Using UMAP to Inspect Audio Data for Unsupervised Anomaly Detection under Domain-Shift Conditions

Andres Fernandez, Mark D. Plumbley

{andres.fernandez, m.plumbley}@surrey.ac.uk

Centre for Vision, Speech and Signal Processing (CVSSP)
University of Surrey, UK

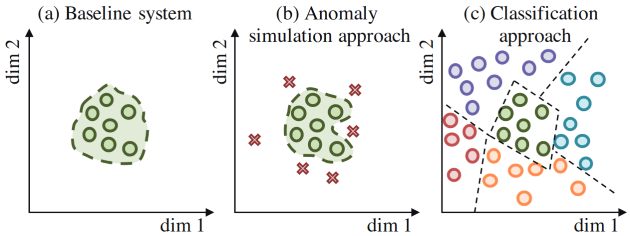
DCASE Workshop, 15-19 November 2021

Contents

Unsupervised Anomaly Detection (UAD)

Limited knowledge about anomalies → **unsupervised task**:

- Only non-anomalous (*normal*) data known.
- Detector provides anomaly score, then threshold.



Different UAD strategies (from [Koizumi20]). Only normal data (green) is known beforehand.
Common assumption: Normal data clustering around modes, anomalies further away.

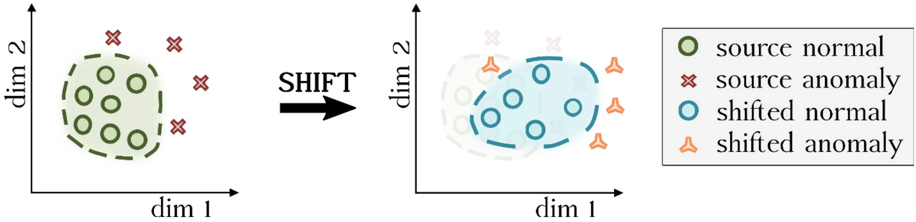
Audio UAD:

- Useful in predictive maintenance of industrial machines[Carvalho19, Grollmisch19].
- Sounds may **overlap** with variable signal-to-noise ratios.
- Anomalies may be given by **short sounds** or **large sequences**.

UAD under domain-shift conditions (UAD-S)

Normal **and** anomalous data exposed to potentially unknown changes:

- Detector must embrace strong environmental changes.
- But still be sensitive to even slight anomalies.



UAD-S illustration. Maintaining the green boundary would result in false positives/negatives.

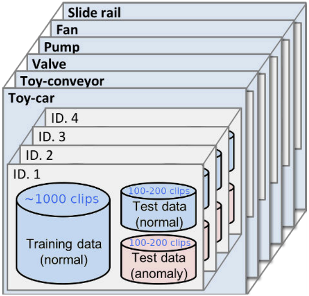
UAD/UAD-S in DCASE

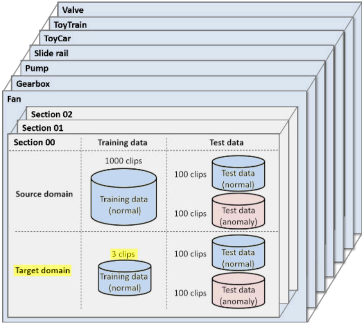
UAD (2020, Task 2)

- Machines intentionally damaged for anomaly eval data[MIMII, ToyAdmos].
- Evaluation: AUC and pAUC, averaged over machines.
- DL dominated. Variety of approaches with $> 90\%$.

UAD-S (2021, Task 2)

- Larger dataset. Added target domain with $\sim 0.3\%$ training data[MIMII_DUE, ToyAdmos2].
- Evaluation: AUC/pAUC, harmonic mean over machines+domains.
- Results drastically lower (1^{st} : 66.8%).





Development datasets for DCASE2021 (left) and 2020 (right). Images from [DCASE website].

Motivation and Proposed Contributions

Why 2021 $\ll$ 2020?

- DL-based reconstruction approach from 2020 was barely present in top ranks (exception: [Kuroyanagi21]).
- Instead, predominance of representation-learning methods.
- Established domain-shift techniques were attempted[Lopez21].
- Relative success of non-parametric inference methods[Morita21].

Literature hints at task difficulty independently of the approach, and relevance of data representation. Plan: **visually inspect data**.

Proposed contributions

- Methodology+software using UMAP projections to **visualize audio data** and find **separability** and **discriminative support**.
- Insights on micro- and macrostructure of the 2021 UAD-S dataset.
- Formulation of verifiable hypotheses to direct future efforts.

Contents

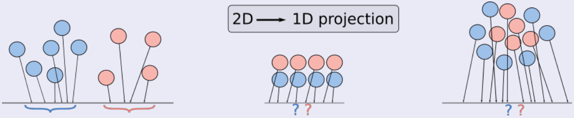
UMAP and Dimensionality Reduction

Uniform Manifold Approximation and Projection (UMAP) is a scalable, stable and meaningful projection technique[McInnes18]:

1. For each point, find radius based on distance to $k \geq 1$ neighbours.
2. Compute radius-dependent distance and $P(neigh)$ for neighbours.
3. Find low-dimensional graph with edges that minimize cross-entropy to high-dimensional graph (e.g. via SGD):

$$\sum_{e \in E} \overbrace{w_h(e) log \left(\frac{w_h(e)}{w_l e} \right)}^{\text{attraction for high } w} + \overbrace{(1 - w_h(e)) log \left(\frac{1 - w_h(e)}{1 - w_l e} \right)}^{\text{repulsion for low } w}$$

Main assumption:

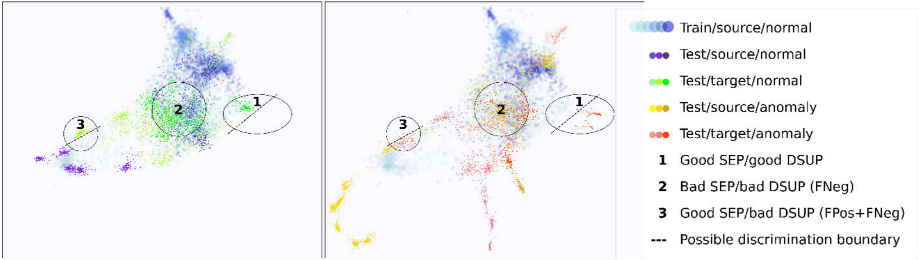


If two regions appear separable on the projection, they are also separable in the original domain. The converse is not necessarily true!

Plotting UAD-S Audio Data

Based on our *main assumption*, 2 beneficial **visual qualities**:

- **Separability (SEP):** Simple boundary between normal and anomalous.
- **Discriminative Support (DSUP):** Training data covers normals, not anomalies.



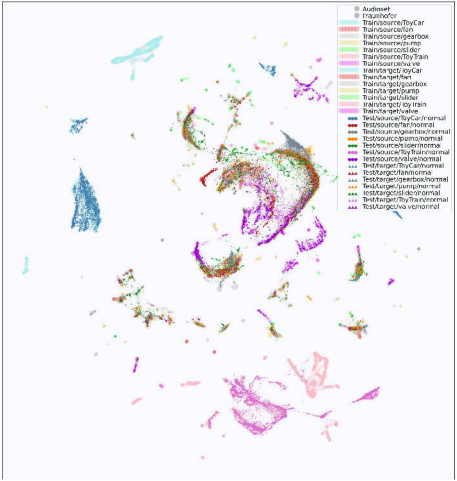
Excerpt from the UMAP plot for *pump*, illustrating different cases of SEP/DSUP.

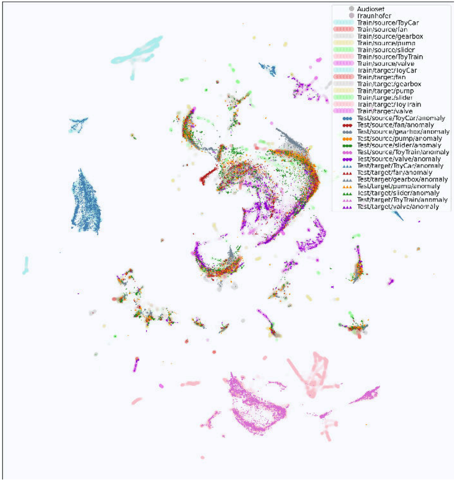
We render global, per-device and per-split plots:

- *Training* data is shown equally on both sides.
- *Normal test* data is shown on the left, *anomalies* on the right.
- Representations: log-STFT, log-mel and L3[Arandjelovic17].

Result Samples: Global Plot

- Clear distinction between *ToyAdmos* (teal, pink) and *MIMII* (rest).
- Generally, **little observable difference** between normals and anomalies, and no evident anomalous patterns or regions.

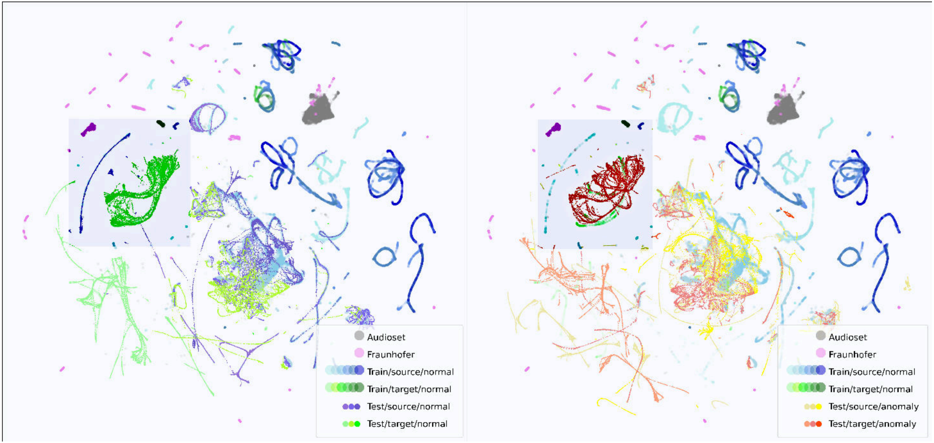




1024-Log-STFT spectrogram→sample→UMAP of all devices.

Result Samples: Device Plot+External Data

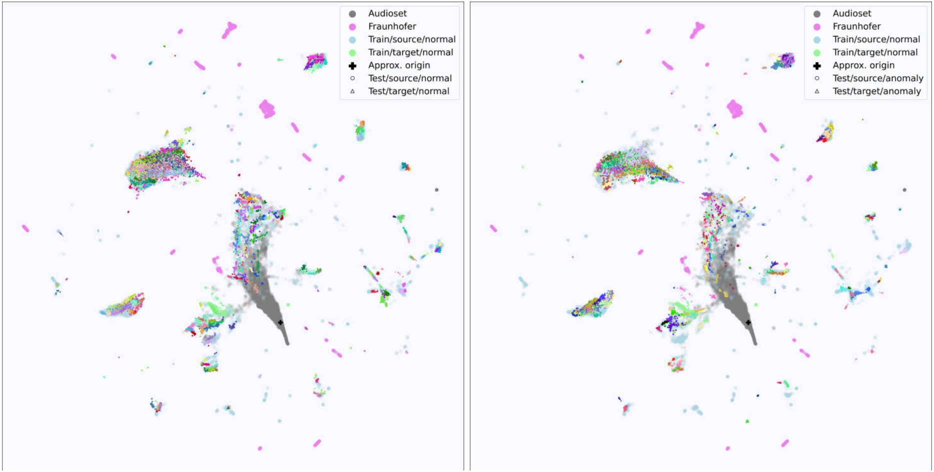
- AudioSet (L3 pretraining data) appears compact, rest scattered.
- ToyAdmos structures simpler than MIMII (e.g. *ToyTrain* below).
- Interesting cycles (circular train trajectory?).



OpenL3→sample→UMAP of the *train* device. Highlighted: noticeable anomalous patterns.

Result Samples: Section Plot

- Individual clips can be inspected.
- Tip of the non-negative cone (i.e. zero amplitude) populated by AudioSet.



128-Log-mel spectrogram→sample→UMAP of *gearbox* section 0.

Contents

Discussion

Based on observations+literature, we propose these **hypotheses**:

1. Mixing ToyAdmos+MIMII may hinder performance[Primus20].
2. Temporal context+pretraining: tradeoff between SEP and DSUP.
3. Normalization is a dominating factor [Lopez21, Wilkinghoff21].
4. Incorporating domain-related priors may help performance (see image).

Machine type	Section	Description	Attribute format in file names of training data
Fan	00	Wind strength variations between domains	strength_<wind_strength_index>_ambient
Fan	01	Two products from the same manufacturer with size variations between domains	strength_1_<product_nickname>_ambient <product_nickname> : big or small
Fan	02	Factory noise variations between domains	Source domain: strength_1_ambient Target domain: strength_1_Fact_D_ambient
Gearbox	00	Voltage variations between domains	<weight>_g_<arm_length>_mm_<voltage>_mV_none

The 2021 DCASE dataset provides labels for the domain shifts.

We also present **shortcomings** and possible ways to address them:

- Computational limitations: Projecting, subsampling and stacking.
- Perceptual biases: Shapes, colors, amount of information.

Thank You

- Conclusions:
  1. Presented methodology+software tool for UAD-S analysis.
  2. Revealed interesting properties of the dataset.
  3. Proposed verifiable hypotheses based on plots and literature.

- Future Work:
  1. Incorporate further representations and temporal relations.
  2. Enhanced, interactive plot (sonification, highlighting).
  3. Extension to supervised scenarios.

Come to our poster session! More info & resources:
`https://ai4s.surrey.ac.uk/2021/dcase_uads`

References (1)

[Arandjelovic17] Relja Arandjelovic and Andrew Zisserman. “Look, Listen and Learn”. In: *ICCV 2017*.

[Carvalho19] Thyago P. Carvalho et al. “A systematic literature review of machine learning methods applied to predictive maintenance”. In: *Computers & Industrial Engineering* (2019).

[Grollmisch19] Sascha Grollmisch et al. “Sounding Industry: Challenges and Datasets for Industrial Sound Analysis”. In: *EUSIPCO 2019*.

[ToyAdmos2] Noboru Harada et al. “ToyADMOS2: Another dataset of miniature-machine operating sounds for anomalous sound detection under domain shift conditions”. In: *arXiv:2106.02369* (2021).

[DCASE website] *Hoomepage of the DCASE Challenge and Workshop*. URL: `http://dcase.community` (visited on 10/12/2021).

[Koizumi20] Yuma Koizumi et al. “Description and Discussion on DCASE2020 Challenge Task2: Unsupervised Anomalous Sound Detection for Machine Condition Monitoring”. In: *DCASE2020 Proceedings*. 2020.

[ToyAdmos] Yuma Koizumi et al. “ToyADMOS: A Dataset of Miniature-machine Operating Sounds for Anomalous Sound Detection”. In: *WASPAA 2019*.

[Kuroyanagi21] Ibuki Kuroyanagi et al. *ANOMALOUS SOUND DETECTION WITH ENSEMBLE OF AUTOENCODER AND BINARY CLASSIFICATION APPROACHES*. Tech. rep. DCASE2021 Challenge, 2021.

[Lopez21] Jose Lopez, Georg Stemmer, and Paulo Lopez-Meyer. *Ensemble of Complementary Anomaly Detectors Under Domain Shifted Conditions*. Tech. rep. DCASE2021 Challenge, 2021.

[McInnes18] Leland McInnes and John Healy. “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction”. In: *arXiv:1802.03426* (2018).

[Morita21] Kazuki Morita, Tomohiko Yano, and Khai Tran. *ANOMALOUS SOUND DETECTION USING CNN-BASED FEATURES BY SELF SUPERVISED LEARNING*. Tech. rep. DCASE2021 Challenge, 2021.

[Primus20] Paul Primus. *Reframing Unsupervised Machine Condition Monitoring as a Supervised Classification Task with Outlier-Exposed Classifiers*. Tech. rep. DCASE2020 Challenge, 2020.

References (2)

[MIMII] Harsh Purohit et al. "MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection". In: *DCASE 2019 Proceedings.*

[MIMII_DUE] Ryo Tanabe et al. "MIMII DUE: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection with Domain Shifts due to Changes in Operational and Environmental Conditions". In: *arXiv:2006.05822, 1–4* (2021).

[Wilkinghoff21] Kevin Wilkinghoff. *Utilizing Sub-Cluster AdaCos for Anomalous Sound Detection under Domain Shifted Conditions*. Tech. rep. DCASE2021 Challenge, 2021.