RF One-Class Classification via Shifted DeepSVDD

Ted Reehorst[†] and Phil Schniter[*]

[†]Air Force Research Laboratory, Wright-Patterson AFB, OH (edward.reehorst@us.af.mil)

[*]Dept. ECE, The Ohio State University (schniter.1@osu.edu)



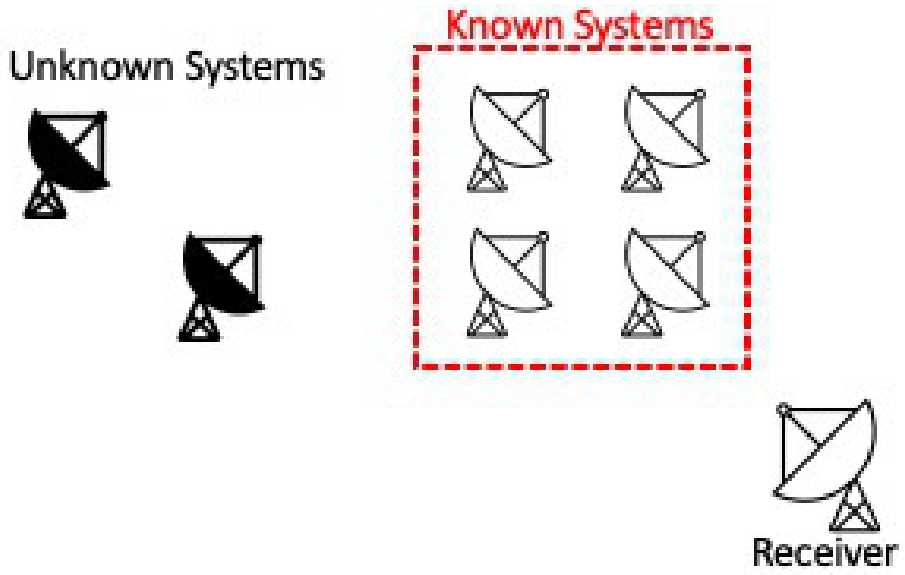
THE OHIO STATE UNIVERSITY

Asilomar Conference on Signals, Systems, and Computers

October 2026

Motivating Application: RF Spectrum Monitoring

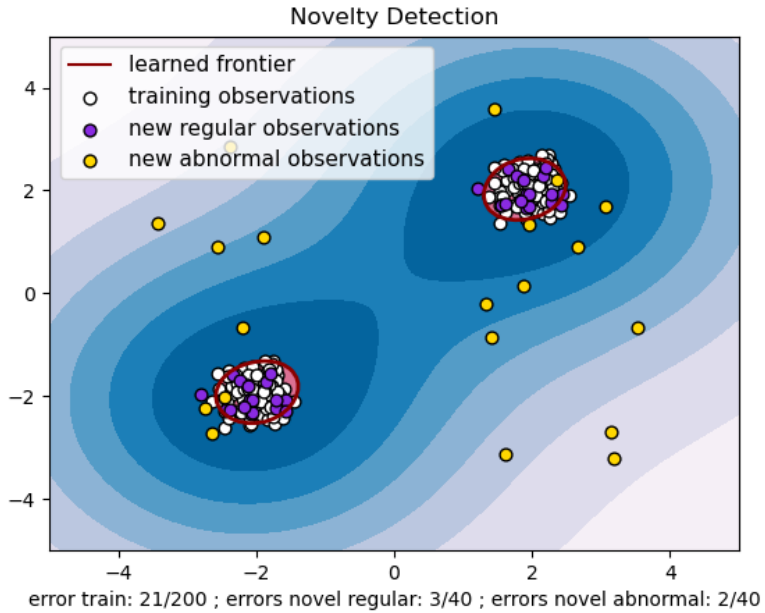
- We have examples of approved transmission types
- Would like to detect non-approved transmissions
- This is known as one-class classification, or novelty detection, or OOD detection, or unsupervised anomaly detection
  - Important: we have training examples of inliers, but not of outliers



One-class classification: low dimensions

For low-dimensional signals, many density-estimation techniques work well:

- GMMs
- K-nearest-neighbors
- kernel-density estimation
- Mahalanobis distance

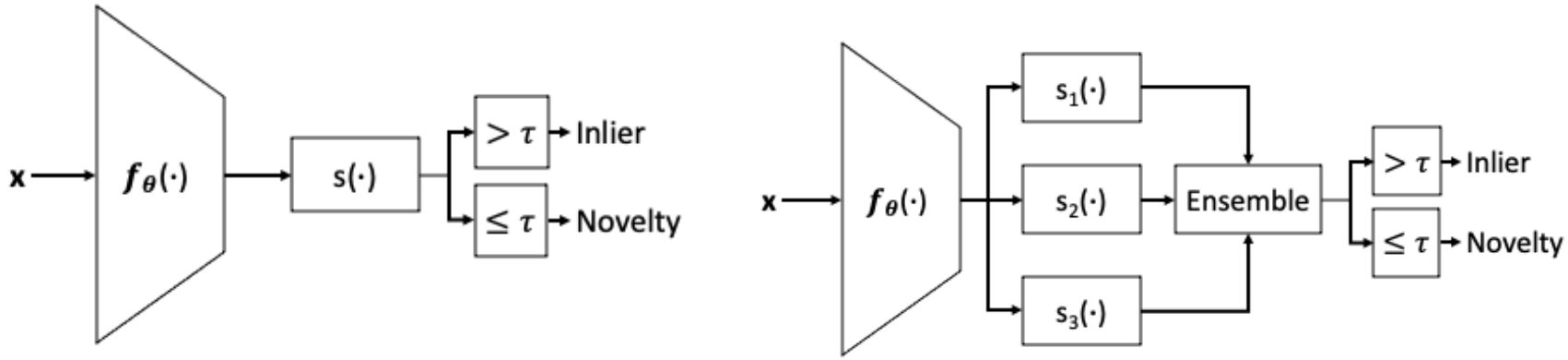


But with high-dimensional signals, we don't have enough data to accurately model the density (the curse of dimensionality)

One-class classification: high dimensions

With high-dimensional signals (e.g., images), novelty detectors include:

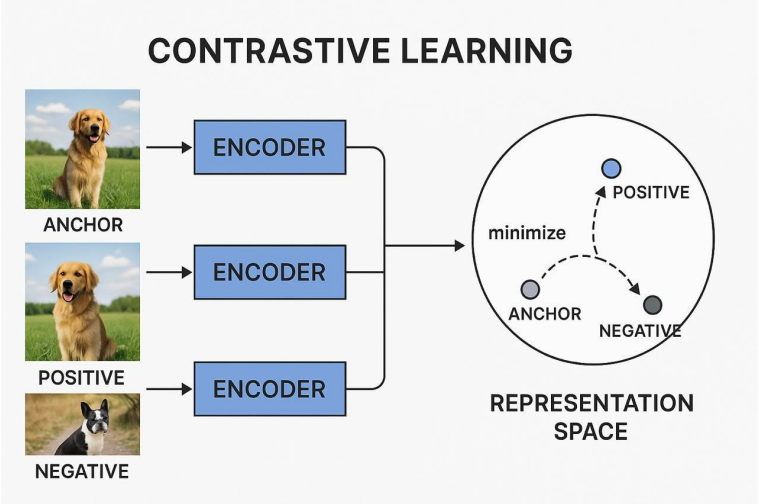
- Kernel-based methods like kernel-PCA, OC-SVM, and SVDD
- Methods that use a deep neural network (DNN) for dimensionality reduction followed by some form of scoring or ensembling



Dimensionality reduction for novelty detection

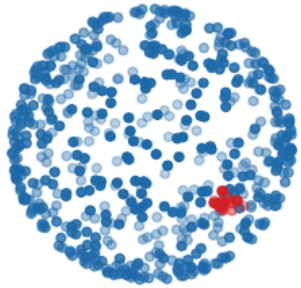
Dimensionality reduction is tricky, since we have no outlier data!

- Want to reduce dimensionality w/o throwing away discriminative information
- If we had multiple inlier classes ("open set" classification), we could train a DNN for classification and use the penultimate features for novelty detection
  - Unfortunately we can't do this because we have only a single inlier class
- Instead, we will train our dimensionality reducer in a self-supervised manner, e.g., by learning to perform an auxiliary task like rotation prediction, or by contrastive learning



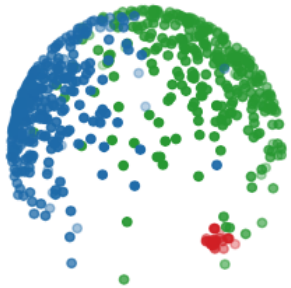
Distribution augmentation via "shifting" transforms

- With image data, Sohn et al. (2021) noticed that contrastive dimensionality reducers tend to over-crowd the feature space with inliers



- To remedy this, they proposed to augment the inlier data using the outputs of $p > 1$ "shifting" transformations $\boldsymbol{R}_j(\cdot)$:

$$\{\mathbf{x}_i\}_{i=1}^n \rightarrow \bigcup_{j=1}^{p} \{\boldsymbol{R}_j(\mathbf{x}_i)\}_{i=1}^n$$

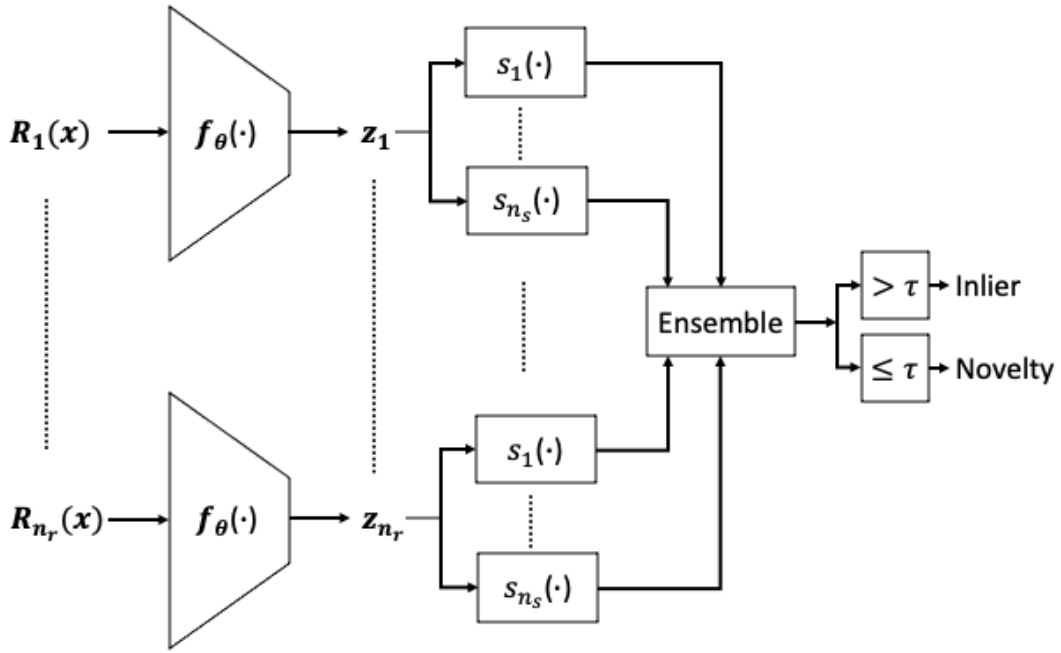


- The shifted samples are *semantically distinct* from the original images, but not completely unrelated:



Ensembling over shifted test samples

- Similar shifts were used in the CSI paper by Tack et al. (2020)
- But they went further by applying shifting transforms to the test sample (in addition to the training) to make novelty detection more robust:



- However, CSI used ad-hoc scores and simple sum-based ensembling:

$$\sum_{l=1}^{m}\sum_{j=1}^{p} s_l(\boldsymbol{f}_\theta(\boldsymbol{R}_j(\boldsymbol{x}))) \gtrless \tau$$

An improved ensembler

Reehorst & Schniter (2022) proposed an improved ensembling method:

- Given score functions $\{s_l(\cdot)\}_{l=1}^m$ designed to produce larger outputs for inliers than outliers . . .
- From the *inlier* training data, we learn a Gaussianizing nonlinear transformation that gives the z-scores

$$z_l(\boldsymbol{x}) \triangleq \Phi^{-1}\big(\widehat{F}_l(s_l(\boldsymbol{x}))\big) \sim \mathcal{N}(0,1) \text{ for inliers } \boldsymbol{x}$$

- The *outlier* z-scores are then expected to be negative (on average):

$$\alpha_l \triangleq \mathbb{E}\{z_l(\boldsymbol{x}) \mid \boldsymbol{x} \text{ is an outlier}\} < 0$$

- For tractability, we modeled the inlier z-scores as uncorrelated, and the outlier z-scores as identity-covariance Gaussian, giving the composite hypothesis test

$$\begin{aligned} \text{(inliers)} \quad & H_0 : \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}) \\ \text{(outliers)} \quad & H_1 : \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{\alpha}, \boldsymbol{I}), \quad \boldsymbol{\alpha} \in \mathbb{R}_-^c, \end{aligned}$$

An improved ensembler (cont.)

- To solve the composite hypothesis test

$$\begin{aligned} &\text{(inliers)} \quad H_0 : \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}) \\ &\text{(outliers)} \quad H_1 : \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{\alpha}, \boldsymbol{I}), \quad \boldsymbol{\alpha} \in \mathbb{R}^c_-, \end{aligned}$$

  we proposed to apply the GLRT

$$t_{\text{glrt}}(\boldsymbol{z}) \underset{H_0}{\overset{H_1}{\gtrless}} \tau \quad \text{for} \quad t_{\text{glrt}}(\boldsymbol{z}) \triangleq \log \frac{\max_{\boldsymbol{\alpha} \in \mathbb{R}^m_-} \mathcal{N}(\boldsymbol{z}; \boldsymbol{\alpha}, \boldsymbol{I})}{\mathcal{N}(\boldsymbol{z}; \boldsymbol{0}, \boldsymbol{I})}$$

- We furthermore showed that the test statistic simplifies to

$$t_{\text{glrt}}(\boldsymbol{z}) = (\tfrac{1}{2}\boldsymbol{z}^- - \boldsymbol{z})^{\mathsf{T}} \boldsymbol{z}^-, \quad \text{where componentwise} \quad z_j^- \triangleq \min\{z_j, 0\}$$

- Finally, we demonstrated that this GLRT-based detector improves on other ensembling-based novelty detectors for *image* data

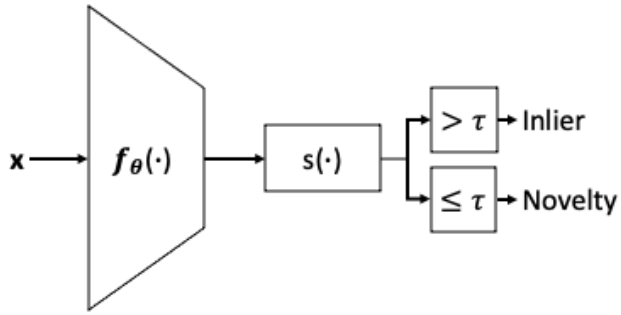
Novelty detection of RF data

Most recent works on novelty detection target *images*. There is much less work on novelty detection of RF waveforms:

- For multi-class inlier data, several "open set" methods have been proposed, but we focus on single-class inliers
- Most existing are adaptations of the OC-SVM, e.g., with hand-crafted kernels based on autocorrelation features (Pavy & Rigling, 2015).
- We were not successful in applying the contrastive methods of Sohn et al. (2021) or Tack et al. (2020) to RF data
- However, we found that DeepSVDD by Ruff et al. (2018) works relatively well with RF data, and we use it as a foundation

Deep SVDD

- In DeepSVDD, a DNN is used to reduce dimensionality and the resulting features are used in an SVDD-like score function



- The DNN f_θ is trained to minimize the loss

$$\mathcal{L} = \frac{1}{n}\sum_{i=1}^{n} \|f_\theta(x_i) - c\|^2 + \frac{\lambda_0}{2}\sum_{k=1}^{K} \|W_\theta^k\|_F^2 \quad \text{with } c = \left[\frac{1}{n}\sum_{i=1}^{n} f_{\theta_0}(x_i)\right]_{|\cdot|>\epsilon}$$

  where W_θ^k is the layer-k weight matrix and $[\cdot]_{|\cdot|>\epsilon}$ denotes elementwise projection away from 0 in magnitude

- Note c is determined from the initial parameters θ_0 and fixed during training

- Once trained, DeepSVDD uses the inlier score $s(x) = -\|f_\theta(x) - c\|$

Our contribution

Now that we've covered the background...

In this work, we propose to combine

- DeepSVDD
- shifting transforms
- GLRT-based score combining

for novelty detection of RF waveforms

The proposed "Ensembled Shifted DeepSVDD"

- Given the p shifting transformations $\{\boldsymbol{R}_j\}_{j=1}^p$, we compute the center vectors

$$\boldsymbol{c}_j = \left[\frac{1}{n}\sum_{i=1}^n \boldsymbol{f}_{\boldsymbol{\theta}_0}(\boldsymbol{R}_j(\mathbf{x}_i))\right]_{|\cdot|>\epsilon} \text{ for } j = 1, \ldots, p$$

  from inliers $\{\mathbf{x}_i\}_{i=1}^n$ and initial DNN parameters $\boldsymbol{\theta}_0$

- Then we train the DNN to minimize the loss

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{j=1}^p \frac{\lambda_j}{n}\sum_{i=1}^n \|\boldsymbol{f}_{\boldsymbol{\theta}}(\boldsymbol{R}_j(\mathbf{x}_i)) - \boldsymbol{c}_j\|^2 + \frac{1}{2}\sum_{k=1}^K \|\boldsymbol{W}_{\boldsymbol{\theta}}^k\|_F^2$$

  with hyperparmeters $\{\lambda_j\}_{j=1}^p$

- Finally we define the p inlier scores

$$s_j(\mathbf{x}) = -\|\boldsymbol{f}_{\boldsymbol{\theta}}(\boldsymbol{R}_j(\mathbf{x}))\| \text{ for } j = 1, \ldots, p$$

  and ensemble them using the GLRT scheme described earlier

Application to RadioML-2018 Data

- The RadioML-2018 dataset contains
  - 24 modulation types: OOK, 32PSK, 16QAM, AM-DSC-WC, FM, etc
  - for each type, there's 106 496 complex-valued RF waveforms of length 1024
  - the waveforms span SNRs between -20 and 30 dB

- We set up 5 experiments. In each, the inliers were a single modulation type in
  $\{$OOK, 32PSK, 16QAM, AM-DSB-WC, FM$\}$
  and the outliers contained the other modulation types

- After trying various shifting transforms (freq shift, time shift, phase shift, time reverse, permutation) on calibration data, we chose to use the $p = 2$ shifts
  $\{$no shift, frequency shift of 0.2π rad/samp $\}$

Methods evaluated

We evaluated the following methods:

- One-class SVM on raw RF waveforms
- One-class SVM on autocorrelation features, as in Pavy & Rigling (2015)
- Standard DeepSVDD, as in Ruff et al. (2018)
- Shifted DeepSVDD with sum-based score combining
- Shifted DeepSVDD with GLRT-based score combining

Each DeepSVDD variant was tested on 10 different random initializations $\boldsymbol{\theta}_0$

AUC performance on RadioML 2018

AUC averaged over 10 random trials (± standard error)

Method	Inlier modulation type					Avg.
	OOK	32PSK	16QAM	AM-DSB-WC	FM	
OCSVM – raw RF	0.719	0.500	0.486	0.790	0.794	0.658
OCSVM – ACF	0.722	0.623	0.600	**0.877**	0.674	0.700
DeepSVDD	0.677±0.082	0.710±0.019	0.673±0.012	0.827±0.007	0.926±0.011	0.763
ShiftedDeepSVDD – sum	0.827±0.007	0.769±0.052	**0.697**±0.016	0.682±0.020	0.816±0.032	0.758
ShiftedDeepSVDD – GLRT	**0.844**±0.008	**0.801**±0.029	**0.697**±0.011	0.842±0.012	**0.944**±0.005	**0.826**

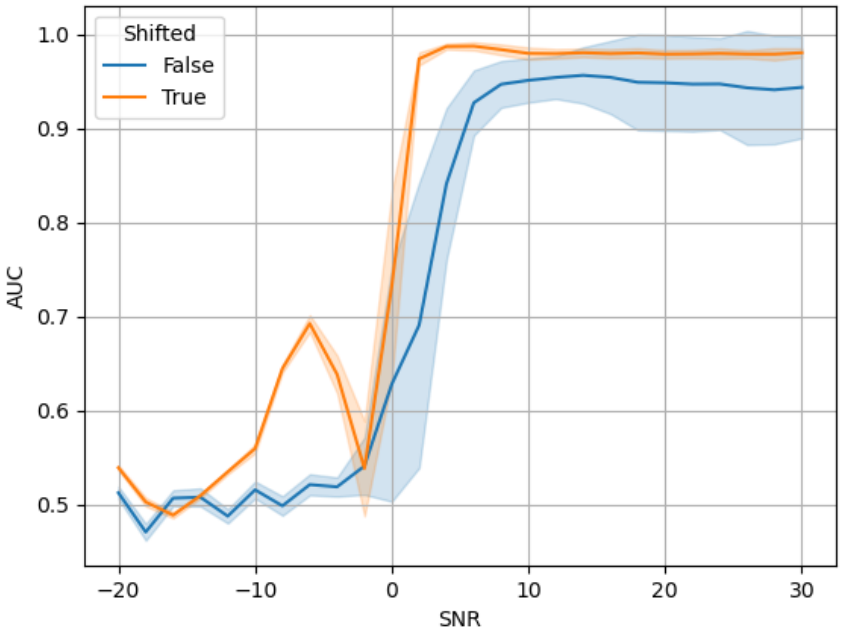
The proposed ShiftedDeepSVDD with GLRT ensembling...

- wins in 4 out of 5 experiments
- wins on average over the 5 experiments
- beats standard DeepSVDD in all experiments
- beats or ties sum-combined ShiftedDeepSVDD in all experiments

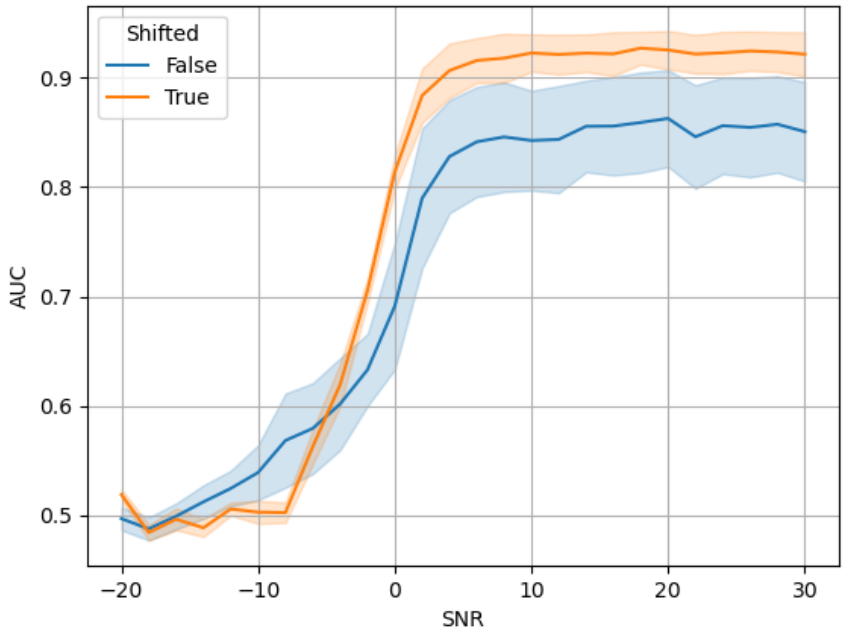
AUC vs SNR

AUC of standard & Shifted DeepSVDD vs SNR:

OOK inliers:



32-PSK inliers:



Conclusion

In this work, we combined

- DeepSVDD
- shifting transforms
- GLRT-based score combining

for novelty detection of RF waveforms

Performance on the RadioML-2018 dataset shows improved AUC over existing OC-SVM methods, and over standard and sum-combined DeepSVDD