Ensemble Confidence Calibration for Sound Event Detection in Open-environment

Yuanjian Chen* and Han Yin†
* School of Computer Science and Technology, Harbin University of Science and Technology, Harbin, China
E-mail: 2010400002@stu.hrbust.edu.cn (corresponding author)
† School of Electrical Engineering, KAIST, Daejeon, Republic of Korea
E-mail: hanyin@kaist.ac.kr

***Abstract*—Sound event detection (SED) has made strong progress in controlled environments with clear event categories. However, real-world applications often take place in open environments. In such cases, current methods often produce predictions with too much confidence and lack proper ways to measure uncertainty. This limits their ability to adapt and perform well in new situations. To solve this problem, we are the first to use ensemble methods in SED to improve robustness against out-of-domain (OOD) inputs. We propose a confidence calibration method called Energy-based Open-World Softmax (EOW-Softmax), which helps the system better handle uncertainty in unknown scenes. We further apply EOW-Softmax to sound occurrence and overlap detection (SOD) by adjusting the prediction. In this way, the model becomes more adaptable while keeping its ability to detect overlapping events. Experiments show that our method improves performance in open environments. It reduces overconfidence and increases the ability to handle OOD situations.**

I. Introduction

Sound event detection (SED), which involves recognizing event categories and identifying their corresponding timestamps [1], [2], is an important method for analyzing and interpreting acoustic information within audio clips. SED has been widely applied across various domains, including industrial monitoring [3], healthcare [4], and smart city development [5].

Recently, with the flourishing development of the Detection and Classification of Acoustic Scenes and Events (DCASE) challenge series [6], a substantial number of excellent research achievements [7]–[13] have continuously emerged. Among these advances, there are ensemble model approaches [14]–[18] that integrate the sound event posterior probabilities from individual models to obtain final prediction results. The success of current methods mostly depends on the assumption that sound events come from a small, known, and clearly defined set. However, as the need for real-world applications continues to grow and as tasks become more complex, this assumption no longer holds [19], [20]. In open-environment, SED system's performance is significantly affected by factors such as ambiguous event definitions, long-tail data distributions [2], [21], and complex environmental noise, making the learning process more dynamic and uncertain.

To address open-environment challenges, researchers have conducted several explorations. Wei et al. [22] employed data manipulation-based methods, handling differences between different acoustic environments through domain adaptation techniques. Xiao et al. [10] specifically addressed incremental learning challenges through an unsupervised class incremental learning framework. Despite the progress these methods have achieved in their respective targeted challenges, significant limitations remain. First, existing methods lack effective mechanisms for handling the coupling relationships among multiple sound events in open-environment. Second, when confronted with out-of-domain (OOD) challenges, models often generate overconfident predictions and lack effective confidence calibration mechanisms, severely constraining their adaptability and generalization performance in OOD scenes.

To overcome these issues, it is necessary to look at ideas that have worked in related audio tasks. Recent research [23] has demonstrated that constructing speech anti-spoofing model ensembles and employing linear fusion algorithms to combine classification scores from multiple candidate models can effectively enhance the robustness of classification models in OOD scenes. Following this line, in this work, we introduce the ensemble methods into open-environment SED to address the OOD challenges. However, confidence calibration, as a key technique for reflecting the reliability of model predictions [24], faces severe challenges in OOD scenarios: OOD inputs often lead models to produce overconfident predictions, thereby compromising the effectiveness of ensemble fusion.

To solve the overconfidence issue in OOD scenes, we propose a confidence calibration method based on Energy-based Open-World Softmax (EOW-Softmax) [25], which models uncertainty in open-world conditions. This method is applied before final decision-making to adjust the magnitude of logit scores, allowing better prediction confidence for each class. Considering the polyphonic nature of SED, we apply this technique to sound occurrence and overlap detection [12], a statistical sub-task of SED that captures overlapping sound patterns. Our work contributes to the field in three ways: (1) we are the first to introduce ensemble methods into SED for open environments, improving robustness and generalization in OOD scenes using interpolation-based fusion; (2) we address the common overconfidence problem in OOD predictions by using EOW-Softmax to calibrate model confidence based on open-world uncertainty; and (3) we design a task-specific strategy for applying EOW-Softmax to polyphonic detection by adjusting logit magnitudes, ensuring that ensemble models remain adaptable without losing their ability to detect overlapping events. The code will be publicly available soon.

979-8-3315-7206-8/25/$31.00

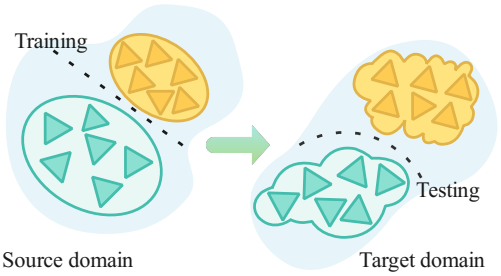

Fig. 1. Overview of OOD challenge for SED. Training in source domain acoustic scenes and testing in target domain acoustic scenes.

II. Methods

A. Out-of-domain challenge for SED

As we mentioned, the SED systems perform exceptionally well in controlled settings but experience considerable difficulties in open environments, such as those encountered in intelligent security surveillance, where acoustic scenes vary dramatically between indoor and outdoor environments, and event definitions are often ambiguous. To model the evolution of the dataset across different domains, we introduce $\mathbb{D}^{1:s}_{train} = \{\mathbb{X}^{1:s}_{train}, \mathbb{Y}^{1:s}_{train}\}$ for data up to domain s during training stage, where $\mathbb{X}^{1:s}_{train}$ denotes the set of audio samples and $\mathbb{Y}^{1:s}_{train}$ represents the corresponding frame-level annotations, $\mathbb{D}^{s+1}_{test} = \{\mathbb{X}^{s+1}_{test}, \mathbb{Y}^{s+1}_{test}\}$ for the subsequent domain $s+1$ during testing stage. The OOD challenge involves the model's ability to generalize effectively across significantly different domains. While leveraging foundational models can partially address OOD challenges, it may lead to overfitting or performance degradation in new domains [26]. As shown in Fig.1, the issue of OOD challenge lies in the distributional gap between training data $\mathbb{X}^{1:s}_{train}$ and testing data $\mathbb{X}^{s+1}_{test}$, represented as $P(\mathbb{X}^{1:s}_{train}) \neq P(\mathbb{X}^{s+1}_{test})$. The desired robustness property, $P(\mathbb{Y}^{1:s}_{train} \mid \mathbb{X}^{1:s}_{train}) = P(\mathbb{Y}^{s+1}_{test} \mid \mathbb{X}^{s+1}_{test})$, captures the expectation that conditional relationships learned from source scenes should generalize to target scenes. Assuming that our training set can be expressed as $\{(x_n, y_n)\}_{n=1}^{N}$, where x_n denotes the audio of the n-th clips $\mathbb{X}^{1:s}_{train}$, with frame-level label $y_n = \{y^1_n, ..., y^t_n, ..., y^T_n\}$ forming a sequence of C-dimensional binary vectors $y^t_n \in \{0,1\}_C$ for C event categories across T temporal frames, N represents the total clips in the training set. The SED model $\mathcal{F}_\phi$ aims to optimize its performance by learning the parameter ϕ. Thus, the optimization objective can be expressed as:

$$min \sum_{n=1}^{N} \mathcal{L}_{\text{SED}}(\mathcal{F}_\phi(x_i), y_i) + \tau \mathcal{L}_{\text{OPEN}} \tag{1}$$

where $\mathcal{L}_{\text{SED}}$ denotes the standard SED objective function, while $\mathcal{L}_{\text{OPEN}}$ represents an auxiliary loss term designed to handle open-environment challenges, with τ serving as the weighting hyperparameter.

B. Candidate Model and EOW-Softmax

We adopt the baseline of DCASE 2023 SED challenge as our candidate model. To better address polyphonic phenomena in open-environment and improve event boundary detection performance, we introduce the sound occurrence and overlap detection (SOD) task [12]. This task focuses on sound activity patterns rather than specific event categories, providing more effective frame-level supervision signals for SED. The overall framework is illustrated in Figure 2. This framework consists of four components: seven CNN blocks, two Bi-GRU blocks, and two prediction heads for SED and SOD tasks, respectively . The prediction head of SED outputs frame-level event predictions, while the SOD branch outputs frame-level event activity states categorized into three classes: $\{0, 1, 2\}$, where 0 indicates no target events, 1 represents monophonic event, and 2 denotes polyphonic events.

In the SED branch, we adopt binary cross entropy (BCE) as $\mathcal{L}_{SED}$ in our equation 1. To address OOD challenges in SED and perform confidence calibration for the candidate model, we employ the EOW-Softmax [25] strategy to train the model in the SOD branch, thereby improving the output fusion of model ensemble. The EOW-Softmax strategy addresses open-world uncertainty as an additional dimension [25], implementing a $(P+1)$-way classification scheme. Here, the primary $P = 3$ dimensions encode the three event activity categories within the SOD branch, whereas the extra fourth dimension quantifies the open-world uncertainty. This uncertainty metric increases when the model exhibits low prediction confidence. Following Figure 2(a), let $\Upsilon_\phi : \mathbb{R}^D \to \mathbb{R}^{P+1}$ denotes the linear mapping in the SOD prediction head, yielding four logits expressed as $\Upsilon_\phi(x)\,\|i\|$, with $i \in \{1, 2, 3, 4\}$, for the input in-domain acoustic representation $x \in \mathbb{R}^D$. Probability distributions are obtained via softmax transformation:

$$\Psi_\phi(x)\,\|i\| = \frac{\exp \Upsilon_\phi(x)\,\|i\|}{\sum_{j=1}^{P+1} \exp \Upsilon_\phi(x)\,\|j\|} \tag{2}$$

where Ψ_ϕ is the concatenation of the linear layer and a softmax transformation layer. Then, the loss function $\mathcal{L}_{OPEN}(x, \mathfrak{o})$ for Eow-Softmax is defined as:

$$\begin{aligned} \mathcal{L}_{\text{OPEN}}(x, \mathfrak{o}) =& \min_{\phi} \mathbb{E}_{p(x)} \left[-\log \Psi_\phi(x)\,\|\mathfrak{o}\|\right] \\ &+ \lambda \mathbb{E}_{p_{\hat{\phi}}(x)} \left[-\log \Psi_\phi(x)\,\|P+1\|\right] \end{aligned} \tag{3}$$

where $\lambda \in \mathbb{R}^+$ is a hyperparameter. The initial component represents the maximum log-likelihood (MLL) criterion for the SOD task utilizing ground truth labels $\mathfrak{o}$; the subsequent component constitutes the MLL objective for identifying acoustic features sampled from normal distributions $p_{\hat{\phi}}(x)$, where $\hat{\phi}$ represents the frozen parameters from $p_\phi(x)$ at the current iteration. These normal distributions are learned through SGLD-based optimizations [25]. As a result, our candidate model should be optimized to generate elevated open-world uncertainty when processing samples from the latter distribution.

C. Model Ensemble Strategy

To enhance SED performance in open-environment, we adopt confidence-based model ensemble approach. As shown in

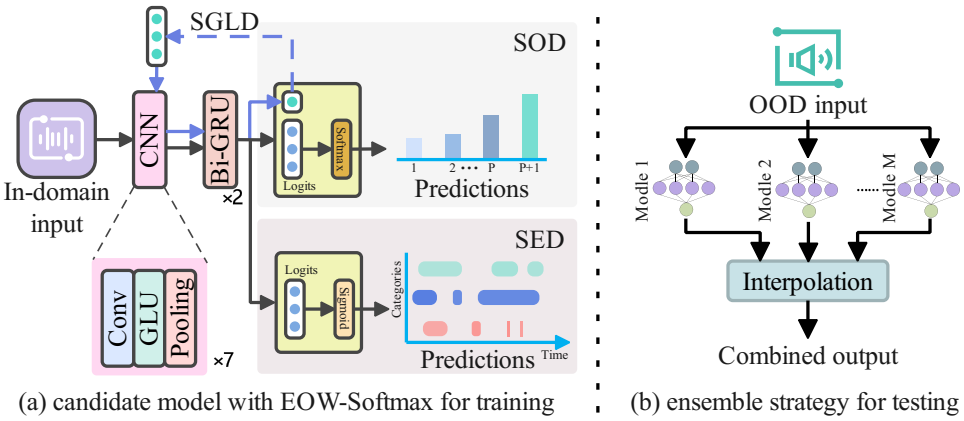

Fig. 2. Pipeline of the proposed approach: (a) dual-branch candidate model for training with SED and SOD tasks, and (b) model ensemble strategy with interpolation for testing.

Figure 2(b), multiple candidate models' results are combined to enhance the ensemble's prediction. Although confidence calibration strategy is incorporated into the SOD sub-task during training, we focus on optimizing SED task performance during the testing stage. As the confidence calibration from the SOD branch encodes the model's frame-level audio comprehension accuracy, we leverage SOD confidence to guide the SED ensemble process. For the c-th event category, the final SED combined prediction $s_{\text{SED}}(x)[c]$ is defined as:

$$s_{\text{SED}}(x)[c] = \frac{\sum_{m=1}^{M} \mathcal{F}_{\phi_m}(x)[c] \cdot c_m(x)}{\sum_{m=1}^{M} c_m(x)} \tag{4}$$

where $\mathcal{F}_{\phi_m}(x)[c]$ denotes the prediction of the SED branch of the m-th candidate model for the c-th event category, $c_m(x)$ is the confidence calibration learned by the SOD branch of the m-th candidate model, and M is the total number of candidate models in the ensemble. The SOD confidence serves as weights to modulate each candidate model's contribution. Models with higher confidence play a more significant role in final predictions, thereby improving SED performance in open-environment.

III. Experiments

A. Dataset and Performance Metric

In this work, we construct an integrated dataset following the methodology described in [2]. The dataset comprises four acoustic scenes ("home", "residential area", "city center", and "office") and 25 distinct sound event categories. We select nine event classes as target events: 'bird singing (bs)', 'car (ca)', 'children (ch)', 'impact (im)', 'large vehicle (lv)', 'people talking (pt)', 'people walking (pw)', 'rustling (ru)', and 'squeaking (sq)'. The remaining event categories are treated as detection-irrelevant events to simulate open-environment. All audio samples are sourced from the TUT 2016 dataset [27] and TUT 2017 dataset [28], each containing annotations for both sound events and acoustic scenes. Detailed annotation information is available on this website [1]. To evaluate the effectiveness of our proposed ensemble confidence calibration method in open environments, we utilize the "home" and "residential area" scenes from the TUT 2016 dataset as the training set (in-domain data), while the "city center" scene from the TUT 2017 dataset and the "office" scene from the TUT 2016 dataset serve as the test set (out-of-domain data). The training set contains a total audio duration of 194 minutes, while the test set comprises 70 minutes of audio. All audio samples are clipped or concatenated into 10-second clips. Figure 3 presents the total duration statistics of each target event, along with the corresponding ratios between training and test sets.

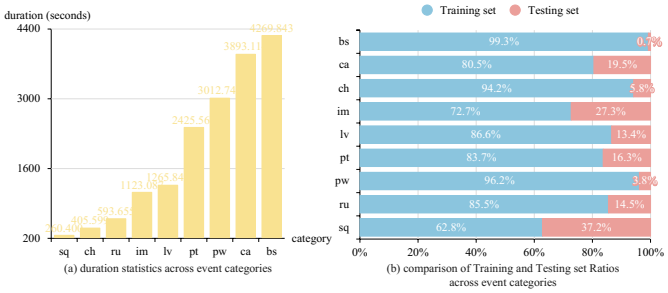

Fig. 3. Overview of event durations and training-testing ratios across categories: (a) event duration statistics, (b) ratios between training and testing sets.

We employ F-score [29] as the evaluation metric, which can be categorized into two types based on evaluation granularity: event-based F-score focuses on the detection performance of complete event instances and their temporal boundaries, while segment-based F-score divides audio recordings into fixed-length segments for frame-level classification evaluation. In this work, we adopt the frame length as the basic unit for segment division. Furthermore, both evaluation approaches utilize macro-average and micro-average strategies. Therefore, the SED performance in open-environment is comprehensively assessed through four metrics: event-based macro-average F-score (Ema-F1), event-based micro-average F-score (Emi-F1), segment-based macro-average F-score (Sma-F1), and segment-based micro-average F-score (Smi-F1).

[1] https://www.ksuke.net/dataset

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT MODELS FOR SED IN OPEN-ENVIRONMENT

ID	Model	Ema-F1 score (%) ↑	Emi-F1 score (%) ↑	Sma-F1 score (%) ↑	Smi-F1 score (%) ↑
P1	CRNN	7.74	10.89	15.18	37.28
P2	EOW-Softmax	9.07	13.00	17.91	43.34
P3	CRNN w/ avg. linear ensemble	8.49	11.98	16.75	40.56
P4	EOW-Softmax w/ calibrated ensemble	**11.80**	**17.18**	**23.51**	**57.92**

B. Implementation Details

For data preprocessing, we use a sliding window to extract log mel-spectrogram features as audio representations. Each frame has a length of 128 milliseconds, and the window moves forward by 16 milliseconds. We apply soft mixup, a data augmentation method where two samples are mixed using a weight drawn from a Beta distribution with $\alpha = 0.2$ and $\beta = 0.2$. During training, we use the Adam optimizer with an initial learning rate of 0.001 and a batch size of 48. The training process runs for 200 epochs and starts with an exponential warmup during the first 50 epochs. We apply early stopping if the loss does not improve for 5 consecutive epochs. To ensure robustness, we train five models using different random seeds. The final model result is the average of the best checkpoints from these five runs, selected based on validation performance. For model ensembling, each candidate model is given an equal interpolation weight of 0.2. During inference, we apply median filtering with a window size of 7 to smooth predictions over time. Finally, we integrate the ensemble with EOW-Softmax for confidence calibration. The EOW-Softmax settings follow the same configuration as used in [25].

IV. RESULTS AND DISCUSSION

We present the results of four F-scores in Table I. All models were trained on acoustic scenes from "home" and "residential area" and tested on scenes from "city center" and "office". Among the 25 event categories, nine were chosen as target events for detection. The other sixteen were used as background interference to simulate open-environment conditions.

We first compare the baseline system (P1) with an individual model trained using EOW-Softmax (P2). As shown in Table I, applying EOW-Softmax to the SOD branch leads to clear improvements in F-scores. This is due to the regularization effect of EOW-Softmax, which helps the model capture uncertainty in predictions and also reduces overfitting. Next, in method P3, we apply an ensemble strategy by averaging the outputs from five candidate models. This approach produces better results than the baseline (P1), but still falls short compared to the EOW-Softmax model (P2). The reason is that without confidence calibration, predictions with larger magnitudes dominate the average output. This imbalance can reduce the reliability of the final prediction, as discussed in prior work [23]. Finally, our proposed method (P4), which combines EOW-Softmax with model ensembling, achieves the best performance. It improves F-scores by 52.45%, 57.76%, 54.87%, and 55.36% over the baseline. These gains show that calibrating confidence allows the ensemble to better handle OOD inputs and improves robustness in open environments.

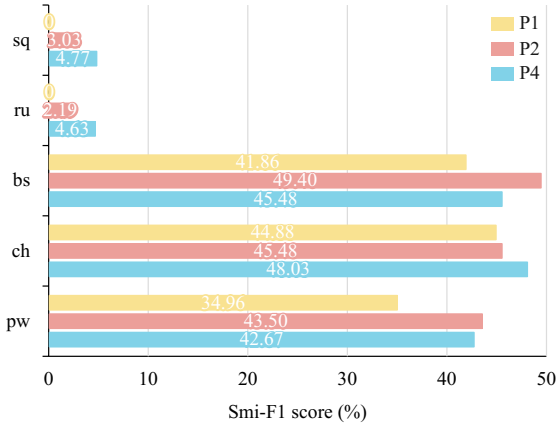

Fig. 4. Smi-F1 score evaluation of different strategies for SED on challenging events, the strategy numbers correspond to the IDs listed in Table I.

To further validate the performance in open-environment settings, we selected several challenging sound events from Figure 4 and compared the results of P1, P2, and P4 (as listed in Table I). These events fall into two groups: (1) events with imbalanced data, such as 'bs', 'ch', 'pw', and (2) events with short durations and unclear features, such as 'ru', 'sq'. These cases reflect common difficulties in real-world scenarios, such as variation in event length, strong class imbalance, and high overlap in acoustic features. Figure 4 shows the Smi-F1 scores for each method. Overall, models using EOW-Softmax outperform the baseline across all cases. Moreover, combining EOW-Softmax with ensembling (P4) gives the highest scores. This improvement comes from two main factors: first, EOW-Softmax makes training more stable by modeling uncertainty; second, calibrated ensembles more effectively combine predictions from multiple models, resulting in higher overall accuracy.

V. CONCLUSIONS

In this work, we innovatively propose the adoption of model ensembles to address SED in open-environment. Through experimental evaluation, we demonstrate that the introduced EOW-Softmax effectively regularizes the training process and performs confidence calibration during the model ensembles stage, thereby significantly improving the overall performance of SED. Future research can extend the application of this technique to more challenging OOD scenarios.

REFERENCES

[1] A. Mesaros, T. Heittola, T. Virtanen, and M. D. Plumbley, "Sound event detection: A tutorial," *IEEE Signal Processing Magazine*, vol. 38, no. 5, pp. 67–83, 2021.

[2] Y. Zhang, R. Togneri, and D. Huang, "A unified loss function to tackle inter-class and intra-class data imbalance in sound event detection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2024, pp. 996–1000.

[3] J. Wu, F. Yang, and W. Hu, "Unsupervised anomalous sound detection for industrial monitoring based on arcface classifier and gaussian mixture model," *Appl. Acoust.*, vol. 203, p. 109 188, 2023.

[4] W. Qiu, C. Quan, L. Zhu, *et al.*, "Heart sound abnormality detection from multi-institutional collaboration: Introducing a federated learning framework," *IEEE Trans. Biomed. Eng.*, vol. 71, no. 10, pp. 2802–2813, 2024.

[5] E.-L. Tan, F. A. Karnapi, L. J. Ng, K. Ooi, and W.-S. Gan, "Extracting urban sound information for residential areas in smart cities using an end-to-end iot system," *IEEE Internet Things J.*, vol. 8, no. 18, pp. 14 308–14 321, 2021.

[6] A. Mesaros, R. Serizel, T. Heittola, T. Virtanen, and M. D. Plumbley, "A decade of dcase: Achievements, practices, evaluations and future challenges," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2025, pp. 1–5.

[7] K. Li, Y. Song, L.-R. Dai, I. McLoughlin, X. Fang, and L. Liu, "Ast-sed: An effective sound event detection method based on audio spectrogram transformer," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2023, pp. 1–5.

[8] S. Bhosale, S. Nag, D. Kanojia, J. Deng, and X. Zhu, "Diffsed: Sound event detection with denoising diffusion," in *Proc. AAAI Conf. Artif. Intell.*, vol. 38, 2024, pp. 792–800.

[9] N. Shao, X. Li, and X. Li, "Fine-tune the pretrained atst model for sound event detection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2024, pp. 911–915.

[10] Y. Xiao and R. Kumar Das, "Ucil: An unsupervised class incremental learning approach for sound event detection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2025, pp. 1–5.

[11] H. Yue, Z. Zhang, D. Mu, Y. Dang, J. Yin, and J. Tang, "Full-frequency dynamic convolution: A physical frequency-dependent convolution for sound event detection," in *Proc. Int. Conf. Pattern Recognit.*, 2025, pp. 260–272.

[12] Y. Guan, J. Han, H. Song, *et al.*, "Sound activity-aware based cross-task collaborative training for semi-supervised sound event detection," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 32, pp. 3947–3959, 2024.

[13] H. Yin, J. Bai, Y. Xiao, *et al.*, "Exploring text-queried sound event detection with audio source separation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2025, pp. 1–5.

[14] Y. Guo, M. Xu, Z. Wu, J. Wu, and B. Su, "Multi-scale convolutional recurrent neural network with ensemble method for weakly labeled sound event detection," in *Proc. 8th Int. Conf. Affect. Comput. Intell. Interact. Workshops Demos*, 2019, pp. 1–5.

[15] Q. Wang, H. Wu, Z. Jing, *et al.*, "A model ensemble approach for sound event localization and detection," in *Proc. 12th Int. Symp. Chinese Spoken Lang. Process.*, 2021, pp. 1–5.

[16] M. A. S. M. Afendi and M. Yusoff, "A sound event detection based on hybrid convolution neural network and random forest," *IAES Int. J. Artif. Intell.*, vol. 11, no. 1, pp. 121–128, 2022.

[17] H. Dinkel, Z. Yan, Y. Wang, M. Song, J. Zhang, and W. Wang, "A large multi-modal ensemble for sound event detection," *IEEE AASP Challenge on DCASE 2022, Tech. Rep.*, 2022.

[18] A. Mukhamadiyev, I. Khujayarov, D. Nabieva, and J. Cho, "An ensemble of convolutional neural networks for sound event detection," *Mathematics*, vol. 13, no. 9, pp. 1502–1527, 2025.

[19] Z.-H. Zhou, "Open-environment machine learning," *National Science Review*, vol. 9, no. 8, nwac123, 2022.

[20] Y. Xiao and R. K. Das, "WildDESED: An LLM-powered dataset for wild domestic environment sound event detection system," in *Proc. Detection Classification Acoust. Scenes Events Workshop*, 2024, pp. 196–200.

[21] Y. Xiao, H. Yin, J. Bai, and R. K. Das, "Mixstyle based domain generalization for sound event detection with heterogeneous training data," *arXiv preprint arXiv:2407.03654*, 2024.

[22] W. Wei, H. Zhu, E. Benetos, and Y. Wang, "A-crnn: A domain adaptation model for sound event detection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2020, pp. 276–280.

[23] C. Y. Kwok, D.-T. Truong, and J. Q. Yip, "Robust audio deepfake detection using ensemble confidence calibration," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2025, pp. 1–5.

[24] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in *Int. Conf. Machine Learning*, 2017, pp. 1321–1330.

[25] Y. Wang, B. Li, T. Che, K. Zhou, Z. Liu, and D. Li, "Energy-based open-world uncertainty modeling for confidence calibration," in *Proc. IEEE/CVF Int. Conf. Computer Vision*, 2021, pp. 9302–9311.

[26] S. Liang, W. Wang, R. Chen, *et al.*, "Object detectors in the open environment: Challenges, solutions, and outlook," *arXiv preprint arXiv:2403.16271*, 2024.

[27] A. Mesaros, T. Heittola, and T. Virtanen, "Tut database for acoustic scene classification and sound event detection," in *24th Eur. Signal Process. Conf.*, 2016, pp. 1128–1132.

[28] A. Mesaros, T. Heittola, A. Diment, *et al.*, "Dcase 2017 challenge setup: Tasks, datasets and baseline system," in *Proc. Detection Classification Acoust. Scenes Events Workshop*, 2017.

[29] A. Mesaros, T. Heittola, and T. Virtanen, "Metrics for polyphonic sound event detection," *Appl. Sci.*, vol. 6, no. 6, pp. 162–178, 2016.