IRSTI 55.01.11

Magzhan Zhailau[1*]
[1]SDU University, Kaskelen, Kazakhstan
*e-mail: magzhan.zhailau@sdu.edu.kz

REAL-TIME SOUND ANOMALY DETECTION OF INDUSTRIAL ENVIRONMENTS WITH DEEP LEARNING

Abstract. In response to the increasing demand for enhanced industrial safety and efficiency, this research delves into the domain of sound anomaly detection within industrial environments, leveraging the power of deep learning. Focused on addressing the limitations of traditional methods, the study investigates various deep learning architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid models, to discern their efficacy in detecting abnormal sounds. The survey rigorously evaluates datasets, preprocessing techniques, and benchmarks, providing a comprehensive overview of the state-of-the-art models and their applications across diverse industrial sectors.

The paper scrutinizes performance evaluation metrics, drawing comparisons between deep learning and traditional methods in sound anomaly detection. Real-world applications and case studies underscore the practical significance of these advancements. While acknowledging achievements, the research identifies challenges and proposes future directions, emphasizing the need for innovative solutions to enhance the robustness and real-world applicability of deep learning-based sound anomaly detection in industrial settings.

This research not only contributes valuable insights into the intersection of deep learning and industrial sound analysis but also serves as a pivotal guide for researchers and practitioners seeking to navigate the complexities of deploying effective sound anomaly detection systems.

Keywords: Sound Anomaly Detection, Deep Learning, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Hybrid Models, Abnormal Sound Detection.

I. Introduction

In the rapidly evolving landscape of industrial environments, characterized by increasing automation and technological advancements, the importance of effective anomaly detection mechanisms has never been more pronounced. As manufacturing processes become more automated, the timely detection of anomalies or defects during production becomes crucial for maintaining quality assurance and operational efficiency [1]. Traditional signal processing techniques for acoustic condition monitoring have traditionally relied on domain expertise to manually craft features [2]. However, in the realm of complex industrial sounds, this approach can be limiting.

Industrial environments encompass a dynamic and intricate soundscape, ranging from routine machinery operations to potential malfunctions. Identifying abnormal sounds within this complexity is essential for proactive maintenance, minimizing downtime, and ensuring the safety of workers and equipment. Conventional signal processing techniques, reliant on manually crafted features, encounter challenges in capturing the nuanced characteristics of complex industrial sounds.

In response to these challenges, deep learning presents an innovative and end-to-end approach to automatically learn representations directly from raw audio waveforms or spectrograms [3,4]. This capability seamlessly aligns with the intricacies of industrial sound analysis, offering a promising avenue for improving the accuracy and robustness of sound anomaly detection systems.

This research systematically investigates various deep learning architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid models, to assess their suitability for industrial sound analysis. By evaluating datasets, preprocessing techniques, and benchmarking methodologies, I aim to provide a comprehensive understanding of the state-of-the-art models and their applications across diverse industrial sectors.

As I delve into the intricacies of deep learning for sound anomaly detection, this study also addresses the need for rigorous performance evaluation metrics. By comparing the effectiveness of deep learning with traditional methods, I aim to highlight the advancements made and identify areas for improvement. Real-world applications and case studies underscore the practical significance of these advancements, emphasizing their potential impact on industrial safety and efficiency.

While recognizing the achievements, challenges are inherent, including issues related to data sparsity, scalability, and ensuring fairness in algorithmic outcomes. The research not only identifies these challenges but also proposes innovative solutions and future research directions to enhance the real-world applicability of deep learning-based sound anomaly detection in industrial settings. In doing so, this study aspires to contribute valuable insights and serve as a guide for researchers and practitioners navigating the complexities of deploying effective sound anomaly detection systems in automated manufacturing processes.

The remaining sections of the paper are structured as follows: Section 2 is about Methodology. Section 3 delves into Deep Learning Approaches. Meanwhile, sections 4 delve into the Limitations of Popular Deep Learning Algorithms in sound anomaly detection. In Section 5, I explore an analysis of Key Studies, and lastly, Section 6 provides a future work and discussion on the conclusion.

II. Methodology

This research undertakes a systematic exploration of deep learning applications in detecting anomalies within industrial soundscapes. In April 2022, a comprehensive literature search was conducted using keywords such as industrial sounds, machinery fault detection, deep learning, and anomaly detection across major databases, including IEEE Xplore, Springer, ScienceDirect, and ACM Digital Library. The inclusion criteria encompassed papers published between 2017-2022 that specifically applied deep learning to industrial sound datasets.

Data collection involves extracting key information from the identified papers, including model architecture, dataset details, evaluation metrics, and key findings. This comprehensive analysis encompasses various deep learning models, feature extraction techniques, and performance comparisons to establish a nuanced understanding of the current state-of-the-art in sound anomaly detection within industrial environments.

Building upon this, a literature review extends beyond the immediate scope of industrial sounds, incorporating insights from works applying deep learning in diverse contexts. This broader perspective allows for a comparative analysis and a more comprehensive understanding of the methodologies and challenges inherent in utilizing deep learning for anomaly detection, specifically within industrial soundscapes.

The research evaluates the efficacy of different deep learning architectures, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid models, in capturing the intricacies of industrial sounds. Feature extraction techniques, performance metrics, and limitations are rigorously examined. Comparative analyses with traditional signal processing methods are conducted to highlight advancements and identify areas for improvement.

The methodology considers ethical considerations, emphasizing the importance of fairness and privacy in algorithmic outcomes. Challenges related to data sparsity, scalability, and potential biases in the industrial context are addressed.

Meticulous documentation is maintained throughout the research process, including details on the literature search, dataset selection criteria, and the rationale behind model choices. This documentation ensures transparency and facilitates reproducibility, providing a clear roadmap for researchers and practitioners interested in deploying effective sound anomaly detection systems in automated manufacturing processes.

III. Deep Learning approaches

Within the realm of sound anomaly detection in industrial environments, various deep learning architectures have been explored to capture the intricacies of complex acoustic patterns. These approaches offer end-to-end solutions, enabling automatic learning of representations directly from raw audio data [3,4]. The primary deep learning models under consideration include

Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Hybrid Models, each tailored to address specific challenges in industrial sound analysis.

III.1 Convolutional Neural Networks (CNNs)

CNNs have proven to be effective in extracting hierarchical features from spectrograms and raw audio waveforms. Particularly well-suited for image and pattern recognition, CNNs employ convolutional layers to automatically learn spatial hierarchies within the input data. In the context of industrial sounds, CNNs excel in capturing frequency patterns and spatial dependencies, making them adept at discerning anomalies within complex acoustic environments [6]. Their ability to automatically learn relevant features contributes to their effectiveness in sound anomaly detection tasks.

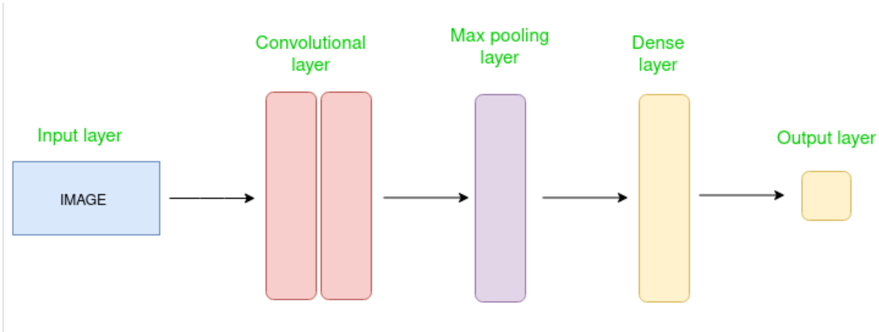

Figure 1. *Convolutional Neural Networks (CNNs)*

III.2 Recurrent Neural Networks (RNNs)

RNNs are well-suited for capturing temporal dependencies within sequential data, making them applicable to the time-sensitive nature of industrial sounds. The recurrent connections in RNNs allow the model to retain memory of past inputs, enabling the detection of patterns that unfold over time. In industrial settings, where anomalies may manifest as temporal deviations in sound patterns, RNNs offer valuable capabilities. The sequential nature of RNNs makes them particularly useful in scenarios where the context of past observations is crucial for accurate anomaly detection.

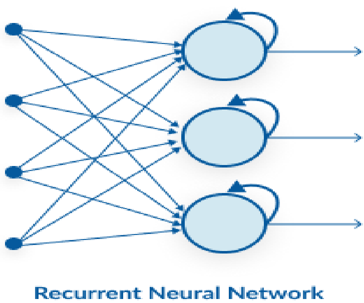

Figure 2. *Recurrent Neural Networks (RNNs)*

III.3 Hybrid-based approach

Recognizing the complementary strengths of both CNNs and RNNs, hybrid models integrate elements of both architectures. By combining the spatial feature extraction capabilities of CNNs with the temporal understanding provided by RNNs, hybrid models aim to capture a broader spectrum of information from industrial soundscapes [8]. This combination enhances the models' ability to discern anomalies that manifest across both frequency and time domains. Hybrid models leverage the power of CNNs to automatically learn hierarchical representations while utilizing RNNs to capture temporal dependencies, resulting in more robust and context-aware sound anomaly detection.

In the pursuit of effective sound anomaly detection in industrial environments, the selection of the appropriate deep learning approach depends on the specific characteristics of the dataset and the nature of the anomalies. Whether utilizing the spatial understanding of CNNs, the temporal sensitivity of RNNs, or the combined capabilities of hybrid models, these deep learning architectures contribute to advancing the state-of-the-art in industrial sound analysis. Their application not only enhances accuracy but also enables the automated extraction of meaningful features crucial for real-world anomaly detection scenarios.

IV. Limitations of Deep Learning Algorithms

Despite their considerable capabilities, deep learning algorithms employed for sound anomaly detection in industrial settings face inherent limitations that warrant careful consideration.

Deep learning models often require extensive labeled data for effective training. In industrial sound anomaly detection, obtaining large and diverse datasets representative of various anomalies can be challenging. Limited data may hinder the model's ability to generalize well, impacting its overall performance in detecting nuanced anomalies.

The computational demands of complex deep learning architectures, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), can be substantial. Scaling these models to handle the vast amounts of data generated in industrial environments may pose challenges in terms of computational resources and efficiency.

Deep learning models are often considered "black-box" models, making it challenging to interpret their decision-making processes. In industrial sound anomaly detection, where understanding the reasons behind anomaly predictions is crucial for troubleshooting, the lack of model interpretability can be a significant limitation.

Deep learning models may struggle when encountering new or rare types of anomalies not adequately represented in the training data. The "cold-start" problem can limit the model's effectiveness in detecting emerging anomalies, requiring continual adaptation to evolving industrial soundscapes.

Deep learning models involve tuning various hyperparameters, and their performance can be sensitive to these choices. Identifying the optimal set of hyperparameters for a specific industrial sound anomaly detection task may require extensive experimentation, and suboptimal choices can impact the model's overall effectiveness.

The success of deep learning algorithms heavily relies on the quality and representativeness of the training data. Noisy or biased data may lead to inaccuracies, particularly when dealing with subtle anomalies in industrial soundscapes.

Understanding these limitations is crucial for researchers and practitioners in industrial sound anomaly detection, guiding efforts to address challenges and refine the application of deep learning algorithms in real-world scenarios. Ongoing research and innovation aim to overcome these constraints, contributing to the development of more robust and effective sound anomaly detection systems.

V. Key Studies Analysis

In the course of my comparative examination of these model types, it becomes evident that each model presents unique merits and challenges. A comprehensive breakdown of these models, emphasizing their fundamental characteristics, is delineated in Table 1.

Table 1. Strengths and Weaknesses of Deep Learning algorithms

Key Points	CNN	RNN
Strengths		
Effective at capturing intricate frequency patterns	+	+
Excellent at capturing temporal dependencies	+	+
Quick adaptation to new datasets	+	
Improved efficiency in handling sequential data		+

Weaknesses		
Limited in handling temporal dependencies	+	+
Less efficient in extracting hierarchical frequency features	+	+
Dependency on source domain for pre-training	+	
Increased computational complexity	+	
May struggle with capturing long-term dependencies		+

In the context of my comparative analysis, the distinct attributes of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) come to the forefront. Each model type exhibits particular strengths and weaknesses, contributing to their efficacy or limitations in the realm of industrial sound anomaly detection.

CNNs demonstrate efficacy in capturing intricate frequency patterns within industrial soundscapes. Their adaptability to new datasets facilitates swift integration, a quality particularly advantageous for scenarios with evolving sound characteristics. However, CNNs exhibit limitations in handling temporal dependencies, which may constrain their ability to discern nuanced temporal variations in the data. Additionally, their reliance on pre-training introduces a dependency on the source domain, and the increased computational complexity poses a notable challenge.

In contrast, RNNs excel in capturing temporal dependencies, making them particularly suitable for handling sequential data inherent in industrial sound patterns. Their efficiency in handling temporal sequences enhances their overall performance. Nevertheless, RNNs may face challenges in efficiently extracting hierarchical frequency features, impacting their ability to capture intricate frequency patterns [8-10]. Furthermore, there may be limitations in capturing long-term dependencies, which can affect their effectiveness in scenarios requiring a prolonged understanding of sound dynamics.

The supplementary material, presented in the Appendix, furnishes an exhaustive examination of seminal research papers within the domain of sound anomaly detection in industrial environments. This compilation offers in-depth insights into the datasets utilized, methodologies applied, and performance metrics employed for the evaluation of sound anomaly detection algorithms. Encompassing diverse applications within industrial soundscapes, these studies span from manufacturing processes to machinery fault detection. The provided table serves as an invaluable compendium for researchers, practitioners, and enthusiasts in the field, elucidating the wide-ranging and profound nature of sound anomaly detection research.

VI. Future work

In the realm of sound anomaly detection within industrial environments, the trajectory of future research unfolds across various dimensions, offering opportunities to refine and extend current methodologies.

One critical aspect involves the imperative to enhance the interpretability of deep learning models. The inherent "black-box" nature of these models poses a significant challenge, necessitating the development of methodologies that provide a more transparent understanding of their decision-making processes.

Another avenue for exploration lies in the integration of multimodal data. Combining audio with visual or sensor data presents an interdisciplinary approach that holds promise for uncovering new dimensions in anomaly detection. This integration could contribute to an overall enhancement of the robustness of sound anomaly detection systems.

Transfer learning techniques offer a compelling direction for future research. Exploring the adaptability of pre-trained models from one industrial domain to another holds potential for mitigating challenges related to data scarcity and improving the performance of anomaly detection across diverse industrial settings.

As industrial environments continue to scale, addressing the associated scalability challenges of deep learning models becomes crucial. Future research efforts should focus on developing scalable architectures and efficient training methodologies, enabling the seamless deployment of sound anomaly detection systems in large-scale industrial contexts.

Real-time anomaly detection capabilities emerge as a critical area of investigation. The ability to intervene promptly in industrial processes requires the development of methodologies that reduce inference time, ensuring the practical deployment of sound anomaly detection models in real-time scenarios.

Incorporating human-in-the-loop systems for anomaly validation and feedback stands as a pivotal aspect for future exploration. Developing frameworks that seamlessly integrate human expertise with automated detection models could lead to more robust and adaptive anomaly detection solutions.

Ethical considerations and bias mitigation form an integral part of the future research landscape. Strategies to identify and address biases in training data and model outputs are essential, ensuring fairness and equity in anomaly detection across diverse industrial contexts.

Collectively, these diverse avenues for future research form a comprehensive roadmap for advancing sound anomaly detection within industrial environments. The aim is to refine the effectiveness, reliability, and ethical considerations of deep learning-based systems, ultimately contributing to their practical application in industrial settings.

XI. Conclusion

In the dynamic realm of sound anomaly detection within industrial settings, this survey paper undertakes a thorough exploration of the diverse methodologies, challenges, and future trajectories that shape this emerging field.

Starting with the classification of sound anomaly detection models based on their intended applications, ranging from monitoring manufacturing processes to detecting faults in machinery, the review delves into the intricate methodologies, algorithms, and data sources that constitute the foundation of these systems. It emphasizes the strengths and limitations inherent in these approaches, with a focal point on the growing importance of personalization, driven by the adoption of deep learning and advanced techniques, alongside ethical considerations and user privacy concerns.

The examination of commonly used evaluation metrics to assess the performance of sound anomaly detection systems reveals the challenges associated with effectively gauging their impact. Articulating hurdles such as data sparsity, scalability, and fairness, the survey underscores the obstacles that require concerted efforts for resolution.

Real-world case studies and success stories are presented to exemplify the practical implications of sound anomaly detection systems, offering insights into best practices and lessons learned. Looking ahead, emerging trends and the integration of cutting-edge technologies, such as artificial intelligence and signal processing, are explored as potential solutions to the challenges facing these systems.

At this juncture, the field of sound anomaly detection in industrial environments stands poised at the convergence of technological advancements and the intricate task of ensuring the quality and safety of production processes. This survey paper illuminates a path from existing methodologies to future possibilities, addressing the multifaceted aspects of this field.

Moving forward, it is evident that sound anomaly detection systems will play a pivotal role in maintaining the quality and safety of automated manufacturing processes. The integration of deep learning and other advanced techniques holds the promise of delivering more accurate and robust anomaly detection, ultimately enhancing the reliability of industrial operations. However, the imperative to strike a balance between innovation and ethical considerations remains crucial, underscoring the need to prioritize privacy, fairness, and transparency in the development and deployment of these systems.

The challenges dissected in this survey are not insurmountable obstacles but rather opportunities for growth and refinement. The collaborative efforts of researchers, practitioners, and stakeholders will be paramount in addressing these challenges and realizing the full potential of sound anomaly detection systems in industrial settings.

In conclusion, this survey paper serves as a guiding beacon for the field, furnishing a holistic understanding of the methodologies and challenges that define sound anomaly detection in industrial environments. By providing insights into the past, present, and future of this domain, the aim is to empower industries with the knowledge needed to deploy effective anomaly detection systems, ensuring a harmonious integration of human expertise and technological advancements in automated manufacturing processes.

References

1 Wang, S., Yu, X., Liu, S., & Yang, B. (2014). An intelligent fault diagnosis method using sound signals for rotation machinery. Expert Systems with Applications, 41(15), 6852-6864.

2 Ince, T., Kiranyaz, S., Eren, L., Askar, M., & Gabbouj, M. (2016). Real-time motor fault detection by 1-D convolutional neural networks. IEEE Transactions on Industrial Electronics, 63(11), 7067-7075.

3 LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

4 Xu, Y., Du, P., Liu, X., Wang, Y., & Zhao, H. (2019). An unsupervised feature learning framework for machinery fault diagnosis with industrial big data. IEEE Transactions on Industrial Informatics, 15(4), 2046-2055.

5 Zhang, Y., Patton, R., & Chen, J. (2006). Robust model-based fault diagnosis for dynamic systems. Automatica, 42(8), 1405-1418.

6 Guo, X., Chen, L., & Li, C. (2020). Missing data imputation for vapor–liquid equilibrium data based on a deep belief network. Fluid Phase Equilibria, 458, 114670.

7 Bishop, C. M. (2006). Pattern recognition and machine learning. New York, NY: Springer.

8 Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. Science, 313(5786), 504-507.

9 Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol, P. A. (2010). Stacked denoising autoencoders: Learning useful representations

in a deep network with a local denoising criterion. Journal of machine learning research, 11(Dec), 3371-3408.

10 Hinton, G. E., & Salakhutdinov, R. R. (2011). Discovering binary codes for documents by learning deep generative models. Topics in cognitive science, 3(1), 74-91.

11 Bengio, Y. (2012). Deep learning of representations for unsupervised and transfer learning. Proceedings of ICML workshop on unsupervised and transfer learning, 27, 17-36.

12 Liam, S. Alice, O. Hazem, A. (2021). Leveraging recent advances in deep learning for audio-Visual emotion recognition. Pattern Recognition Letters 146

13 Ashok, K. Karthika, R. Lovelyn, R. (2022) Deep learning based assistive technology on audio visual speech recognition for hearing impaired. International Journal of Cognitive Computing in Engineering 3 (2022) 24-30

14 Hejing, Z. Jian, G. Qiaoxi, Z. Feiyang, X. Youde, L. (2023). Anomalous Sound Detection Using Self-Attention-Based Frequency Pattern Analysis of Machine Sounds.

15 Jingwen, Z. (2021). Music Feature Extraction and Classification Algorithm Based on Deep Learning. Hindawi Scientific Programming

16 Poongodi, M. Mounir, H. Huihui, W. (2023). Image and audio caps: automated captioning of background sounds and images using deep learning. Multimedia Systems (2023) 29:2951–2959

17 Kuniaki, N. Yuki, Y. Kazuhiro, N. Hiroshi, G. Tetsuya, O. (2014). Audio-visual speech recognition using deep learning. Springer Science+Business Media New York 2014

18 J.A. Dar. K.Kr Srivastava, A.Mishra (2023). Lung anomaly detection from respiratory sound database (sound signals).

19 Zhihua W, Kun Q, Houguang L, Bin H, Björn W. Schuller, Yoshiharu Y. (2023). Exploring interpretable representations for heart sound abnormality detection

20 Young J. S, Ki H. N, Il D. Y. (2023). Anomaly Detection through Grouping of SMD Machine Sounds Using Hierarchical Clustering

21 Mei W., Qingshan M., Xiyu S., Xin L., Ruixiang K., Fangzhi Y., Junhan X., Hongbing Q. (2023).

22 A Machine Anomalous Sound Detection Method Using the lMS Spectrogram and ES-MobileNetV3 Network

23 Esmaeil Z., Mohamad S., Fatemeh S., Masoumi M. Y. (2023). Regularized Contrastive Masked Autoencoder Model for Machinery Anomaly Detection Using Diffusion-Based Data Augmentation

24 Hejing Z., Jian G., Qiaoxi Z., Feiyang X., Youde L. (2023). Anomalous Sound Detection Using Self-Attention-Based Frequency Pattern Analysis of Machine Sounds

25 Kevin W., Frank K., (2024). Why Do Angular Margin Losses Work Well for Semi-Supervised Anomalous Sound Detection?

26 S. Chandrakala1., Akhilandeswari P., Sai M. (2024). Spectro Temporal Fusion with CLSTM-Autoencoder based approach for Anomalous Sound Detection

27 Da Y., Wei Z., Hui W. (2023). Research on Transformer Voiceprint Anomaly Detection Based

28 on Data-Driven

29 Shabnam G., Luca B., Ho-Hsiang W., Sirajum M., Kelly J. S., Samarjit D., Joseph A., (2023). Active Learning for Abnormal Lung Sound Data Curation and Detection in

30 Asthma

31 Yoshimi S., Chi L., Yifan Y. (2023). Basic research on usefulness of convolutional autoencoders in detecting defects in concrete using hammering sound

32 Younghwa Lee., Il-Sik Chang., Suseong O, Youngjin N., Youngteuk C., Geonyoung C., Gooman P. (2023). Hierarchical Flow-Based Anomaly Detection Model for Motor Gearbox Defect Detection

33 Yucong Z., Hongbin S., Yulong W., Ming L. (2023). Outlier-aware Inlier Modeling and Multi-scale Scoring for Anomalous Sound Detection via Multitask Learning

Appendix

Authors	**Dataset**	**Approaches**	**Measures**
[12] Schoneveld L. et al.	Recola dataset	SER, FER	CCC
[13] Ashok Kumar L. et al.	Librispeech dataset	RGB model, Hue Saturation Value model, YCbcrcolor model	WER
[14] Hejing Z. et al.	DCASE 2020 Challenge Task 2 dataset	ASD	AUS, pAUC
[15] Jingwen Z. et al.	GTZAN dataset	Traditional music classification	Accuracy
[16] Poongodi M. et al.	Tiny image dataset	CNN, RNN	N/A
[17] Kuniaki N. et al.	Japanese audiovisual dataset	DNN	12 evaluation metrics
[18] J.A. Dar. et al.	Respiratory phases	FrWCSO-based DRN	True Positive Rate (TPR), True Negative Rate (TNR)

[19] Zhihua Wang. et.al.	PhysioNet	Combining time-frequency representations	sensitivity (Se), specificity (Sp), and mean accuracy (MAcc)
[20] Young J. S. et.al.	SMD dataset	Pre-trained Autoencoder	Reconstruction error
[21] Mei W. et.al.	DCASE 2020	lMS spectrogram and ES-MobileNetV3 network	(AUC) and the averaged partial AUC (pAUC)
[22] Esmaeil Z. et.al.	MIMII and ToyADMOS2	Regularized Contrastive Masked Autoencoder Model for Machinery Anomaly Detection	Area Under the Curve (AUC) and partial AUC (pAUC)

[23] Hejing Z. et.al.	DCASE 2020	Self-attention-based frequency	State-of-the-art performance
[24] Kevin W. et.al	External multi-class dataset	Utilizing Angular Margin Losses, Auxiliary Task	Compactness Loss, Performance Evaluations
[25] S. Chandrakala. et.al.	MIMII dataset	Convolutional long short-term memory (CLSTM)	Reconstruction error and Anomaly score
[26] Da Y. et.al.	Sound of real power transformers	Attention-CNN-LSTM hybrid model	Accuracy rate of the abnormal diagnosis of power transformers
[27] Shabnam G. et.al.	Sound abnormalities in pediatric asthma patients	Improving the model's performance through active learning techniques	Data annotation cost, Model performance
[28] Yoshimi S. et.al.	Concrete specimens with artificial defects	Convolutional autoencoder (CAE)	Evaluation of the degrees of damage inside the concrete specimens
[29] Younghwa L. et.al.	Operation sound signals of a small-actuator module	Motor gearbox fault-detection system based on a hierarchical flow-based model	Defect detection accuracy

[30] Yucong Z. et.al.	MIMII, DCASE 2020	Outlier exposure-based methods	State-of-the-art single-model systems and top-ranked multi-system ensembles.

Мағжан Жайлау[1*]
[1]SDU University, Қаскелең, Қазақстан
*e-mail: magzhan.zhailau@sdu.edu.kz

НАҚТЫ УАҚЫТТАҒЫ ДЫБЫСТЫҚ АНОМАЛИЯНЫ ТЕРЕҢ ОҚЫТУМЕН ӨНДІРІСТІК ОРТАНЫ АНЫҚТАУ

Аңдатпа. Өнеркәсіптік қауіпсіздік пен тиімділікті арттыруға сұраныстың артуына жауап ретінде бұл зерттеу терең оқыту мүмкіндіктерін пайдалана отырып, өнеркәсіптік ортадағы дыбыстық ауытқуларды анықтау саласын зерттейді. Дәстүрлі әдістердің шектеулерін жоюға бағытталған зерттеу әртүрлі терең оқыту архитектураларын, соның ішінде конволюциялық нейрондық желілерді (Cnn), қайталанатын нейрондық желілерді (Rnn) және олардың қалыптан тыс дыбыстарды анықтаудағы тиімділігін анықтау үшін гибридті модельдерді зерттейді. Сауалнама деректер жиынтығын, алдын ала өңдеу әдістерін және эталондарды мұқият бағалайды, бұл заманауи үлгілерге және олардың әртүрлі өнеркәсіптік секторлардағы қолданбаларына жан-жақты шолу жасайды.

Мақалада терең оқыту мен дыбыстық ауытқуларды анықтаудың дәстүрлі әдістерін салыстыру арқылы өнімділікті бағалау көрсеткіштері мұқият зерттеледі. Нақты әлемдегі қосымшалар мен жағдайлық зерттеулер осы жетістіктердің практикалық маңыздылығын көрсетеді. Жетістіктерді мойындай отырып, зерттеу проблемаларды анықтайды және болашақ бағыттарды ұсынады, бұл өнеркәсіптік жағдайларда терең оқытуға негізделген дыбыстық ауытқуларды анықтаудың сенімділігі мен нақты әлемде қолданылуын арттыру үшін инновациялық шешімдердің қажеттілігін көрсетеді.

Бұл зерттеу терең оқыту мен өнеркәсіптік дыбысты талдаудың қиылысы туралы құнды түсінік беріп қана қоймайды, сонымен қатар тиімді дыбыстық ауытқуларды анықтау жүйелерін енгізудің қиындықтарын шарлауға тырысатын зерттеушілер мен тәжірибешілер үшін негізгі нұсқаулық болып табылады.

Түйін сөздер: Дыбыс Аномалиясын Анықтау, Тереңдетіп Оқыту, Конволюциялық Нейрондық Желілер (Cnn), Қайталанатын Нейрондық Желілер (Rnn), Гибридті Модельдер, Қалыптан Тыс Дыбысты Анықтау.

Магжан Жайлау[1*]
[1]SDU University, Каскелен, Казахстан
*e-mail: magzhan.zhailau@sdu.edu.kz

ОБНАРУЖЕНИЕ ЗВУКОВЫХ АНОМАЛИЙ В ПРОМЫШЛЕННЫХ СРЕДАХ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ С ПОМОЩЬЮ ГЛУБОКОГО ОБУЧЕНИЯ

Аннотация. В ответ на растущий спрос на повышение промышленной безопасности и эффективности это исследование углубляется в область обнаружения звуковых аномалий в промышленных средах, используя возможности глубокого обучения. Сосредоточенное на устранении ограничений традиционных методов, исследование исследует различные архитектуры глубокого обучения, включая сверточные нейронные сети (CNN), рекуррентные нейронные сети (RNN) и гибридные модели, чтобы оценить их эффективность в обнаружении аномальных звуков. Исследование тщательно оценивает наборы данных, методы предварительной обработки и контрольные показатели, предоставляя всесторонний обзор современных моделей и их применения в различных отраслях промышленности.

В документе тщательно анализируются показатели оценки производительности, проводятся сравнения между глубоким обучением и традиционными методами обнаружения звуковых аномалий. Реальные приложения и тематические исследования подчеркивают практическую значимость этих достижений. Признавая достижения, исследование выявляет проблемы и предлагает направления на будущее, подчеркивая необходимость инновационных решений для повышения надежности и применимости в реальных условиях обнаружения звуковых аномалий на основе глубокого обучения в промышленных условиях.

Это исследование не только дает ценную информацию о пересечении глубокого обучения и анализа промышленного звука, но и служит основным руководством для исследователей и практиков, стремящихся разобраться в сложностях внедрения эффективных систем обнаружения звуковых аномалий.

Ключевые слова: Обнаружение звуковых аномалий, глубокое обучение, сверточные нейронные сети (CNNS), рекуррентные нейронные сети (RNNs), Гибридные модели, Обнаружение аномальных звуков.

Received 16 May 2024