Boundary of Distribution Support Generator (BDSG): Sample Generation on the Boundary

ICIP 2020

The University of Edinburgh, UK
School of Engineering
Institute for Digital Communications (IDCOM)

Dr Nikolaos Dionelis | Dr Mehrdad Yaghoobi | Prof. Sotirios A. Tsaftaris

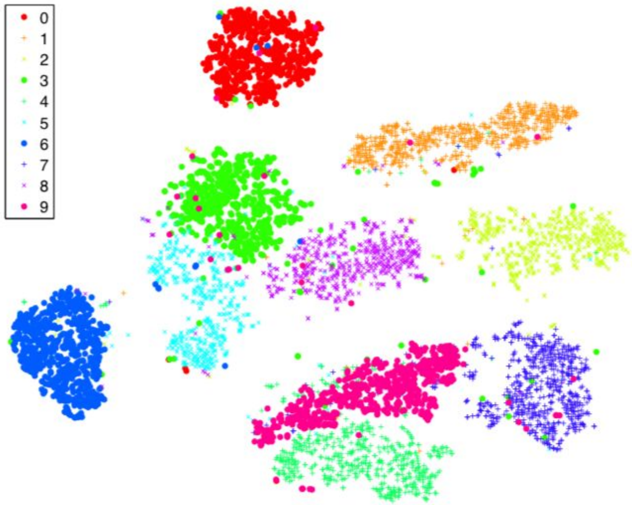


Boundary of Distribution Support Generator

- **Focus on**: Generative Models for Anomaly Detection (AD)
- Create the Boundary of Distribution Support Generator (BDSG) model
- Address limitations of current state-of-the-art:
  - Multimodal distributions
  - Support with disjoint components
    - Mode collapse; Probability density
  - Boundary of distribution support
- Improve the AD methodology
- Create an objective cost function to force the generated samples to the boundary of the data distribution

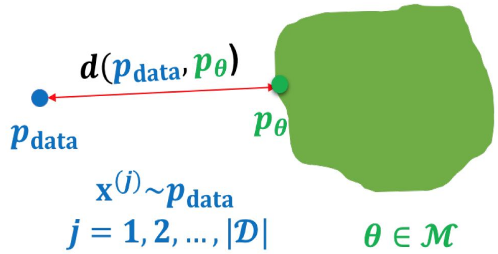
Generative Models for AD

- Promising framework:
  - GAN-AEs; VAEs; Invertible Generative Models
- Compute probability at any point in the data space
- Model definition and training:
  - Architecture; Loss function
  - Minimization; Convergence
- Reduce false negative errors
  - Mises, Type II errors
  - Address false positives
    - False Alarms, Type I errors
- Leave-one-out evaluation
  - AUROC; AUPRC; F1 score



L. van der Maaten and G. Hinton, Visualizing Data using t-SNE, (http://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf)

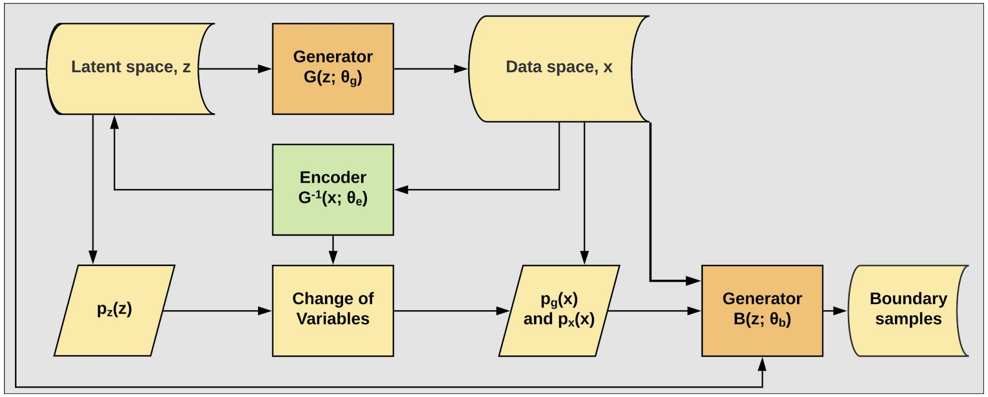
Discernible Limitations for Practical AD



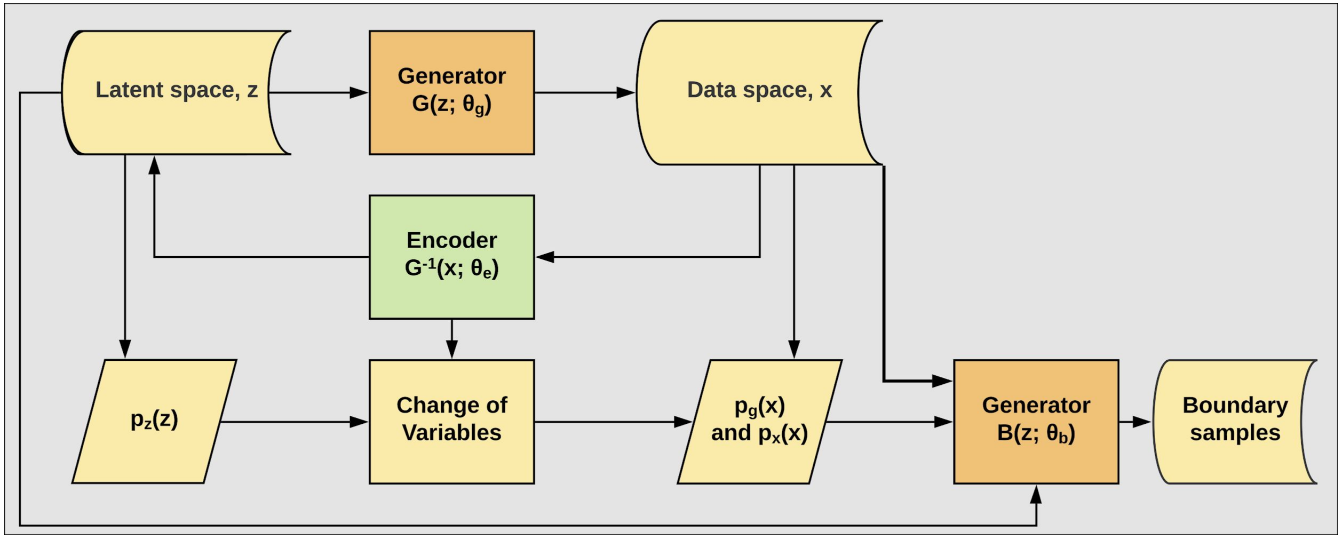
- Improve performance on benchmark datasets
- Shortcomings of current methodologies:
  - Dataset not normal or partially labelled
    - Fit model: Learn normal data distribution
  - Leave-one-out evaluation
    - Anomalies not confined to a finite annotated set
    - Complement of support; Lack of strong anomalies
  - Rarity problem: Sampling complexity
- Current methodologies are problematic:
  - For detecting the boundary of multimodal distributions
- **Aim**: Address such challenges

Flowchart of Proposed Boundary Generator

- Perform sample generation on the boundary
- Generate samples on the boundary of the data distribution:
  - Train an invertible model to fit the normal data distribution
  - Invertible Residual Networks for density estimation
    - Learn Generator G(**z**) and G^{-1}(**x**)
  - Given data distribution, $p_{\mathbf{x}}(\mathbf{x})$: Approximate with $p_g(\mathbf{x})$
  - Create and train B(**z**; $\boldsymbol{\theta}_b$) to generate boundary samples
- B(**z**; $\boldsymbol{\theta}_b$) = Mapping from latent space, **z**, to data space, **x**



Proposed BDSG Boundary Generator



- Run Gradient Descent on proposed loss function
  - Penalize probability and distance from normality
  - Avoid mode collapse: Dispersion, scattering
- Create a cost function that forces the generated samples to the boundary of the support of the data distribution

Cost Function of Boundary Generator

$$L(\boldsymbol{\theta}_b, \mathbf{z}, \mathbf{x}, G, \lambda_1, \lambda_2) = L_0(\boldsymbol{\theta}_b, \mathbf{z}, G) + \lambda_1 L_1(\boldsymbol{\theta}_b, \mathbf{z}, \mathbf{x}) + \lambda_2 L_2(\boldsymbol{\theta}_b, \mathbf{z})$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left[p_g(B(\mathbf{z}_i;\boldsymbol{\theta}_b)) + \lambda_1 \min_{j=1}^{M} ||B(\mathbf{z}_i;\boldsymbol{\theta}_b) - \mathbf{x}_j||_2 + \lambda_2 \frac{1}{N-1}\sum_{j=1,\, j\neq i}^{N} \frac{||\mathbf{z}_i - \mathbf{z}_j||_2}{||B(\mathbf{z}_i;\boldsymbol{\theta}_b) - B(\mathbf{z}_j;\boldsymbol{\theta}_b)||_2}\right]$$

- N = Batch size; M = Sample size
- L_0: Penalize probability density to find the boundary
- L_1: Distance from a point to a set
    - Penalize distance from normality
- L_2: Scattering, dispersion, and diversity
    - Avoid mode collapse

First Term of BDSG Model

- Use change of variables formula:

$$L_0(\boldsymbol{\theta}_b, \mathbf{z}, G) = \frac{1}{N}\sum_{i=1}^{N} p_g(B(\mathbf{z}_i;\boldsymbol{\theta}_b))$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left[p_{\mathbf{z}}(G^{-1}(B(\mathbf{z}_i;\boldsymbol{\theta}_b)))\,|\det \mathbf{J}_G(B(\mathbf{z}_i;\boldsymbol{\theta}_b))|^{-1}\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left[\exp(\log(p_{\mathbf{z}}(G^{-1}(B(\mathbf{z}_i;\boldsymbol{\theta}_b)))) - \log(|\det \mathbf{J}_G(B(\mathbf{z}_i;\boldsymbol{\theta}_b))|))\right]$$

- L_0: $B(\mathbf{z})$, $G^{-1}(\mathbf{x})$, det $\mathbf{J}_G(\mathbf{x})$, $p_{\mathbf{z}}(\mathbf{z})$
- Standard Gaussian distribution, $\mathbf{z} \sim N(\mathbf{0}; \mathbf{I})$
- Inference: Anomaly if $p_g(\mathbf{x}) < \varepsilon$ and normal o/w

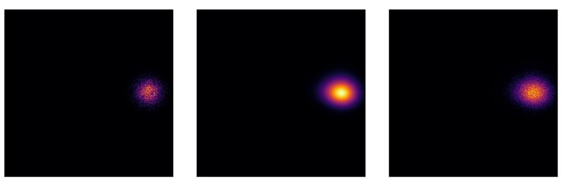
Evaluation of BDSG Model

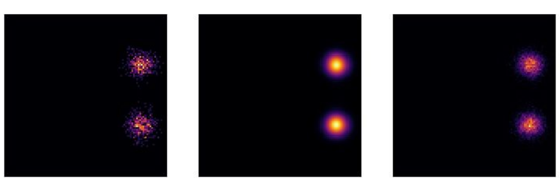
- **Datasets:** Synthetic data; MNIST; CIFAR-10
- **Evaluation for AD:** Leave-one-out; OoD data
- **Baselines:**
  - GANomaly, EGBAD, VAE, AnoGAN, FenceGAN
  - WGAN, MinLGAN
- **Evaluation metrics:**
  - Algorithm convergence criteria; AUROC; AUPRC
- **OoD data:**
    - Fashion-MNIST; KMNIST; QMNIST
    - CIFAR-100; SVHN; STL-10

Boundary Formation of BDSG Model

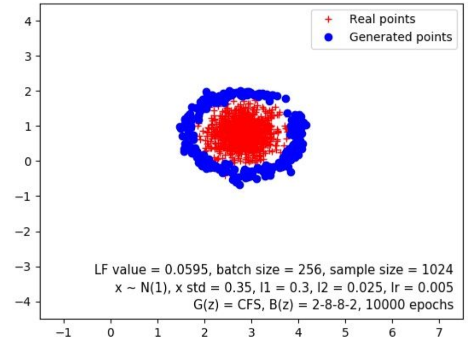
- Synthetic data: Unimode and multimodal distributions

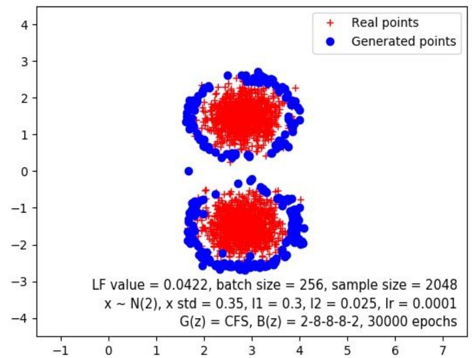
(a) IResNet: Input samples to IResNet (left), output probability density (middle), and output samples (right)



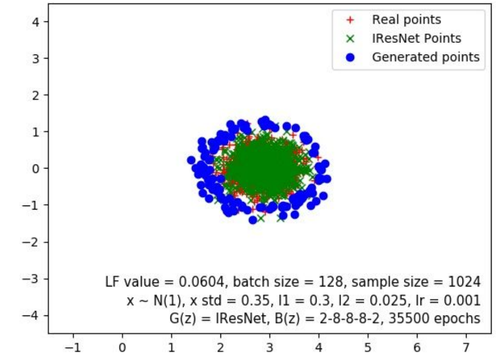


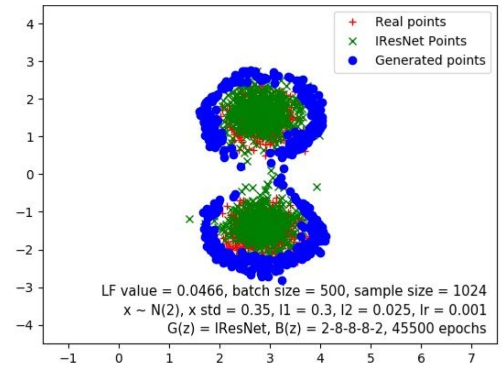
(b) Closed-Form Solution for BDSG Model





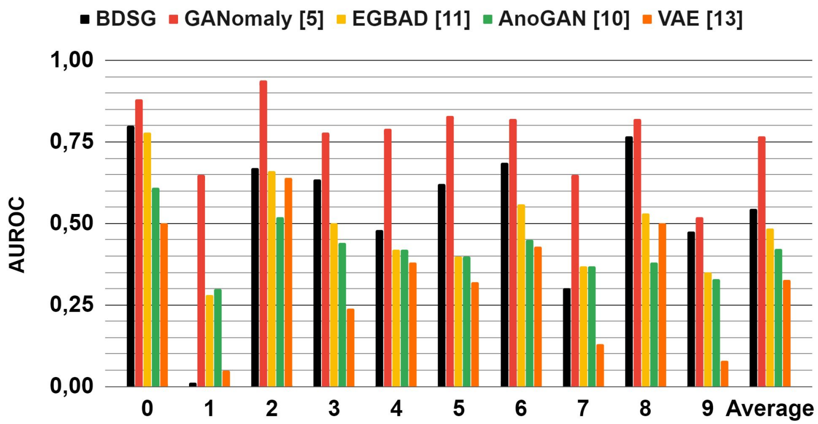
(c) BDSG Model; B(z)

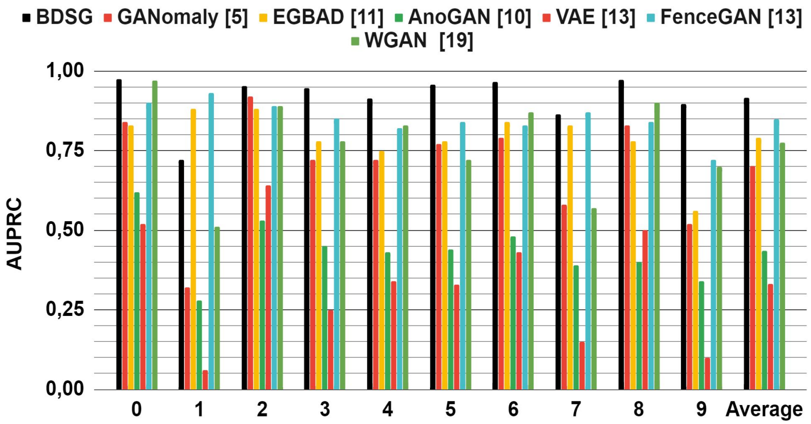




AD Performance of BDSG Model

- Evaluation on MNIST: Leave-one-out methodology
  - Leave-out class (horizontal axis) is the anomaly class
  - BDSG: Competitive performance

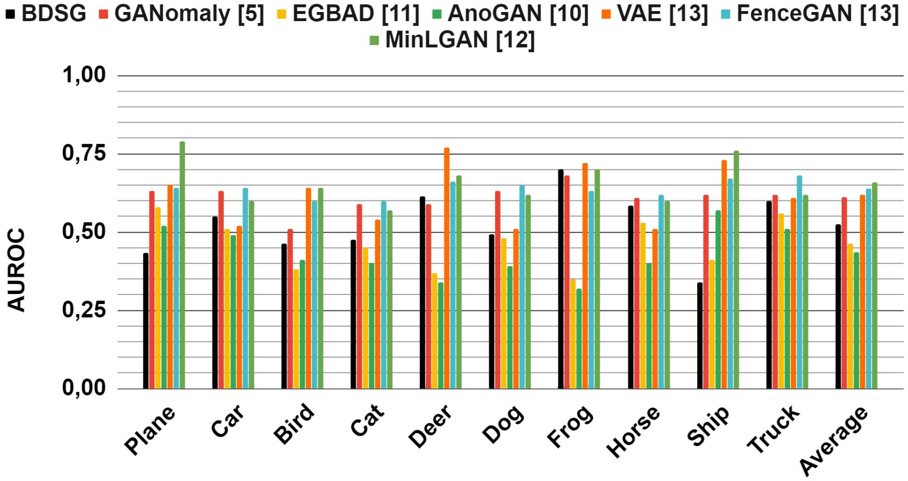


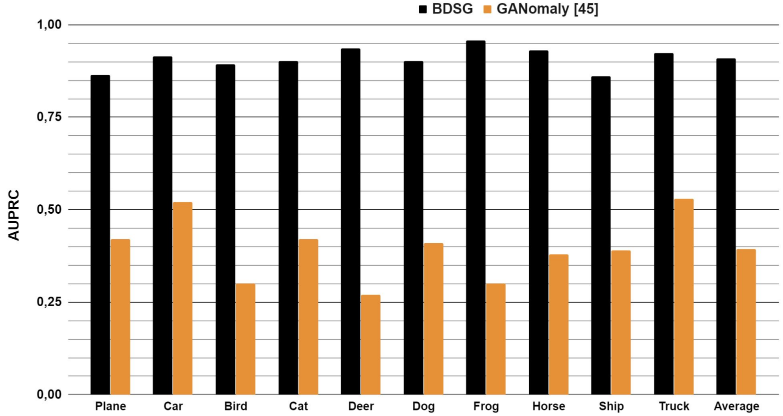


Performance of Proposed BDSG

- Evaluation on CIFAR-10: Leave-one-out methodology

- Leave-out class (horizontal axis) is the anomaly class
- BDSG: Good performance



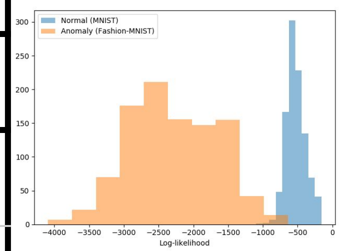


Performance of Proposed BDSG

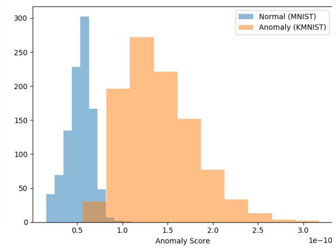
- Evaluation on MNIST: OoD data

MNIST	Loss	L1	L2
MNIST Digits 1-9	**0,74**	**0,93**	**18,26**
MNIST Digits 0	20,36	66,32	18,40
Fashion-MNIST	9,92	31,44	19,44
KMNIST	9,28	29,37	18,73

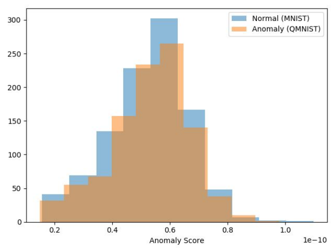
- Histograms:



Fashion-MNIST
AUROC: 0.9996
AUPRC: 0.9997



KMNIST
AUROC: 0.9973
AUPRC: 0.9977

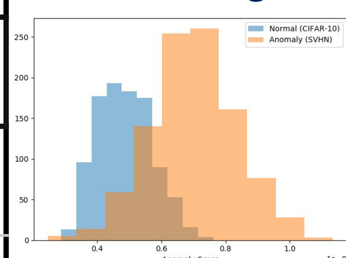


QMNIST
AUROC = 0.5171
AUPRC = 0.5126

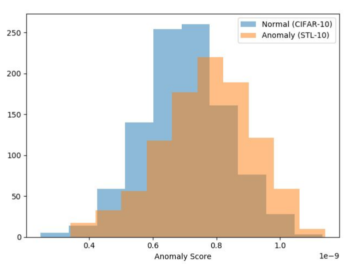
- Evaluation on CIFAR-10: OoD data

CIFAR-10	Loss	L1	L2
CIFAR-10 Digits 0-9	**3,16**	**8,94**	**19,28**
CIFAR-100	7,50	23,43	18,77
SVHN	7,18	22,36	19,07
STL-10	10,00	31,75	19,03

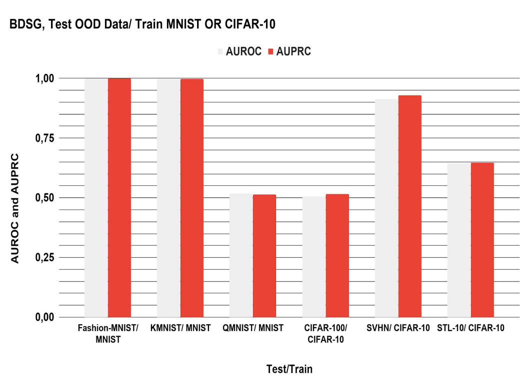
- Histograms:



SVHN
AUROC: 0.9120
AUPRC: 0.9295



STL-10
AUROC: 0.6439
AUPRC: 0.6472



AUROC, AUPRC: Test OOD/ Train MNIST or CIFAR-10

Conclusion

- Determination of the boundary of the data distribution for AD
- Create the BDSG model for AD:
  - Learn the mapping from **z** to **x** concentrating the images of **z** on the support boundary
  - Minimize a cost function to force the generated samples to the boundary of the data distribution
    - Support with disconnected components
  - Address the problem of detecting strong anomalies
  - Create an algorithm for sample generation on the boundary obviating the rarity and sampling complexity problem
- Achieve competitive performance on (i) synthetic data from multimodal distributions, and (ii) MNIST and CIFAR-10