Adaptive Framework for Anomaly Detection in Time Series Audio-Visual Data

Technical Brief

Ref No. 2025/AWaDH/OITC/026

Technology Summary

This technology presents a deep learning-based teacher-student network that fuses audio and visual data for adaptive anomaly detection in multimedia streams. The teacher model distills knowledge from a pre-trained visual network to lightweight student audio and image networks. Features are combined and compressed via PCA, then modeled using an adaptive Gaussian mixture model (AGMM) to capture scene dynamics with concept drift. The joint audio-visual representation significantly outperforms single-modality methods, enhancing accuracy and robustness in real-world, dynamic environments.

Background

Audio anomaly detection traditionally struggles with concept drift—changes in data over time—limiting effectiveness in dynamic environments. Existing methods like AGMM require prior knowledge and lose long-term context. This work introduces a dynamic Huffman coding approach with node merging, enabling adaptive, memory-efficient detection of audio anomalies despite evolving data distributions, improving performance without heavy computational demands.

Technology Description

We present a novel anomaly detection method for audio data using dynamic Huffman coding to handle concept drift. Unlike adaptive Gaussian mixture models (AGMM), our approach is parameter-free, does not require predefined clusters, and adapts dynamically to environmental changes. A key innovation is a node-merging strategy that prevents forgetting past data while controlling tree size. Experiments on long-duration audio datasets with natural concept drift show that our method outperforms AGMM in detection accuracy, achieving higher AUC scores. Additionally, we contribute new datasets tailored for evaluating adaptive anomaly detection in realistic, non-stationary audio environments.

Market Potential / Proposed Deployment

- High demand in smart surveillance and public safety systems
- Applications in smart cities (e.g., noise, traffic, incident detection)
- Use in industrial monitoring for machinery faults and predictive maintenance
- Healthcare monitoring (e.g., elderly care, emergency detection)
- Valuable for defense and border security (e.g., gunshot, intrusion detection)
- Fits into the growing audio analytics market (CAGR > 20%)
- Edge deployment for real-time, low-latency detection
- Cloud/hybrid systems for large-scale learning and analysis

Applications

- Smart surveillance: Detect gunshots, shouting, glass breaking
- Smart cities: Monitor traffic incidents, urban noise
- Industrial monitoring: Identify machinery faults or failures
- Healthcare: Detect falls, distress sounds in patient care
- Transportation: Monitor for mechanical issues or emergencies
- Defense/security: Detect intrusions or suspicious activity

Value Proposition

- Adaptive to Concept Drift: Continuously learns and adapts to changing audio environments
- Parameter-Free Design: No need for prior knowledge of cluster count or distribution
- Low Resource Requirement: Lightweight and efficient for real-time edge deployment
- Improved Accuracy: Higher anomaly detection performance (AUC) than AGMM
- Retains Long-Term Context: Node merging avoids forgetting rare but important events

Technology Status

- TRL 4 – Technology Validated in Lab
- The proposed dynamic Huffman coding-based anomaly detection has been validated through experiments on curated audio datasets with concept drift in a controlled research environment.