Correlation-aware Deep Generative Model for Unsupervised Anomaly Detection

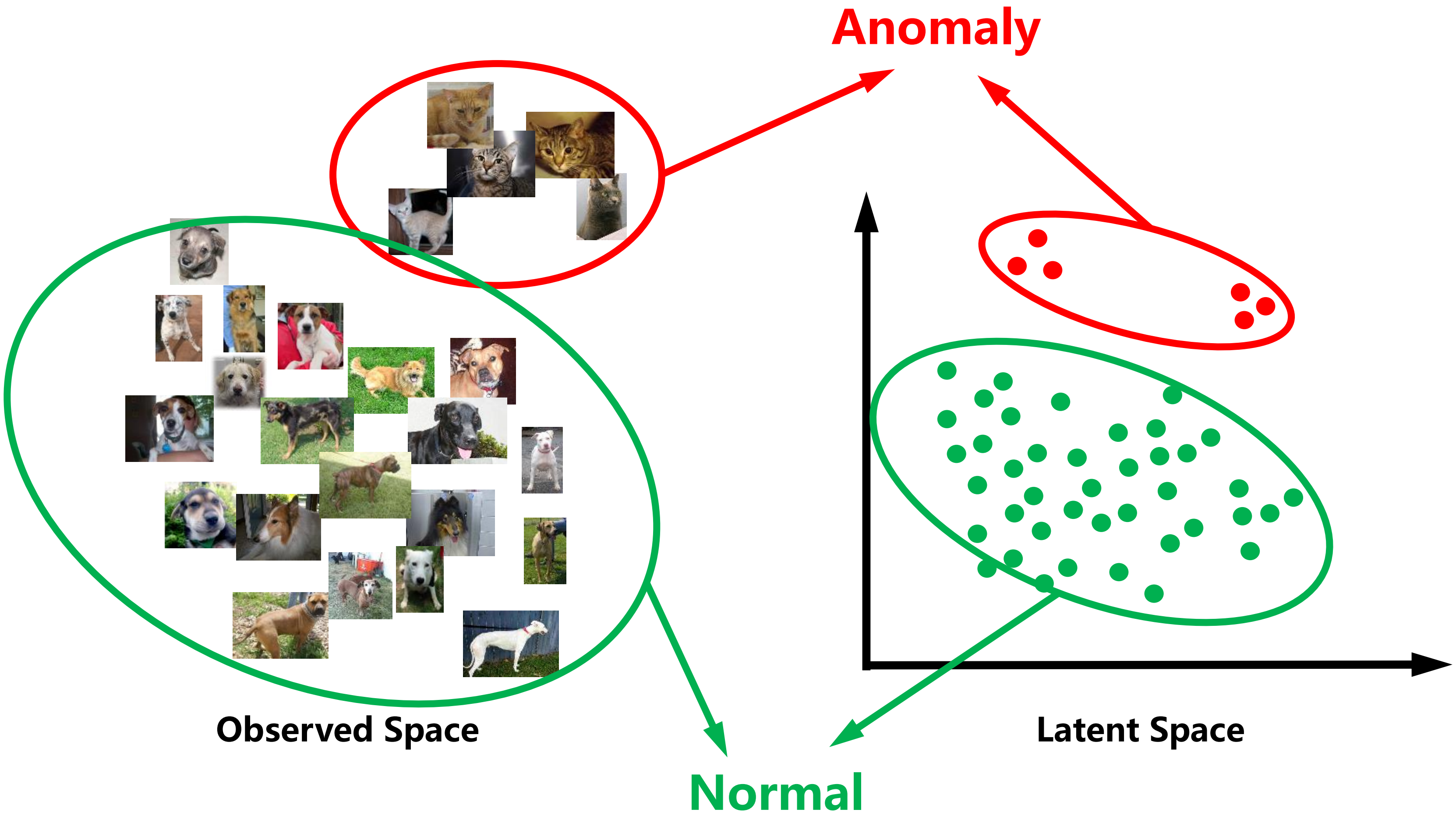
Haoyi Fan [1], Fengbin Zhang [1], Ruidong Wang [1], Liang Xi [1], Zuoyong Li [2]

Harbin University of Science and Technology [1]
Minjiang University [2]

isfanhy@hrbust.edu.cn

Background

Anomaly



Background


https://www.explosion.com/135494/5-effective-strategies-of-fraud-detection-and-prevention-for-ecommerce/

Fraud Detection

https://towardsdatascience.com/building-an-intrusion-detection-system-using-deep-learning-b9488332b321

Intrusion Detection

Anomaly Detection Tasks are Ubiquitous !

https://planforgermany.com/switching-private-public-health-insurance-germany/

Disease Detection


https://blog.exporthub.com/working-with-chinese-manufacturers/

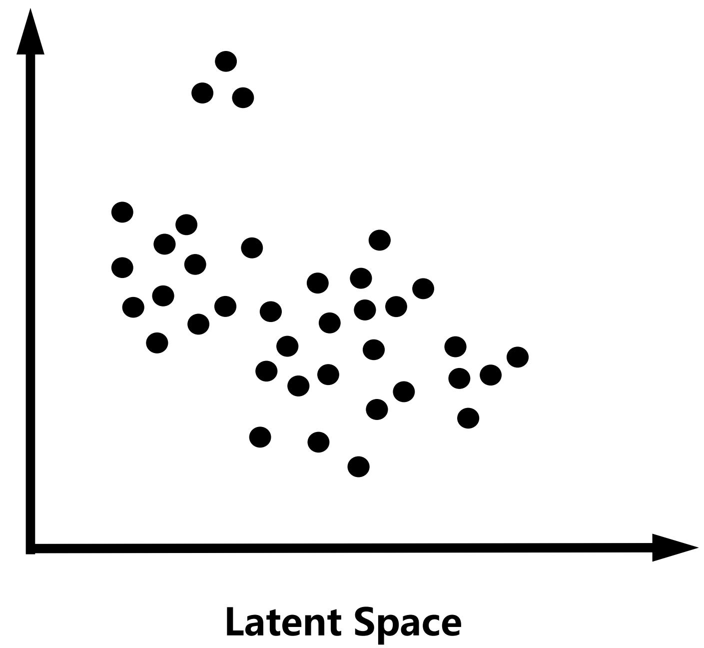
Fault Detection

Background

Unsupervised Anomaly Detection

– From the Density Estimation Perspective

Data samples: $X_{train} = \{x_1, x_2, x_3, \dots, x_n\}$,
x_i is assumed normal.



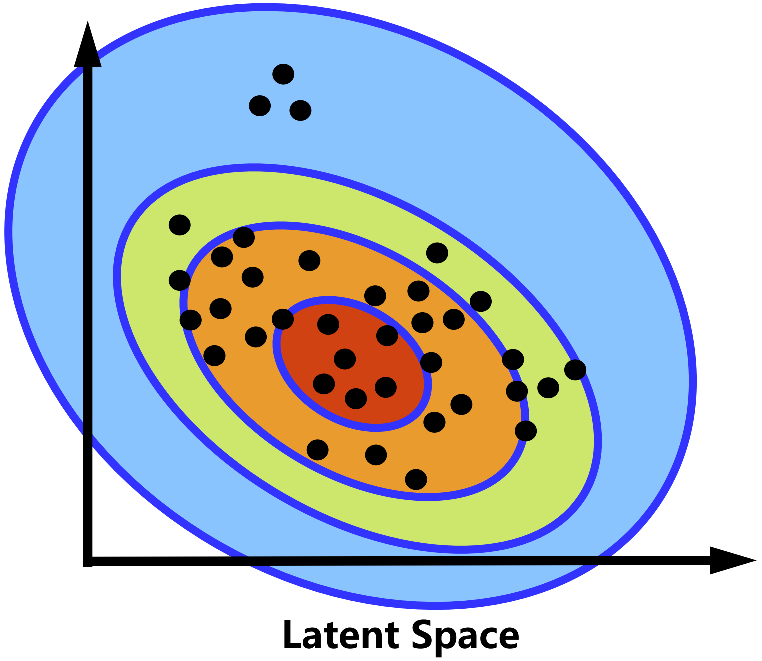
Background

Unsupervised Anomaly Detection

– From the Density Estimation Perspective

Data samples: $X_{train} = \{x_1, x_2, x_3, \dots, x_n\}$,
x_i is assumed normal.

Model: $p(x)$



Background

Unsupervised Anomaly Detection

– From the Density Estimation Perspective

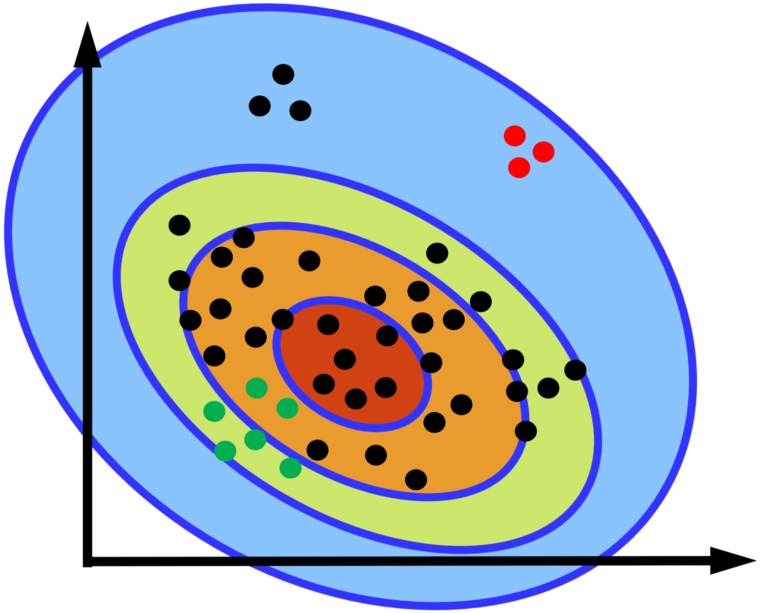
Data samples: $X_{train} = \{x_1, x_2, \ldots, x_n\}$,
x_i is assumed normal.

Model: $p(x)$

Test samples: $X_{test} = \{x_1, x_2, \ldots, x_n\}$,
x_t is unknow.

if $p(x_t) < \lambda$, x_t is **abnormal**.

if $p(x_t) \geq \lambda$, x_t is **normal**.



Latent Space

Background

Unsupervised Anomaly Detection

– From the Density Estimation Perspective

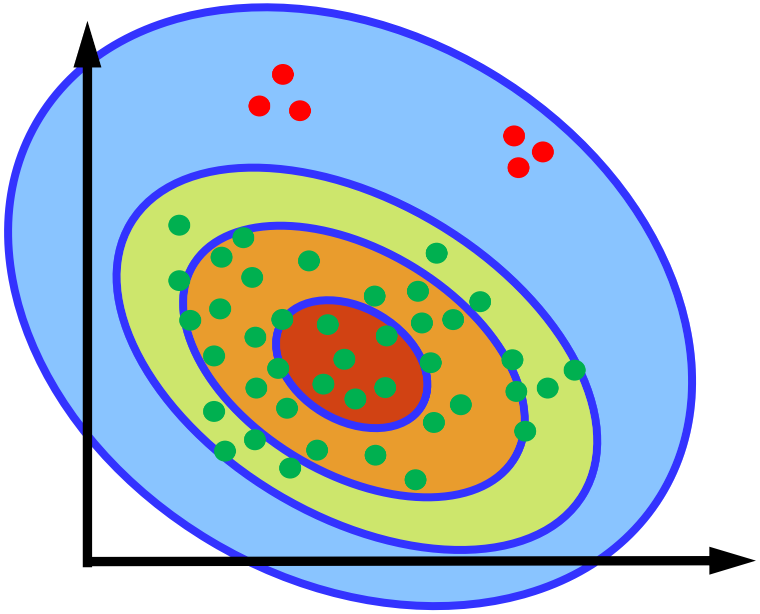
Data samples: $X_{train} = \{x_1, x_2, \dots, x_n\}$,
x_i is assumed normal.

Model: $p(x)$

Test samples: $X_{test} = \{x_1, x_2, \dots, x_n\}$,
x_t is unknow.

if $p(x_t) < \lambda$, x_t is **abnormal**.

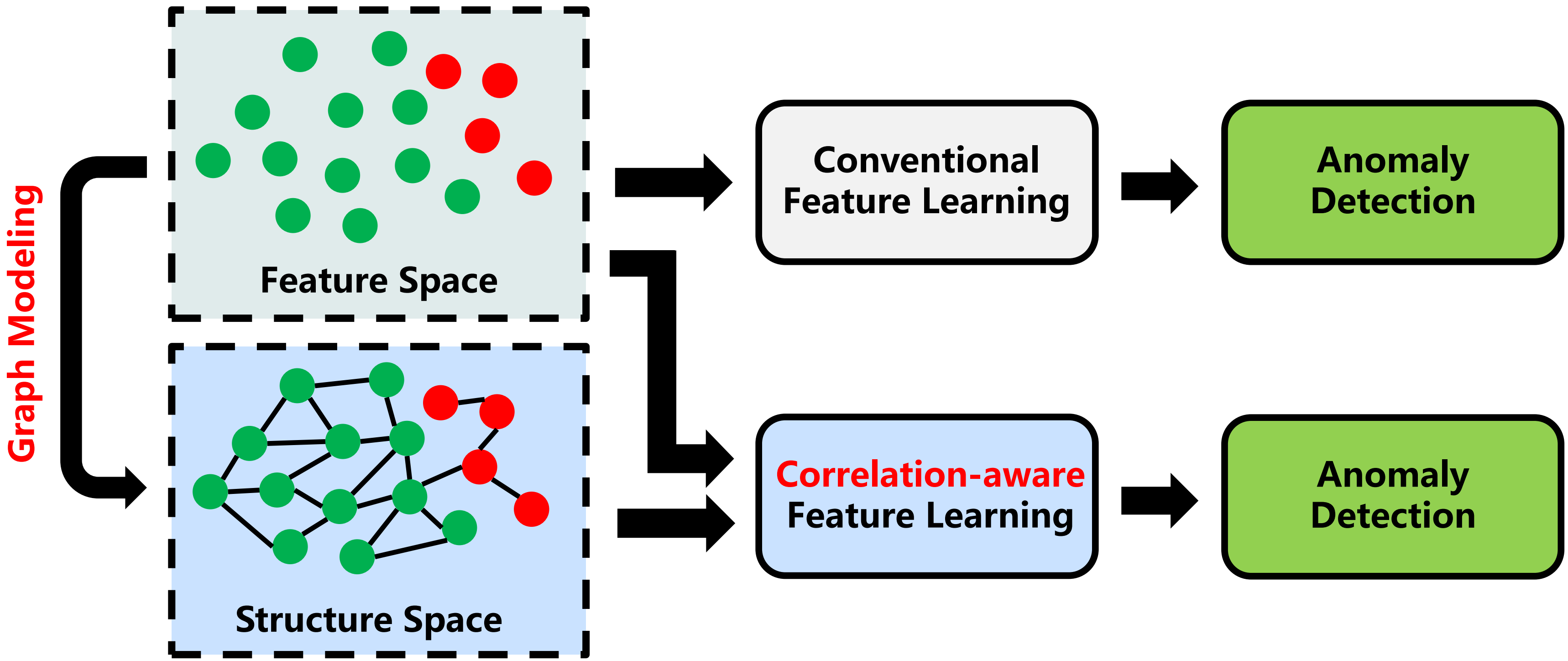
if $p(x_t) \geq \lambda$, x_t is **normal**.



Anomalies reside in the low probability density areas.

Background

Correlation among data samples



How to discover the normal pattern from both the feature level and structural level ?

Problem Statement

Anomaly Detection

Given a set of input samples $\mathcal{X} = \{x_i | i = 1, \ldots, N\}$, each of which is associated with a F dimension feature $\mathbf{X}_i \in \mathbb{R}^F$, we aim to learn a score function $u(\mathbf{X}_i)\colon \mathbb{R}^F \mapsto \mathbb{R}$, to classify sample x_i based on the threshold λ:

$$y_i = \begin{cases} 1, & if\ u(\mathbf{X}_i) \geq \lambda, \\ 0, & otherwise. \end{cases}$$

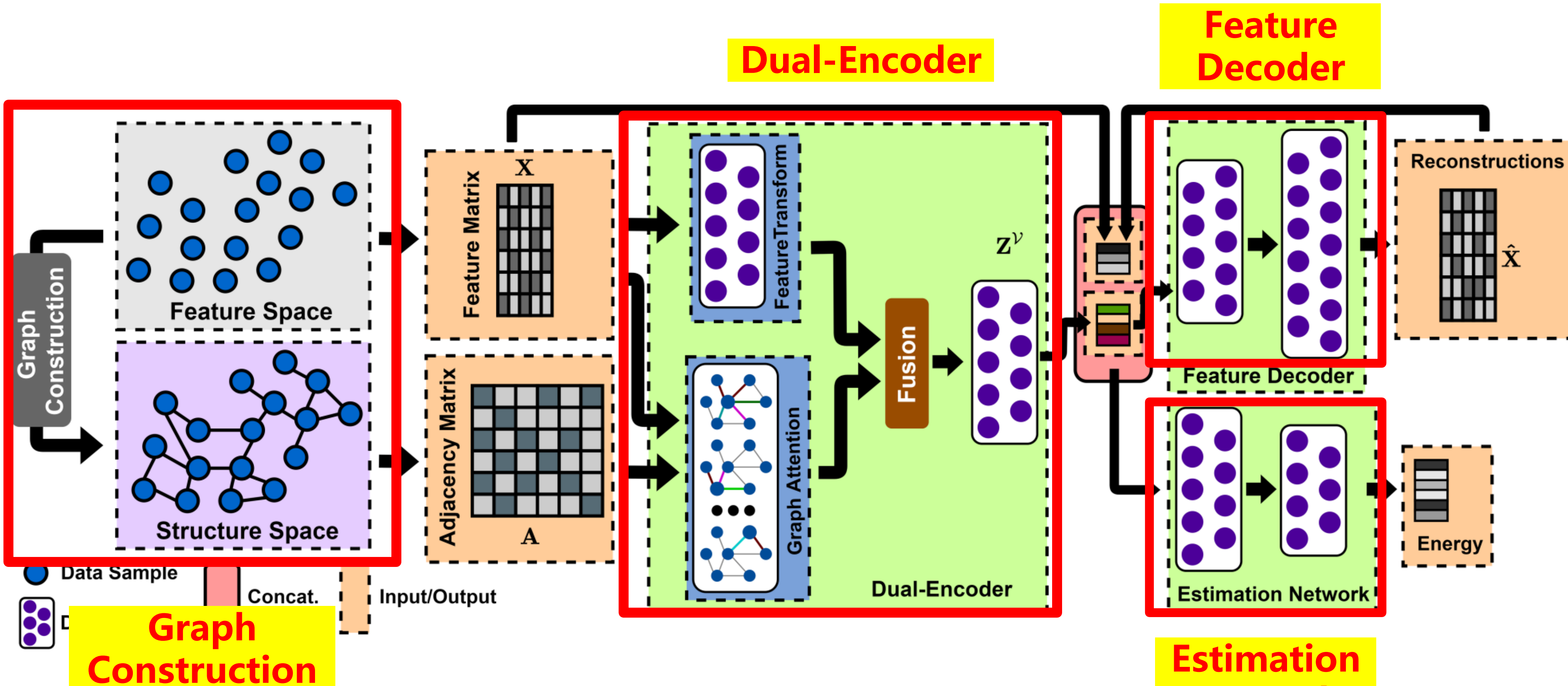
where y_i denotes the label of sample x_i, with 0 being the normal class and 1 the anomalous class.

Notations

$\mathcal{G}$: Graph.
$\mathcal{V}$: Set of nodes in a graph.
$\mathcal{E}$: Set of edges in a graph.
N: Number of nodes.
F : Dimension of attribute.
$\mathbf{A} \in \mathbb{R}^{N \times N}$: Adjacency matrix of a network.
$\mathbf{X} \in \mathbb{R}^{N \times F}$: Feature matrix of all nodes.

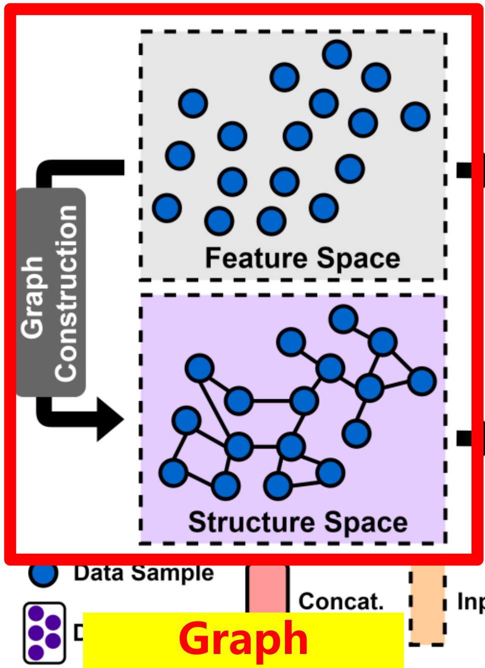
Method

CADGMM



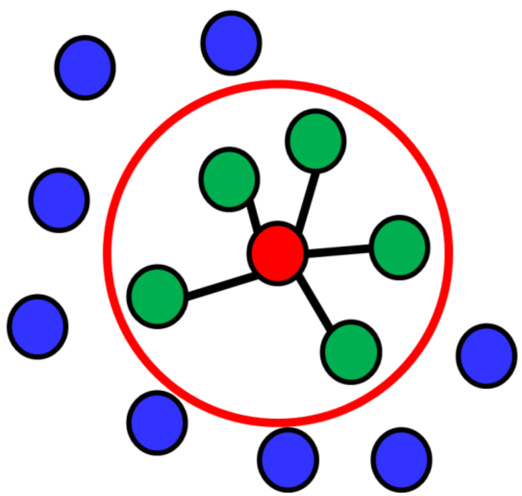
Method

CADGMM



Graph Construction

K-Nearest Neighbor
e.g. K=5



Original feature:

$$\mathcal{X} = \{x_i | i = 1, \ldots, N\}$$

Find neighbors by K-NN:

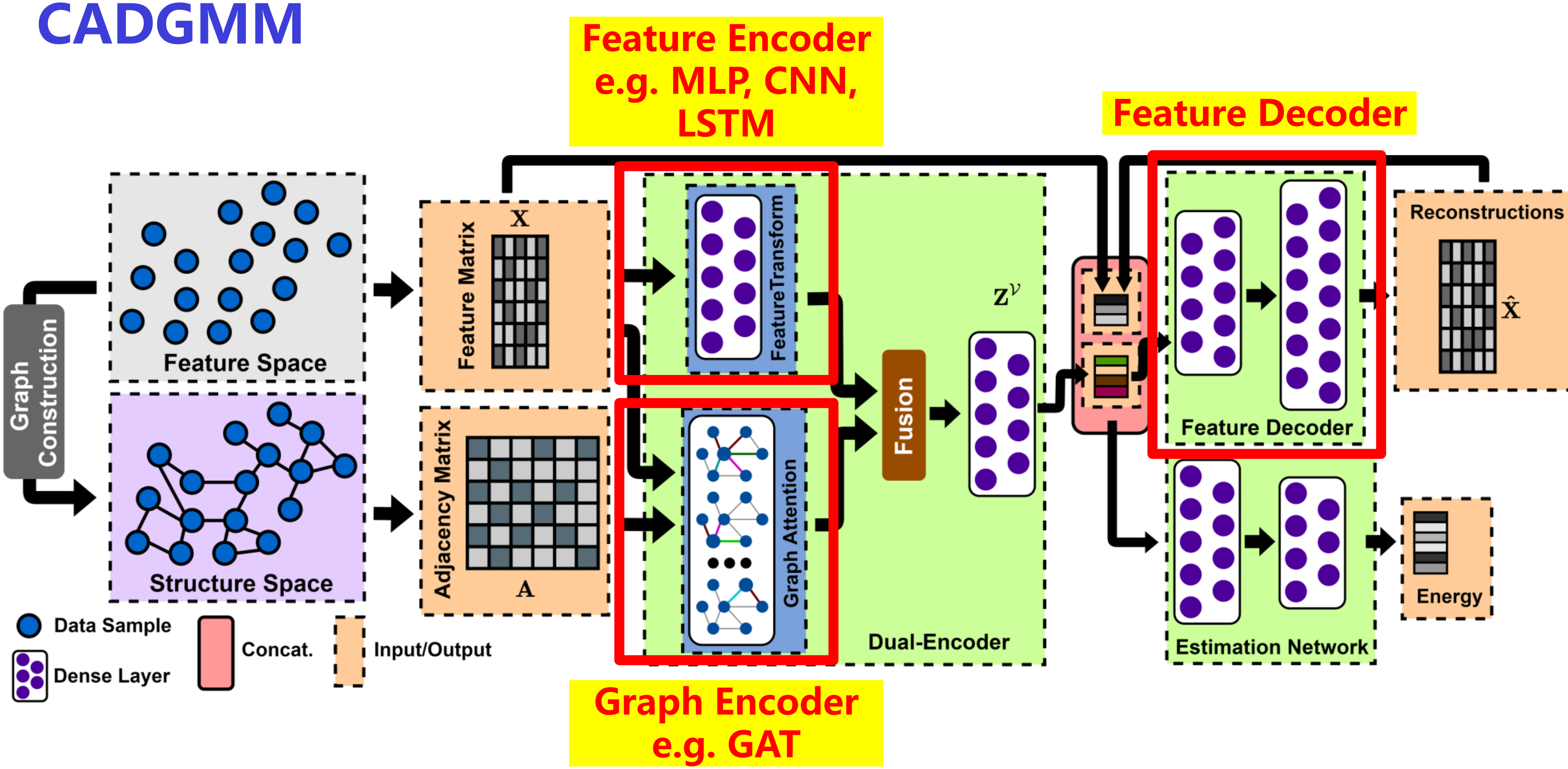
$$\mathcal{N}_i = \{x_{i_k} | k = 1, \ldots, K\}$$

Model correlation as graph:

$$\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{X}\}$$
$$\mathcal{V} = \{v_i = x_i | i = 1, \ldots, N\}$$
$$\mathcal{E} = \{e_{i_k} = (v_i, v_{i_k}) | v_{i_k} \in \mathcal{N}_i\}$$

Method

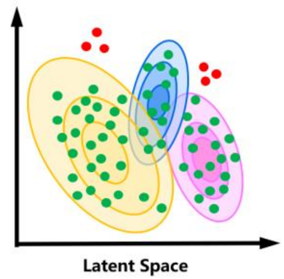
CADGMM



Method

CADGMM

Gaussian Mixture Model



Initial embedding:

Z

Membership:

$$Z^{\mathcal{M}(l_{\mathcal{M}})} = \sigma\left(Z^{\mathcal{M}(l_{\mathcal{M}}-1)}W^{\mathcal{M}(l_{\mathcal{M}}-1)} + b^{\mathcal{M}(l_{\mathcal{M}}-1)}\right), Z^{\mathcal{M}(0)} = Z$$

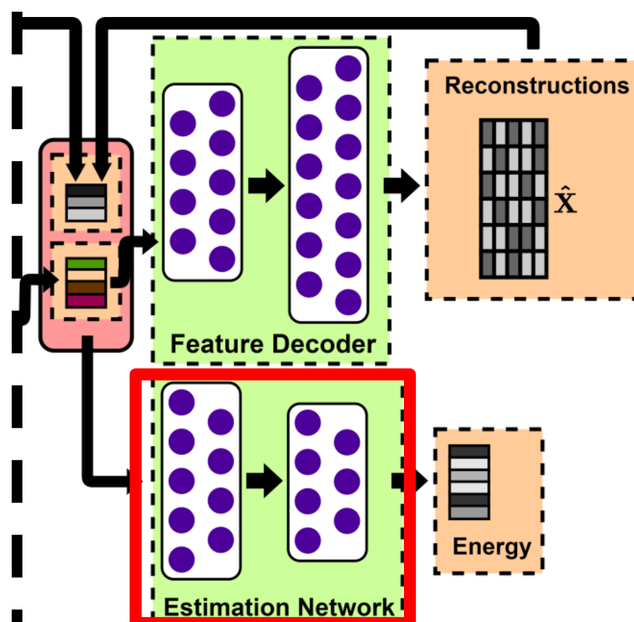
$$\mathcal{M} = \mathrm{Softmax}(Z^{\mathcal{M}(L_{\mathcal{M}})}), \mathcal{M} \in \mathbb{R}^{N\times M}$$

Parameter Estimation:

$$\boldsymbol{\mu}_{\boldsymbol{m}} = \frac{\sum_{i=1}^{N} \mathcal{M}_{i,m} Z_i}{\sum_{i=1}^{N} \mathcal{M}_{i,m}}, \ \boldsymbol{\Sigma}_{\boldsymbol{m}} = \frac{\sum_{i=1}^{N} \mathcal{M}_{i,m}(Z_i - \boldsymbol{\mu}_{\boldsymbol{m}})(Z_i - \boldsymbol{\mu}_{\boldsymbol{m}})^{\mathrm{T}}}{\sum_{i=1}^{N} \mathcal{M}_{i,m}}$$

Energy:

$$\mathbf{E}_Z = -\log\left(\sum_{m=1}^{M} \sum_{i=1}^{N} \frac{\mathcal{M}_{i,m}}{N} \frac{\exp(-\frac{1}{2}(Z - \boldsymbol{\mu}_{\boldsymbol{m}})^{\mathrm{T}} \boldsymbol{\Sigma}_m^{-1}(Z - \boldsymbol{\mu}_{\boldsymbol{m}}))}{|2\pi\boldsymbol{\Sigma}_m|^{\frac{1}{2}}}\right)$$



Estimation network

Method

Loss and Anomaly Score

Loss Function:

$$\mathcal{L} = \boxed{||\mathrm{X} - \widehat{\mathrm{X}}||_2^2} + \boxed{\lambda_1 \mathbf{E}_Z} + \boxed{\lambda_2 \sum_{m=1}^{M} \sum_{i=1}^{N} \frac{1}{(\boldsymbol{\Sigma_m})_{ii}}} + \boxed{\lambda_3 ||\mathrm{Z}||_2^2}$$

Rec. Error

Energy

Covariance Penalty

Embedding Penalty

Anomaly Score:

$$Score = \mathbf{E}_Z$$

Solution for Problem:

$$y_i = \begin{cases} 1, & if\ u(\mathbf{X}_i) \geq \lambda, \\ 0, & otherwise. \end{cases}$$

λ=*Distribution*(*Score*)

Experiment

Datasets

Table 1. Statistics of the public benchmark datasets.

Database	# Dimensions	# Instances	Anomaly ratio
KDD99	120	494,021	0.2
Arrhythmia	274	452	0.15
Satellite	36	6,435	0.32

Baselines

OC-SVM Chen et al. 2001
IF Liu et al. 2008
DSEBM Zhai et al. 2016
DAGMM Zong et al. 2018
AnoGAN Schlegl et al. 2017
ALAD Zenati et al. 2018

Evaluation Metrics

Precision
Recall
F1-Score

Experiment

Results

Table 3. Anomaly Detection Performance on KDD99, Arrhythmia, and Satellite datasets. Better results are marked in **bold**.

Method	KDD99			Arrhythmia			Satellite		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
OC-SVM [4]	74.57	85.23	79.54	53.97	40.82	45.81	52.42	59.99	61.07
IF [8]	92.16	93.73	92.94	51.47	54.69	53.03	60.81	**94.89**	75.40
DSEBM-r [21]	85.21	64.72	73.28	15.15	15.13	15.10	67.84	68.61	68.22
DSEBM-e [21]	86.19	64.66	73.99	46.67	45.65	46.01	67.79	68.56	68.18
DAGMM [23]	92.97	94.42	93.69	49.09	50.78	49.83	80.77	81.6	81.19
AnoGAN [13]	87.86	82.97	88.65	41.18	43.75	42.42	71.19	72.03	71.59
ALAD [20]	94.27	95.77	95.01	50	53.13	51.52	79.41	80.32	79.85
CADGMM	**96.01**	**97.53**	**96.71**	**56.41**	**57.89**	**57.14**	**81.99**	82.75	**82.37**

Consistent performance improvement!

Experiment

Results

Table 4. Anomaly Detection Performance on KDD99 with different ratios of anomalies during training.

Radio	CADGMM			DAGMM			OC-SVM		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
1%	95.53	97.04	96.28	92.01	93.37	92.68	71.29	67.85	69.53
2%	95.32	96.82	96.06	91.86	93.40	92.62	66.68	52.07	58.47
3%	94.83	96.33	95.58	91.32	92.72	92.01	63.93	44.70	52.61
4%	94.62	96.12	95.36	88.37	89.89	89.12	59.91	37.19	45.89
5%	94.35	96.04	95.3	85.04	86.43	85.73	11.55	33.69	17.20

Less sensitive to noise data!
More robust!

Experiment

Results

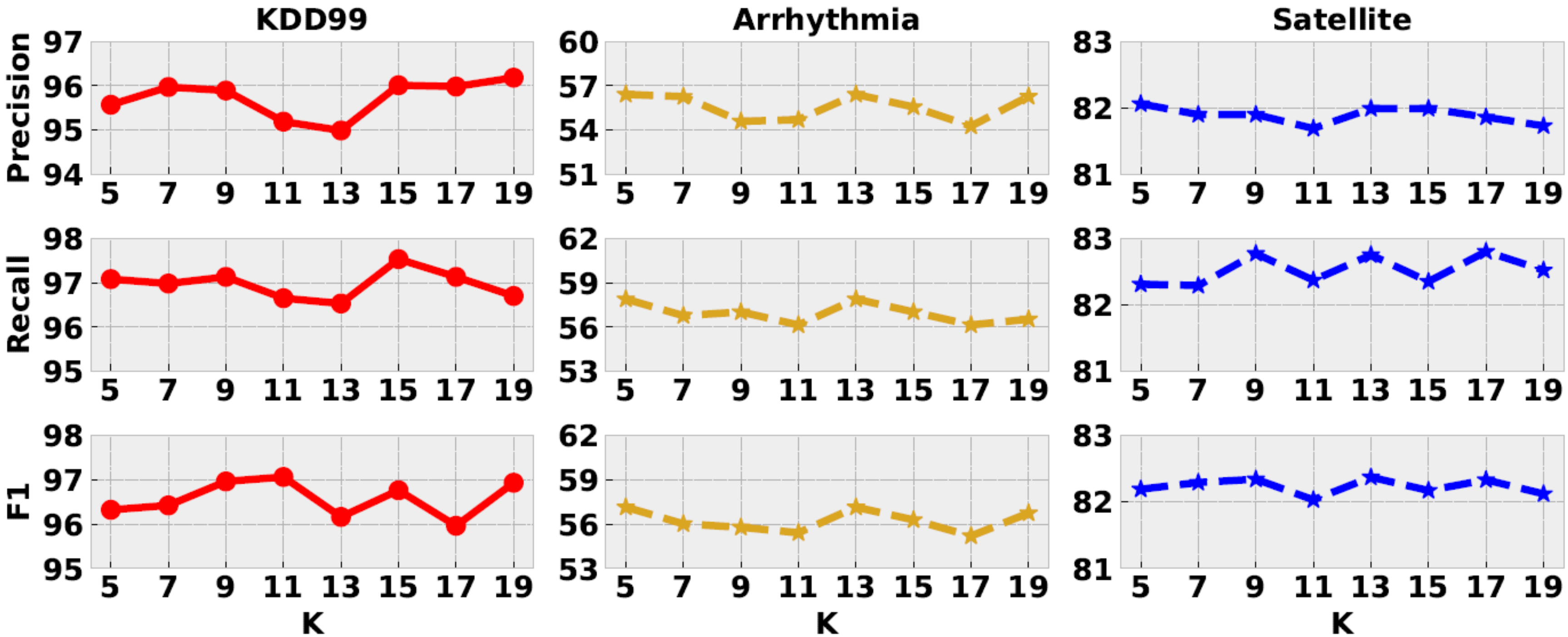
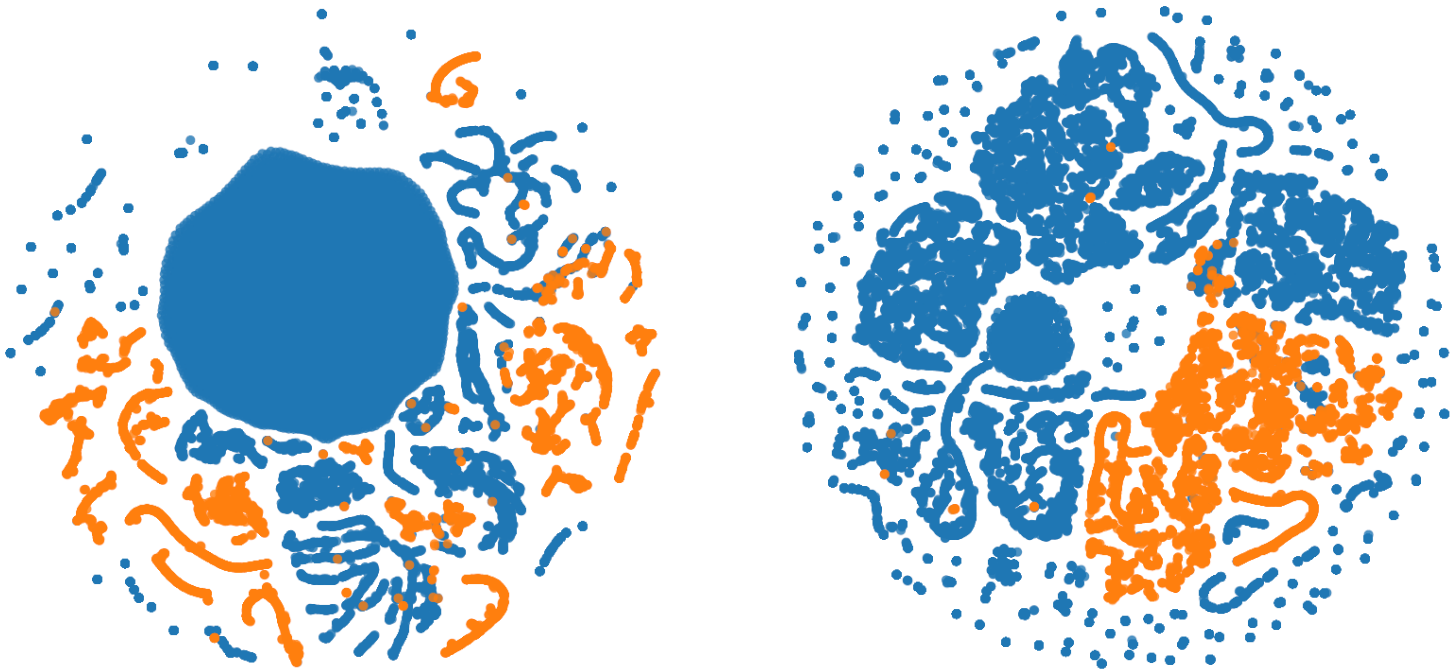


Fig. Impact of different K values of K-NN algorithms in graph construction.

Less sensitive to hyper-parameters! Easy to use!

Experiment

Results



(a). DAGMM (b). CADGMM

Fig. Embedding visualization on KDD99 (Blue indicates the normal samples and orange the anomalies).

Explainable and Effective!

Conclusion and Future Works

- **Conventional feature learning models cannot effectively capture the correlation among data samples for anomaly detection.**

- **We propose a general representation learning framework to model the complex correlation among data samples for unsupervised anomaly detection.**

- **We plan to explore the correlation among samples for extremely high-dimensional data sources like image or video.**

- **We plan to develop an adaptive and learnable graph construction module for a more reasonable correlation modeling.**

Reference

- **[OC-SVM]** Chen, Y., Zhou, X.S., Huang, T.S.: One-class svm for learning in image retrieval. **ICIP**. 2001
- **[IF]** 8. Liu, F.T., Ting, K.M., Zhou, Z.H.: Isolation forest. **ICDM**. 2008.
- **[DSEBM]** Zhai, S., Cheng, Y., Lu, W., Zhang, Z.: Deep structured energy based models for anomaly detection. **ICML**. 2016.
- **[DAGMM]** Zong, B., Song, Q., Min, M.R., Cheng, W., Lumezanu, C., Cho, D., Chen, H.: Deep autoencoding gaussian mixture model for unsupervised anomaly detection. **ICLR**. 2018.
- **[AnoGAN]** Schlegl, T., Seeb□ck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G.: Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. **IPMI**. 2017.
- **[ALAD]** Zenati, H., Romain, M., Foo, C.S., Lecouat, B., Chandrasekhar, V.: Adversarially learned anomaly detection. **ICDM**. 2018.

Thanks

Thanks for listening!

Contact: isfanhy@hrbust.edu.cn
Home Page: https://haoyfan.github.io/