IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received November 5, 2021, accepted November 24, 2021, date of publication November 30, 2021,
date of current version December 13, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3131949

Applications of Generative Adversarial Networks in Anomaly Detection: A Systematic Literature Review

MIKAEL SABUHI[1], (Graduate Student Member, IEEE),
MING ZHOU[1], (Graduate Student Member, IEEE),
COR-PAUL BEZEMER[1], (Member, IEEE),
AND PETR MUSILEK[1,2], (Senior Member, IEEE)
[1]Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2R3, Canada
[2]Department of Applied Cybernetics, University of Hradec Králové, 500 03 Hradec Kralove, Czech Republic

Corresponding author: Mikael Sabuhi (sabuhi@ualberta.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Project ALLRP549804-19 and Project CRD543899-19, and in part by the Government of Alberta under Major Innovation Fund Project RCP-19-001-MIF.

ABSTRACT Anomaly detection has become an indispensable tool for modern society, applied in a wide range of applications, from detecting fraudulent transactions to malignant brain tumors. Over time, many anomaly detection techniques have been introduced. However, in general, they all suffer from the same problem: lack of data that represents anomalous behaviour. As anomalous behaviour is usually costly (or dangerous) for a system, it is difficult to gather enough data that represents such behaviour. This, in turn, makes it difficult to develop and evaluate anomaly detection techniques. Recently, generative adversarial networks (GANs) have attracted much attention in anomaly detection research, due to their unique ability to generate new data. In this paper, we present a systematic review of the literature in this area, covering 128 papers. The goal of this review paper is to analyze the relation between anomaly detection techniques and types of GANs, to identify the most common application domains for GAN-assisted and GAN-based anomaly detection, and to assemble information on datasets and performance metrics used to assess them. Our study helps researchers and practitioners to find the most suitable GAN-assisted anomaly detection technique for their application. In addition, we present a research roadmap for future studies in this area. In summary, GANs are used in anomaly detection to address the problem of insufficient amount of data for the anomalous behaviour, either through data augmentation or representation learning. The most commonly used GAN architectures are DCGANs, standard GANs, and cGANs. The primary application domains include medicine, surveillance and intrusion detection.

INDEX TERMS Anomaly detection, data augmentation, generative adversarial networks, outlier detection, representation learning.

I. INTRODUCTION

In modern society, many systems depend on and generate enormous amounts of data. This data is important for many decision-making processes. Normally, systems operate under the expected conditions. However, in rare cases, anomalies may occur. Such anomalies can have a disastrous impact on the system itself or on its environment. Therefore, to lower the impact, it is important to be able to detect such anomalies as early as possible. For example, cancer is an anomaly in human tissue. Breast cancer is the second leading cause of cancer death in women [1]. According to a recent study by the American Cancer Society [2], breast cancer alone accounts for 30% of female cancers. Early detection and treatment of breast cancer would highly increase the chance of survival [3]. Similarly, with an increasing need to ensure public safety in crowded areas, the development of real-time video surveillance systems becomes unavoidable. It is critical to

The associate editor coordinating the review of this manuscript and approving it for publication was Baker Mohammad.

seamlessly monitor the crowd to immediately detect anomalous (or *abnormal*) movements to help prevent theft [4], vandalism [5], and terrorist attacks [6].

The process of finding the anomalous behaviour of a system is referred to as *anomaly detection*. The primary objective of anomaly detection is to differentiate between the expected and unexpected behaviour of a system. Considering the importance of anomaly detection, it has received widespread attention in research. Despite the progress in this research area, there is still an important open challenge: the acquisition of data about anomalies that can be used to test anomaly detection techniques.

A recent trend in anomaly detection is the use of generative adversarial networks (GANs). Proposed by Ian Goodfellow *et al.* [7] in 2014, GANs are a type of unsupervised generative model which gained much attention from the research community. A well-trained GAN can generate realistic-looking data by sampling from a learned data distribution. A GAN consists of a generator and a discriminator model. These two models are pitted against each other in a two-player zero-sum game situation, iteratively improving their capabilities to generate and discriminate data.

The ability of GANs to generate data makes them attractive for anomaly detection research from two perspectives. First, they can potentially help generate the hard-to-acquire anomalous data points. Second, they can be used to learn the distribution of the data for the normal operating condition of a system and act as an anomaly or outlier detector.

In this paper, we conduct a systematic literature review of the applications of GANs for anomaly detection. We address the following research questions (RQs):

- **RQ1: What is the role of GANs in anomaly detection?** We identified two roles that GANs play in anomaly detection: data augmentation and representation learning. In contrast to the remarkable ability of GANs to generate realistic-looking data, most of the reviewed papers use them for representation learning rather than data augmentation. The reason for this inclination is that, despite the improvement in the anomaly detection accuracy after data augmentation, the reported improvements are not substantial. When GANs are used for data augmentation in anomaly detection, we refer to it as GAN-assisted anomaly detection. The other role of GANs in anomaly detection is representation learning. In this case, the examined papers use the data from the normal class for training a GAN to learn the distribution of the normal data. A score is assigned to the new data by defining a score function, and the anomalous data in the test stage is identified based on a specific threshold. We refer to these techniques as GAN-based anomaly detection.
- **RQ2: What are the application domains of anomaly detection with GANs?** The primary application areas where GANs are used for anomaly detection are medicine (19%), surveillance (15%) and intrusion detection (13%).
- **RQ3: Which GAN architecture is used most often in anomaly detection systems?** We identified 21 architectures of GANs that are used for anomaly detection. Among these architectures, deep convolutional GANs (DCGANs) (32%), standard GANs (23%), and conditional GANs (16%) are the most commonly used.
- **RQ4: Which type of data instance and datasets are most commonly used for anomaly detection with GANs?** 50% of the proposed GAN-based anomaly detection techniques use image datasets for anomaly detection purposes. Before being fed to the anomaly detection algorithms, the data are usually preprocessed. The most common preprocessing methods are resizing (23%), normalization (19%), and cropping (13%).
- **RQ5: Which metrics are used to evaluate the performance of GANs in generating data and anomaly detection?** Only 21% of the studied papers evaluated the GAN's performance in generating synthetic data, either in data augmentation or representation learning. Structural similarity indices (SSIM) (26%) and peak signal-to-noise ratio (PSNR) (26%) are the most commonly used metrics. Visual inspection to evaluate the quality of the generated data was reported in 5% of the studied papers. To evaluate the performance of GANs in anomaly detection applications, 53% of the primary studies used the area under the receiver operating characteristic curve (AUROC).
- **RQ6: Which anomaly detection techniques are used along with GANs?** GAN-based anomaly detection is mostly done in a semi-supervised manner. DCGANs and standard GANs are the most popular architectures in semi-supervised anomaly detection using GANs.

  In supervised learning-based anomaly detection, GANs are used to augment the dataset for the anomalous class. However, the studied papers report only minor improvements in the performance of anomaly detection techniques after augmenting the dataset with GANs. Only a few primary studies focused on pure unsupervised anomaly detection based on GANs, most using the standard version of GANs. Similar to semi-supervised techniques, unsupervised GAN-based anomaly detection techniques are mostly compared with autoencoder-based approaches.

The findings presented in this survey will help researchers and practitioners to find the most suitable GAN-based anomaly detection techniques for their applications.

The rest of this paper is organized as follows. Section II provides a brief introduction to GANs. Section III describes the methodology used for conducting this systematic literature review. Section IV presents the results of the review. Section V discusses the open challenges and provides directions for future research. Section VI identifies the threats to validity of the review, and Section VII concludes the paper.

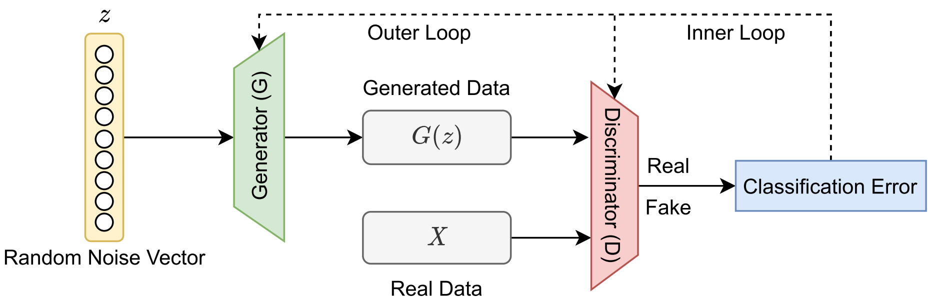


FIGURE 1. **The building blocks of GANs. The classification error is used to update the parameters of the discriminator and generator models (shown by dashed lines).**

II. GENERATIVE ADVERSARIAL NETWORKS

In 2014, Goodfellow *et al.* [7] introduced a framework for estimating generative models based on an adversarial process. This framework consists of two deep neural network-based models: a generative model G and a discriminator model D. Model G learns the training data distribution and uses it to generate new samples. Model D determines whether a sample comes from the training data or was generated by the generative model. The power of GANs comes from the adversarial process, in which the two models are competing against each other to improve their accuracy in the designated task.

The diagram in Figure 1 shows the building blocks of a GAN [8]:

- The *Real Data (X)*, or the training dataset, contains the instances that the generator G should learn to generate, usually in the form of a batch.
- *Random Noise Vector (z)* is the raw input to the generator. It is a vector of random numbers which the generator uses to generate fake examples.
- The *Generator model (G)* is trained to learn the distribution of the input data. This model uses the input (z) to generate fake examples ($G(z)$) that are indistinguishable from the real data.
- The *Discriminator model (D)* tries to distinguish the data that is generated by the generator from the real data. The inputs to this model are the real data (X) and the generated data ($G(z)$). The output of this model is a binary decision for each data instance, i.e. real/fake.
- *Iterative Training:* The GAN is trained using the classification error of the discriminator. The error is used to tune the parameters (weights and biases) of the discriminator, and then the parameters of the generator. Backpropagation [9] is commonly used as the training algorithm. This iterative training consists of two loops:
  - An *inner loop* where the discriminator's parameters are tuned to maximize the classification accuracy of predicting correct labels for real data and generated data.
  - An *outer loop* where the generator's parameters are tuned to generate data that has a minimal chance of being distinguished from the real data by the discriminator.

The adversarial training of the generator and the discriminator model is a zero-sum game problem: when one model gets better the other one gets worse in equal proportions [8]. For all zero-sum games, there is a point where neither of the players can improve their situation. This point is referred to as the Nash equilibrium. The goal of a GAN is to reach this equilibrium, as then the fake data produced by the generator model is indistinguishable from the real data by the discriminator model. The output of the discriminator is then a random guess on whether the input data is real or fake.

III. METHODOLOGY

The planning, conducting, and reporting of this systematic literature review (SLR) were based on the guidelines proposed by Kitchenham [10]. The planning stage of the SLR includes three steps: identification of the need for the systematic review, development of the review protocol, and evaluation of the protocol [10]. In the conducting stage, based on the review protocol that was developed during planning stage, we search for and select the primary studies, extract data from the primary studies, and synthesize the data. The set of primary studies contains all individual studies that contribute to the SLR [10]. In the last stage, we conclude the systematic review by reporting the collected data and findings. Figure 2 summarizes the required steps for each stage of the review. In the following, each step is explained in more detail.

A. THE NEED FOR A SYSTEMATIC REVIEW

Recently, GANs have become a hot research topic in many application domains. One of these domains is anomaly detection. The ability of GANs to generate realistic looking data and to perform representation learning makes them attractive for anomaly detection research. Basically, GANs are trained in an unsupervised manner to learn the distribution of the data. However, they are highly flexible and can be used in semi-supervised fashion as well (e.g. [11]). In addition, GANs are implicit density models which do not require any explicit hypothesis on the distribution of the data [12]. Considering all these advantages, GANs can be leveraged to

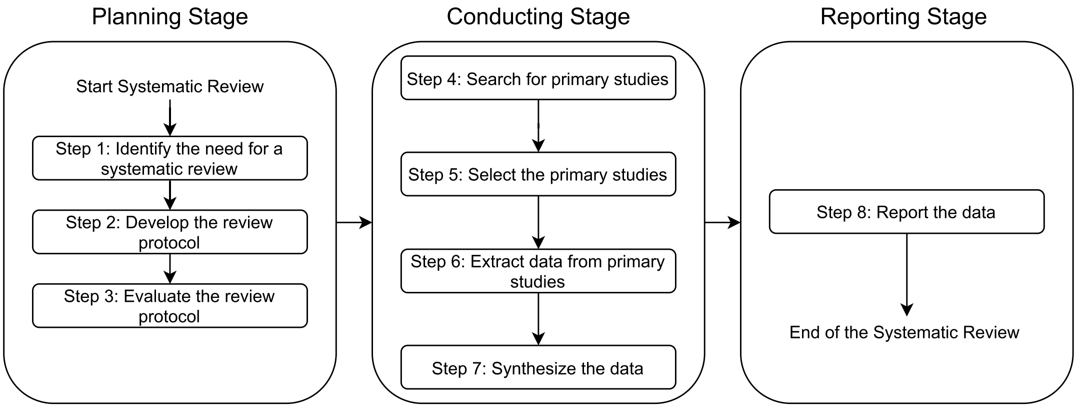

Planning Stage
Start Systematic Review
Step 1: Identify the need for a systematic review
Step 2: Develop the review protocol
Step 3: Evaluate the review protocol
Conducting Stage
Step 4: Search for primary studies
Step 5: Select the primary studies
Step 6: Extract data from primary studies
Step 7: Synthesize the data
Reporting Stage
Step 8: Report the data
End of the Systematic Review

FIGURE 2. **The steps of our systematic literature review, based on Kitchenham's guidelines [10].**

address some existing problems in anomaly detection, such as the lack of a sufficient amount of data for anomalous behaviour of the system. Therefore, a study summarizing existing research on applications of GANs in anomaly detection would be of a high value to the research community.

When we started this systematic literature review, we identified only one survey paper [13] reviewing applications of GANs in anomaly detection. However, this paper only covered 11 papers on anomaly detection with GANs. In addition, it did not follow a systematic approach to conduct the review. This confirmed the need for a systematic literature review on applications of GANs in anomaly detection, which covers a vast number of papers. To reduce researcher bias [10], we followed a systematic approach for designing, executing and reporting our findings.

B. DEVELOPING THE REVIEW PROTOCOL

To reduce the possibility of researcher bias in a systematic manner, a review protocol is required to specify the method for conducting the systematic review [10]. This protocol includes definition of the following elements: 1) research questions, 2) search strategy, 3) study selection criteria (including study quality assessment), 4) data extraction strategy, and 5) synthesis of the extracted data.

1) OUR RESEARCH QUESTIONS

In this systematic literature review, we address the following research questions (RQs):

1) *RQ1: What is the role of GANs in anomaly detection?* (Section IV-A)
Motivation: It is important to learn how GANs are used in anomaly detection. One intuitive way is to generate anomalous data to address the problem of the imbalanced dataset. Still, there might be more opportunities. Moreover, we will investigate what are the alternative, non-GAN approaches to handle these identified roles.
2) *RQ2: What are the application domains of anomaly detection with GANs?* (Section IV-B)
Motivation: The use of GANs in anomaly detection may be more common in certain domains. Here, we look into which domains and which types of GANs work together well.
3) *RQ3: Which GAN architecture is used most often in anomaly detection systems?* (Section IV-C)
Motivation: There exist many architectures of GANs. Each one attempts to handle a specific type of data or to address an existing problem in the previous architectures. Some architectures may be better suitable for anomaly detection than others. Therefore, we look into which architectures of GANs are commonly used.
4) *RQ4: Which type of data instance and datasets are most commonly used for anomaly detection with GANs?* (Section IV-D)
Motivation: Identifying which datasets are used to evaluate anomaly detection with GANs in certain domains can reveal the "standard benchmarks" in specific domains and which domains require benchmarks in general.
5) *RQ5: Which metrics are used to evaluate the performance of GANs in generating data and anomaly detection?* (Section IV-E)
Motivation: Evaluating GANs in anomaly detection systems is not a straightforward task as their goal is to create realistic looking data that is different enough from known anomalies, yet still representative of real anomalies. Therefore, one cannot just compare the generated data with the real data. We study which approaches are commonly used to evaluate the quality of the generated data, and support practitioners in deciding which metrics to use for evaluating data in specific anomaly detection problems.
6) *RQ6: Which anomaly detection techniques are used along with GANs?* (Section IV-F)
Motivation: GANs are often used together with more traditional anomaly detection techniques, especially when they are used in a supervised manner. In this

question, we identify the anomaly detection techniques that are based on or assisted by GANs.

2) SEARCH STRATEGY

To find relevant papers for this systematic review, we searched the IEEE Xplore,[1] ACM Digital Library,[2] Science Direct[3] and Scopus[4] digital libraries. The focus of this study is the application of GANs in anomaly detection. Therefore, we combined the keywords related to anomaly detection with keywords and abbreviations for generative adversarial networks. To find closely relevant papers for this study, we searched the title and the abstract of the papers for the following query: ("anomaly" `OR` "anomalies" `OR` "anomalous" `OR` "outlier" `OR` "abnormal") `AND` ("generative adversarial network" `OR` "generative adversarial networks" `OR` "GAN" `OR` "GANs"). The list of primary studies was collected on 3rd June, 2020.

We conducted a pilot study to ensure that the well-known primary studies were included in the query results. During this study, we searched for the matched papers and their shared references on Google Scholar to ensure that the most cited papers were covered by the query. After several iterations of improving the query, we were confident that it returned important and well-known studies.

3) STUDY SELECTION CRITERIA

We defined the following criteria for the inclusion of a paper in our study. The last two criteria in the list are included to assess the quality of the study.

- The paper must be in the specified digital libraries.
- The primary study should focus on anomaly detection while leveraging GANs.
- The developed methods should be evaluated on at least one real dataset, not only on simulated data, to ensure the practical relevance of the study.
- The primary study should be available online to ensure accessibility.
- The article should be written in the English language.

All types of papers, including journal, conference, workshop, and symposium papers are considered in this review. The procedure for searching (described in Section III-B2) and selecting the primary studies is shown in Figure 3. The final list of papers used for data extraction and synthesis consists of 128 primary studies. Not every selected primary study provides answers to all six research questions.

4) DATA EXTRACTION STRATEGY

To facilitate the data extraction, we devised a data extraction form to collect the required information for each RQ from the primary studies. This form (shown in Table 1) was refined through several iterations with randomly selected papers on the subject. This refinement was accomplished by comparing the data extraction forms of the two first authors and addressing the potential ambiguities in the data extraction form.

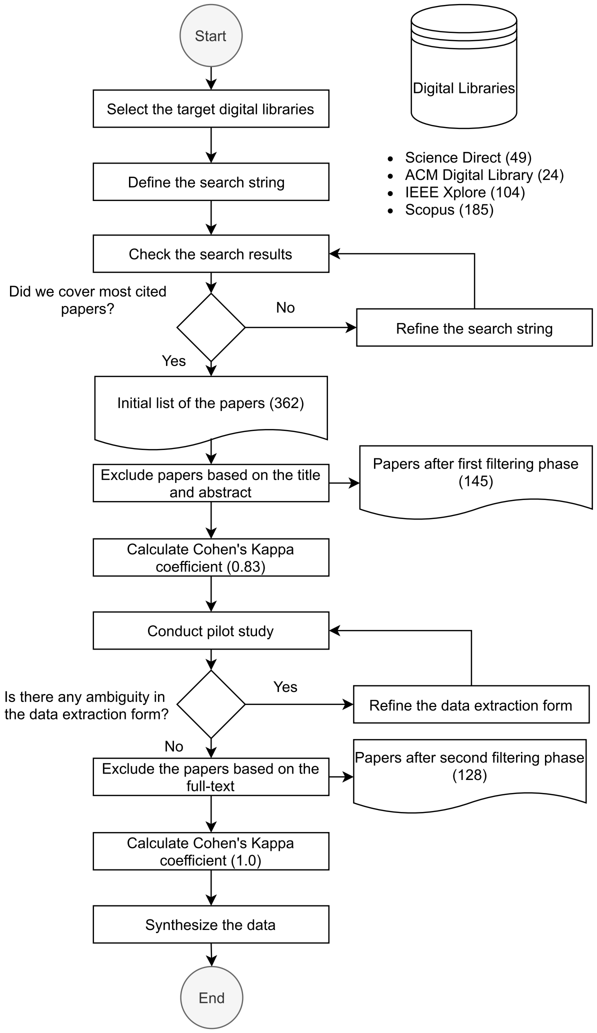

Start
Digital Libraries
Select the target digital libraries
Define the search string
Science Direct (49)
ACM Digital Library (24)
IEEE Xplore (104)
Scopus (185)
Check the search results
Did we cover most cited papers?
No
Refine the search string
Yes
Initial list of the papers (362)
Exclude papers based on the title and abstract
Papers after first filtering phase (145)
Calculate Cohen's Kappa coefficient (0.83)
Conduct pilot study
Is there any ambiguity in the data extraction form?
Yes
Refine the data extraction form
No
Exclude the papers based on the full-text
Papers after second filtering phase (128)
Calculate Cohen's Kappa coefficient (1.0)
Synthesize the data
End

FIGURE 3. The procedure for searching, selecting, and extracting data from the primary studies for conducting the SLR.

5) DATA SYNTHESIS

During the data synthesis step, we aggregate the data collected from the data extraction forms to answer the research questions. Putting all this data together gives invaluable information concerning the current best practices and architectures for anomaly detection with GANs.

C. EVALUATION OF THE REVIEW PROTOCOL

The protocol is a critical part of the SLR. It was evaluated by the last two authors and, after several iterations, the final version of the protocol was approved and used throughout the conducting stage of the SLR.

D. CONDUCTING THE SLR

The conducting stage of the SLR includes the following four steps: searching for primary studies, selecting the primary

[1] https://ieeexplore.ieee.org
[2] https://dl.acm.org
[3] https://sciencedirect.com
[4] https://scopus.com

TABLE 1. The data extraction form.

Reviewer Name / Review Date / Study Identifier / Publication Title / Names of Authors / Publication Source / Type of Study
RQ1: What is the role of GANs in anomaly detection?
  Role(s) of the GAN:
  Is the GAN used for generating new data or learning the distribution of the data?
  What is the type of generated/learned data (normal/abnormal)?
RQ2: What are the application domains of anomaly detection with GANs?
  The application domain(s):
RQ3: Which GAN architecture is used most often in anomaly detection systems?
  The GAN architecture(s) used in the study:
RQ4: Which type of data instance and datasets are most commonly used for anomaly detection with GANs?
  The type of input data to the GAN (main input): (e.g. image, text, etc.)
  The preprocessing technique used on the input data:
  Datasets that are used for the study:
  Usage of the dataset: (e.g., addressing the unbalanced dataset, training on normal/abnormal, etc.)
RQ5: Which metrics are used to evaluate the performance of GANs in generating data and anomaly detection?
  The type of performance metrics used for evaluating the performance of GAN:
RQ6: Which anomaly detection techniques are used along with GANs?
  The type of anomaly detection techniques used: (e.g., classification-based, clustering-based, etc)
  The anomaly detection techniques: (e.g. K- Nearest neighborhood, Neural networks)

studies, extracting the primary studies, and synthesizing the data.

We identified 362 papers that matched our search query (see Section III-B2): 49 papers in Science Direct, 24 in ACM Digital Library, 104 in IEEE Xplore, and 185 in Scopus.

We organized the papers for further analysis using Mendeley as a reference manager.

We filtered out irrelevant and duplicated papers according to the study selection criteria introduced in Section III-B3. Figure 3 shows the procedure for selecting the primary studies. We filtered the papers in two steps. In the first step, the first two authors independently read the abstract and the title of the primary studies and decided if they were related to anomaly detection with GANs. The Cohen's kappa coefficient [14] for this binary classification (relevant vs. irrelevant) was 0.83, which shows a satisfactory agreement between the researchers. There were 32 papers from the initial list of papers on which the first two authors disagreed. For those papers, the third author was asked to make the final decision regarding inclusion or exclusion. After the first phase of filtering papers, we ended up with 145 papers. The second filtering step was performed while reading the full text: the first two authors decided to include or exclude the paper in the data extraction step. In this phase, the first two authors made the same decision regarding the excluded papers and excluded 17 papers. Finally, 128 primary studies were included and analysed in this SLR.

Based on the data extraction strategy introduced in Section III-B4, we examined these 128 primary studies to collect the data that contributes to addressing the RQs of this SLR. The primary studies were randomly divided into 10 batches. For each batch, the first two authors extracted the data from the primary studies and filled in the data extraction forms. After extracting the data from each batch, the data extraction forms were randomly distributed between the first two authors, and then the disagreements were identified and discussed in a meeting. If they failed to reach a consensus, one of the last two authors made the final decision. After extracting data from all primary studies and addressing the discrepancies in the data extraction forms, we created a spreadsheet for each data extraction form.

We summarize and report the extracted data in the following section.

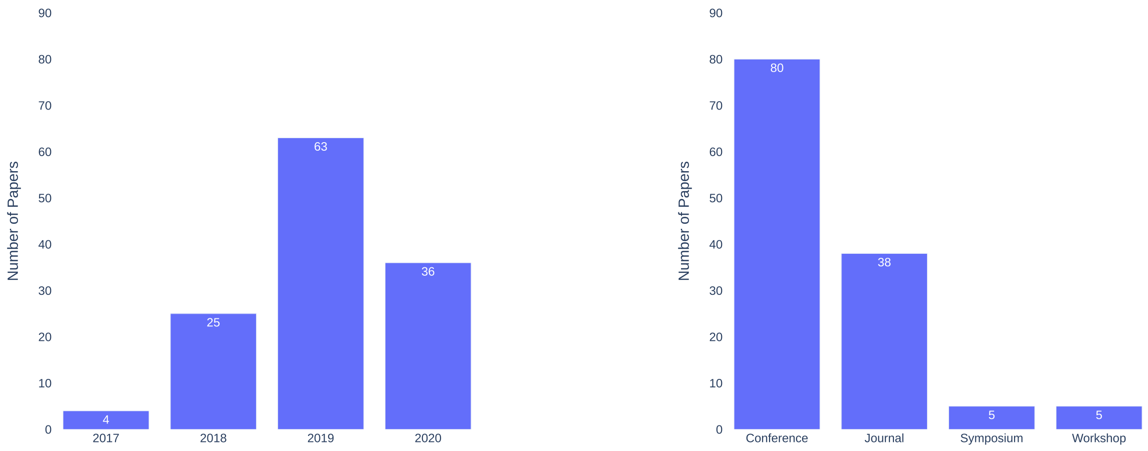
IV. RESULTS

This systematic literature review covers 128 primary studies that describe applications of GANs in anomaly detection. As shown in Figure 4(a) these primary studies were published between 2017 and early 2020. The number of studies per year is increasing, suggesting that interest in this research area is growing rapidly. Figure 4(b) shows that the majority of the reviewed papers (63%) appeared in conference proceedings, 29% of the papers were published in journals, and 4% in workshops and 4% in symposia.

A. RQ1: WHAT IS THE ROLE OF GANs IN ANOMALY DETECTION?

We identified four types of GAN applications in anomaly detection: (1) generating abnormal data instances, (2) generating normal and abnormal instances, (3) learning the normal behaviour of a system, (4) learning both normal and abnormal behaviour of a system. Applications 1 and 2 can be classified as data augmentation with GANs and 3 and 4 as representation learning with GANs.

Generative models, such as GANs, are mainly designed for data augmentation, i.e., to generate new data and use it to augment the existing data. They can also be used for representation learning, i.e., to learn representations of the data to support information extraction for use when building classifiers or other predictors [15]. In this case, the generator and the discriminator of a GAN can be used to learn the distribution of a specific class of data, i.e. normal or abnormal data. In turn, the learned distribution can be used to identify nonconforming or irregular data. Table 2 shows the two main roles of GANs in anomaly detection, along with the types of data used in each role. Most of the primary studies opted



(a) Publication years.

(b) Publication types.

FIGURE 4. Distribution of the primary studies according to the publication types and years.

to use GANs to learn the representation of the data rather than augmenting the datasets. Moreover, this representation learning is mainly performed on normal data. The rationale behind this preference is that, due to the data imbalance, it is usually easier to learn a model of normality rather than abnormality. In addition, by learning only the normal data distribution, the need for data from the abnormal condition of the target system is eliminated.

1) REPRESENTATION LEARNING WITH GENERATIVE ADVERSARIAL NETWORKS

The main goal of GANs is to learn a generative model that produces realistic-looking data by sampling from the learned distribution. This generative power of GANs was highlighted by Goodfellow *et al.* [7] and Radford *et al.* [144]. Representation learning with GANs for anomaly detection exploits the ability of GANs in learning the distribution of a specific class of data, as shown in Figure 5. Several anomaly detection techniques are proposed that use this representation learning ability of GANs (shown in Table 3). We will explain the concept of anomaly detection using representation learning through an examples of a well-known GAN-based anomaly detection techniques (AnoGAN). All other anomaly detection techniques that rely on representation learning through a GAN are variations to some extent of the AnoGAN technique.

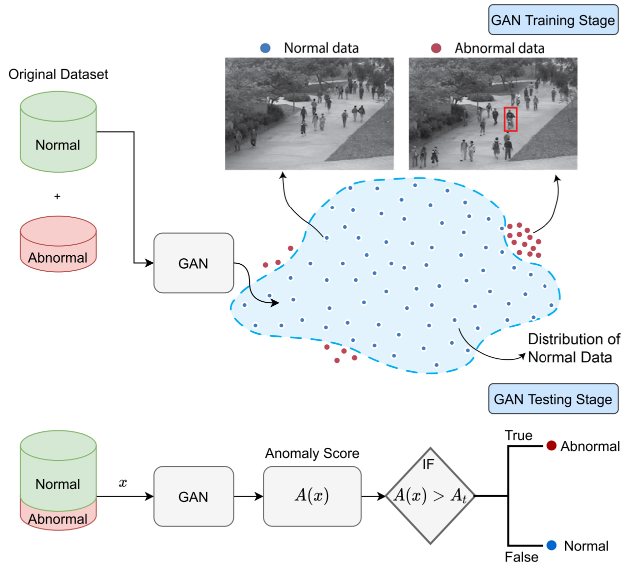


FIGURE 5. Overview of representation learning with GANs (GAN-based anomaly detection). In the training stage, a GAN is trained to learn the distribution of the normal data (shown as the blue area) and also to minimize the output of an anomaly score [$A(x)$] for every generated data. In the testing stage, since the GAN is only trained on the normal data distribution, the anomaly score will be higher than a threshold (A_t) for abnormal data, indicating anomalies.

Schlegl *et al.* [104] introduced the first GAN-based anomaly detection technique, called AnoGAN, taking advantage of the representation learning ability of GANs. Schlegl *et al.* put forth AnoGAN, which employs the DCGAN architecture, to learn the distribution of normal anatomical variability. The idea comes from the concept of a smooth transition in the latent space of the data, i.e., that sampling from two close points in the latent space should lead to similar data in the data space [145]. Schlegl *et al.* hypothesize that the latent vector of the GANs represents the distribution of the trained data. Therefore, one can learn the representation of the normal data by training GANs only on normal data. From an anomaly detection view, learning the representation of the normal data is useful as one can decide for new (potentially anomalous) data points how likely they are part of that normal data. During the training of a GAN, the generator learns the mapping from latent space to data space $G(z) = z \rightarrow x$ (i.e., the representation of the data). However, the inverse mapping, which is necessary to decide whether a data point is anomalous, is not straightforward to obtain [104]. To address this problem, Schlegl *et al.* proposed an additional step after training the GAN on normal data. For an image x, they

TABLE 2. List of studies using normal, abnormal, and normal and abnormal data together for different tasks of GANs.

Task	Data type	List of references
Representation Learning	Normal	[16]–[105]
	Normal and Abnormal	[106]–[109]
Data Augmentation	Abnormal	[110]–[135]
	Normal and Abnormal	[136]–[143]

TABLE 3. Representation learning with GANs.

Type of GAN	List of references
DCGAN	[18], [22], [25], [30]–[33], [35], [40], [44], [49], [53], [55], [60], [66], [69], [77], [78], [81], [84], [86], [88], [92], [95]–[97], [99], [103]–[105], [108], [109]
Standard GAN	[16], [21], [24], [36], [38], [43], [45], [46], [51], [52], [54], [58], [59], [61], [70], [75], [79], [83], [87], [93], [94], [100], [107]
cGAN	[23], [26], [37], [39], [56], [57], [64], [65], [72]–[74], [76], [85], [90]
BiGAN	[28], [34], [47], [50], [62], [67], [87], [91]
WGAN	[17], [27], [67], [80], [87], [95]
WGAN-GP	[42], [63], [102], [106]
VAE-GAN	[19], [20], [98], [101]
O-GAN	[29], [71]
Cycle-GAN	[48]
EBGAN	[56]
GAN-QP	[71]
OCGAN	[89]
PatchGAN	[41]
RaSGAN	[82]
TextGAN	[68]

proposed to find a point z in the latent space that corresponds to an image $G(z)$, which is the most similar to the image x on the learned manifold χ. Schlegel *et al.* proposed an iterative process to find the most similar image $G(z_\Gamma)$ to x using residual and discrimination loss. The similarity of images x and $G(z)$ depends on how closely x follows the distribution of the data learned by the generator (p_g). After identifying the most similar image, AnoGAN computes an anomaly score that is related to the similarity of x and $G(z)$. Finally, based on a threshold for the anomaly score, AnoGAN decides whether x is an anomaly.

2) DATA AUGMENTATION WITH GENERATIVE ADVERSARIAL NETWORKS

Machine learning techniques, especially deep learning methods [146], require a massive amount of data to perform well in their designated task [147]. Data augmentation, also known as oversampling, is carried out to compensate for an insufficient amount of data in the dataset to prevent model overfitting. It can also be used to address the problem of data imbalance, which occurs when the sizes of the classes in a dataset differ considerably. For instance, in a binary classification task, the class with fewer samples is called the minority class, and the other class is called the majority class. The corresponding training process would be biased towards the majority class, hence a classifier trained using this dataset would have a better accuracy for this class [148]. To address the imbalanced dataset problem, one can either randomly remove samples from the majority class to balance the class size (undersampling), or augment the minority class by adding artificially generated instances (oversampling) using suitable techniques.

The problem of the imbalanced dataset is more critical in anomaly detection since it is hard and expensive to collect data on anomalous behaviour of the system under study. Often, there are very few or no examples of anomalous data available. In this situation, GANs can help by generating more samples for the anomalous class, as shown in Figure 6.

Table 2 summarizes the main roles of GANs in anomaly detection and the type of data used for that purpose. Data augmentation with GANs is mostly used to generate data that represents anomalous behaviour of the system. There was no primary study augmenting only the normal condition of the system under study for anomaly detection. This is due to the fact that there is usually an abundance of data for the normal condition. However, some studies augmented both normal and abnormal data, e.g., using Cycle-GANs, by learning the transformation from abnormal to normal and from normal to abnormal to generate new data. After augmenting the dataset, it is ready to be used for anomaly detection, usually performed by a classifier (as discussed in Section IV-F). Most primary studies that use GANs for data augmentation report a slight improvement in classification accuracy compared either to traditional techniques or without data augmentation. For instance, Madani *et al.* [135] report that, using data augmentation with GANs, the test accuracy for cardiovascular abnormality detection improved from 81.93% to 84.19%. In comparison, using traditional augmentation methods, they achieved only 83.12% test accuracy. This improvement is significant when dealing with large amounts of data, especially

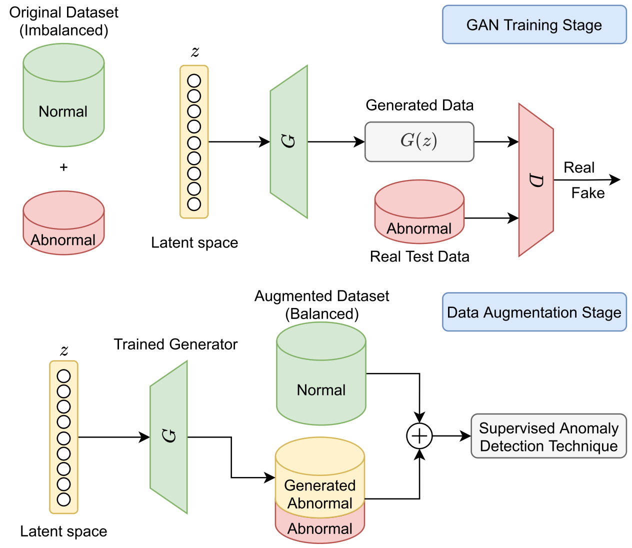


FIGURE 6. Overview of data augmentation with GANs (GAN-assisted anomaly detection). In the training stage, a GAN learns to generate realistic-looking samples for the minority (abnormal) class. In the data augmentation stage, the generator of the trained GAN is used to generate samples to augment the minority class.

in medical applications. However, in some studies, it has been reported that GAN did not meet their expectations in improving the classification accuracy after augmenting the data, e.g. [141]. The list of different GANs used for data augmentation is shown in Table 4.

In the examined primary studies, we identified several traditional techniques for addressing the problem of imbalanced and insufficient amounts of data. The effects of adopting these techniques are compared to the GANs in terms of improving the classification accuracy. For example, random undersampling was evaluated in two primary studies [110], [126], where samples were randomly removed from the majority class. Using this approach, some important and critical data may be lost that could otherwise be beneficial for learning a robust decision boundary [149]. Random oversampling was investigated in four primary studies [53], [110], [111], [126]. In this case, some samples from the minority class are copied to increase its size. However, this approach is likely to cause over-fitting [126]. All these studies confirmed the superiority of GANs in data augmentation compared to random over/undersampling.

Chawla *et al.* [150] proposed synthetic minority oversampling (SMOTE) to improve the random oversampling by synthesizing new samples from the neighbourhood of the minority class samples. This improvement is accomplished by interpolating between several minority class instances. SMOTE and its variants (e.g., borderline-SMOTE [151]) were compared with GANs in several studies [53], [110], [111], [117], [126], [132]. The ADAptive SYNthetic (ADASYN) sampling approach for imbalanced datasets [152] was compared with GANs in two studies [53], [126]. ADASYN uses a weighted distribution for different minority class instances based on their difficulty level, i.e., the more difficult to learn instances are synthesized more frequently. In addition, several other traditional techniques for data augmentation, such as adding Gaussian noise to the dataset [118], Gaussian smoothing, unsharp masking, minimum filtering [138], and affine transforms [141] were compared to GANs. Most of these studies show that data augmentation with GANs results in training datasets that improve the anomaly detection.

B. RQ2: WHAT ARE THE APPLICATION DOMAINS OF ANOMALY DETECTION WITH GANs?

Table 5 shows the different domains where GANs were applied in the primary studies. The table reveals that a vast number of primary studies (24 papers) perform anomaly detection in medical applications, closely followed by surveillance and intrusion detection with 19 and 17 papers, respectively.

1) MEDICAL ANOMALY DETECTION

Anomaly detection in medicine deals with analyzing patients' health conditions using medical records and images [153]. Specific applications include retinal optical coherence tomography (OCT) anomaly detection [17], [26], [29], [104], seizure detection [18], cardiovascular disease detection [30], lung nodule detection [42], abnormal chest X-ray identification [59], [97], [135], [138], polyp detection [80], [123], metastatic bone tumor detection [78], lesion detection [101], [137], laparoscopy anomaly detection [85], breast cancer detection [132], [143], MRI quality control [98], diabetic retinopathy detection [133], brain tumor detection [134] and hemorrhage detection [105]. One of the challenges in this domain is the difficulty of obtaining expert labels for medical data, such as clinical images, since annotation is an exhaustive and time-consuming task.

2) SURVEILLANCE ANOMALY DETECTION

To improve public safety, surveillance cameras are widely used in public places such as streets, stores, and banks. The goal of video surveillance is to identify suspicious activity, unusual traffic patterns, or accidents by automatically analyzing the behaviour of the surveillance target. In video surveillance anomaly detection, this can be accomplished by identifying the out-of-ordinary behaviours that differ from dominant (normal) behaviours in the scene [153]. Automated video surveillance can reduce the dependence on human workers and reduce the risk of late detection of anomalous behaviour. Most primary studies in this application domain leverage GANs for video anomaly detection to find irregularities in the crowds. However, traffic anomaly detection [24] and threat object recognition with X-ray imaging [91] have also been studied.

3) INTRUSION DETECTION

Intrusion detection systems are defined as software and/or hardware components that monitor and analyze events in computer systems to identify signs of intrusion [154]. Any

TABLE 4. Different types of GANs used for data augmentation.

Type of GAN	List of references	Type of GAN	List of references
AAE	[132]	D2GAN	[124]
AC-GAN	[116], [118], [140]	DCGAN	[122], [125], [127], [129], [135], [136], [141]–[143]
BGAN	[116]	PG-GAN	[116], [138]
BiGAN	[115]	SeqGAN	[121]
cGAN	[110], [111], [123], [131], [133], [134], [139]	Standard GAN	[112]–[114], [119], [120], [128], [130]
Cycle-GAN	[117], [124], [137]	WGAN	[126]

TABLE 5. The application domains of GANs for anomaly detection.

Application domain	List of references
Medical	[17], [18], [26], [29], [30], [42], [59], [78], [80], [85], [97], [98], [101], [104], [105], [123], [132]–[135], [137], [138], [141], [143]
Surveillance	[20]–[22], [24], [31], [37], [39], [41], [43], [54], [64], [65], [74], [75], [81], [90], [91], [109], [122]
Intrusion Detection	[19], [46], [76], [82], [83], [87], [93], [106], [110], [112], [114], [117], [119], [121], [128], [129], [131]
Various	[38], [50]–[52], [62], [67], [71], [79], [96], [107], [111], [120]
System Health	[16], [28], [53], [57], [58], [92], [100], [116], [125], [139]
Image Recognition	[25], [47], [51], [55], [63], [66], [86], [88], [89], [95], [136]
Manufacturing	[33], [40], [44], [69], [77], [99], [108], [124]
Autonomous Systems	[23], [31], [56], [60], [72], [73], [84], [118]
Power/Energy	[32], [35], [48], [49], [102], [127]
Fraud Detection	[45], [103], [113], [126], [130]
Hyperspectral Images	[27], [36], [70], [140]
Trajectory Detection	[61], [115]
Software Systems	[34], [94]
Text	[68]
Climate Changes	[142]

malicious intrusion or attack on network vulnerabilities, computers or information systems may result in a serious predicament and violate the confidentiality, integrity and availability of the systems [155]. The examined primary studies are mainly focused on network intrusion detection [83], [87], [106], [110], [117], [119], [121], [128], [129], [131]. Other applications of GANs in intrusion detection are smartphone lock pattern intrusion detection [112], presentation attack detection [82], phishing detection [114], cognitive radio intrusion detection [76], cyber-physical system intrusion detection [93], and IoT security [19], [46].

4) GENERAL APPROACHES

Some primary studies do not focus on a single application domain. Instead, they evaluate the proposed approaches in different application domains (shown as *Various* in Table 5). For example, three primary studies [38], [50], [51] investigate their proposed GAN-based anomaly detection for intrusion detection and image recognition. Two evaluated studies [79], [107] apply GAN-based anomaly detection in image recognition and video surveillance applications. Other primary studies evaluate their anomaly detection approach in intrusion detection, medical and image recognition domains [62], [71]. Zhu *et al.* [52] investigate the application of their proposed GAN-based anomaly detection in medicine and on trajectory anomaly detection. Oh *et al.* [111] evaluate their proposed technique for image recognition in addition to medical and trajectory anomaly detection. Wang *et al.* [67] study the application of GANs in fraud and intrusion detection, and Liu *et al.* [120] evaluate their approach in medicine, image recognition, aviation, human activity, spam identification, and waveform anomaly detection.

5) SYSTEM HEALTH ANOMALY DETECTION

System health monitoring is a way to identify anomalous behaviour in large (often industrial) systems. In industrial processes, the anomalous behaviour can represent, for example, wear or damage to the industrial equipment after continuous use. It is critical that such degradations in a system's performance are detected before they escalate and cause loss of revenue or endanger human life. Examples of industrial applications of system health anomaly detection with GANs include industrial process anomaly detection [16], [28], electrical insulator anomaly detection [116], rolling bearing anomaly detection [53], steam turbine anomaly detection [58], magnetic flux leakage detection [139], fused magnesium furnace anomaly detection [125], railway turnout anomaly detection [92] and communication system anomaly detection [100].

6) IMAGE RECOGNITION

Image anomaly detection refers to finding images with abnormal patterns that do not comply with other images in the same set. Most primary studies in this application domain use public image datasets, such as MNIST or CIFAR-10, to prove the concept of their proposed anomaly detection techniques. However, Moussa and Lim [136] evaluate the application of GANs for object recognition in images, such as finding an airplane in the picture. Bergmann *et al.* [66] propose a dataset of high-resolution color images of different object and texture categories suitable for anomaly detection. They evaluate several anomaly detection techniques, including GANs, to process their dataset. The proposed dataset aims to provide more challenging images than the commonly used datasets mentioned above.

7) MANUFACTURING ANOMALY DETECTION

This anomaly detection application refers to the quality inspection of manufactured products to identify defective products. These defects reveal themselves as irregularities on metal or wood surfaces, electronic parts, and so on. For example, an application of visual surface defect detection is studied in four primary studies [33], [44], [69], [124] and industrial quality inspection is investigated in three studies [40], [77], [99].

8) ANOMALY DETECTION IN AUTONOMOUS SYSTEMS

Autonomy is defined as self-governance or freedom from external influences [156]. An autonomous system is referred to as a system that can perceive the environment, make decisions based on the sensed information, and then react to internal/external changes using actuators. However, a fault may occur in each of these steps. For example, in an autonomous robot, faults can occur in sensors, software, or after physical damage to the actuator. This domain includes driving anomaly detection [23], [84], [118] to assist the driver or to identify abnormalities in the driver's behaviour. Autonomous surveillance with moving agents is addressed in three primary studies [31], [72], where an autonomous moving agent, such as a patrol robot, scans the environment to find abnormal activity. Two primary studies focused on controller anomaly detection [56], [73], to identify abnormal decision making by a controller in a closed-loop control system. Sun *et al.* [60] study autonomous vehicle anomaly detection.

9) POWER/ENERGY ANOMALY DETECTION

This application domain is concerned with identifying abnormalities in the power/energy consumption and power/energy infrastructure. Examples include catenary support component anomaly detection [32], [35], [49], power plant anomaly detection [48], [127] and power consumption anomaly detection [102].

10) FRAUD DETECTION

Fraud is defined as exploiting one's occupation for personal enrichment by willful misuse or application of their employer's resources or assets without authorization [157]. Fraud detection refers to uncovering these illegal activities. Examples of applications in this domain include click advertisement fraud [113], stock market manipulation [45], credit card fraud [126], health care insurance providers fraud [130], and satellite image forgery [103].

11) OTHER DOMAINS OF ANOMALY DETECTION WITH GANs

There are several additional application domains for anomaly detection using GANs that are less common: trajectory anomaly detection [61], [115], human mobility anomaly detection [115], climate change [142], text anomaly detection [68], and software systems anomaly detection [34], [94].

C. RQ3: WHICH GAN ARCHITECTURE IS USED MOST OFTEN IN ANOMALY DETECTION SYSTEMS?

Many types of GANs have been proposed to tackle the deficiencies of the first type of GAN proposed by Goodfellow *et al.* [7] or to handle specific tasks. In most cases, they modify the GAN architecture or the cost function of the generator and discriminator. According to the GAN Zoo GitHub repository,[5] more than 500 types of GAN were identified from 2014 to 2018.

We identified 21 different types of GANs used for anomaly detection purposes (see Table 6 for a list of primary studies using each of these architectures). DCGANs, standard GANs, and cGANs are the most commonly used GAN architectures. These were among the first proposed GAN architectures, and there are many new ones which are not (yet) used for anomaly detection purposes. The correspondence between the identified GAN architectures and their application domains is shown in Table 7. DCGANs, standard GANs, and BiGANs have been used in various application domains, indicating their flexibility. A variety of GAN architectures have also been used for applications in medicine, intrusion detection, and system health. However, some of the application domains are not well researched regarding GAN architectures, such as text anomaly detection and fraud detection.

Since the anomaly detection techniques examined in this review are either based on or assisted by GANs, any deficiency in the networks used for anomaly detection directly impacts the performance of the corresponding anomaly detection techniques. Therefore, the improvements in anomaly detection techniques using GANs are strongly correlated with the advances in the GAN architecture and training strategies. There are many studies in the literature describing the challenges of existing GANs and available solutions [12], [177]–[179]. The most crucial problem with GANs is the problem of mode collapse. When this happens, the generator of a GAN always generates samples from a highly concentrated distribution (partial collapse) [12], or simply a single sample (complete collapse) [180], [181]. Therefore, the generated data lacks the expected diversity. There have been several treatments proposed to lessen the effect of mode collapse during GAN training, such as WGAN [160], and Unrolled GAN [182]. Another challenge

[5] https://github.com/hindupuravinash/the-gan-zoo

TABLE 6. Type of GAN.

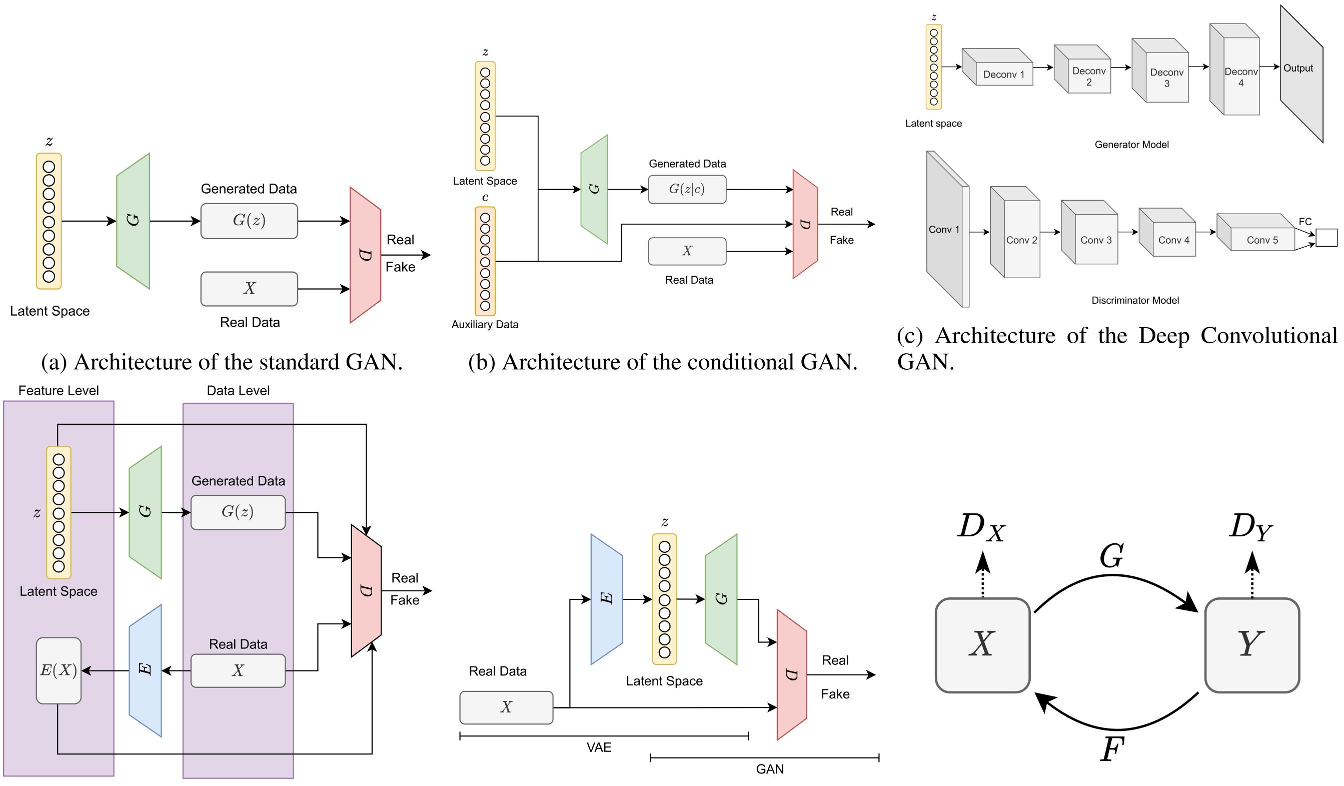
Type of GAN	List of papers using this type of GAN
Deep Convolutional GANs (DCGAN) [144]	[18], [22], [25], [30]–[33], [35], [40], [44], [49], [53], [55], [60], [66], [69], [77], [78], [81], [84], [86], [88], [92], [95]–[97], [99], [103]–[105], [108], [109], [122], [125], [127], [129], [135], [136], [141]–[143]
Standard GANs [7]	[16], [21], [24], [36], [43], [45], [46], [51], [52], [54], [58], [59], [61], [70], [75], [79], [83], [87], [93], [94], [100], [107], [112]–[114], [119], [120], [128], [130]
Conditional GANs (cGAN) [158]	[23], [26], [37], [39], [56], [57], [64], [65], [72]–[74], [76], [85], [90], [110], [111], [123], [131], [133], [134], [139]
Bi-directional GANs (BiGAN) [159]	[28], [34], [47], [50], [62], [67], [87], [91], [115]
Wasserstein GANs (WGAN) [160]	[17], [27], [67], [80], [87], [95], [126]
Wasserstein GANs with Gradient Penalty (WGAN-GP) [161]	[42], [63], [102], [106]
Variational Auetoencoder GANs (VAE-GAN) [162]	[19], [20], [98], [101]
Cycle-GAN [163]	[48], [117], [124], [137]
Auxiliary GANs (AC-GAN) [164]	[116], [118], [140]
Progressive Growing GANs (PG-GAN) [165]	[116], [138]
Orthogonal GAN (O-GAN) [166]	[29], [71]
Adversarial AutoEncoders (AAE) [167]	[132]
Balancing GANs (BGAN) [168]	[116]
Energy-Based GANs (EBGAN) [169]	[56]
Dual Discriminator GANs (D2GAN) [170]	[124]
GANs with Quadratic Potential (GAN-QP) [171]	[71]
One-Class GAN (OCGAN) [172]	[89]
Patch GANs (PatchGAN) [173]	[41]
Relativistic Discriminator GANs (RaSGAN) [174]	[82]
Sequence GANs (SeqGAN) [175]	[121]
Text GANs (TextGAN) [176]	[68]

TABLE 7. Type of GAN used in each application domain. AS: autonomous systems, CC: climate changes, FD: fraud detection, HI: hyperspectral images, IR: image recognition, ID: intrusion detection, MA: manufacturing, ME: medical, PE: power/energy, SS: software systems, SU: surveillance,SH: system health, TE: text, TD: trajectory detection, VA: various.

Type of GAN	AS	CC	FD	HI	IR	ID	MA	ME	PE	SS	SU	SH	TE	TD	VA
DCGANs	✓	✓	✓		✓	✓	✓	✓	✓		✓	✓			
Standard GANs			✓	✓		✓		✓		✓	✓	✓		✓	✓
BiGANs					✓	✓				✓	✓	✓		✓	✓
cGANs	✓					✓		✓			✓	✓			✓
WGANs			✓	✓	✓	✓		✓							✓
WGANs-GP					✓	✓		✓	✓						
VAE-GANs						✓		✓			✓				
Cycle-GAN						✓	✓	✓	✓						
AC-GANs	✓			✓								✓			
PG-GANs								✓				✓			
O-GANs								✓							✓
AAEs								✓							
BGANs												✓			
EBGANs	✓														
D2GANs							✓								
GANs-QP															✓
OCGANs								✓							
PatchGANs											✓				
RaSGANs						✓									
SeqGANs						✓									
TextGANs													✓		

of training of GANs is the instability of the training process and its failure to converge to a Nash equilibrium. Methods proposed to address this problem include Two Time-Scale Update Rule (TTUR) [183], WGAN [160] and feature matching [182].

The architectures of the top five GAN variants are shown in Figure 7. We elaborate on the most used architectures, and what makes them suitable for anomaly detection in the following subsections. In addition, we discuss if and how they deal with the challenges mentioned above.



(a) Architecture of the standard GAN.

(b) Architecture of the conditional GAN.

(c) Architecture of the Deep Convolutional GAN.

(d) Architecture of the Bi-directional GAN.

(e) Architecture of the VAE-GAN.

(f) Architecture of the Cycle-GAN.

FIGURE 7. The architectures of the most commonly used generative adversarial networks: (a) the standard GAN, where *G* and *D* denote the generator and discriminator; note that the architecture of the Wasserstein GAN is the same; (b) conditional GAN; the only difference is the addition of auxiliary data, shown as *c*; (c) deep convolutional GAN (DCGAN), where `Deconv` denotes a deconvolutional layer, `Conv` a convolutional layer, and `FC` a fully connected layer; (d) BiGAN, with *E* denoting the encoder; (e) VAE-GAN, where a variational autoencoder is combined with a GAN; (f) Cycle-GAN, which contains two mapping functions $G : X \rightarrow Y$ and $F : Y \rightarrow X$; the associated adversarial discriminators are shown as D_Y and D_X.

1) STANDARD GENERATIVE ADVERSARIAL NETWORKS (GAN)

In the first GAN architecture, proposed by Goodfellow *et al.* [7], the generator and discriminator models are defined by fully connected multilayer perceptrons. For the generator model, to learn the distribution of the generator p_g over data x, a prior on input noise variable $p_z(z)$ must be defined. This mapping is represented as $G(z; \theta_g)$, where G is a differentiable function represented by a multilayer perceptron with parameters θ_g. For the discriminator model, another multilayer perceptron $D(z; \theta_d)$ is defined. Its single scalar output represents the probability that x comes from the data rather than p_g. The training goal for the discriminator model D is to maximize the probability of assigning the correct label to both training examples and samples from the generator model G. Simultaneously, G is trained to minimize $\log(1 - D(G(z))$. D and G are pitted against each other following a two-player minimax game.

The Standard GAN optimizes the Jensen-Shannon (JS) divergence to learn the distribution of the data. Consequently, it suffers from an unstable, weak signal when the discriminator is approaching a local optimum, known as the problem of gradient vanishing [176]. This can also lead to mode collapse. Another problem of the standard GAN is that it does not provide any inference model to directly capture the inverse mapping. Hence, further training is needed to attain this inference model, adding to the computational cost of the GAN training. Moreover, as standard GAN uses MLP in the generator and discriminator models, it is not suitable for high dimensional data such as images. This is because MLPs are fully connected networks that require optimization of many parameters. Therefore, more efficient GAN architectures (such as DCGANs) are preferred for images and other high-dimensional data.

2) CONDITIONAL GENERATIVE ADVERSARIAL NETWORKS (cGAN)

Mirza and Osindero proposed the conditional GAN [158] as an extension to the standard GAN that can control what type of data is generated. For example, a condition can be specified to generate only data of a certain class or type. The conditional model of GAN can be obtained if both the generator and the discriminator are conditioned on some additional information y fed through additional input layers. There is no limitation on the type of the data; for example, it can contain class labels or data from different sources [158]. The conditional data generation is advantageous for anomaly detection purposes since cGANs can better generate data from different sources, i.e. multimodal data generation, or it can be used for multimodal anomaly detection.

3) DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS (DCGAN)

Striving to bridge the rift between the success of Convolutional Neural Networks (CNNs) for supervised learning and unsupervised learning, Radford *et al.* [144] introduced DCGANs, which integrate convolutional neural networks into the standard GAN. DCGANs provide a better network topology for more stable GAN training. The optimization and training processes are the same as for the standard GANs. However, Radford *et al.* proposed several improvements to the CNNs and Standard GANs.

These modifications are: (1) using all convolutional nets [184] in the generator and discriminator, (2) removing fully connected layers on top of the convolutional layer, and (3) using batch normalization [185]. These changes result in a better model and training stability with deeper gradient flow through the network, preventing mode collapse.

DCGANs were originally designed for image processing since they employ CNNs. The CNNs allow DCGANs to learn a hierarchy of representations from object parts to scenes in both the generator and discriminator, which makes DCGANs well suited for image anomaly detection.

4) BI-DIRECTIONAL GENERATIVE ADVERSARIAL NETWORKS (BiGAN)

The bi-directional GAN [159] adds an autoencoder that learns the mapping of data x to the latent representation z (inference), which makes it well suited for anomaly detection. BiGANs do not make any assumptions about the nature or structure of data. As a result, they provide a general, robust approach for unsupervised representation learning capable of capturing semantic attributes of the data [159]. Donahue *et al.* empirically show that, despite their generality, BiGANs are competitive with the state-of-the-art approaches to perform self-supervised and weakly supervised feature learning tasks. Comparing BiGANs with the standard GAN, the inference mechanism, i.e., feature learning, of BiGANs makes it suitable for anomaly detection techniques since they can be immediately used to generate anomaly scores.

5) WASSERSTEIN GENERATIVE ADVERSARIAL NETWORKS (WGAN)

In an attempt to alleviate the problem of mode collapse and the challenges of standard GANs to converge to the Nash equilibrium, Arjovsky *et al.* [160] suggested using the Earth-Mover (EM) distance or Wasserstein-1 distance instead of JS divergence used in the standard GAN. Unlike DCGAN, WGAN attempts to enhance the stability of GANs by modifying the adversarial cost function. Arjovsky *et al.* show that these distances provide gradients that are more useful for updating the generator than the JS divergence function [160].

Although WGAN better handles the problem of mode collapse compared to standard GANs and DCGANs, the weight clipping used in its discriminator made it difficult to converge. Gulrajani *et al.* [161] proposed an improved version WGAN-GP introducing a gradient penalty to the discriminator model of WGAN instead of weight clipping. This results in better convergence, training speed, and sample quality by forcing the discriminator to learn relatively smoother decision boundaries [179]. This improved version of WGAN is already used by several studies for anomaly detection (see Table 6).

6) VARIATIONAL AUTOENCODER GANs (VAE-GANs)

Larsen *et al.* [162] proposed a new GAN architecture that combines a variational autoencoder (VAE) with a GAN. A VAE-GAN replaces the decoder of a VAE with a generator of a GAN and modifies the loss function to be computed by a discriminator (see Figure 7e). The rationale behind this modification is to use the discriminator model to evaluate the similarity of the reconstructed image and the original one [98]. A VAE-GAN outperforms VAEs in terms of visual fidelity, and Larsen *et al.* demonstrate how to use the learned feature representation in the GAN discriminator as the basis for the VAE reconstruction objective. They replaced the element-wise errors with feature-wise errors, leading to a better performance in learning the distribution of the data.

7) CYCLE-GAN

Zhu *et al.* [163] proposed the Cycle-GAN to learn the mapping (G) between a source domain (X) to a target domain (Y) in the absence of paired examples. The goal of a Cycle-GAN is to learn a mapping $G : X \rightarrow Y$ in a way that the distribution of images from $G(X)$ is indistinguishable from the distribution of Y, using adversarial loss. Since this mapping is extremely under-constrained, a Cycle-GAN uses an inverse mapping $F : Y \rightarrow X$ and introduces the cycle consistency loss to enforce $F(G(X)) \approx X$ (and vice versa). A Cycle-GAN uses two different GANs coupled together to perform this transformation. It uses cycle-consistency loss to preserve the original image after a cycle of translation between two domains. In the training stage, the first discriminator tells whether the original image belongs to the source domain. The same thing happens for the target domain discriminator. In the investigated primary studies, the Cycle-GAN is only used as a data augmentation technique.

D. RQ4: WHICH TYPE OF DATA INSTANCE AND DATASETS ARE MOST COMMONLY USED FOR ANOMALY DETECTION WITH GANs?

We identified six types of input data used for anomaly detection with GANs. As shown in Table 8, image is by far the most common type, appearing in 50% of the examined papers. The two most common application domains for anomaly detection with GANs are related to images: medicine and surveillance. Tabular data is second (26%), followed by video, time series, text, and frequency data.

Data preprocessing is a key element that determines the success or failure of many deep learning models [41], [110], [112]. We identified 22 types of data preprocessing

TABLE 8. List of different data types.

Type of input data	Image	Tabular	Video	Time series	Text	Frequency
Number of papers	67	34	19	10	2	1
% of papers	50%	26%	14%	7%	2%	1%

TABLE 9. List of different preprocessing types with corresponding application to different data types.

Type of preprocessing	#papers	Image	Tabular	Video	Time series	Text	Frequency
Resizing	30	✓		✓	✓		
Normalization	24	✓	✓		✓	✓	
Cropping	16	✓					
Feature extraction	12	✓	✓		✓		
Augmentation (*e.g.* flipping)	8	✓	✓		✓		
Patch extraction	7	✓		✓			
End-to-end (no preprocessing)	6	✓	✓	✓	✓		
Frame extraction	6			✓			
Scaling	6	✓	✓			✓	
One-hot representation	5		✓				
Down-sampling	5	✓	✓	✓			
Data cleaning	5	✓	✓				
Mapping	2	✓	✓				
Manually labeling	2		✓		✓		
Label re-encoding	2		✓				
Denoising	2	✓					
Transforming to image	2		✓				✓
Splitting	1					✓	
Edging	1	✓					
Dimension reduction	1	✓					

techniques, summarized in Table 9. Owing to images being the most common data type, resizing, normalization, and cropping are the most prevalent preprocessing techniques. These techniques make data more uniform by changing its range and scale. A normalized dataset also speeds up learning. Preprocessing is commonly applied to image, tabular and time series data, but rarely to other types of data. It is also worth noting that some studies [76], [117] first transform tabular or frequency data to images before applying other types of preprocessing techniques.

Tables 10-a and 10-b show the datasets used in the primary studies as well as their associated application domains. The ''custom dataset'' stands for a dataset that was either constructed by the study authors or that contains proprietary data not released to the public domain. These tables show that the majority of the utilized datasets are custom, while *UCSD pedestrian* [186], *MNIST* [187], and *CIFAR-10* [188] are the most commonly used publicly available datasets.

The *UCSD anomaly detection* dataset was acquired by a stationary camera that captures pedestrian walkways. It includes two subsets: *Ped1* with 34 training and 36 testing video sequences, and *Ped2* with 16 training and 12 testing video sequences. In the normal setting, the video in this dataset contains only pedestrians. Abnormal events occur when either nonpedestrian entities are in the walkway, or there are anomalous pedestrian motion patterns, such as people walking across a walkway or in the grass that surrounds it. This dataset is challenging due to the low-resolution images, different types of moving objects, and the presence of one or more anomalies in the scene.

The *MNIST* and *CIFAR-10* datasets both appear in 7% of the examined studies. The *MNIST* database of handwritten digits has a training set of 60,000 examples and a test set of 10,000 examples. The *CIFAR-10* dataset is a collection of 60,000 colour images arranged in 10 object classes of equal size. When MNIST and CIFAT-10 are used in anomaly detection studies, one class is simulated as abnormal and removed from the training class, while the remaining classes are treated as normal.

The *UMN crowd* dataset [189] is used in 4% of the examined papers. It contains normal and abnormal crowd behaviour captured in indoor and outdoor scenes of the University of Minnesota. The dataset contains 11 videos with a total of 7,736 frames that were captured under several scenarios at three different indoor and outdoor scenes.

From the perspective of application domain, we found that studies on fraud detection use the *Credit Card Fraud Detection* dataset, *real world credit (RWC) dataset*, *TalkingData AdTracking*, *UCI dataset* and a custom dataset. For surveillance anomaly detection purposes, 74% studies use the *CUHK avenue*, *ShanghAaiTech*, *UCSD*, and *UMN* datasets. Among all datasets identified in the primary studies, twelve are used for intrusion detection, seven for manufacturing anomaly detection, fifteen for medical anomaly detection, and twelve for image anomaly detection. While most datasets are not used across all domains, eighteen are used in multiple domains as highlighted in Table 10-a and Table 10-b. For

TABLE 10. a: The list of identified datasets from primary studies. AS: autonomous systems, CC: climate changes, FD: fraud detection, HI: hyperspectral images, IR: image recognition, ID: intrusion detection, MA: manufacturing, ME: medical, PE: power/energy, SS: software systems, SU: surveillance, SH: system health, TE: text, TD: trajectory detection, VA: various. Note that the datasets highlighted in bold are applied in more than one domain. b: The list of identified datasets from primary studies. AS: autonomous systems, CC: climate changes, FD: fraud detection, HI: hyperspectral images, IR: image recognition, ID: intrusion detection, MA: manufacturing, ME: medical, PE: power/energy, SS: software systems, SU: surveillance, SH: system health, TE: text, TD: trajectory detection, VA: various. Note that the datasets highlighted in bold are applied in more than one domain.

Name of Dataset	AS	CC	FD	HI	IR	ID	MA	ME	PE	SS	SU	SH	TE	TD	VA
20Newsgroups: [68]													✓		
ABU: [36]				✓											
ADFA-LD: [117], [121]						✓									
ADNI: [134]								✓							
AI city challenge 2019: [24]											✓				
ARRHYTHMIA: [62], [71], [120]															✓
BratS18: [134], [137]								✓							
CALTECH-256: [25], [79], [108]					✓		✓								✓
CARDIO: [16]												✓			
Cardiotocography: [132]								✓							
CCSD-NL: [124]							✓								
CDnet2014: [81]											✓				
CelebA: [89]					✓										
CICIDS2017: [83], [87], [110], [129], [131]						✓									
CIFAR-10: [38], [50], [51], [55], [62], [63], [71], [86], [89], [95], [96]					✓										✓
CIFAR-100: [55]					✓										
COIL-100: [63], [89]					✓										
CRACK: [33]							✓								
Credit card fraud detection: [126]			✓												
CUHK avenue: [41], [43], [65], [75]											✓				
Custom: [18], [31], [32], [34], [35], [45], [48], [51], [53], [56]–[58], [60], [69], [72], [73], [76]–[78], [80], [84], [85], [92], [98]–[105], [112], [114], [116], [118], [127], [130], [136], [139], [140]	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓
CVC-Clinic: [123]								✓							
CVC-ClinicVideoDB: [123]								✓							
CWRU: [53]												✓			
DAD: [23]	✓														
DDSM: [141], [143]								✓							
ECG time_series: [52]															✓
El segundo: [27]				✓											
Fashion-MNIST: [63], [89]					✓										
Faster R-CNN: [49]									✓						
FFOB: [96]															✓
GEFCom2012: [19]						✓									
Geolife GPS trajectory: [115]														✓	
HYDICE: [36]				✓											
IONOSPHERE: [16], [120]												✓			✓
IRIS: [82]						✓									
IR-MNIST: [107], [108]							✓								✓
Joint european torus: [44]							✓								
KDD-Cup99 10%: [38], [50], [51], [62], [67], [71], [119]					✓	✓									✓
LIDC-IDRI: [42]								✓							
LiTS: [137]							✓								
LSUN: [89], [95]					✓										
Lymphography&Mammography: [132]								✓							
MIT-BIH: [30]								✓							
MNIST: [25], [44], [47], [51], [63], [71], [79], [86], [88], [89], [95], [96], [108], [115]					✓		✓								✓
MVTec: [66]					✓										
NIH chest X-ray: [59], [97]								✓							
NIH PLCO: [135]								✓							
NSL-KDD: [106], [110], [128], [129]						✓									
NYC_TAXI: [52]															✓
PIMA: [111], [120]															✓
OCT: [29]							✓								
RSNA: [138]								✓							
RWC: [126]			✓												

TABLE 10. ***(Continued.)*** **a: The list of identified datasets from primary studies. AS: autonomous systems, CC: climate changes, FD: fraud detection, HI: hyperspectral images, IR: image recognition, ID: intrusion detection, MA: manufacturing, ME: medical, PE: power/energy, SS: software systems, SU: surveillance, SH: system health, TE: text, TD: trajectory detection, VA: various. Note that the datasets highlighted in bold are applied in more than one domain. b: The list of identified datasets from primary studies. AS: autonomous systems, CC: climate changes, FD: fraud detection, HI: hyperspectral images, IR: image recognition, ID: intrusion detection, MA: manufacturing, ME: medical, PE: power/energy, SS: software systems, SU: surveillance, SH: system health, TE: text, TD: trajectory detection, VA: various. Note that the datasets highlighted in bold are applied in more than one domain.**

Name of Dataset	AS	CC	FD	HI	IR	ID	MA	ME	PE	SS	SU	SH	TE	TD	VA
San diego: [27], [36], [70]				✓											
San Francisco cabspotting: [115]					✓										
SBHAR: [46]						✓									
SD-OCT: [17]								✓							
Sentence polarity: [68]													✓		
ShanghAaiTech: [41], [75]											✓				
SIXray: [91]											✓				
Spectralis OCT: [26]								✓							
SWaT system: [93]						✓									
SVHN: [50], [62]															✓
TalkingData AdTracking: [67], [113]			✓												✓
Tennessee eastman: [16], [28]												✓			
Texas coast: [27]				✓											
Thyroid: [132]							✓								
UBA: [96]															✓
UCI: [38], [126]			✓												✓
UCSD: [21], [22], [37], [39], [41], [54], [64], [65], [74], [75], [79], [90], [107], [109], [122]											✓				✓
Udacity: [56], [61]														✓	
UMN: [39], [43], [64], [65], [74], [90], [107]											✓				✓
UNSW-NB15: [110]						✓									
VIRAT: [81]											✓				
WADI test-bed: [93]						✓									
WOA13 monthly: [142]		✓													
WOOD: [33]							✓								

instance, *KDD-cup99 10%* is used in both image recognition [50] and intrusion detection [51] while *MNIST* is used in image recognition [44], [47], [51], [63], [86], [96], [108], [115] and manufacturing anomaly detection [71].

E. RQ5: WHICH METRICS ARE USED TO EVALUATE THE PERFORMANCE OF GANs IN GENERATING DATA AND ANOMALY DETECTION?

We found that 27 out of 128 primary studies evaluated the quality of the generated samples using 9 different performance evaluation metrics. Most studies evaluated data quality quantitatively, while six papers implemented visual inspection to evaluate the quality of the generated samples [17], [40], [60], [69], [135], [139]. During the inspections, the generated samples were examined by application domain experts, or simply the authors of the individual studies. Quantitative evaluation was performed using eight performance metrics, most commonly the structural similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR) that were each used in 26% of the studies that evaluated performance.

SSIM, adopted by seven papers [26], [55], [56], [63], [65], [98], [99], quantifies the relative perceptual similarity between two images. This metric ranges from -1 to 1, with 1 indicating a perfect pixel match between the original and generated samples, -1 corresponding to inverted images, and 0 marking no similarity [190]. Seven papers [21], [22], [24], [41], [44], [85], [137] used PSNR as a metric to measure the quality of the generated images. This metric evaluates the similarity of two samples through the ratio of the total number of pixels divided and the mean squared error between the original and generated images. A higher value of the PSNR indicates that the generated sample is closer to the original.

The Fréchet inception distance (FID), adopted by two papers [82], [124], is a widely used evaluation method for evaluating the diversity and similarity of generated images [183]. By calculating and comparing the feature vectors of a collection of real and generated images, FID can measure the distance between the real and generated distribution.

We also studied whether performance metrics were used with specific input data types. As shown in Table 11, we observed that most performance metrics have been applied to image and video data, while only two papers utilized metrics for time series data [53], [60]. It is also worth mentioning that frame data is usually extracted from the video before being fed into the GAN. Therefore, the performance metrics are essentially used only to evaluate the image data. We also examined the relationship between the performance metrics and application domains. Table 12 shows that most metrics were adopted in surveillance anomaly detection to evaluate the generated samples, while most domains such as fraud detection, power/energy anomaly detection, and software systems did not evaluate the quality of the generated samples at all. In addition, SSIM, PSNR, FID, and visual inspection were used in various domains, while other metrics were only applied to one specific domain.

TABLE 11. The performance metrics that are used to evaluate generated samples with correspondent input types. Note that there are no performance metrics used for evaluating tabular, text and frequency data.

Type of performance metrics used	Image	Tabular	Video	Time series	Text	Frequency
Structural similarity indices metrics (SSIM): [55], [56], [63], [65], [98], [99], [116]	✓		✓			
Peak signal to noise ratio (PSNR): [21], [22], [24], [41], [44], [85], [137]	✓		✓			
Visual inspection: [17], [40], [60], [69], [135], [139]	✓			✓		
Fréchet inception distance (FID): [82], [124]	✓					
Signal to noise ratio (SNR): [53]				✓		
L2-norm distance: [109]			✓			
Fully convolutional network (FCN)-score: [65]			✓			
Earth mover's distance: [21]			✓			
Cosine similarity: [109]			✓			

TABLE 12. The performance metrics that are used to evaluate generated samples with correspondent application domains. **AS**: autonomous systems, **CC**: climate changes, **FD**: fraud detection, **HI**: hyperspectral images, **IR**: image recognition, **ID**: intrusion detection, **MA**: manufacturing, **ME**: medical, **PE**: power/energy, **SS**: software systems, **SU**: surveillance, **SH**: system health, **TE**: text, **TD**: trajectory detection, **VA**: various.

Type of Performance Metrics Used	AS	CC	FD	HI	IR	ID	MA	ME	PE	SS	SU	SH	TE	TD	VA
Structural similarity indices (SSIM): [55], [56], [63], [65], [98], [99], [116]	✓				✓		✓	✓			✓	✓			
Peak signal to noise ratio (PSNR): [21], [22], [24], [41], [44], [85], [137]								✓			✓				
Visual inspection: [17], [40], [60], [69], [135], [139]	✓						✓	✓				✓			
Fréchet inception distance (FID): [82], [124]						✓	✓								
Signal to noise ratio (SNR): [53]												✓			
L2-norm distance: [109]											✓				
Fully convolutional network (FCN)-score: [65]											✓				
Earth mover's distance (EMD): [21]											✓				
Cosine similarity: [109]											✓				

The actual anomaly detection performance in the studies is evaluated using different (usually more traditional) metrics. The area under the receiver operating characteristic curve (AUROC) is the most frequently used metric to evaluate anomaly detection (in 68 primary studies). Precision, F1-score, accuracy, recall, sensitivity, equal error rate (EER), and specificity are also frequently used as metrics in this area.

F. RQ6: WHICH ANOMALY DETECTION TECHNIQUES ARE USED ALONG WITH GANs?

This section discusses the different types of anomaly detection techniques that either use GANs or are compared with GANs. Based on the labelled data availability, anomaly detection techniques are divided into three classes: supervised, semi-supervised, and unsupervised anomaly detection. During data synthesis for RQ6, we noticed that not all primary studies use consistent definitions for these classes. Therefore, we use Chandola *et al.*'s [153] definition of supervised, semi-supervised, and unsupervised anomaly detection. In addition, as there is a wide variety of anomaly detection techniques, we only considered the techniques used in more than three primary studies.

1) SUPERVISED ANOMALY DETECTION

Anomaly detection techniques trained in a supervised fashion assume that labelled data for the normal and anomaly classes is available. A training dataset is used to train the model for predicting the class labels, and then the predictive model is evaluated on an unseen test dataset. Dependence of the supervised anomaly detection techniques on the data labels makes them more vulnerable to the problem of an imbalanced dataset. Therefore, these anomaly detection techniques are biased toward classifying the majority class.

In 26.3% of the investigated primary studies, GANs are applied to address the problem of the imbalanced dataset for anomaly detection by augmenting it with the data generated by GANs. The list of primary studies that use GANs for data augmentation is shown in Table 2. The list of papers that used GANs along with supervised anomaly detection techniques, i.e. GAN-assisted approaches, is shown in Table 13.

2) SEMI-SUPERVISED ANOMALY DETECTION

Anomaly detection techniques trained in a semi-supervised manner assume that the labelled data is available only for the normal class. The main benefit of semi-supervised anomaly detection techniques is that they do not require data for the anomalous class. In the reviewed primary studies, GANs are mostly used in a semi-supervised manner. By training a GAN to learn the distribution of the normal class, a deviation from the normal distribution is identified using an anomaly scoring technique. The list of GAN-based semi-supervised anomaly detection techniques is shown in Table 14. From all

TABLE 13. GAN-assisted supervised anomaly detection techniques.

Techniques	References
Support vector machines	[109], [110], [113], [114], [118], [119], [125]–[127], [129], [136]
Neural network-based methods	[91], [106], [109], [110], [113], [117], [121], [124], [125], [129], [134], [139]–[141], [143]
Nearest neighbors	[109], [114], [119], [121], [136]
Naive Bayes	[110], [119], [121], [129]
Ensemble methods	[109], [110], [113], [114], [119], [121], [129], [131], [136]
Linear model	[113], [114], [119], [121], [121], [126]
Decision trees	[109], [110], [114], [119], [121], [136]

TABLE 14. Semi-supervised GAN-based anomaly detection techniques.

Type of GAN	List of references	Type of GAN	List of references
AC-GAN	[118]	PatchGAN	[41]
BiGAN	[28], [34], [50], [62], [67], [87]	RaSGAN	[82]
cGAN	[26], [37], [39], [56], [57], [64], [65], [72]–[74], [76], [85], [90], [111]	Standard GAN	[24], [38], [43], [45], [46], [51], [52], [54], [58], [59], [61], [75], [79], [83], [87], [93], [94], [100], [107], [128], [130]
Cycle-GAN	[48]	TextGAN	[68]
DCGAN	[22], [30]–[33], [35], [40], [44], [49], [53], [55], [60], [66], [69], [77], [78], [81], [84], [86], [88], [92], [95]–[97], [99], [103]–[105], [108]	VAE-GAN	[98]
EBGAN	[56]	WGAN	[27], [67], [80], [87], [95]
GAN-QP	[71]	WGAN-GP	[42], [63]
O-GAN	[29], [71]		

TABLE 15. Semi-supervised anomaly detection techniques compared to GANs.

Techniques	List of references
Dynamic texture	[22], [39], [41], [64], [65], [74], [79], [90], [107]
Generative probabilistic model	[39], [54], [64], [65], [74], [75], [90]
Sparse dictionary learning	[39], [64], [65], [74], [75], [90]
Autoencoder-based	[16], [17], [21], [22], [26], [27], [38], [39], [41], [43], [54], [61], [64], [65], [67], [74], [75], [77], [79], [80], [84], [87], [88], [90], [93], [95], [96], [98], [128]
Recurrent neural networks	[83], [84], [92], [128]

GAN-based techniques, AnoGAN [104] is the GAN-based technique most often used as a baseline for comparison with newly proposed methods. There are several other techniques that can be used for anomaly detection purposes in a semi-supervised manner, as listed in Table 15. In investigated primary studies, the performance of these anomaly detection techniques is evaluated and compared to the GAN-based techniques. The experimental results of the primary studies show that the GAN-based anomaly detection techniques have as good as or superior performance over these competing techniques.

The table shows that several papers used Mixture of Dynamic Texture (MDT) [191] as an anomaly detection technique in crowded scenes. Mehran *et al.* [189] use a generative probabilistic model called Social Force for semi-supervised anomaly detection. Two primary studies investigate sparse dictionary learning-based anomaly detection techniques, detection at 150 FPS [192] and sparse reconstruction [193]. In the primary studies, the performance of different types of autoencoders (AEs) were compared to the performance of GANs in anomaly detection such as standard AEs [194], Variational AEs (VAEs), convolutional AEs (CVAEs) [195], Denoising AEs (DAEs) [196], and Adversarial AEs (AAEs) [167]. Moreover, some of the primary studies compared the proposed anomaly detection techniques with a Long Short Term Memory (LSTM) based approach [197].

3) UNSUPERVISED ANOMALY DETECTION

These types of anomaly detection techniques do not require a labelled dataset. This is based on the central assumption that normal instances are far more frequent than anomalies in the test data [153]. However, if this assumption is not valid, the anomaly detection will significantly suffer from false alarms. Assuming that the unlabeled dataset contains very few anomalous instances and the model is robust against these few anomalies, we can adapt a semi-supervised anomaly detection technique to work in an unsupervised manner by training the model on a portion of the unlabeled dataset. A list of GAN-based unsupervised anomaly detection techniques

TABLE 16. Unsupervised anomaly detection techniques based on GANs.

GAN architecture	List of references	GAN architecture	List of references
AAE	[132]	Standard GAN	[16], [21], [36], [70], [112], [120]
BiGAN	[47], [115]	VAE-GAN	[19], [20], [101]
cGAN	[23]	WGAN	[17], [95]
DCGAN	[18], [25], [95]	WGAN-GP	[102]

TABLE 17. Unsupervised anomaly detection techniques compared to GANs.

Techniques	List of references
Support vector machines	[16], [34], [38], [40], [50]–[52], [55], [61], [62], [69], [71], [87], [89], [93], [95], [98], [103], [112], [120]
Ensemble methods	[16], [38], [50], [52], [61], [62], [94], [95], [112], [132]
Principal component analysis	[22], [25], [28], [39], [41], [54], [63]–[65], [74], [75], [77], [79], [90], [93], [94], [108]
Linear model	[25], [63], [79], [108]
Probabilistic model	[40], [66], [95], [118], [120]
Stochastic model	[63], [79], [108]
Density-based model	[24], [40], [52], [120]
Energy-based model	[50], [51], [62], [71]
Autoencoder-based	[50], [51], [62], [71]

is presented in Table 16. In addition, Table 17 presents a list of unsupervised anomaly detection techniques that have been considered for anomaly detection and compared to GANs in the literature. The results of the investigated primary studies demonstrate that their proposed GAN-based anomaly detection techniques achieve superior performance compared to the techniques presented in Table 17.

From Table 17, we can observe that one-class classifiers [198] have been of great interest from an unsupervised anomaly detection perspective. Isolation forest [199] is another unsupervised technique competing with GANs in this area. Several variants of principal component analysis [200] have also been compared often to GANs in terms of performance. Several linear models have also been applied, namely REAPER [201] and Low Rank Representation [202]. Other techniques compared to GANs for anomaly detection include Gaussian mixture models [203], R-graph [204], local outlier factor [205], deep structured energy-based models [206], and deep autoencoding Gaussian mixture models [207].

V. FUTURE RESEARCH DIRECTIONS

Generative adversarial network-based anomaly detection is in its early stage of development with many research opportunities. However, most of these opportunities lie in the field of GANs itself. In this section, we present possible directions for the future work of applying GANs in anomaly detection.

A. FUTURE DIRECTION 1: SPEEDING UP THE GAN TRAINING PROCESS

Training GANs is a computationally demanding task. As reported in almost all primary studies, it takes a long time and powerful GPUs to train GANs to the point of satisfactory performance. Consequently, future studies need to explore GAN architectures that are lightweight and efficient in terms of resource consumption [22], [28], [83], [89], [127]. For instance, the effects of selecting GAN hyperparameters on the anomaly detection performance have not been studied in the literature. There is also a need to consider the use of emerging GAN optimization and training methods, e.g. [180], [182], for better training stability and faster convergence.

B. FUTURE DIRECTION 2: ACCOUNTING FOR CHANGING BEHAVIOUR OF A SYSTEM

In most industrial anomaly detection applications, behaviour of the target system varies over time. Therefore, it is crucial to examine the temporal behaviour of the system to find anomalies. RQ4 showed that only 7% of the primary studies used GANs for anomaly detection in time series data. Therefore, more studies are required on anomaly detection using GANs for time series data to make them suitable for industrial applications, especially for multivariate time series data. In many industrial applications, data is collected online. Huang and Lei [110] suggest to take advantage of this data via online training of GAN-based anomaly detection techniques. This approach might be adaptable for more real-time anomaly detection tasks.

C. FUTURE DIRECTION 3: IMPROVING SUPPORT FOR MULTIMODAL, DISCRETE AND NOISY DATA

Another open challenge of using GANs for anomaly detection is the lack of studies on multimodal anomaly detection using GANs. In real-world cases, data often comes from multiple sources with different types. For instance, Qiu *et al.* [23] propose a GAN-based driving anomaly detection technique using physiological and CAN-bus data. Qiu *et al.* suggest incorporating other information such as results obtained from vision-based object detection systems applied to the road. Many other GAN-based anomaly detection approaches could benefit from using multimodal data. In addition, GANs were initially created to generate continuous data. As a result, they have limited ability to deal with discrete data, as it hinders the backpropagation process [68]. Fadhel and Nyarko [68] point out this problem and study GAN architectures suitable for discrete data. Despite the promising results they report, this study is the only example of anomaly detection for discrete data in our review. Finally, Lei *et al.* [22] used optical flow as foreground shape information for video anomaly detection. Lei *et al.* point out that when the optical flow is inaccurate, it will affect the robustness of their proposed GAN-based anomaly detection technique. Thus, a potential direction for future research is the study of the effects of measurement inaccuracies and noise in the data on the performance of GAN-based anomaly detection techniques and the development of solutions to alleviate their impact.

D. FUTURE DIRECTION 4: SEARCHING FOR BETTER ANOMALY SCORING METHODS

As mentioned earlier, GAN-based anomaly detection techniques require an anomaly scoring method to distinguish between normal and abnormal samples. The selection of anomaly metrics for anomaly scores is still a challenging task. Further investigation is needed to improve the scoring methods and to identify the best fit for each application domain or for a specific case [28], [63], [99].

E. FUTURE DIRECTION 5: IMPROVING THE PERFORMANCE EVALUATION OF GANs

It is essential to evaluate the performance of GANs in generating data before using them for anomaly detection, either in a GAN-assisted or GAN-based setup. By doing so, one can ensure that GAN has learned the distribution of the data correctly. The results of RQ5 revealed that most primary studies do not evaluate the performance of their GAN-generated data. Additionally, almost all cases that assess data performance use image data. For other types of data, such as tabular, text and time series, there is no performance indicator of the generated data quality. Therefore, additional research is needed to identify the most suitable metrics for assessing the performance of GANs for each data type.

F. FUTURE DIRECTION 6: EMPLOYING IMPROVED GAN ARCHITECTURES FOR ANOMALY DETECTION

We observed in RQ3 that the 'older' GAN architectures are by far the most popular in anomaly detection studies. However, many improved GAN architectures were proposed recently, which could improve anomaly detection as well. For example, it is desirable to generate high-resolution images with GANs. However, it is a challenging task. High-resolution images make it easier for the discriminator to tell apart the generated images from the training samples [164]. Also, high-resolution images require more memory storage, which leads to smaller minibatches and may compromise training stability [165]. Several primary studies highlighted the need to generate high-resolution images for better anomaly detection, e.g. [55], [67], [89]. Future studies may examine the effect of using improved GAN architectures, e.g., to improve the resolution of images using SRGAN [208], ESRGAN [209], or BigGAN [210], on the performance of anomaly detection techniques.

VI. THREATS TO VALIDITY

One threat to the validity of our review is that of missing papers. The source of this threat is selecting the search keywords and the search engines. To mitigate this threat, we iteratively added keywords to our search query until no relevant new papers were found.

Moreover, the list of papers was finalized on the 3rd of June, 2020, and no papers that were published afterwards were added. In a fast-moving field in which many papers are published, such as anomaly detection using GANs, it is impossible to include all literature up to the date of submission of the review. Due to the amount of necessary work to conduct the SLR, there will always be a (considerable) amount of time between the cutoff date for the data collection and the submission of the review. Hence, it is possible that there are new GAN-based anomaly detection techniques that address some of the issues or challenges identified in this review.

Some recent works on the applications of GANs in anomaly detection which were published after our data collection ended include work in the following areas: medical applications [211]–[213], surveillance [214]–[216], intrusion detection [217]–[219], hyperspectral imaging [220]–[223], and manufacturing [224]. In the next literature review on the applications of GANs in anomaly detection, these studies need to be covered as well. Our paper is the first to systematically review the applications of GANs in anomaly detection (up to June 3, 2020) and should be used as the first building block towards building a concise survey of the available work on this topic.

Also, the data synthesis of the RQs was divided between the two first authors. To reduce bias in the data synthesis step, the first two authors met regularly to address disagreements. If a disagreement could not be resolved, one of the last two authors made the final decision.

VII. CONCLUSION

This systematic literature review presents an extensive study on the applications of GANs for anomaly detection, covering 128 primary studies. We define and answer several RQs to capture the current best practices and available techniques to employ GANs for anomaly detection purposes. We also identify the existing challenges and provide six future research directions in this area.

The results reveal that GANs are used for GAN-assisted (data augmentation) and GAN-based (representation learning) anomaly detection. In both cases, the problem of insufficient amount of data for the anomalous behaviour of the system is addressed. In a GAN-assisted approach, the goal is to augment the minority class using the generative ability of GANs. In GAN-based anomaly detection, the goal is to use the representation learning ability of GANs, eliminating the need for minority class data. The most commonly used GAN architectures in the primary studies are DCGANs, standard GANs, and cGANs. GANs are applied for anomaly detection in a wide range of application domains. The primary areas in which GANs are used for anomaly detection are medicine, surveillance and intrusion detection. However, their application in many other domains, such as anomaly detection in sensor networks, smart grids, and cloud computing shows that GANs are a suitable solution for anomaly detection.

We identified six important directions for future research. Some of these directions are related to fundamental GAN research. For example, our study reveals that only 21% of the primary studies evaluated the quality of the data that was generated with GANs. Hence, an important direction for future research is to investigate how the performance of GANs can be evaluated, as better performing GANs will also result in better performing anomaly detection approaches. Another important future research direction is speeding up the training process of GANs. In addition, we identified several important future research directions for anomaly detection researchers. In particular, GAN-assisted anomaly detection approaches should improve their support for multimodal, discrete and noisy data, and account for the changing behaviour of a system. Finally, researchers should investigate how recent improvements to GAN architectures can help improve their role in anomaly detection.

This systematic review of GAN literature has examined the fundamentals and recent advances in the area of GAN applications in anomaly detection. With hundreds of new articles published every year, it can be expected that there will be an influx of new architectures and improved learning methods. They promise to provide a powerful tool to generate realistic data across a broad range of problem domains. In the context of anomaly detection, this is extremely valuable as anomalous data is scarce and expensive to acquire. To ensure the quality of the generated data, it is necessary to expand the variety of standard datasets for use in GAN training, as well as to develop new performance metrics that are currently missing for many important domains.

REFERENCES

[1] Y. Sun, Z. Zhao, Z. Yang, F. Xu, H. Lu, Z. Zhu, W. Shi, J. Jiang, P. Yao, and H. Zhu, "Risk factors and preventions of breast cancer," *Int. J. Biol. Sci.*, vol. 13, no. 11, p. 1387, 2017.

[2] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA A, Cancer J. Clinicians*, vol. 70, no. 4, pp. 7–30, 2020.

[3] O. Ginsburg, C. H. Yip, A. Brooks, A. Cabanes, M. Caleffi, J. A. D. Yataco, B. Gyawali, V. McCormack, M. M. de Anderson, R. Mehrotra, and A. Mohar, "Breast cancer early detection: A phased approach to implementation," *Cancer*, vol. 126, pp. 2379–2393, May 2020.

[4] R. Mandal and N. Choudhury, "Automatic video surveillance for theft detection in ATM machines: An enhanced approach," in *Proc. 3rd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, Mar. 2016, pp. 2821–2826.

[5] M. Ghazal, C. Vazquez, and A. Amer, "Real-time automatic detection of vandalism behavior in video sequences," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, Oct. 2007, pp. 1056–1060.

[6] N. Mould, J. L. Regens, C. J. Jensen, and D. N. Edger, "Video surveillance and counterterrorism: The application of suspicious activity recognition in visual surveillance systems to counterterrorism," *J. Policing, Intell. Counter Terrorism*, vol. 9, no. 2, pp. 151–175, Jul. 2014.

[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[8] J. Langr and V. Bok, *GANs in Action: Deep Learning With Generative Adversarial Networks*. Shelter Island, NY, USA: Manning Publications, 2019.

[9] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.

[10] B. Kitchenham, *Procedures for Performing Systematic Reviews*, vol. 33. Keele, U.K.: Keele Univ., 2004, pp. 1–26.

[11] J. T. Springenberg, "Unsupervised and semi-supervised learning with categorical generative adversarial networks," 2015, *arXiv:1511.06390*.

[12] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process.*, vol. 35, no. 1, pp. 53–65, Jan. 2017.

[13] B. J. B. Rani, "Survey on applying GAN for anomaly detection," in *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, Jan. 2020, pp. 1–5.

[14] M. L. McHugh, "Interrater reliability: The Kappa statistic," *Biochem. Med.*, vol. 22, no. 3, pp. 276–282, 2012.

[15] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, 2013.

[16] S. Plakias and Y. S. Boutalis, "Exploiting the generative adversarial framework for one-class multi-dimensional fault detection," *Neurocomputing*, vol. 332, pp. 396–405, Mar. 2019.

[17] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, "f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks," *Med. Image Anal.*, vol. 54, pp. 30–44, May 2019.

[18] S. You, B. H. Cho, S. Yook, J. Y. Kim, Y.-M. Shon, D.-W. Seo, and I. Y. Kim, "Unsupervised automatic seizure detection for focal-onset seizures recorded with behind-the-ear EEG using an anomaly-detecting generative adversarial network," *Comput. Methods Programs Biomed.*, vol. 193, Sep. 2020, Art. no. 105472.

[19] Y. Zhang, Q. Ai, H. Wang, Z. Li, and X. Zhou, "Energy theft detection in an edge data center using threshold-based abnormality detector," *Int. J. Electr. Power Energy Syst.*, vol. 121, Oct. 2020, Art. no. 106162.

[20] H. Wu, J. Shao, X. Xu, F. Shen, and H. T. Shen, "A system for spatiotemporal anomaly localization in surveillance videos," in *Proc. 25th ACM Int. Conf. Multimedia*, Oct. 2017, pp. 1225–1226.

[21] A. Jamadandi, S. Kotturshettar, and U. Mudenagudi, "PredGAN: A deep multi-scale video prediction framework for detecting anomalies in videos," in *Proc. 11th Indian Conf. Comput. Vis., Graph. Image Process.*, Dec. 2018, pp. 1–8.

[22] Z. Lei, F. Deng, and X. Yang, "Spatial temporal balanced generative adversarial AutoEncoder for anomaly detection," in *Proc. Int. Conf. Image, Video Signal Process. (IVSP)*, 2019, pp. 1–7.

[23] Y. Qiu, T. Misu, and C. Busso, "Driving anomaly detection with conditional generative adversarial network using physiological and CAN-bus data," in *Proc. Int. Conf. Multimodal Interact.*, Oct. 2019, pp. 164–173.

[24] K.-T. Nguyen, D.-T. Dinh, M. N. Do, and M.-T. Tran, "Anomaly detection in traffic surveillance videos with GAN-based future frame prediction," in *Proc. Int. Conf. Multimedia Retr.*, Jun. 2020, pp. 457–463.

[25] L. Xu and Z. Xu, "One-class classification with deep adversarial learning," in *Proc. 3rd Int. Conf. Comput. Sci. Artif. Intell.*, Dec. 2019, pp. 103–106.

[26] K. Zhou, S. Gao, J. Cheng, Z. Gu, H. Fu, Z. Tu, J. Yang, Y. Zhao, and J. Liu, "Sparse-GAN: Sparsity-constrained generative adversarial network for anomaly detection in retinal OCT image," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1227–1231.

[27] K. Jiang, W. Xie, Y. Li, J. Lei, G. He, and Q. Du, "Semisupervised spectral learning with generative adversarial network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 5224–5236, Jul. 2020.

[28] X. Yang and D. Feng, "Generative adversarial network based anomaly detection on the benchmark Tennessee Eastman process," in *Proc. 5th Int. Conf. Control, Autom. Robot. (ICCAR)*, Apr. 2019, pp. 644–648.

[29] C. Zhang, Y. Wang, X. Zhao, Y. Guo, G. Xie, C. Lv, and B. Lv, "Memory-augmented anomaly generative adversarial network for retinal OCT images screening," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1971–1974.

[30] F. Luer, D. Mautz, and C. Bohm, "Anomaly detection in time series using generative adversarial networks," in *Proc. Int. Conf. Data Mining Workshops (ICDMW)*, Nov. 2019, pp. 1047–1048.

[31] W. Lawson, E. Bekele, and K. Sullivan, "Finding anomalies with generative adversarial networks for a patrolbot," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 12–13.

[32] P. Yang, W. Jin, and P. Tang, "Anomaly detection of railway catenary based on deep convolutional generative adversarial networks," in *Proc. IEEE 3rd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Oct. 2018, pp. 1366–1370.

[33] W. Zhai, J. Zhu, Y. Cao, and Z. Wang, "A generative adversarial network based framework for unsupervised visual surface inspection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 1283–1287.

[34] T. Kumarage, S. Ranathunga, C. Kuruppu, N. D. Silva, and M. Ranawaka, "Generative adversarial networks (GAN) based anomaly detection in industrial software systems," in *Proc. Moratuwa Eng. Res. Conf. (MERCon)*, Jul. 2019, pp. 43–48.

[35] Y. Lyu, Z. Han, J. Zhong, C. Li, and Z. Liu, "A generic anomaly detection of catenary support components based on generative adversarial networks," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 5, pp. 2439–2448, May 2020.

[36] T. Jiang, Y. Li, W. Xie, and Q. Du, "Discriminative reconstruction constrained generative adversarial network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4666–4679, Jul. 2020.

[37] T. Golda, N. Murzyn, C. Qu, and K. Kroschel, "What goes around comes around: Cycle-consistency-based short-term motion prediction for anomaly detection using generative adversarial networks," in *Proc. 16th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Sep. 2019, pp. 1–8.

[38] C. Wang, Y.-M. Zhang, and C.-L. Liu, "Anomaly detection via minimum likelihood generative adversarial networks," in *Proc. 24th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2018, pp. 1121–1126.

[39] T. Ganokratanaa, S. Aramvith, and N. Sebe, "Anomaly event detection using generative adversarial network for surveillance videos," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2019, pp. 1395–1399.

[40] Y. Lai, J. Hu, Y. Tsai, and W. Chiu, "Industrial anomaly detection and one-class classification using generative adversarial networks," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2018, pp. 1444–1449.

[41] F. Dong, Y. Zhang, and X. Nie, "Dual discriminator generative adversarial network for video anomaly detection," *IEEE Access*, vol. 8, pp. 88170–88176, 2020.

[42] Y. Kuang, T. Lan, X. Peng, G. E. Selasi, Q. Liu, and J. Zhang, "Unsupervised multi-discriminator generative adversarial network for lung nodule malignancy classification," *IEEE Access*, vol. 8, pp. 77725–77734, 2020.

[43] M. Yan, X. Jiang, and J. Yuan, "3D convolutional generative adversarial networks for detecting temporal irregularities in videos," in *Proc. 24th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2018, pp. 2522–2527.

[44] R. Skilton and Y. Gao, "Visual detection of generic defects in industrial components using generative adversarial networks," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2019, pp. 489–494.

[45] T. Leangarun, P. Tangamchit, and S. Thajchayapong, "Stock price manipulation detection using generative adversarial networks," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Nov. 2018, pp. 2104–2111.

[46] A. Ferdowsi and W. Saad, "Generative adversarial networks for distributed intrusion detection in the Internet of Things," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[47] K. Gray, D. Smolyak, S. Badirli, and G. Mohler, "Coupled IGMM-GANs for improved generative adversarial anomaly detection," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2018, pp. 2538–2541.

[48] Y. Choi, H. Lim, H. Choi, and I.-J. Kim, "GAN-based anomaly detection and localization of multivariate time series data for power plant," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Feb. 2020, pp. 71–74.

[49] Y. Lyu, Z. Han, J. Zhong, C. Li, and Z. Liu, "A GAN-based anomaly detection method for isoelectric line in high-speed railway," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, May 2019, pp. 1–6.

[50] Y. Hou, Z. Chen, M. Wu, C.-S. Foo, X. Li, and R. M. Shubair, "Mahalanobis distance based adversarial network for anomaly detection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 3192–3196.

[51] P. C. Ngo, A. A. Winarto, C. K. L. Kou, S. Park, F. Akram, and H. K. Lee, "Fence GAN: Towards better anomaly detection," in *Proc. IEEE 31st Int. Conf. Tools With Artif. Intell. (ICTAI)*, Nov. 2019, pp. 141–148.

[52] G. Zhu, H. Zhao, H. Liu, and H. Sun, "A novel LSTM-GAN algorithm for time series anomaly detection," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Qingdao)*, Oct. 2019, pp. 1–6.

[53] W. Jiang, Y. Hong, B. Zhou, X. He, and C. Cheng, "A GAN-based anomaly detection approach for imbalanced industrial time series," *IEEE Access*, vol. 7, pp. 143608–143619, 2019.

[54] Y. Yang, Z. Fu, and S. M. Naqvi, "Enhanced adversarial learning based video anomaly detection with object confidence and position," in *Proc. 13th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Dec. 2019, pp. 1–5.

[55] H. Cheng, H. Liu, F. Gao, and Z. Chen, "ADGAN: A scalable GAN-based architecture for image anomaly detection," in *Proc. IEEE 4th Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, Jun. 2020, pp. 987–993.

[56] N. Patel, A. N. Saridena, A. Choromanska, P. Krishnamurthy, and F. Khorrami, "Learning-based real-time process-aware anomaly monitoring for assured autonomy," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 4, pp. 659–669, Dec. 2020.

[57] T. Jiang, J. Zeng, K. Zhou, P. Huang, and T. Yang, "Lifelong disk failure prediction via GAN-based anomaly detection," in *Proc. IEEE 37th Int. Conf. Comput. Design (ICCD)*, Nov. 2019, pp. 199–207.

[58] Z. J. Que, Y. Xiong, and Z. G. Xu, "A semi-supervised approach for steam turbine health prognostics based on GAN and PF," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage. (IEEM)*, Dec. 2019, pp. 1476–1480.

[59] Y.-X. Tang, Y.-B. Tang, M. Han, J. Xiao, and R. M. Summers, "Abnormal chest X-ray identification with generative adversarial one-class classifier," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1358–1361.

[60] Y. Sun, W. Yu, Y. Chen, and A. Kadam, "Time series anomaly detection based on GAN," in *Proc. 6th Int. Conf. Social Netw. Anal., Manage. Secur. (SNAMS)*, Oct. 2019, pp. 375–382.

[61] P. R. Roy and G.-A. Bilodeau, "Adversarially learned abnormal trajectory classifier," in *Proc. 16th Conf. Comput. Robot Vis. (CRV)*, May 2019, pp. 65–72.

[62] H. Zenati, M. Romain, C.-S. Foo, B. Lecouat, and V. Chandrasekhar, "Adversarially learned anomaly detection," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2018, pp. 727–736.

[63] J. Bian, X. Hui, S. Sun, X. Zhao, and M. Tan, "A novel and efficient CVAE-GAN-based approach with informative manifold for semi-supervised anomaly detection," *IEEE Access*, vol. 7, pp. 88903–88916, 2019.

[64] M. Ravanbakhsh, E. Sangineto, M. Nabi, and N. Sebe, "Training adversarial discriminators for cross-channel abnormal event detection in crowds," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2019, pp. 1896–1904.

[65] T. Ganokratanaa, S. Aramvith, and N. Sebe, "Unsupervised anomaly detection and localization based on deep spatiotemporal translation network," *IEEE Access*, vol. 8, pp. 50312–50329, 2020.

[66] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, "MVTec AD—A comprehensive real-world dataset for unsupervised anomaly detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 9592–9600.

[67] C. Wang, Y. Dai, and W. Dai, "Deep embedding GAN-based model for anomaly detection on high-dimensional sparse data," in *Proc. Chin. Control Conf. (CCC)*, Jul. 2019, pp. 8718–8722.

[68] M. B. Fadhel and K. Nyarko, "GAN augmented text anomaly detection with sequences of deep statistics," in *Proc. 53rd Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2019, pp. 1–5.

[69] K. Liu, A. Li, X. Wen, H. Chen, and P. Yang, "Steel surface defect detection using GAN and one-class classifier," in *Proc. 25th Int. Conf. Autom. Comput. (ICAC)*, Sep. 2019, pp. 1–6.

[70] T. H. Emerson, J. A. Edelberg, T. Doster, N. Merrill, and C. C. Olson, "Generative and encoded anomaly detectors," in *Proc. 10th Workshop Hyperspectral Imag. Signal Process., Evol. Remote Sens. (WHISPERS)*, Sep. 2019, pp. 1–5.

[71] S. Mao, J. Guo, and Z. Li, "Discriminative autoencoding framework for simple and efficient anomaly detection," *IEEE Access*, vol. 7, pp. 140618–140630, 2019.

[72] M. Baydoun, M. Ravanbakhsh, D. Campo, P. Marin, D. Martin, L. Marcenaro, A. Cavallaro, and C. S. Regazzoni, "A multi-perspective approach to anomaly detection for self—Aware embodied agents," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 6598–6602.

[73] K. Vatanparvar and M. A. Al Faruque, "Self-secured control with anomaly detection and recovery in automotive cyber-physical systems," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, Mar. 2019, pp. 788–793.

[74] M. Ravanbakhsh, M. Nabi, E. Sangineto, L. Marcenaro, C. Regazzoni, and N. Sebe, "Abnormal event detection in videos using generative adversarial nets," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 1577–1581.

[75] H. Song, C. Sun, X. Wu, M. Chen, and Y. Jia, "Learning normal patterns via adversarial attention-based autoencoder for abnormal event detection in videos," *IEEE Trans. Multimedia*, vol. 22, no. 8, pp. 2138–2148, Aug. 2020.

[76] A. Toma, A. Krayani, M. Farrukh, H. Qi, L. Marcenaro, Y. Gao, and C. S. Regazzoni, "AI-based abnormality detection at the PHY-layer of cognitive radio by learning generative models," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 1, pp. 21–34, Mar. 2020.

[77] H. Yan, H.-M. Yeh, and N. Sergin, "Image-based process monitoring via adversarial autoencoder with applications to rolling defect detection," in *Proc. IEEE 15th Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2019, pp. 311–316.

[78] H. Watanabe, R. Togo, T. Ogawa, and M. Haseyama, "Bone metastatic tumor detection based on AnoGAN using CT images," in *Proc. IEEE 1st Global Conf. Life Sci. Technol. (LifeTech)*, Mar. 2019, pp. 235–236.

[79] M. Sabokrou, M. Khalooei, M. Fathy, and E. Adeli, "Adversarially learned one-class classifier for novelty detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3379–3388.

[80] Y. Liu, Y. Tian, G. Maicas, L. Z. Cheng Tao Pu, R. Singh, J. W. Verjans, and G. Carneiro, "Photoshopping colonoscopy video frames," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1–5.

[81] S. Ammar, T. Bouwmans, N. Zaghden, and M. Neji, "Deep detector classifier (DeepDC) for moving objects segmentation and classification in video surveillance," *IET Image Process.*, vol. 14, no. 8, pp. 1490–1501, Jun. 2020.

[82] S. Yadav, C. Chen, and A. Ross, "Relativistic discriminator: A one-class classifier for generalized iris presentation attack detection," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2020, pp. 2635–2644.

[83] M. Labonne, A. Olivereau, B. Polve, and D. Zeghlache, "Unsupervised protocol-based intrusion detection for real-world networks," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Feb. 2020, pp. 299–303.

[84] J.-J. Chung and H.-J. Kim, "An automobile environment detection system based on deep neural network and its implementation using IoT-enabled in-vehicle air quality sensors," *Sustainability*, vol. 12, no. 6, p. 2475, Mar. 2020.

[85] W. Reiter, "Video anomaly detection in post-procedural use of laparoscopic videos," in *Bildverarbeitung Für Die Medizin 2020*. Wiesbaden, Germany: Springer Fachmedien Wiesbaden, 2020, pp. 101–106.

[86] S. Anno and Y. Sasaki, "GAN-based abnormal detection by recognizing ungeneratable patterns," in *Pattern Recognition*. Cham, Switzerland: Springer, 2020, pp. 401–411.

[87] Y. Sun, L. Guo, Y. Li, L. Xu, and Y. Wang, "Semi-supervised deep learning for network anomaly detection," in *Algorithms and Architectures for Parallel Processing*. Cham, Switzerland: Springer, 2020, pp. 383–390.

[88] R. Anirudh, J. J. Thiagarajan, B. Kailkhura, and P. T. Bremer, "MimicGAN: Robust projection onto image manifolds with corruption mimicking," *Int. J. Comput. Vis.*, vol. 128, no. 10, pp. 2459–2477, 2020.

[89] N. Tuluptceva, B. Bakker, I. Fedulova, and A. Konushin, "Perceptual image anomaly detection," in *Proc. Asian Conf. Pattern Recognit.*, Springer, 2019, pp. 164–178.

[90] Y. Mu and B. Zhang, "Abnormal event detection and localization in visual surveillance," in *Communications, Signal Processing, and Systems*. Singapore: Springer, 2020, pp. 1217–1225.

[91] J. K. Dumagpi, W.-Y. Jung, and Y.-J. Jeong, "A new GAN-based anomaly detection (GBAD) approach for multi-threat object classification on large-scale X-ray security images," *IEICE Trans. Inf. Syst.*, vol. E103.D, no. 2, pp. 454–458, 2020.

[92] L. Xue and S. Gao, "Unsupervised anomaly detection system for railway turnout based on GAN," *J. Phys., Conf. Ser.*, vol. 1345, no. 3, Nov. 2019, Art. no. 032069.

[93] D. Li, D. Chen, B. Jin, L. Shi, J. Goh, and S.-K. Ng, "MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks," in *Artificial Neural Networks and Machine Learning—ICANN 2019: Text and Time Series*. Cham, Switzerland: Springer, 2019, pp. 703–716.

[94] B. Xia, J. Yin, J. Xu, and Y. Li, "LogGAN: A sequence-based generative adversarial network for anomaly detection based on system logs," in *Proc. Int. Conf. Sci. Cyber Secur.*, Springer, 2019, pp. 61–76.

[95] L. Deecke, R. Vandermeulen, L. Ruff, S. Mandt, and M. Kloft, "Image anomaly detection with generative adversarial networks," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Springer, 2018, pp. 3–17.

[96] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, "GANomaly: Semi-supervised anomaly detection via adversarial training," in *Computer Vision—ACCV 2018*. Cham, Switzerland: Springer, 2019, pp. 622–637.

[97] Y. Tang, Y. Tang, M. Han, J. Xiao, and R. M. Summers, "Deep adversarial one-class learning for normal and abnormal chest radiograph classification," *Proc. SPIE*, vol. 10950, Mar. 2019, Art. no. 1095018.

[98] M. Mostapha, J. Prieto, V. Murphy, J. Girault, M. Foster, A. Rumple, J. Blocher, W. Lin, J. Elison, J. Gilmore, S. Pizer, and M. Styner, "Semi-supervised VAE-GAN for out-of-sample detection applied to MRI quality control," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019*. Cham, Switzerland: Springer, 2019, pp. 127–136.

[99] E. A. Donahue, T. T. Quach, K. Potter, C. Martinez, M. Smith, and C. D. Turner, "Deep learning for automated defect detection in high-reliability electronic parts," *Proc. SPIE*, vol. 11139, Sep. 2019, Art. no. 1113907.

[100] D. Zhang, Q. Niu, and X. Qiu, "Detecting anomalies in communication packet streams based on generative adversarial networks," in *Wireless Algorithms, Systems, and Applications*. Cham, Switzerland: Springer, 2019, pp. 470–481.

[101] C. Baur, B. Wiestler, S. Albarqouni, and N. Navab, "Deep autoencoding models for unsupervised anomaly segmentation in brain MR images," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Cham, Switzerland: Springer, 2019, pp. 161–169.

[102] Z. Huang, W. Mao, M. Chen, Q. Wu, B. Xiong, and W. Xu, "An intelligent operation and maintenance system for power consumption based on deep learning," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 486, no. 1, 2019, Art. no. 012107.

[103] S. K. Yarlagadda, D. Güera, P. Bestagini, F. M. Zhu, S. Tubaro, and E. J. Delp, "Satellite image forgery detection and localization using GAN and one-class classifier," *Electron. Imag.*, vol. 2018, no. 7, pp. 1–214, 2018.

[104] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," in *Information Processing in Medical Imaging*. Cham, Switzerland: Springer, 2017, pp. 146–157.

[105] A. Singanamalli, J. Mitra, K. Wallace, P. Venugopal, S. Smith, L. Mo, L. Y. Leung, J. Morrison, T. Rasmussen, and L. Marinelli, "Blood flow anomaly detection via generative adversarial networks: A preliminary study," *Proc. SPIE*, vol. 11315, Mar. 2020, Art. no. 1131522.

[106] J.-T. Wang and C.-H. Wang, "High performance WGAN-GP based multiple-category network anomaly classification system," in *Proc. Int. Conf. Cyber Secur. Emerg. Technol. (CSET)*, Oct. 2019, pp. 1–7.

[107] M. Sabokrou, M. Pourreza, M. Fayyaz, R. Entezari, M. Fathy, J. Gall, and E. Adeli, "AVID: Adversarial visual irregularity detection," in *Computer Vision—ACCV 2018*. Cham, Switzerland: Springer, 2019, pp. 488–505.

[108] M. Kimura and T. Yanagihara, "Anomaly detection using GANs for visual inspection in noisy training data," in *Computer Vision—ACCV 2018 Workshops*. Cham, Switzerland: Springer, 2019, pp. 373–385.

[109] W. Shin and S.-B. Cho, "CCTV image sequence generation and modeling method for video anomaly detection using generative adversarial network," in *Intelligent Data Engineering and Automated Learning—IDEAL 2018*. Cham, Switzerland: Springer, 2018, pp. 457–467.

[110] S. Huang and K. Lei, "IGAN-IDS: An imbalanced generative adversarial network towards intrusion detection system in ad-hoc networks," *Ad Hoc Netw.*, vol. 105, Aug. 2020, Art. no. 102177.

[111] J.-H. Oh, J. Y. Hong, and J.-G. Baek, "Oversampling method using outlier detectable generative adversarial network," *Expert Syst. Appl.*, vol. 133, pp. 1–8, Nov. 2019.

[112] S.-Y. Shin, Y.-W. Kang, and Y.-G. Kim, "Android-GAN: Defending against Android pattern attacks using multi-modal generative network as anomaly detector," *Expert Syst. Appl.*, vol. 141, Mar. 2020, Art. no. 112964.

[113] G. S. Thejas, K. G. Boroojeni, K. Chandna, I. Bhatia, S. S. Iyengar, and N. R. Sunitha, "Deep learning-based model to fight against ad click fraud," in *Proc. ACM Southeast Conf.*, Apr. 2019, pp. 176–181.

[114] A. AlEroud and G. Karabatis, "Bypassing detection of URL-based phishing attacks using generative adversarial deep neural networks," in *Proc. 6th Int. Workshop Secur. Privacy Anal.*, Mar. 2020, pp. 53–60.

[115] D. Smolyak, K. Gray, S. Badirli, and G. Mohler, "Coupled IGMM-GANs with applications to anomaly detection in human mobility data," *ACM Trans. Spatial Algorithms Syst.*, vol. 6, no. 4, pp. 1–14, Aug. 2020.

[116] L. Luo, W. Hsu, and S. Wang, "Data augmentation using generative adversarial networks for electrical insulator anomaly detection," in *Proc. 2nd Int. Conf. Manage. Sci. Ind. Eng.*, Apr. 2020, pp. 231–236.

[117] M. Salem, S. Taheri, and J. S. Yuan, "Anomaly generation using generative adversarial networks in host-based intrusion detection," in *Proc. 9th IEEE Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Nov. 2018, pp. 683–687.

[118] H. Kim, J. Park, K. Min, and K. Huh, "Anomaly monitoring framework in lane detection with a generative adversarial network," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1603–1615, Mar. 2021.

[119] M. Usama, M. Asim, S. Latif, and J. Qadir, "Generative adversarial networks for launching and thwarting adversarial attacks on network intrusion detection systems," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2019, pp. 78–83.

[120] Y. Liu, Z. Li, C. Zhou, Y. Jiang, J. Sun, M. Wang, and X. He, "Generative adversarial active learning for unsupervised outlier detection," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 8, pp. 1517–1528, Aug. 2020.

[121] S. Shin, I. Lee, and C. Choi, "Anomaly dataset augmentation using the sequence generative models," in *Proc. 18th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2019, pp. 1143–1148.

[122] U. Muneeb, E. Koyuncu, Y. Keshtkarjahromi, H. Seferoglu, M. F. Erden, and A. E. Cetin, "Robust and computationally-efficient anomaly detection using powers-of-two networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 2992–2996.

[123] Y. Shin, H. A. Qadir, and I. Balasingham, "Abnormal colon polyp image synthesis using conditional adversarial networks for improved detection performance," *IEEE Access*, vol. 6, pp. 56007–56017, 2018.

[124] S. Niu, B. Li, X. Wang, and H. Lin, "Defect image sample generation with GAN for improving defect recognition," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 3, pp. 1611–1622, Jul. 2020.

[125] S. Lu and H. Gao, "Deep learning based fusion of RGB and infrared images for the detection of abnormal condition of fused magnesium furnace," in *Proc. IEEE 15th Int. Conf. Control Autom. (ICCA)*, Jul. 2019, pp. 987–993.

[126] M. Shao, N. Gu, and X. Zhang, "Credit card transactions data adversarial augmentation in the frequency domain," in *Proc. 5th IEEE Int. Conf. Big Data Anal. (ICBDA)*, May 2020, pp. 238–245.

[127] S. Jintao, G. Chaoyue, S. Hui, S. Jiangang, and L. Zhe, "Data expansion for foreign object detection in power grid," in *Proc. IEEE PES Asia–Pacific Power Energy Eng. Conf. (APPEEC)*, Dec. 2019, pp. 1–4.

[128] B. Mohammadi and M. Sabokrou, "End-to-end adversarial learning for intrusion detection in computer networks," in *Proc. IEEE 44th Conf. Local Comput. Netw. (LCN)*, Oct. 2019, pp. 270–273.

[129] J. Yang, T. Li, G. Liang, W. He, and Y. Zhao, "A simple recurrent unit model based intrusion detection system with DCGAN," *IEEE Access*, vol. 7, pp. 83286–83296, 2019.

[130] K. Naidoo and V. Marivate, "Unsupervised anomaly detection of healthcare providers using generative adversarial networks," in *Proc. 19th Conf. e-Bus., e-Services e-Soc. (I E)*, in Responsible Design, Implementation and Use of Information and Communication Technology, Skukuza, South Africa. Springer, Apr. 2020, pp. 419–430.

[131] J. Lee and K. Park, "AE-CGAN model based high performance network intrusion detection system," *Appl. Sci.*, vol. 9, no. 20, p. 4221, Oct. 2019.

[132] S. K. Lim, Y. Loo, N.-T. Tran, N.-M. Cheung, G. Roig, and Y. Elovici, "DOPING: Generative data augmentation for unsupervised anomaly detection with GAN," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2018, pp. 1122–1127.

[133] R. Zheng, L. Liu, S. Zhang, C. Zheng, F. Bunyak, R. Xu, B. Li, and M. Sun, "Detection of exudates in fundus photographs with imbalanced learning using conditional generative adversarial network," *Biomed. Opt. Exp.*, vol. 9, no. 10, pp. 4863–4878, Oct. 2018.

[134] H. C. Shin, N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, K. P. Andriole, and M. Michalski, "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," in *Simulation and Synthesis in Medical Imaging*. Cham, Switzerland: Springer, 2018, pp. 1–11.

[135] A. Madani, M. Moradi, A. Karargyris, and T. Syeda-Mahmood, "Chest X-ray generation and data augmentation for cardiovascular abnormality classification," *Proc. SPIE*, vol. 10574, Mar. 2018, Art. no. 105741M.

[136] M. Moussa and C. H. Lim, "Interpreting abnormality of a complex static scene using generative adversarial network," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2019, pp. 2003–2007.

[137] L. Sun, J. Wang, Y. Huang, X. Ding, H. Greenspan, and J. Paisley, "An adversarial learning approach to medical image synthesis for lesion detection," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 8, pp. 2303–2314, Aug. 2020.

[138] P. Ganesan, S. Rajaraman, R. Long, B. Ghoraani, and S. Antani, "Assessment of data augmentation strategies toward performance improvement of abnormality classification in chest radiographs," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 841–844.

[139] M. Fu, J. Liu, H. Zhang, and S. Lu, "Multisensor fusion for magnetic flux leakage defect characterization under information incompletion," *IEEE Trans. Ind. Electron.*, vol. 68, no. 5, pp. 4382–4392, May 2021.

[140] D. Wang, R. Vinson, M. Holmes, G. Seibel, A. Bechar, S. Nof, and Y. Tao, "Early detection of tomato spotted wilt virus by hyperspectral imaging and outlier removal auxiliary classifier generative adversarial nets (OR-AC-GAN)," *Sci. Rep.*, vol. 9, no. 1, pp. 1–14, Dec. 2019.

[141] S. Guan, "Breast cancer detection using synthetic mammograms from generative adversarial networks in convolutional neural networks," *J. Med. Imag.*, vol. 6, no. 3, p. 1, Mar. 2019.

[142] X. Li, Y. Liang, M. Zhao, C. Wang, and Y. Jiang, "Few-shot learning with generative adversarial networks based on WOA13 data," *Comput., Mater. Continua*, vol. 60, no. 3, pp. 1073–1085, 2019.

[143] S. Guan and M. Loew, "Using generative adversarial networks and transfer learning for breast cancer detection by convolutional neural networks," *Proc. SPIE*, vol. 10954, Mar. 2019, Art. no. 109541C.

[144] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.

[145] R. A. Yeh, C. Chen, T. Y. Lim, A. G. Schwing, M. Hasegawa-Johnson, and M. N. Do, "Semantic image inpainting with deep generative models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5485–5493.

[146] G. Marcus, "Deep learning: A critical appraisal," 2018, *arXiv:1801.00631*.

[147] Z. Obermeyer and E. J. Emanuel, "Predicting the future-big data, machine learning, and clinical medicine," *New England J. Med.*, vol. 375, no. 13, p. 1216, 2016.

[148] H. Kaur, H. S. Pannu, and A. K. Malhi, "A systematic review on imbalanced data challenges in machine learning: Applications and solutions," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–36, Sep. 2019.

[149] T. R. Hoens and N. V. Chawla, *Imbalanced Datasets: From Sampling to Classifiers*. Hoboken, NJ, USA: Wiley, 2013, ch. 3, pp. 43–59.

[150] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.

[151] H. Han, W.-Y. Wang, and B.-H. Mao, "Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning," in *Advances in Intelligent Computing*. Berlin, Germany: Springer, 2005, pp. 878–887.

[152] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IEEE World Congr. Comput. Intell.)*, Jun. 2008, pp. 1322–1328.

[153] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, Jul. 2009.

[154] A. Lazarevic, V. Kumar, and J. Srivastava, "Intrusion detection: A survey," in *Managing Cyber Threats: Issues, Approaches, and Challenges*. Boston, MA, USA: Springer, 2005, pp. 19–78.

[155] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," *J. Netw. Comput. Appl.*, vol. 36, no. 1, pp. 16–24, 2013.

[156] C. D. Nguyen, S. Miles, A. Perini, P. Tonella, M. Harman, and M. Luck, "Evolutionary testing of autonomous software agents," *Auton. Agent Multi-Agent Syst.*, vol. 25, no. 2, pp. 260–283, 2012.

[157] Y. Kou, C.-T. Lu, S. Sirwongwattana, and Y.-P. Huang, "Survey of fraud detection techniques," in *Proc. IEEE Int. Conf. Netw., Sens. Control*, Mar. 2004, pp. 749–754.

[158] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.

[159] J. Donahue, P. Krähenbühl, and T. Darrell, "Adversarial feature learning," 2016, *arXiv:1605.09782*.

[160] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, *arXiv:1701.07875*.

[161] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in *Adv. neural Inf. Process. Syst.*, 2017, pp. 5767–5777.

[162] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, "Autoencoding beyond pixels using a learned similarity metric," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1558–1566.

[163] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.

[164] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier GANs," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 2642–2651.

[165] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," 2017, *arXiv:1710.10196*.

[166] J. Su, "O-GAN: Extremely concise approach for auto-encoding generative adversarial networks," 2019, *arXiv:1903.01931*.

[167] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial autoencoders," 2015, *arXiv:1511.05644*.

[168] G. Mariani, F. Scheidegger, R. Istrate, C. Bekas, and C. Malossi, "BAGAN: Data augmentation with balancing GAN," 2018, *arXiv:1803.09655*.

[169] J. Zhao, M. Mathieu, and Y. LeCun, "Energy-based generative adversarial network," 2016, *arXiv:1609.03126*.

[170] T. Nguyen, T. Le, H. Vu, and D. Phung, "Dual discriminator generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 2670–2680.

[171] J. Su, "GAN-QP: A novel GAN framework without gradient vanishing and Lipschitz constraint," 2018, *arXiv:1811.07296*.

[172] P. Perera, R. Nallapati, and B. Xiang, "OCGAN: One-class novelty detection using GANs with constrained latent representations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 2898–2906.

[173] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1125–1134.

[174] A. Jolicoeur-Martineau, "The relativistic discriminator: A key element missing from standard GAN," 2018, *arXiv:1807.00734*.

[175] L. Yu, W. Zhang, J. Wang, and Y. Yu, "SeqGAN: Sequence generative adversarial nets with policy gradient," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 2852–2858.

[176] Y. Zhang, Z. Gan, K. Fan, Z. Chen, R. Henao, D. Shen, and L. Carin, "Adversarial feature matching for text generation," 2017, *arXiv:1706.03850*.

[177] Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan, and Y. Zheng, "Recent progress on generative adversarial networks (GANs): A survey," *IEEE Access*, vol. 7, pp. 36322–36333, 2019.

[178] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, "A review on generative adversarial networks: Algorithms, theory, and applications," 2020, *arXiv:2001.06937*.

[179] A. Jabbar, X. Li, and B. Omar, "A survey on generative adversarial networks: Variants, applications, and training," 2020, *arXiv:2006.05132*.

[180] M. Arjovsky and L. Bottou, "Towards principled methods for training generative adversarial networks," 2017, *arXiv:1701.04862*.

[181] S. Arora, R. Ge, Y. Liang, T. Ma, and Y. Zhang, "Generalization and equilibrium in generative adversarial nets (GANs)," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 224–232.

[182] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training GANs," 2016, *arXiv:1606.03498*.

[183] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local Nash equilibrium," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 6626–6637.

[184] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for simplicity: The all convolutional net," 2014, *arXiv:1412.6806*.

[185] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*.

[186] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, "Anomaly detection in crowded scenes," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1975–1981.

[187] Y. LeCun, C. Cortes, and C. Burges. MNIST Handwritten Digit Database. Florham Park, NJ, USA. Accessed: 2010. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[188] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," M.S. thesis, Dept. Comput. Sci., Univ. Toronto, Toronto, ON, Canada, 2009.

[189] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 935–942.

[190] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[191] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 1, pp. 18–32, Jan. 2014.

[192] C. Lu, J. Shi, and J. Jia, "Abnormal event detection at 150 FPS in MATLAB," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2720–2727.

[193] Y. Cong, J. Yuan, and J. Liu, "Sparse reconstruction cost for abnormal event detection," in *Proc. CVPR*, Jun. 2011, pp. 3449–3456.

[194] M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," in *Proc. MLSDA 2nd Workshop Mach. Learn. Sensory Data Anal. (MLSDA)*, 2014, pp. 4–11.

[195] J. An and S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability," *Special Lecture IE*, vol. 2, pp. 1–18, Dec. 2015.

[196] D. Xu, E. Ricci, Y. Yan, J. Song, and N. Sebe, "Learning deep representations of appearance and motion for anomalous event detection," 2015, *arXiv:1510.01553*.

[197] M. Du, F. Li, G. Zheng, and V. Srikumar, "DeepLog: Anomaly detection and diagnosis from system logs through deep learning," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2017, pp. 1285–1298.

[198] B. Schölkopf, R. C. Williamson, A. J. Smola, J. Shawe-Taylor, and J. C. Platt, "Support vector method for novelty detection," *Adv. Neural Inf. Process. Syst.*, vol. 12, 1999, pp. 582–588.

[199] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Proc. 8th IEEE Int. Conf. Data Mining*, Dec. 2008, pp. 413–422.

[200] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *J. Roy. Statist. Soc. B*, vol. 61, no. 3, pp. 611–622, 1999.

[201] G. Lerman, M. B. McCoy, J. A. Tropp, and T. Zhang, "Robust computation of linear models by convex relaxation," *Found. Comput. Math.*, vol. 15, no. 2, pp. 363–410, Apr. 2015.

[202] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, 2010, pp. 663–670.

[203] X. Yang, L. J. Latecki, and D. Pokrajac, "Outlier detection with globally optimal exemplar-based GMM," in *Proc. SIAM Int. Conf. Data Mining*. Philadelphia, PA, USA: SIAM, 2009, pp. 145–154.

[204] C. You, D. P. Robinson, and R. Vidal, "Provable self-representation based outlier detection in a union of subspaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3395–3404.

[205] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying density-based local outliers," *ACM SIGMOD Rec.*, vol. 29, no. 2, pp. 93–104, Jun. 2000.

[206] S. Zhai, Y. Cheng, W. Lu, and Z. Zhang, "Deep structured energy based models for anomaly detection," 2016, *arXiv:1605.07717*.

[207] B. Zong, Q. Song, M. R. Min, W. Cheng, C. Lumezanu, D. Cho, and H. Chen, "Deep autoencoding Gaussian mixture model for unsupervised anomaly detection," in *Proc. Int. Conf. Learn. Represent.*, 2018, pp. 1–19.

[208] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4681–4690.

[209] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "ESRGAN: Enhanced super-resolution generative adversarial networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*, Sep. 2018, pp. 63–79.

[210] A. Brock, J. Donahue, and K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," 2018, *arXiv:1809.11096*.

[211] C. Han, L. Rundo, K. Murao, T. Noguchi, Y. Shimahara, Z. Á. Milacski, S. Koshino, E. Sala, H. Nakayama, and S. Satoh, "MADGAN: Unsupervised medical anomaly detection GAN using multiple adjacent brain MRI slice reconstruction," *BMC Bioinf.*, vol. 22, no. S2, pp. 1–20, Apr. 2021.

[212] C. Han, L. Rundo, R. Araki, Y. Nagano, Y. Furukawa, G. Mauri, H. Nakayama, and H. Hayashi, "Combining noise-to-image and image-to-image GANs: Brain MR image augmentation for tumor detection," *IEEE Access*, vol. 7, pp. 156966–156977, 2019.

[213] C. Baur, R. Graf, B. Wiestler, S. Albarqouni, and N. Navab, "SteGANomaly: Inhibiting CycleGAN steganography for unsupervised anomaly detection in brain MRI," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020*. Cham, Switzerland: Springer, 2020, pp. 718–727.

[214] M. Pourreza, B. Mohammadi, M. Khaki, S. Bouindour, H. Snoussi, and M. Sabokrou, "G2D: Generate to detect anomaly," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 2003–2012.

[215] X. Feng, D. Song, Y. Chen, Z. Chen, J. Ni, and H. Chen, "Convolutional transformer based dual discriminator generative adversarial networks for video anomaly detection," in *Proc. 29th ACM Int. Conf. Multimedia*, Oct. 2021, pp. 5546–5554.

[216] K. Doshi and Y. Yilmaz, "Any-shot sequential anomaly detection in surveillance videos," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 934–935.

[217] P. F. de Araujo-Filho, G. Kaddoum, D. R. Campelo, A. G. Santos, D. Macedo, and C. Zanchettin, "Intrusion detection for cyber–physical systems using generative adversarial networks in fog environment," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6247–6256, Apr. 2021.

[218] Y. Zhong, Y. Zhu, Z. Wang, X. Yin, X. Shi, and K. Li, "An adversarial learning model for intrusion detection in real complex network environments," in *Wireless Algorithms, Systems, and Applications*. Cham, Switzerland: Springer, 2020, pp. 794–806.

[219] M. H. Shahriar, N. I. Haque, M. A. Rahman, and M. Alonso, "G-IDS: Generative adversarial networks assisted intrusion detection system," in *Proc. IEEE 44th Annu. Comput., Softw., Appl. Conf. (COMPSAC)*, Jul. 2020, pp. 376–385.

[220] S. Arisoy, N. M. Nasrabadi, and K. Kayabol, "GAN-based hyperspectral anomaly detection," in *Proc. 28th Eur. Signal Process. Conf. (EUSIPCO)*, Jan. 2021, pp. 1891–1895.

[221] Y. Li, T. Jiang, W. Xie, J. Lei, and Q. Du, "Sparse coding-inspired GAN for hyperspectral anomaly detection in weakly supervised learning," *IEEE Trans. Geosci. Remote Sens.*, early access, Aug. 12, 2021, doi: 10.1109/TGRS.2021.3102048.

[222] C. Zhao, C. Li, S. Feng, and N. Su, "Hyperspectral anomaly detection using bilateral-filtered generative adversarial networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2021, pp. 4408–4411.

[223] T. Jiang, W. Xie, Y. Li, J. Lei, and Q. Du, "Weakly supervised discriminative learning with spectral constrained generative adversarial network for hyperspectral anomaly detection," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, May 31, 2021, doi: 10.1109/TNNLS.2021.3082158.

[224] C. Cooper, J. Zhang, R. X. Gao, P. Wang, and I. Ragai, "Anomaly detection in milling tools using acoustic signals and generative adversarial networks," *Proc. Manuf.*, vol. 48, pp. 372–378, Jan. 2020.

MIKAEL SABUHI (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in control systems from the Iran University of Science and Technology, in 2016 and 2018, respectively. He is currently pursuing the Ph.D. degree in software engineering and intelligent systems with the University of Alberta. He is also a member of the ENergy digiTizAtIon Lab (ENTAIL) and the Analytics of Software, GAmes And Repository Data (ASGAARD) Lab, working on anomaly detection and empirical performance analysis of cloud software systems. While he has different artificial intelligence-related research interests, his main focus is on applications of deep learning architectures for anomaly detection with a particular emphasis on applications of generative adversarial networks (GANs) for anomaly detection.

MING (CHLOE) ZHOU (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Alberta, Edmonton, AB, Canada, in 2020, where she is currently pursuing the M.Sc. degree with the Department of Electrical and Computer Engineering. She is also a member of the ENergydigiTizAtIon Lab (ENTAIL). Her research interests include anomaly detection, smart grids, and time series forecasting.

COR-PAUL BEZEMER (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in computer science from the Delft University of Technology, The Netherlands, in 2007, 2009, and 2014, respectively. He is currently an Assistant Professor with the Electrical and Computer Engineering Department, University of Alberta. He also heads the Analytics of Software, GAmes And Repository Data (ASGAARD) Lab. Before that, he was a Postdoctoral Research Fellow with the Software Analysis and Intelligence Lab (SAIL), Queen's University, Kingston, Canada. His work has been published at premier software engineering venues, such as the TSE and EMSE journals and the ESEC-FSE, ICSME, and ICPE conferences. His research interests include software engineering and performance engineering-related topics. He is one of the vice-chairs of the SPEC Research Group on DevOps Performance.

PETR MUSILEK (Senior Member, IEEE) received the Ing. degree (Hons.) in electrical engineering and the Ph.D. degree in cybernetics from the Military Academy in Brno, Czech Republic, in 1991 and 1995, respectively. In 1995, he was appointed the Head of the Computer Applications Group, Institute of Informatics, Military Medical Academy, Hradec Kralove, Czech Republic. From 1997 to 1999, he was an NATO Science Fellow with the Intelligent Systems Research Laboratory, University of Saskatchewan, Canada. In 1999, he joined the Department of Electrical and Computer Engineering, University of Alberta, Canada, where he is currently a Full Professor. Since 2016, he has been serving as the Director of the Computer Engineering Program and an Associate Chair (Undergraduate). He is also an Associate Chair (Research and Planning). His research interests include artificial intelligence and energy systems. He developed a number of innovative solutions in the areas of renewable energy systems, smart grids, wireless sensor networks, and environmental monitoring and modeling.