Comprehensive Review of One-Class Classification Approaches for Anomaly Detection

Divya Challa Software Engineer, Sift

ABSTRACT

Anomaly detection is a crucial task in various domains, including cybersecurity, healthcare, and finance, where identifying rare and abnormal events is of paramount importance. One-Class Classification (OCC) methods have emerged as a powerful approach for this task, effectively distinguishing between normal and anomalous data when only the normal class is wellrepresented. This paper provides a comprehensive review of OCC techniques, categorizing them into four main approaches: Reconstruction-Based Methods, Variational Autoencoders (VAEs), Convolutional Approaches, and Hybrid Models. Additionally, it delves into the diverse applications of OCC, including the detection of rare diseases in healthcare, financial fraud prevention, and industrial fault detection. The review also addresses key challenges such as data imbalance, model interpretability, and scalability, while highlighting recent trends and advancements in the field.

Keywords

One-class classifcation (OCC), Anomaly Detection, Variational Autoencoders (VEs), Convolutional Neural Networks (CNN), Outlier detection, Convolutional Autoencoders (CAE), Generative Adversarial Networks (GANs), Long Short-Term Memory (LSTM), Principal components analysis (PCA), Self-Organizing Map (SOM), One-Class Support Vector Machine (OCSVM), Deep learning.

1. INTRODUCTION

Anomaly detection is an essential component in a variety of critical applications, such as fraud detection in financial transactions, early diagnosis of rare diseases in healthcare, and identifying operational faults in industrial systems. Unlike traditional classification problems, anomaly detection often deals with scenarios where only normal data is available during training, making the task challenging due to the absence of labeled anomalies. One-Class Classification (OCC) methods have gained prominence for their ability to learn from this imbalanced data by effectively modeling the normal class and identifying deviations as anomalies.

Over the years, OCC approaches have evolved significantly, incorporating diverse methodologies ranging from traditional machine learning techniques to advanced deep learning models. This review paper provides a structured analysis of OCC methods, categorizing them into four main types: Reconstruction-Based Methods, which identify anomalies through reconstruction errors; Variational Autoencoders (VAEs), which model data distributions probabilistically; Convolutional Approaches, suitable for image and multimodal data; and Hybrid Models, which combine various techniques to enhance robustness and performance.

The paper is organized as follows: Section 2 provides a detailed discussion of each OCC category, explaining their core principles and methodologies. Section 3 explores the practical applications of OCC methods across different domains. Section 4 highlights the key challenges faced by OCC techniques, such as data imbalance and interpretability. Section 5 goes into emerging trends in one-class classification for anomaly detection. Finally, Section 6 concludes the paper with insights into future directions.

2. OVERVIEW OF OCC AND ITS METHODOLOGIES

In many real-world scenarios, datasets often consist of labeled instances from only one class, with either very few or no labeled examples from the other classes. This situation presents a challenge for standard classification algorithms that typically require a balanced representation of all classes. One-Class Classification (OCC) addresses this issue by focusing exclusively on learning the characteristics of the available class, referred to as the normal class. The goal of OCC is to identify whether new, unseen instances belong to this learned class or deviate significantly from it, thus being considered anomalies or outliers.

Fig 1: Illustration of One-Class Classification

OCC is particularly useful in applications such as fraud detection, network security, and medical diagnostics, where obtaining labeled examples of abnormal behavior is difficult, costly, or even impossible. By modeling only the normal class, OCC can effectively identify deviations, classifying them as anomalous without requiring extensive negative samples. However, this also makes OCC inherently more complex, as it must ensure accurate classification with limited data, while also being robust against variations within the normal class.

This approach is a type of classification that contrasts with traditional multi-class or binary classification tasks, where multiple class samples are available and necessary for training. OCC methods are designed to work with highly imbalanced datasets, making them essential for domains where the occurrence of anomalous events is rare but critical to identify.

OCC methods for anomaly detection are classified based on their

underlying model architecture and data handling strategies into four distinct categories: Reconstruction-Based Approaches, Variational Autoencoders (VAEs), Convolutional Approaches and Hybrid Methods. This classification captures the diversity of techniques used to tackle the challenge of detecting outliers in varying data contexts and aids in systematically evaluating and applying the most appropriate method for different anomaly detection challenges.

2.1 Reconstruction Based Approaches

Reconstruction-based methods rely on an autoencoder's ability to reconstruct normal data patterns, identifying anomalies through high reconstruction errors. Anomalies are detected by measuring the reconstruction error. If the reconstructed data significantly deviates from the input (high reconstruction error), it is considered an anomaly. They are well suited for structured data like tabular datasets, time series, and simple image data where the normal patterns are well-represented and deviations are easy to capture.

[1,2] are some of the earliest reconstruction-based one-class classification methods. [1] proposed a method for model identification and error covariance matrix estimation using PCA, which is used for anomaly detection by identifying deviations from the learned normal data distribution. Teuvo et al. [2] introduced self-organizing maps as a method for mapping highdimensional data onto a low-dimensional grid, which can be used to identify outliers as deviations from the learned topological structure of normal data.

Fabrizio et al [3] propose utilizing a few outlying examples during training to improve autoencoder performance in detecting anomalies. [4] introduces a Latent-Insensitive Autoencoder (LIS-AE) that leverages unlabeled data from similar domains as negative examples to refine the latent space, improving the robustness and accuracy of anomaly detection. [5] proposes an adaptive encoder network model for anomaly detection in tabular datasets using reconstruction errors. [6] uses progressive reconstruction and hierarchical feature fusion (PRFF-AD) model to enhance defect detection in industrial applications. To address the issue of reconstructing anomalous data too well, [7] introduces a Autoencoder based on the robust adversarial training for novelty detection.

2.2 Variational Autoencoder (VAE)

Variational Autoencoders extend reconstruction approaches by modeling the probability distribution of data, which enables a more nuanced identification of anomalies based on likelihood. Anomalies are detected by comparing the likelihood of the input data under the learned distribution. Inputs with low likelihood (not fitting well into the learned normal data distribution) are considered anomalies. They can better represent multimodal distributions & complex time-series data, where understanding the distribution of normal data is important.

Jinwon et all [8] introduced Variational Autoencoder for anomaly detection by leveraging reconstruction probability to identify anomalous data. [9] introduces VESC, an unsupervised deep learning model utilizing a recursive reconstruction strategy for anomaly detection. [10] Combines VAEs and Normalizing Flows to enhance the robustness and accuracy of anomaly detection. [11] Explores the application of VAEs for detecting anomalies in time-series data using reconstruction-based methods. [12] Proposes a one-class VAE approach called OC-FakeDECT for detecting deepfakes by learning the distribution of authentic data. Daehyung et al. [13] proposed using an LSTM-based Variational Autoencoder to detect anomalies in robot-assisted feeding scenarios.

2.3 Convolutional Approaches

Convolutional approaches leverage convolutional layers to extract spatial features from data and then identify deviations from learned normal patterns. This capability makes convolutional approaches especially powerful for anomaly detection tasks in visual data such as images and multimodal data.

[14] suggests using a convolutional autoencoder for multimodal anomaly detection by effectively capturing features across different data types. [15] introduces a One-Class Learned Encoder-Decoder (OLED) model to learn robust representations for OCC. [16] combines deep networks with one-class objectives to tightly enclose normal data. [17] employs a CNN-based OCC approach that utilizes pseudo-negative Gaussian data in feature space and binary cross-entropy loss for training. [18] presents the Fully Convolutional Data Description (FCDD), an explainable deep OCC model that provides detection performance along with interpretable explanations. Lastly, [19] explores a 2D-CNNbased method for time-series anomaly detection, specifically for identifying water-level anomalies.

2.4 Hybrid Methods

Hybrid methods combine multiple OCC strategies or integrate traditional models with deep learning techniques to leverage their complementary strengths.. They are effective in scenarios where single models struggle.

[20] Combines a Convolutional Neural Network with a transformer backbone, to enhance anomaly detection in medical imaging. Ruff et al.[21] proposed a hybrid model combining autoencoder-based feature extraction and support vector data description for anomaly detection. Chen et al. [22] developed Interpolated Gaussian descriptor (IGD) model, that uses a oneclass Gaussian anomaly classifier trained with adversarially interpolated training samples. Tax et al. [23] proposed a one-class ensemble technique that combines multiple one-class classifiers to improve accuracy of OCC. Krawczyk et al. [24] proposed a cluster-based ensemble for OCC by splitting data into clusters and applying OCSVMs for each cluster, then combining all models to create a robust OCC model.

Tuble 11 Currentles of GCC			
#	Category	Paper	Methods
	Reconstruction	Model identification and error covariance matrix estimation from noisy data using PCA $[1]$	PCA
	Reconstruction	Self-Organizing Maps [2]	SOM
3	Reconstruction	Reconstruction Error-based Anomaly Detection with Few Outlying Examples $[3]$	AE-SAD, AE
4	Reconstruction	Latent-Insensitive autoencoders for Anomaly Detection [4]	LIS-AE, AE
	Reconstruction	Reconstruction-Based One-Class classification for tabular data [5]	Adaptive Encoder

Table 1: Categories of OCC

3. APPLICATIONS OF ONE-CLASS CLASSIFICATION

One-Class Classification (OCC) based anomaly detection is a powerful technique widely used across various fields to identify unusual patterns or behaviors that deviate from the norm. This section goes over some of its applications in different domains.

3.1 Cybersecurity

In cybersecurity, OCC is used to detect anomalies that may indicate cyber-attacks or unauthorized access. For instance, a study presented a Spiking One-Class Anomaly Detection Framework (SOCCADF) for Industrial Control Systems (ICS). [25] uses evolving Spiking Neural Networks to detect Advanced Persistent Threats (APTs) by learning from normal operational data and identifying deviations.

3.2 Industrial Monitoring

In industrial settings, OCC helps in monitoring the health of machinery and detecting faults. SOCCADF framework [25] can also detect anomalies in the operation of industrial equipment, potentially preventing damage or downtime.

3.3 Healthcare

In healthcare, OCC is used for patient monitoring and early detection of diseases. For example, anomaly detection can be applied to electrocardiogram (ECG) data to identify abnormal heartbeats that may indicate cardiac issues [26]. A comprehensive review of OCC applications [27] highlights its potential in detecting rare diseases and monitoring patient health.

3.4 Financial Fraud Detection

OCC is also crucial in financial fraud detection. By learning from legitimate transaction data, OCC models can identify fraudulent activities that deviate from normal transaction patterns. [28] shows that the use of LSTM-based autoencoders combined with Generative Adversarial Networks (GANs) can provide robust performance in detecting financial fraud.

4. CHALLENGES IN ONE-CLASS CLASSIFICATION

One-Class Classification (OCC) approaches face several challenges that can impact their effectiveness across different applications. Here are some key challenge:

4.1 Imbalanced Data

One of the primary challenges in OCC is dealing with imbalanced datasets. Since OCC models are trained on data from only one class, they may struggle to identify anomalies in highly imbalanced datasets where the normal class vastly outnumbers the anomalies. This issue is particularly prevalent in big data applications [27].

4.2 Feature Selection

Choosing the right features for OCC is crucial but challenging. The effectiveness of an OCC model heavily depends on the features used during training. Poor feature selection can lead to high false positive rates, where normal instances are incorrectly classified as anomalies [27].

4.3 Model Generalization

OCC models often face difficulties in generalizing well to unseen data. This is because they are trained on a single class,

making it hard to capture the full variability of the data. Research shows that many state-of-the-art OCC algorithms fail to generalize effectively, especially when applied to diverse datasets.

4.4 Scalability

Scalability is another significant challenge, especially when dealing with large datasets. OCC models can become computationally expensive and slow when applied to big data, making them less practical for real-time applications [27].

4.5 Evaluation Metrics

Evaluating the performance of OCC models is complex due to the lack of a clear distinction between normal and anomalous instances. Traditional metrics like accuracy are not always suitable, and alternative metrics such as precision, recall, and F1 score are often used. However, these metrics can still be misleading in highly imbalanced scenarios [29].

4.6 Handling Noise

OCC models are sensitive to noise in the training data. Since they rely on learning the characteristics of the normal class, any noise or outliers in the training data can significantly degrade their performance.

4.7 Interpretability

Interpreting the results of OCC models can be challenging. Unlike multi-class classification models, OCC models do not provide clear explanations for why a particular instance is classified as an anomaly, making it difficult for users to trust and understand the model's decisions [29].

5. RECENT TRENDS AND ADVANCEMENTS

The field of One-Class Classification (OCC) is rapidly evolving,

with a strong focus on enhancing model robustness, adaptability, and interpretability. Recent advancements include the use of adversarial training, which leverages Generative Adversarial Networks (GANs); refining optimization objectives to align decision boundaries effectively; and enhancing explainability to make models more transparent and interpretable. Some of these recent trends are discussed below.

5.1 Adversarial Learning

Researchers have been applying adversarial training techniques to OCC by utilizing Generative Adversarial Networks (GANs). These methods involve training a generator to create pseudoanomalous data while a discriminator tries to distinguish it from normal data, thus learning a more robust boundary for the normal class. OCGAN [30] integrates adversarial learning with OCC frameworks to enhance robustness and accuracy in anomaly detection.

5.2 Optimization Objectives in OCC

Researchers are exploring new ways to set optimization objectives that better align the learned feature space with the desired decision boundary. These advancements aim to prevent trivial solutions and improve the effectiveness of OCC models by fine-tuning the optimization criteria used during training [31].

5.3 Explainability in OCC

There is a growing emphasis on developing explainable OCC models that provide clear reasons for classifying certain data points as anomalies [18]. This trend is crucial for domains such as healthcare and finance, where understanding the decisionmaking process of the model is as important as the detection itself.

5.4 Domain Adaptation

Recent studies have focused on using self-supervised learning techniques to generate feature representations that can be transferred across different domains and tasks. This enables the development of more generalized OCC models that perform well even in settings with limited labeled data [32].

5.5 Dynamic Learning

New research is being conducted to develop OCC models that can adapt dynamically to evolving data distributions, especially in streaming or time-series data [11]. These models are designed to handle concept drift and changing patterns in real-time environments without requiring full re-training.

5.6 Graph-Based OCC Models

Researchers are increasingly applying Graph Neural Networks (GNNs) for OCC to capture relational and structural dependencies in data [33]. This approach is particularly useful for detecting anomalies in social media and biological systems, where the relationships between data points are critical.

5.7 Meta Learning

Recent advancements include the development of OCC models that can generalize effectively from very few or even zero labeled examples of the normal class. Techniques like meta-learning are being explored to enable these models to adapt quickly to new tasks with minimal training data [34].

6. CONCLUSION & FUTURE DIRECTIONS

This paper provides a comprehensive analysis of One-Class Classification (OCC) methods for anomaly detection, categorizing them into four primary approaches: Reconstruction-
Based Methods, Variational Autoencoders (VAEs), Based Methods, Variational Autoencoders (VAEs), Convolutional Approaches, and Hybrid Models. Each category

has demonstrated unique strengths in tackling various anomaly detection challenges across diverse applications, including healthcare, finance, and industrial monitoring. The review highlighted the pivotal role of autoencoders and their variants in advancing OCC, particularly in their ability to model complex data distributions and identify anomalies through reconstruction errors and probabilistic modeling.

Looking ahead, several key areas offer promising avenues for future research. Improving Generalization and Transferability is crucial for the development of OCC models that can be effectively applied across different domains and datasets. Future work should focus on creating more generalized frameworks that can adapt to varying data distributions and transfer learned knowledge to new, unseen environments.

Additionally, advancements in Semi-Supervised and Weakly Supervised Learning are essential to leverage limited labeled data more effectively. By incorporating a small number of labeled examples or exploiting unlabeled data, OCC models can be made more robust and accurate, even in scenarios where labeled anomalies are scarce. This is particularly relevant for applications such as rare disease detection and fraud identification, where labeled data is often insufficient. Finally, Enhancing Interpretability remains a critical challenge for the widespread adoption of OCC models in high-stakes domains like healthcare and finance. Developing techniques that provide transparent and understandable explanations for model decisions will be vital for building trust and enabling effective decision-making by human stakeholders.

7. REFERENCES

- [1] S. Narasimhan and S. L. Shah, "Model identification and error covariance matrix estimation from noisy data using PCA," Control Engineering Practice, vol. 16, no. 1, pp. 146– 155, Jun. 2007.
- [2] T. Kohonen, Self-Organizing Maps. 2001. doi: 10.1007/978-3-642-56927-2.
- [3] F. Angiulli, F. Fassetti, and L. Ferragina, "Reconstruction Error-based Anomaly Detection with Few Outlying Examples," arXiv.org, May 17, 2023.
- [4] M. S. Battikh and A. A. Lenskiy, "Latent-Insensitive autoencoders for Anomaly Detection," arXiv.org, Oct. 25, 2021.
- [5] J. Lu, W. Li, and Y. Zhao, "Reconstruction-Based One-Class classification anomaly detection for tabular data," in Communications in computer and information science, 2022, pp. 313–326. doi: 10.1007/978-981-19-4109-2_29.
- [6] F. Liu, X. Zhu, P. Feng, and L. Zeng, "Anomaly detection via progressive reconstruction and hierarchical feature fusion," Sensors, vol. 23, no. 21, p. 8750, Oct. 2023, doi: 10.3390/s23218750.
- [7] M. Salehi et al., "ARAE: Adversarially robust training of autoencoders improves novelty detection," Neural Networks, vol. 144, pp. 726–736, Sep. 2021, doi: 10.1016/j.neunet.2021.09.014.
- [8] An, Jinwon and Sungzoon Cho. "Variational Autoencoder based Anomaly Detection using Reconstruction Probability." (2015).
- [9] C. Zhang et al., "VESC: a new variational autoencoder based model for anomaly detection," International Journal of Machine Learning and Cybernetics, vol. 14, no. 3, pp. 683–696, Oct. 2022, doi: 10.1007/s13042-022-01657-w.
- [10] N. Najari, S. Berlemont, G. Lefebvre, S. Duffner, and C. Garcia, "Robust variational autoencoders and normalizing flows for unsupervised network anomaly detection," in Lecture notes in networks and systems, 2022, pp. 281–292. doi: 10.1007/978-3-030-99587-4_24.
- [11] C. Zhang, S. Li, H. Zhang, and Y. Chen, "VELC: a new variational AutoEncoder based model for Time Series anomaly detection," arXiv.org, Jul. 03, 2019.
- [12] "OC-FakeDECT: Classifying deepfakes using one-class variational autoencoder," IEEE Conference Publication | IEEE Xplore, Jun. 01, 2020.
- [13] Daehyung Park, Yuuna Hoshi, and Charles C Kemp. 2018. A multimodal anomaly detector for robot-assisted feeding using an lstm-based variational autoencoder. IEEE Robotics and Automation Letters 3, 3 (2018), 1544–1551.
- [14] F. Laakom, F. Sohrab, J. Raitoharju, A. Iosifidis, and M. Gabbouj, "Convolutional autoencoder-based multimodal one-class classification," arXiv.org, Sep. 25, 2023.
- [15] "One-Class Learned Encoder-Decoder Network with Adversarial Context Masking for Novelty Detection," IEEE Conference Publication | IEEE Xplore, Jan. 01, 2022.
- [16] R. Chalapathy, A. K. Menon, and S. Chawla, "Anomaly Detection using One-Class Neural Networks," arXiv.org, Feb. 18, 2018.
- [17] P. Oza and V. M. Patel, "One-Class Convolutional Neural Network," IEEE Signal Processing Letters, vol. 26, no. 2, pp. 277–281, Feb. 2019, doi: 10.1109/lsp.2018.2889273.
- [18] P. Liznerski, L. Ruff, R. A. Vandermeulen, B. J. Franks, M. Kloft, and K.-R. Müller, "Explainable deep One-Class classification," arXiv.org, Jul. 03, 2020.
- [19] I. T. Nicholaus, J.-S. Lee, and D.-K. Kang, "One-Class convolutional neural networks for Water-Level anomaly detection," Sensors, vol. 22, no. 22, p. 8764, Nov. 2022, doi: 10.3390/s22228764.
- [20] Z. Zhang, B. Patel, B. Patel, and I. Banerjee, "Unsupervised Hybrid framework for ANomaly Detection (HAND) - applied to Screening Mammogram," arXiv.org, Sep 2024.
- [21] L. Ruff et al., "Deep One-Class Classification," PMLR, Jul. 03, 2018.
- [22] Y. Chen, Y. Tian, G. Pang, and G. Carneiro, "Deep One-Class Classification via Interpolated Gaussian Descriptor", AAAI, vol. 36, no. 1, pp. 383-392, Jun. 2022.
- [23] D. M. J. Tax and R. P. W. Duin, "Combining One-Class classifiers," in Lecture notes in computer science, 2001, pp. 299–308. doi: 10.1007/3-540-48219-9_30.
- [24] B. Krawczyk, M. Woźniak, and B. Cyganek, "Clusteringbased ensembles for one-class classification," Information Sciences, vol. 264, pp. 182–195, Jan. 2014, doi: 10.1016/j.ins.2013.12.019.
- [25] K. Demertzis, L. Iliadis, and S. Spartalis, "A spiking One-Class anomaly detection framework for Cyber-Security on industrial control systems," in Communications in computer and information science, 2017, pp. 122–134. doi: 10.1007/978-3-319-65172-9_11.
- [26] E. Šabić, D. Keeley, B. Henderson, and S. Nannemann, "Healthcare and anomaly detection: using machine learning to predict anomalies in heart rate data," AI & Society, vol. 36, no. 1, pp. 149–158, May 2020, doi: 10.1007/s00146-

International Journal of Computer Applications (0975 – 8887) Volume 186 – No.45, October 2024

020-00985-1.

- [27] N. Seliya, A. A. Zadeh, and T. M. Khoshgoftaar, "A literature review on one-class classification and its potential applications in big data," Journal of Big Data, vol. 8, no. 1, Sep. 2021, doi: 10.1186/s40537-021-00514-x.
- [28] P. Zheng, S. Yuan, X. Wu, J. Li, and A. Lu, "One-Class adversarial nets for fraud detection," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, no. 01, pp. 1286–1293, Jul. 2019, doi: 10.1609/aaai.v33i01.33011286.
- [29] P. Perera, P. Oza, and V. M. Patel, "One-Class Classification: a survey," arXiv.org, Jan. 08, 2021. https://arxiv.org/abs/2101.03064
- [30] P. Perera, R. Nallapati, and B. Xiang, "OCGAN: One-class Novelty Detection Using GANs with Constrained Latent Representations,", CVPR, Mar. 20, 2019.
- [31] H. Gao, H. Luo, F. Shen, and Z. Zhang, "Exploring the optimization objective of One-Class classification for anomaly detection," arXiv.org, Aug. 23, 2023.
- [32] G. Michau and O. Fink, "Domain adaptation for One-Class classification: Monitoring the health of critical systems under limited information," International Journal of Prognostics and Health Management, vol. 10, no. 4, Dec. 2019, doi: 10.36001/ijphm.2019.v10i4.2613.
- [33] X. Wang, B. Jin, Y. Du, P. Cui, Y. Tan, and Y. Yang, "Oneclass graph neural networks for anomaly detection in attributed networks," Neural Computing and Applications, vol. 33, no. 18, pp. 12073–12085, Mar. 2021, doi: 10.1007/s00521-021-05924-9.
- [34] A. Oladosu et al., "Meta-Learning for One-Class Classification with Few Examples using Order-Equivariant Network," arXiv.org, Jul. 08, 2020. https://arxiv.org/abs/2007.04459