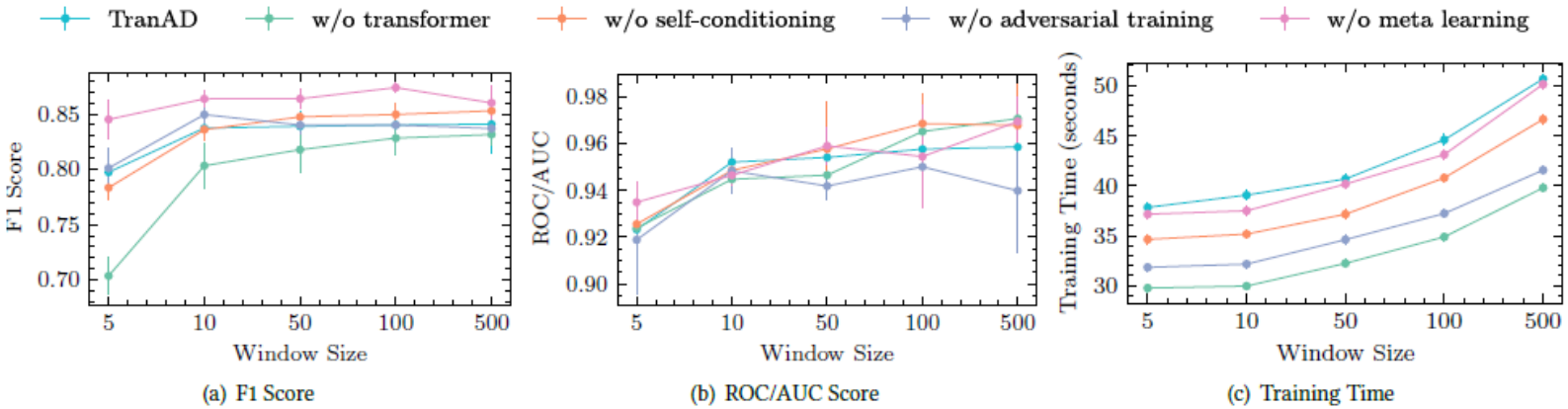
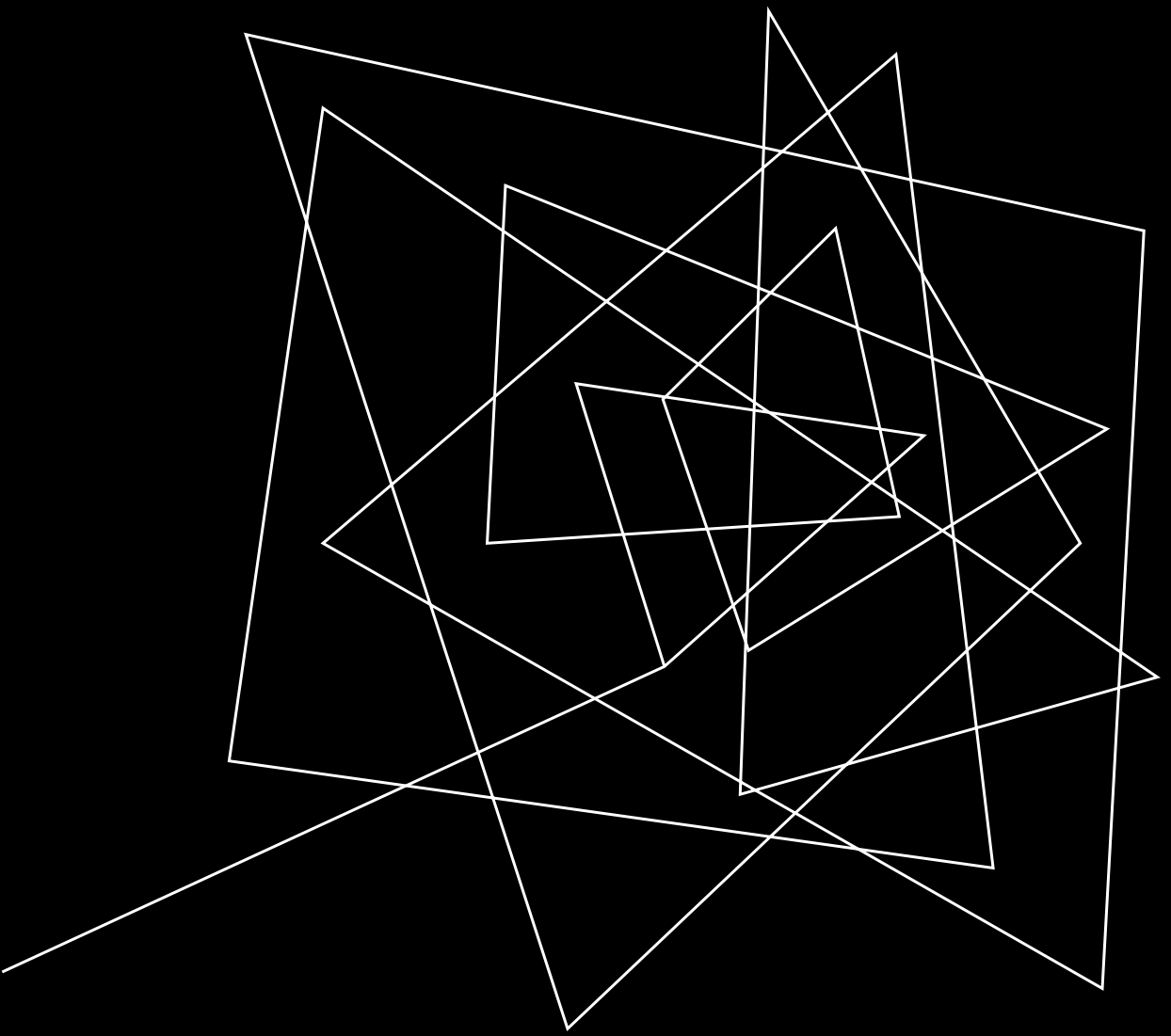


Figure 7: F1 score, ROC/AUC score and training times with window size.



Q&A

[Github]
https://github.com/microsoft/Swin-Transformer