Computational Analysis of Sound and Music

Environmental Sound Analysis – Acoustic Anomaly Detection

Dr.-Ing. Jakob Abeßer

Fraunhofer IDMT

jakob.abesser@idmt.fraunhofer.de

Acoustic Scene Classification

Outline

- Introduction & Application Scenarios
- Traditional Approaches
- Deep Learning-based Approaches

Acoustic Scene Classification

Introduction

- Goal
  - Detect deviations from "normal" state
  - Only using normal state examples for training
  - Is emitted sound from target object normal or anomalous?

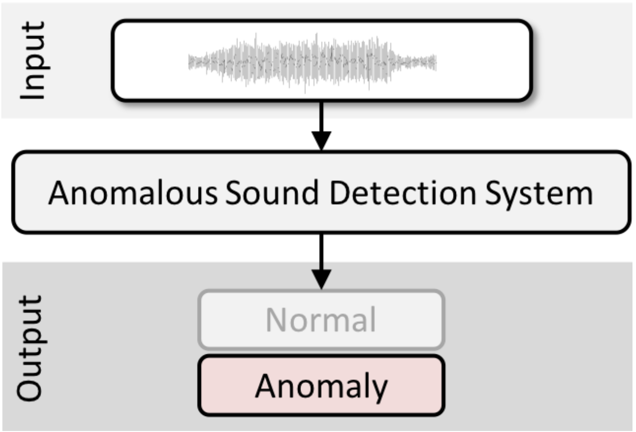


Fig-E4-1

Acoustic Scene Classification

Introduction

- Categories
  - Point anomalies
    - Individual instances deviate from remaining dataset
  - Group/pattern anomalies
    - Subset of instances are anomalous
  - Contextual anomalies
    - Data instance is anomalous only in specific contexts or conditions

Acoustic Scene Classification

Introduction

- Challenges
  - No anomaly examples known during training
  - Robustness & Adaptivity towards changing background sounds/acoustic conditions
  - Temporal dynamics of anomalies (short spikes vs. prolonged anomalies)
  - Real-time processing constraints (industry/surveillance)
  - Acoustic anomalies are often subtle compared to background noise (e.g., loud machines in factory settings)

Acoustic Scene Classification

Application Scenarios

- Security/Surveillance
  - Unusual sounds can indicate security threads or criminal activity
- Healthcare
  - Abnormal sounds can indicate medical conditions or emergencies (e.g., abnormal heart or lung sounds)
- Predictive maintenance of machines in industrial settings
  - Possible anomalies caused by abnormal vibrations, friction, or mechanical failures
- Smart City & Environmental Monitoring
  - Detect critical events (car accident),
  - Detect illegal poaching & deforestation (chainsaw, vehicles, gunshots)

Acoustic Scene Classification

Traditional Approaches

- Distribution outlier detection
  - Modelling normal state distribution
  - Detect distribution outliers
  - E.g.: One-class GMM / SVM

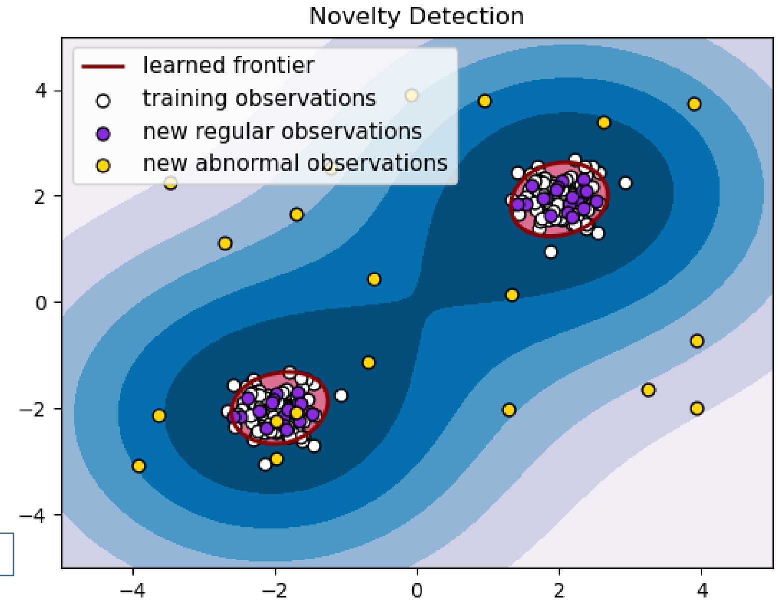


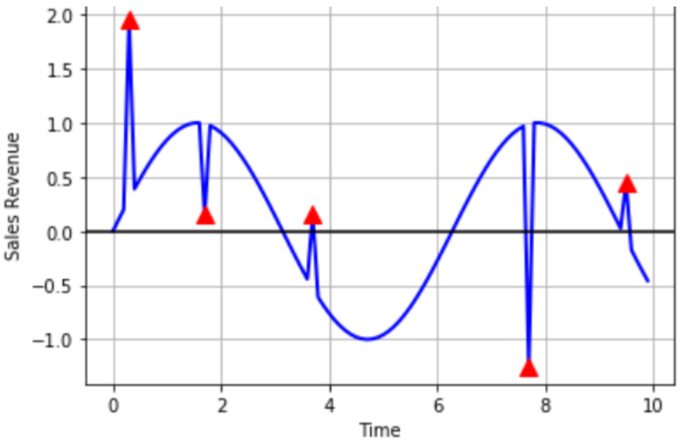
Fig-E4-2

Acoustic Scene Classification

Traditional Approaches

- Time-series analysis
    - AD via prediction error
    - Examples
        - Autoregressive models
        - Hidden-Markov-Models (HMM)

Fig-E4-3



Acoustic Scene Classification

Deep Learning-based Approaches

- Autoencoder (encoder → decoder) models

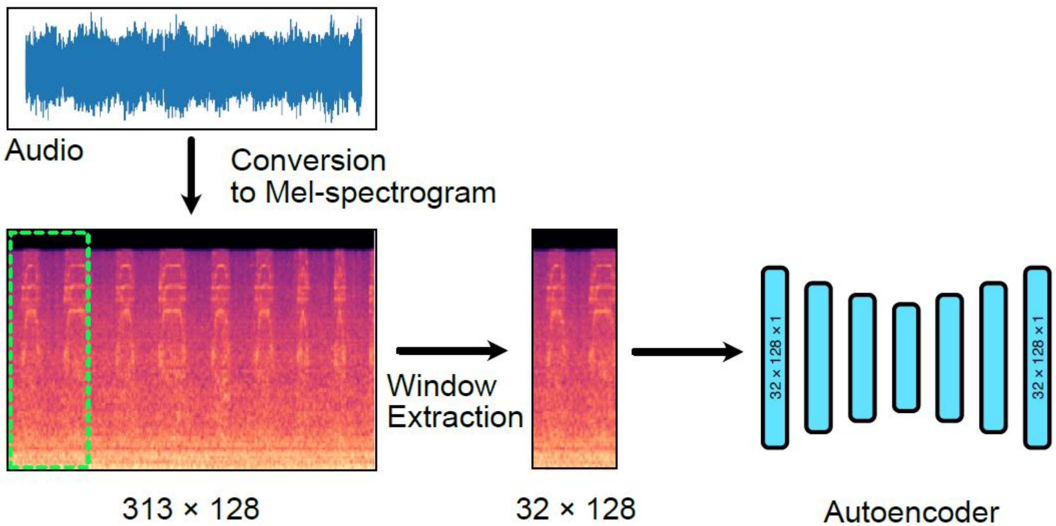


Fig-E4-4

Acoustic Scene Classification

Deep Learning-based Approaches

- Idea: Normal sounds can be better reconstructed than anomalous sounds
- Reconstruction error
  - $\mathcal{L} = \|x - D(E(x))\|_2^2$
- Anomaly detection by thresholding $\mathcal{L}$

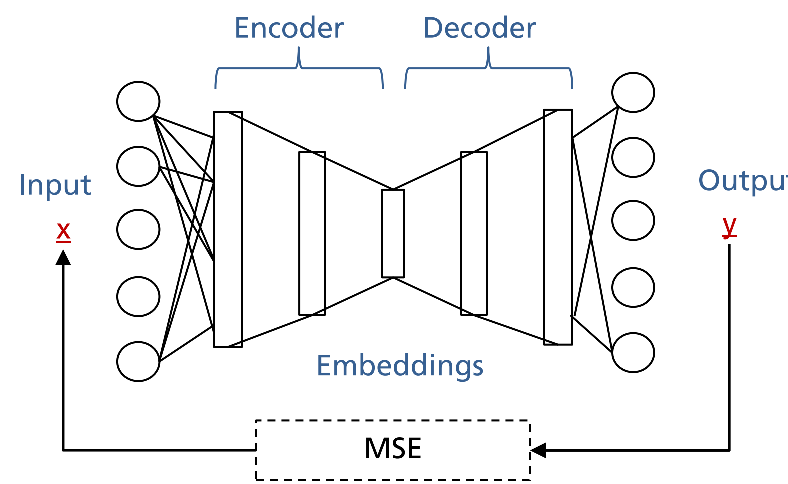
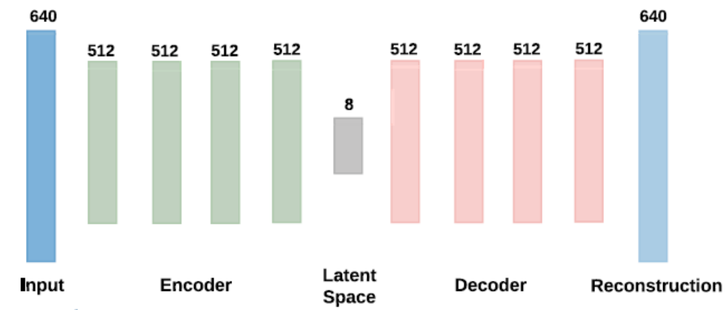


Fig-E4-5

Acoustic Scene Classification

Deep Learning-based Approaches

- [Coelho, 2022]
  - 3 AE architectures
    - Dense AE
    - Convolutional AE
    - Recurrent AE



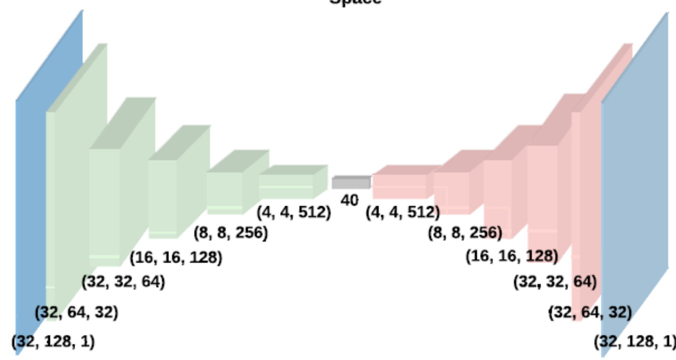


Fig-E4-6

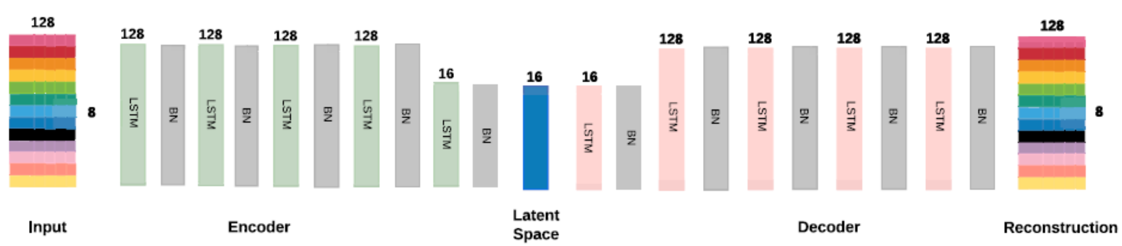


Fig-E4-7

Acoustic Scene Classification

Deep Learning-based Approaches

- Performance depends on threshold applied to reconstruction error
- Evaluation metric: Area under the receiver-operating curve (ROC)

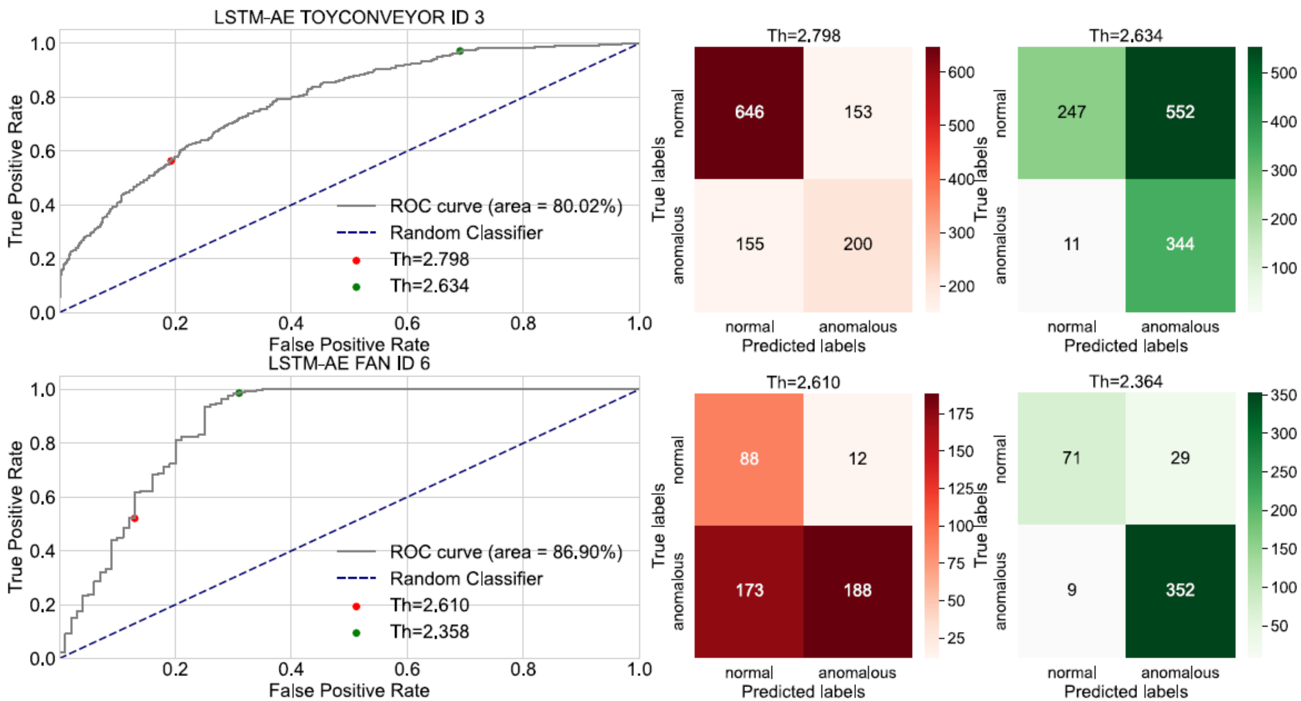


Fig-E4-8

Acoustic Scene Classification

Deep Learning-based Approaches

- [Suefusa 2020]
    - Interpolation Autoencoder
        - Interpolate frames from neighbor frames

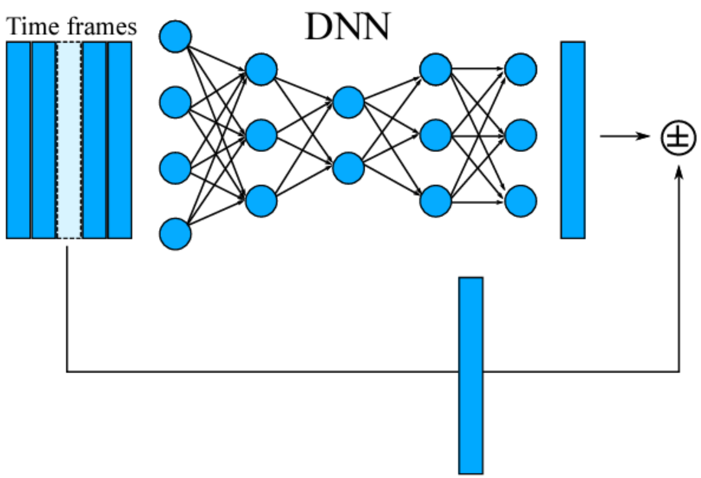


Fig-E4-9

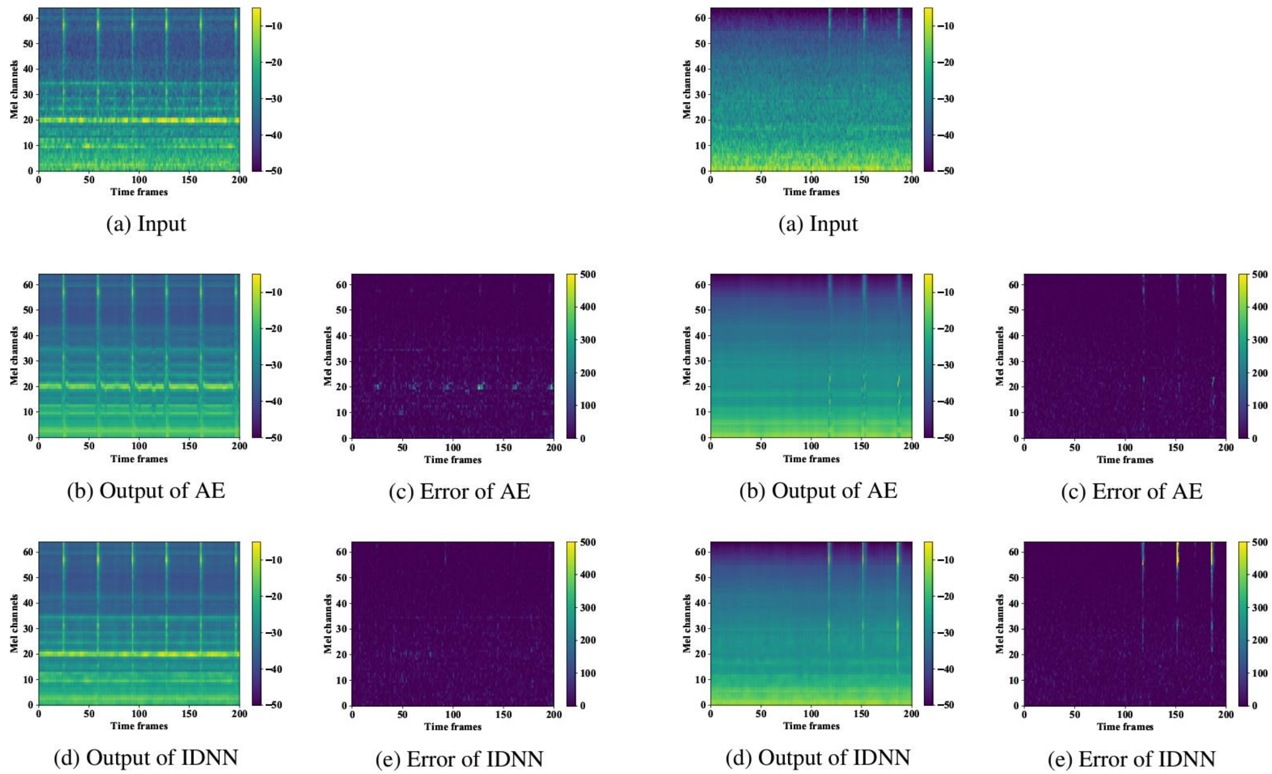
Acoustic Scene Classification

Deep Learning-based Approaches

- Example: Valve sound

Normal sound

Anomaly sound



Original sounds

Reconstruction + Error by regular AE

Reconstruction + Error by interpolation AE

Fig-E4-10

Programming session

Fig-A2-13

References

Images

Fig-E4-1: http://dcase.community/challenge2020/task-unsupervised-detection-of-anomalous-sounds (Figure 1)

Fig-E4-2: https://scikit-learn.org/stable/_images/sphx_glr_plot_oneclass_0011.png

Fig-E4-3: https://miro.medium.com/max/722/1*TvZ9jI9vGX-fWwc3AHwNDw.png

Fig-E4-4: [Abbasi, 2021], p.4, Fig. 1

Fig-E4-5: Own

Fig-E4-6: [Coelho, 2022], p. 19492, Fig. 5

Fig-E4-7: [Coelho, 2022], p. 19492, Fig. 6

Fig-E4-8: [Coelho, 2022], p. 19494, Fig. 7 (top two out of three subplots)

Fig-E4-9: [Suefusa, 2020], p. 272, Fig. 3a

Fig-E4-8: [Suefusa, 2020], p. 274, Fig. 7 + 8 (parts thereof)

References

References

Abbasi, S., Famouri, M., Shafiee, M. J., & Wong, A. (2021). OutlierNets: Highly Compact Deep Autoencoder Network Architectures for On-Device Acoustic Anomaly Detection. Sensors, 21(14), 4805. https://doi.org/10.3390/s21144805

Coelho, G., Matos, L. M., Pereira, P. J., et al. (2022). Deep autoencoders for acoustic anomaly detection: experiments with working machine and in-vehicle audio. Neural Computing & Applications, 34(19485-19499). https://doi.org/10.1007/s00521-022-07375-2

Suefusa, K., Nishida, T., Purohit, H., Tanabe, R., Endo, T., & Kawaguchi, Y. (2020). Anomalous Sound Detection Based on Interpolation Deep Neural Network. In ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 271-275). Barcelona, Spain. https://doi.org/10.1109/ICASSP40776.2020.9054344