Leveraging GANs, Autoencoders and their Combined Deep Learning Architectures for Anomaly Detection in Multi-Domain Image Data

1st Ashim Chandra Das
dept. Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
ashim.chandra.das@g.bracu.ac.bd

2nd Anjishnu Banik
dept. Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
anjishnu.banik@g.bracu.ac.bd

3rd Noshin Sayara
dept. Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
noshin.sayara@g.bracu.ac.bd

4th Sazzad Hossain Reaz
dept. Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
sazzad.hossain.reaz@g.bracu.ac.bd

5th Dr. Amitabha Chakrabarty
dept. Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
amitabha@bracu.ac.bd

6th Fahim-Ul-Islam
dept. Computer Science and Engineering
BRAC University
Dhaka, Bangladesh

***Abstract*—The study aims to identify anomalies in multiple domains through deep learning and it evaluates the performance of 9 state-of-the-art models for anomaly detection in industrial quality control, plant disease diagnosis, and medical imaging. It compares them and proposes a modification to the Bi-GAN model because this modification is expected to improve the model's performance. The 9 models are: FCAE, CAE, VAE, CVAE, Beta-VAE, VQ-VAE, AnoGAN, GANomaly, and Bi-GAN, and some are based on autoencoders (mirror) and others are based on GANs (adversarial competition between two agents). The authors modified the Bi-GAN model and they employed 3 datasets: brain MRI scans to detect abnormalities, industrial components to detect defects, and plant leaves to detect diseases, thus allowing them to test the models in different domains. The modified Bi-GAN model was the most effective model for identifying anomalies through the comparison of images, and this makes the model a good fit for anomaly detection in a variety of domains. All nine models showed promise, but each model also had its limitations, because by modifying the Bi-GAN model, the researchers were able to create a more robust and reliable approach.The study contributes and an evaluation and comparison of nine anomaly detection types. An introduction of a novel anomaly detection type that performs well in general, and a comparison of all anomaly detection types. This highlights the strengths and weaknesses of each type, and this provides good insights. The study contributes to our understanding of anomaly detection, and it also provides a basis for the design of novel anomaly detection types.**

***Index Terms*—Image Processing, Deep Learning, Generative Adversarial Networks, Autoencoder Neural Networks, Image Reconstruction, Anomaly Detection**

Identify applicable funding agency here. If none, delete this.

I. Introduction

Anomaly detection in images plays a crucial role in healthcare, manufacturing, and security, and it is critical for anomaly detection to have a very high accuracy rate. The previous anomaly detection methods have many limitations, and they usually fail to deal with unbalanced data. They also have difficulties dealing with data with complicated distributions, and they require lots of manual feature extraction efforts. Recently, deep learning models such as Generative Adversarial Network (GAN) and Autoencoder have been shown to be useful in detecting anomalies, and GANs generate synthetic data and detect anomalies. Autoencoders (AE) can learn to compress data and then reconstruct it, so this reconstruction can be used for detecting anomalies. By combining GANs and AE, we can create better anomaly detection algorithms, therefore this thesis uses GANs, AE, and hybrid models for anomaly detection in images from various domains. We compare several GAN and AE models, and we analyze their loss functions, thus we propose methods for visualizing what the models have learned, which can be used for aiding anomaly detection. We evaluate our models on medical, industrial, and security images, and we compare our models to conventional methods, because the goal of this thesis is to provide an effective and convenient solution for anomaly detection in various domains.

II. Related Works

Autoencoder models (e.g., FCAE, CAE, VAE, Beta-VAE, CVAE, VQ-VAE) reconstruct data and detect anomalies based on the reconstruction error, and autoencoder models learn the pattern of the data and detect anomalies based on the reconstruction error (i.e., the difference between the original data and the reconstructed data). Autoencoders can learn the data patterns and are good for structured and low-dimensional data, but they may not well detect anomalies from high-dimensional and complex data. VAE and its variants introduce probabilistic modeling to enhance the latent space representation, but they still struggle to achieve good performance on highly imbalanced datasets, and VQ-VAE, with vector quantization, can enhance the latent space representation, but it may not achieve high precision and recall for anomaly detection. Generative adversarial networks (GANs) have also been applied for anomaly detection, and AnoGAN is a representative GAN-based model for anomaly detection, because GANs consist of a generator and a discriminator to learn the distribution of input data. AnoGAN is trained adversarially to learn the anomaly detection model by comparing the input and generated data, thus GANs can effectively model complex data distributions, but their training is known to be unstable and requires careful tuning to balance the generator and discriminator. Thus, the instability of GANs can lead to unreliable anomaly detection, and therefore, hybrid models that combine autoencoders and GANs have been proposed to address the limitations of autoencoder-based and GAN-based models, such as GANomaly and BiGAN. GANomaly uses an encoder-decoder-encoder architecture that yields better feature learning and the ability to detect subtle anomalies, and BiGAN performs bidirectional mapping between the data domain and latent space that enables better anomaly detection through reconstruction and adversarial training, so these hybrid models have achieved promising results in various applications including medical imaging, industrial inspection, and agriculture.
However, there still exist some challenges, and the models must achieve the right compromise between accuracy and recall. This is more challenging when the data are highly imbalanced or difficult, because autoencoder-based, GAN-based, and hybrid models have performed well so far, but they still require further research. The training of GANs should be further stabilized; thus, the autoencoders must operate better on challenging data. The right compromise between accuracy and recall must be found, which is crucial for applications of anomaly detection in medicine, industry, and agriculture, so we need to do more research to fix the problems and make the models better.

III. Proposed Approach

This BiGAN is not the same as classical BiGAN, and it introduces additional components which help to improve the quality of reconstruction and make training easier. The most important difference is the presence of Squeeze-and-Excitation (SE) attention in the encoder and generator, because these are not classical convolutional layers like in classical BiGAN. They contain global average pooling, dense layers and multiply, thus allowing the model to focus on the most important feature channels. Another difference is the presence of residual skip connections in the generator with Add layers, which makes it easier for the gradients to flow and preserves the most important features, therefore improving the overall performance. This is better than the simple, direct conv pathways in normal BiGANs, so it provides a more effective way of training the model.

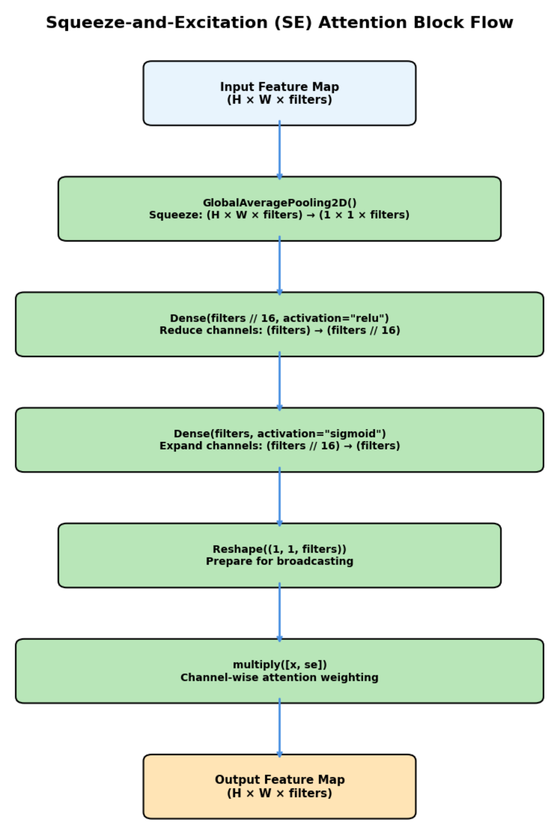

Fig. 1. Squeeze-and-Excitation (SE) Attention Block

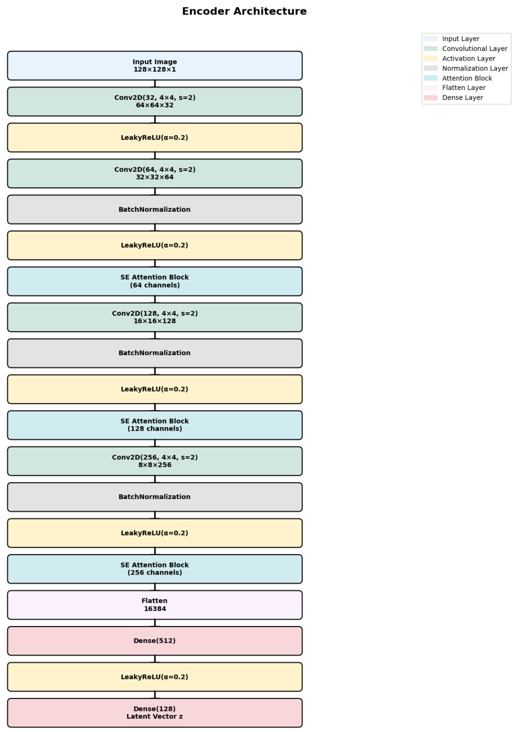

Fig. 2. Architecture of Encoder

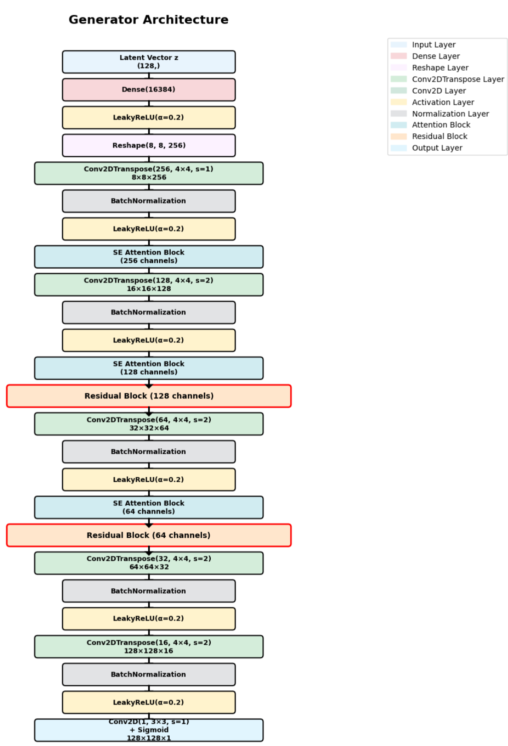

Fig. 3. Architecture of Generator

The training process for this BiGAN is slightly different, and it consists two phases. In the first phase, the encoder and the generator are trained for 150 epochs with only the reconstruction loss, so the encoder and generator learn to reconstruct the image well without any adversarial training. Next, the adversarial training begins and continues for 500 epochs, therefore the discriminator is trained every 5 epochs to distinguish real image-latent pairs from fake ones. The generator and encoder are trained adversarially to generate fake image-latent pairs to fool the discriminator, thus the reconstruction model is trained 3 times per epoch to ensure that it does not become weak during the adversarial training. The reconstruction quality is prioritized over the adversarial performance, because the loss function is defined as where and are perceptual loss using VGG16 features and reconstruction loss. The discriminator is also an image-latent pair discriminator, and it concatenates the flattened image features with processed latent codes, which is more sophisticated than the standard discriminator, therefore it provides a more accurate distinction between real and fake image-latent pairs.

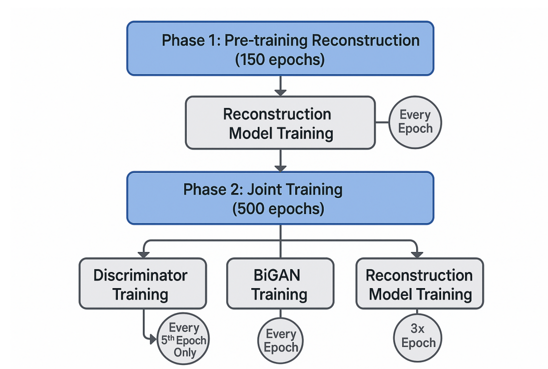

Fig. 4. Overview of Training Process

The anomaly score is based on a combination of pixel-level reconstruction error (mean squared error, MSE, between the original and reconstructed images) and SSIM scores, and the final anomaly score is a weighted combination of 70% MSE error and 30% SSIM error (SSIM error is 1 - SSIM). This provides a good measure of the reconstruction quality that takes into account both the pixel-level differences between the original and reconstructed images, and the structural differences, because images with larger reconstruction errors are treated as more anomalous, as the model trained mainly on normal samples cannot reconstruct well patterns that are not similar to the normal data distribution. This harnesses the model's knowledge of the normal data to detect anomalies, thus the model's knowledge is utilized to identify anomalies.

Batch normalization layers and LeakyReLU activations are used to stabilize training, and the loss function is a mixture of adversarial loss and reconstruction loss. The reconstruction loss is defined as a mixture of MSE (0.5), MAE (0.2), perceptual VGG16 loss (0.2) and SSIM loss (0.1), because it is designed to capture various aspects of the data. The encoder relies on global pooling with attention to obtain meaningful latent representations, thus allowing the model to focus on the most important features. The architecture of the generator relies on transposed convolutions with attention-guided refinement, therefore enabling the model to produce high-quality outputs. These modifications shift the focus of the model from adversarial training (e.g. BiGAN) towards reconstruction and feature learning, so the model is more reconstruction-focused, capable of capturing details and structures.

IV. METHODOLOGY

A. Data Collection

Brain MRI dataset is collected from Kaggle (Brain Tumor MRI Images (predominantly Glioma Tumor)) and it is used to detect abnormalities in brain MRI scans. It contains healthy and diseased brain images because it is used to evaluate the presence of tumors.The MVTecAD dataset is obtained from its official website and it is designed to detect anomalies in industrial applications. Images of industrial items are included, and these items can be metal nuts, etc. These items may have defects, such as scratches or dents, because they are used in various industrial processes. The Plant Disease dataset is collected from Kaggle, and it is designed to detect anomalies in agriculture, thus it contains pictures of healthy and diseased plant leaves. We use these datasets to test the model in different fields, including medical, industrial and agriculture, therefore we can evaluate its performance in various domains. It also decreases the risk of overfitting to a particular dataset, so the model can be applied to different areas.

B. Data Preparation

- **Data Augmentation:** Data augmentation by translation, flipping and scaling to generate synthetic normal and abnormal images to increase the size of the normal and abnormal images dataset, so that the model can generalize better and perform better, therefore it can handle a wider range of images.
- **Grayscale Conversion:** Grayscale transformation was implemented and in order to omit color information from the dataset and simplify the dataset, a grayscale transformation was applied on the images. This makes the model focus on the intensity features, which are more important in medical images, because the color information is not as relevant in this case.
- **Resize:** The images were resized to 128x128 pixels and this ensures that all images are the same size for batch processing. This simplifies the task for the neural network and allows for more efficient training.
- **Normalization:** Pixel values were normalized, which means that the brightness level (pixel values) of the images were standardized, and this makes the training process easier. It also eliminates the problem of data variance, which makes the model more robust, therefore making it more reliable.
- **Split:** Only healthy/good images have been splited 80% for training and 20% for testing.

This data preparation process we have applied for Brain MRI, MVTecAD and Plant Disease datasets. These preprocessing were performed to ensure that the data were of good quality, and they also enabled the model to be effectively trained and tested. They helped to make the data more balanced and reliable, so the model would then learn more easily and more effectively.

C. Training Phase

- Used only Healthy/Good data for training.

D. Testing Phase

- Merged Healthy/Good test samples with Diseased/Defected samples.
- Pass all test images through the trained model.
- Calculate anomaly scores based on reconstruction or latent distance.

E. Decision

- Apply a pre defined threshold on anomaly scores.
- Score $<$ Threshold $\implies$ Healthy/Good
- Score $\geq$ Threshold $\implies$ Diseased/Defected

F. Analysis

The reconstruction error was used to identify and localize anomalies in the samples, and both healthy and diseased samples were fed to the trained model. The error maps showed minimal errors for healthy samples and different regions were highlighted for the corresponding anomalies in diseased samples, because the error maps were effective in identifying and localizing diseased areas, thus anomaly scores were computed using reconstruction error or latent space distance.

V. RESULT ANALYSIS

The anomaly detection models were evaluated on three datasets, namely Brain MRI, MVTecAD and Plant Disease, and the performance metrics were chosen to be AUC, accuracy, precision, recall, F1 score and specificity. In this paper, we aimed to investigate the anomaly detection models to detect the anomaly and to understand the advantages and disadvantages of each model, because the models were tested on three different domains, such as medical images, industrial and agriculture, thus allowing for a comprehensive evaluation.

Table I In all of the tests, the new model achieved the highest AUC value, which was 0.892, significantly higher than that of the second best model, CVAE, which was 0.812, and the accuracy of the new model was 85.4%, higher than that of VAE (77.4%) and CVAE (77.3%). The precision of the new model was 95.9%, indicating that the detection result was highly credible, because the recall of the new model was 86.2%, implying that most of the actual anomalies can be detected. The F1-score of the proposed model is 90.5%, which outperforms the best baseline model with F1-score of 85.4%, therefore the specificity of the proposed model is 81.5%, which

Model	AUC	Accuracy	Precision	Recall	F1	Specificity
Improved BiGAN	0.892	0.854	0.959	0.862	0.905	0.815
BiGAN	0.852	0.820	0.950	0.820	0.890	0.780
GANomaly	0.840	0.810	0.947	0.810	0.875	0.760
AnoGAN	0.837	0.800	0.941	0.800	0.860	0.740
VAE	0.835	0.774	0.934	0.766	0.844	0.695
Beta-VAE	0.830	0.766	0.931	0.751	0.832	0.720
FC-AE	0.796	0.787	0.939	0.785	0.855	0.696
CNN-AE	0.409	0.555	0.840	0.653	0.726	0.371
VQ-VAE	0.411	0.446	0.908	0.402	0.571	0.484
CVAE	0.810	0.773	0.928	0.789	0.853	0.695

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT MODELS ON BRAIN MRI DATASET

outperforms the specificity of Beta VAE with 72.0%. This indicates that the proposed model is superior in detecting anomalies, thus the overall performance of the proposed model is better than the other models.

Model	AUC	Accuracy	Precision	Recall	F1	Specificity
Improved BiGAN	0.921	0.823	0.981	0.813	0.892	0.782
BiGAN	0.872	0.792	0.953	0.784	0.864	0.753
GANomaly	0.854	0.781	0.941	0.771	0.847	0.732
AnoGAN	0.832	0.773	0.932	0.762	0.834	0.715
VAE	0.803	0.743	0.903	0.735	0.813	0.673
Beta-VAE	0.791	0.736	0.892	0.724	0.802	0.662
FC-AE	0.764	0.751	0.911	0.752	0.823	0.684
CNN-AE	0.382	0.523	0.803	0.604	0.683	0.347
VQ-VAE	0.395	0.427	0.872	0.374	0.526	0.438
CVAE	0.781	0.739	0.888	0.748	0.808	0.669

TABLE II
PERFORMANCE COMPARISON OF DIFFERENT MODELS ON MVTEC AD DATASET

In Table II We evaluated the proposed BiGAN-based model and we trained 10 models. We used bottle dataset and the proposed model outperformed the other models. It achieved 0.921 AUC, 0.823 accuracy, and 0.981 precision, thus it can identify almost all defective bottles. It has a low number of false alarms because it achieved 0.813 recall and 0.892 F1-score. This indicates that it achieves a good trade-off between precision and recall, therefore it is a reliable model. By contrast, other models such as BiGAN, GANomaly, AnoGAN, and AE models performed poorly, so they are not suitable for this task. In particular, VQ-VAE and CNN-Autoencoder showed very poor performance, and all scores were low. This indicates that GAN models with encoders are superior, because they are good at detecting defects. They are good at detecting defects, thus this can help make things better and cheaper.

Model	AUC	Accuracy	Precision	Recall	F1	Specificity
Improved BiGAN	0.939	0.919	0.969	0.931	0.949	0.850
BiGAN	0.852	0.820	0.949	0.820	0.879	0.780
GANomaly	0.840	0.810	0.946	0.810	0.872	0.770
AnoGAN	0.837	0.800	0.943	0.800	0.865	0.760
VAE	0.835	0.754	0.926	0.766	0.839	0.695
Beta-VAE	0.830	0.746	0.928	0.751	0.832	0.720
FC-AE	0.796	0.769	0.927	0.785	0.849	0.690
CNN-AE	0.409	0.606	0.838	0.653	0.731	0.370
CVAE	0.410	0.711	0.896	0.773	0.831	0.550
VQ-VAE	0.789	0.889	0.938	0.928	0.933	0.695

TABLE III
PERFORMANCE COMPARISON OF DIFFERENT MODELS ON PLANT DISEASE DATASET

Table III The best-performing model for the detection of diseased plants was Improved BiGAN with an accuracy of 91.9% (precision, 96.9%; recall, 93.1%; F1 score, 94.9%; specificity, 85.0%) and VQ-VAE showed the second-best performance with an accuracy of 88.9% (recall, 92.8%; precision, 93.8%; specificity, 69.5%), but it had the lowest specificity among all models. BiGAN, GANomaly, and AnoGAN showed relatively good performance with an accuracy of 80%-82% but was worse than Improved BiGAN, thus VAE, Beta-VAE, and FC-Autoencoder showed the worst performance with an accuracy of around 75%. CNN-Autoencoder showed the worst performance with an accuracy of 60.6% and the lowest specificity, and it also showed the worst precision among all models. CVAE showed relatively good accuracy and recall but bad precision and specificity, therefore Better BiGAN was the highest performer because of its higher sensitivity (better at identifying diseased plants) and specificity (lower false-positives, i.e. healthy plants identified as diseased).

This demonstrates the superiority of Improved BiGAN over the other methods, and Improved BiGAN outperforms all other methods for the detection of anomalies Brain MRI, MVTecAD and Plant Disease. BiGAN and GANomaly perform comparably to Improved BiGAN, but CNN-Autoencoder and VQ-VAE yield lower performance. This indicates that further development of these methods is required, because in the medical domain anomaly detection, it is extremely important to apply a robust model, therefore a more robust model should be used.

In this study, we leveraged BiGAN (Bidirectional Generative Adversarial Network) model and proposed a improved version of it for image reconstruction and anomaly detection, and it was tested on 3 datasets. A medical dataset (Brain MRI), a MVTec AD dataset (for industrial anomaly detection), and a plant disease dataset are the datasets used.

For the medical data, the model was evaluated by comparing the original and reconstructed images pixel-by-pixel, to evaluate how well it was able to identify diseased regions, and because the model learned normal anatomy, it was able to reconstruct healthy images with low pixel-error across the entire image. However, the model was unable to reconstruct diseased samples well (e.g., those containing tumors), and thus, had higher pixel-errors in the tumor regions, therefore, the results showed a significant difference between the healthy and diseased image reconstructions, so the model's performance was affected by the presence of tumors.

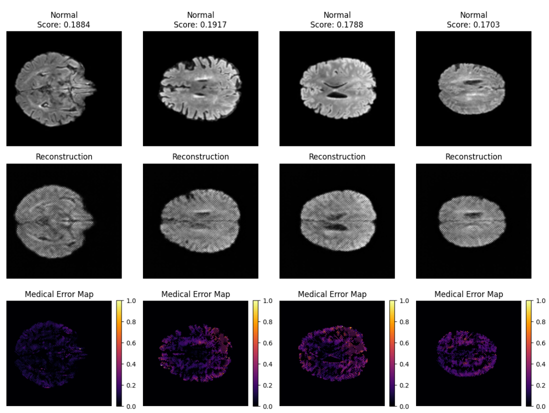

Fig. 5. Error mapping for Healthy Samples

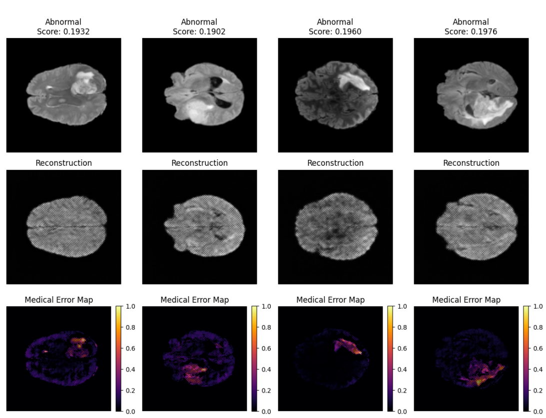

Fig. 6. Error mapping for Diseased Samples

This difference in reconstruction error forms the basis for automatic localization of diseased regions. As shown in Figure 5 and Figure 6, the error maps for healthy samples show minimal and diffuse error, while the error maps for diseased samples display clear, localized highlights corresponding to the abnormal regions.

The advantages of this method include: (1) obvious and non-annotation abnormalities; (2) early and accurate detection; (3) high sensitivity to detect diseased regions and high specificity to identify healthy regions, and these results suggest the potential applications of this method in computer-aided diagnostic systems.

In addition, we have conducted more experiments and we used MVTec AD dataset and plant disease dataset. MVTec AD dataset contains various products with defects, because the dataset is designed to test the performance of anomaly detection models. The plant disease dataset is a set of images of healthy leaves and leaves with diseases, thus providing a comprehensive dataset for disease classification.
In both additional data sets, the model was always able to reconstruct normal samples with low error, and the model was also able to detect defects or diseased parts, as it led to high pixel-wise errors in the reconstruction. This indicates that the model is capable of detecting anomalies outside the medical field as well, because it also works well on other types of images.

In summary, our improved BiGAN model and pixel-level error evaluation provides a robust and generic framework for unsupervised anomaly detection and localization, and based on the results of medical data and the consistent satisfactory performance on MVTec AD and plant disease dataset, our proposed model is effective and can be applied to various applications in and beyond medical field.

VI. Conclusion

This paper validates the effectiveness of reconstruction-based anomaly detection, especially the proposed Enhanced Bi-GAN, and the performance superiority of our proposed architecture over the existing models in various evaluation metrics validates the reliability and effectiveness of the proposed model in anomaly detection for complex data. The improved accuracy, stability and consistency of the Enhanced Bi-GAN in detecting the subtle deviations are essential for practical applications, so the proposed model can be applied to early warning systems, preventive maintenance and better-informed decision making. As anomaly detection is a critical task in cybersecurity, industrial monitoring, healthcare diagnostics and fraud detection, by identifying the abnormal patterns effectively, the proposed model can be applied to early warning systems, preventive maintenance and better-informed decision making, therefore, it can lead to more efficient operations. Besides, the proposed model can be deployed in resource-limited settings due to its ability to adapt to different types of data and its improved reconstruction ability, thus making it a suitable solution for a wide range of applications. In summary, this research provides solid ground for the future developments in reconstruction-based anomaly detection and motivates further exploration and refinement of generative adversarial frameworks, because it has the potential to improve the field significantly.

VII. Acknowledgment

We acknowledge the blessing of God for the strength, endurance, and persistence he has given us to complete this thesis, and we acknowledge our supervisor, DR. Amitabha Chakrabarty for his mentorship, assistance, and input on this research.

We acknowledge Dr. Fahim-Ul-Islam, Research Assistant, BRAC University, for his assistance and providing a conducive learning and research environment, because he has played a significant role in our academic journey, and thus we acknowledge our parents and family for their care, prayers, and support throughout our studies.

We acknowledge all individuals, groups, and resources that provided data, tools, or assistance that contributed to our ability to complete this thesis, therefore we are grateful for their contributions.

References

[1] S. C. Jinwon An, "Variational autoencoder based anomaly detection using reconstruction probability," 27 2015. [Online]. Available: http://dm.snu.ac.kr/static/docs/TR/SNUDM-TR-2015-03.pdf.

[2] J. Donahue, P. Kr¨ahenb¨uhl, and T. Darrell, Adversarial feature learning, en, May 2016. [Online]. Available: https://arxiv.org/abs/1605.09782.

[3] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, "Ganomaly: Semi-supervised anomaly detection via adversarial training", 2018, arXiv: 1805.06725 [cs.CV], and it is available online at https://arxiv.org/abs/1805.06725.

[4] S. S. Hristina Uzunova, "Unsupervised pathology detection in medical images using conditional variational autoencoders", International Journal of Computer Assisted Radiology and Surgery, vol. 14, pp. 451-461, 2018, therefore the DOI for this article is 10.1007/s11548-018-1898-0, and it is available at https://doi.org/10.1007/s11548-018-1898-0.

[5] Lu and P. Xu, "Anomaly detection for skin disease images using variational autoencoder", 2018, thus the DOI for this article is 10.48550/arXiv.1807.01349, and it is available at https://doi.org/10.48550/arXiv.1807.01349, arXiv: 1807.01349 [cs.LG].

[6] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, "T-anogan: Fast unsupervised anomaly detection with generative adversarial networks", Medical Image Analysis, vol. 54, pp. 30-44, 2019, because the ISSN for this journal is 1361-8415, and the DOI is https://doi.org/10.1016/j.media.2019.01.010, so it is available at https://www.sciencedirect.com/science/article/pii/S1361841518302640.

[7] Y. L. Kai Jiang Weiying Xie, "Semisupervised spectral learning with generative adversarial network for hyperspectral anomaly detection", vol. , no. , pp. 5224-5236, Feb. 2020, and the DOI for this article is 10.1109/TGRS.2020.2975295, thus it is available at https://doi.org/10.1109/TGRS.2020.2975295.

[8] Z. Niu, K. Yu, and X. Wu, "Time-series anomaly detection using lstm-based vae-gan", Sensors, vol. 20, no. 13, 2020, therefore the ISSN for this journal is 1424-8220, and the DOI is 10.3390/s20133738, so it is available at https://www.mdpi.com/1424-8220/20/13/3738.

[9] W. Shin, S.-J. Bu, and S.-B. Cho, "3d-convolutional neural network with generative adversarial network and autoencoder for robust anomaly detection in video surveillance", International Journal of Neural Systems, vol. 30, no. 06, p. 2050034, 2020, because the PMID for this article is 32466693, thus the DOI is 10.1142/S0129065720500343, and it is available at https://doi.org/10.1142/S0129065720500343.

[10] V. Skandarani, N. Painchaud, P.-M. Jodoin, and A. Lalande, "On the effectiveness of gan generated cardiac mris for segmentation." 2020, and it is available online at http://arxiv.org/abs/2005.09026. doi: 10.48550/arXiv.2005.09026.

[11] L. Sun, J. Wang, Huang, et al.: "Adversarial learning for medical image synthesis for lesion detection," IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 9, pp. 2303-2314, 2020, because it can be found online at http://ieeexplore.ieee.org/document/9092709/. doi: 10.1109/JBHI.2020.2964016.

[12] K. T. Xue Wang, "Cvae-2: A conditional variational autoencoder with adversarial training for hyperspectral image classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 8, pp. 5676-5692, Aug. 2020, thus it is available online at http://ieeexplore.ieee.org/document/9096553/. doi: 10.1109/TGRS.2020.2968304.

[13] F. Carrara, G. L. Amato, F. Brombin, F. Falchi and C. Gennaro, "Combining gans and autoencoders for efficient anomaly detection," 2021, and it can be accessed at http://ieeexplore.ieee.org/document/9412253/, therefore doi: 10.1109/ICPR48806.2021.9412253.

[14] P. Huang, X. Liu, and Y. Huang, "Data Augmentation for Medical MR Image Using Generative Adversarial Networks," 2021, so it is available online at http://arxiv.org/abs/2111.14297. doi: 10.48550/arXiv.2111.14297

[15] Y. Li, W. Huang, M. Luo, P. Zhang, and Y. Zha, "A novel vq-gan model to synthesize arterial spin labeling images from structural mri," Displays, vol. 70, p. 102079, 2021, because it can be found online at http://www.sciencedirect.com/science/article/pii/S0141938221000885.

[16] F. D. Matlija, P. Calzone, M. D. Simoni and E. Ghelfi, "GANs for anomaly detection: A survey," 2021, and it is available at http://arxiv.org/abs/1906.11632, thus doi: 10.48550/arXiv.1906.11632.

[17] H. Y. X. L. S. H. W. Z. L. J. Y. P. Qian, "Generating multi-contrast mr images using new 3d conditional variational auto-encoding gan," Mobile Netw Appl, vol. 26, issue 26, pp. 415-422, 2021, therefore it is available at http://link.springer.com/10.1007/s11036-020-01678-1. doi: 10.1007/s11036-020-01678-1.

[18] V. N. Rohit Sahoo, "Generative adversarial networks and variational auto encoders as techniques for synthetic data generation and augmentation to improve heart disease prediction," Dec. 2021, so it can be accessed online at http://www.ijrte.com/docs/IJRTE-B3263-1211221.pdf. doi: 10.35940/ijrte.B3263.1211221.

[19] M. Abedi, I. L. Jensen, S. Sadeghi, and T. Kirsten, "Gan-based methods for generating structured data in the medical field," Applied Sciences, vol. 12, no. 14, 2022, and it is available online at http://www.mdpi.com/2076-3417/12/14/7075, therefore doi: 10.3390/app121472075.

[20] B. Ahmad, J. Sun, Q. You, V. Palade, and Z. Mao: Brain tumor classification using variational autoencoder and generative adversarial networks, and this research is available at [Link](https://www.mdpi.com/2227-9059/10/2/223).

[21] A. Zayat, M. A. Hasabelnaby, M. Obeed, and A. Chaaban, proposed Transformer masked autoencoders for next-generation wireless communications architecture and opportunities, therefore, their work was published in 2023, with a doi: [10.3390/s23010355](https://doi.org/10.3390/s23010355).

[22] L. Zhang, Y. Dai, F. Fan, and C. He discussed attentional feature fusion for anomaly detection in gan industrial images, thus, their research is available at [Link](https://doi.org/10.48550/arXiv.2401.03322).

[23] S. A. Najafi, M. H. Asenami, and P. Setoudeh proposed an attention and autoencoder hybrid model for unsupervised online anomaly detection, and their work is also available at [Link](https://doi.org/10.48550/arXiv.2401.03322).