Screw Anomaly Detection Comparison of YoloV8 with Variational Auto Encoders and Generative Adversarial Networks

Manoj Hudnurkar[1], Geeta Sahu[2], Suhas Ambekar[1], Janhavi Vadke[2] and Kartik Kulbhaskar Singh[1]
[1]*Symbiosis Centre for Management and Human Resource Development, (SCMHRD), Symbiosis International (Deemed University), Pune, India*
[2]*Department of Information Technology and Data Science, Vidyalankar School of Information Technology, (VSIT), Mumbai, India*

Keywords: YoloV8, Generative Adversarial Networks, Variational Auto Encoder, Latent Space Exploration, Data Augmentation.

Abstract: This research introduces a novel approach to anomaly detection in screw manufacturing processes by synergising YoloV8 (You Only Look Once) and hybrid Variational Auto encoders (VAE) and Generative Adversarial Networks (GAN). In our present undertaking, we are utilizing a thoughtfully curated dataset from Kaggle. Our primary emphasis is accurately detecting anomalies, particularly subtle irregularities in specific image areas of the screws. Our research underscores the importance of authentic datasets and involves the assessment of advanced methods, explicitly focusing on analysing the MVTec Anomaly Detection dataset for screws. The YoloV8 model showcases its ability to accurately reconstruct images and detect anomalies, showing great potential for applications in maintaining high manufacturing quality standards. Also, VAE and GAN results are acceptable. When YoloV8 is compared against VAE & GAN, the results in YoloV8 provide the highest accuracy with precision & recall. A comprehensive quantitative evaluation of the overall framework's performance in distinguishing between normal and abnormal cases is achieved by including a classification report that provides precision, recall, and F1-score metrics. According to the results, the accuracy attained while applying VAE-GAN is approximately 90%, while the accuracy attained when employing YoloV8 is between 95% and 97%, with high-speed performance. As a result, YoloV8 performs well and processes information more quickly than other traditional methods. These results highlight the importance of using customized datasets and suggest exciting opportunities for improving anomaly detection techniques in the manufacturing industry.

1 INTRODUCTION

Human ability to recognize novel or anomalous images surpasses current machine learning capabilities. Unsupervised algorithms for detecting anomalies, crucial in applications like manufacturing optical inspection, face challenges due to limited defective samples (Dlamini, Kao, et al. , 2021). Recent interest focuses on novelty detection using modern machine learning architectures in natural image data. In classification settings, existing algorithms often prioritize outlier detection, where inlier and outlier distributions differ significantly. The evaluation involves labelling classes arbitrarily from object classification datasets as outliers and using others as inliers. However, assessing state-of-the-art methods for anomaly detection tasks and identifying subtle deviations in con-fined regions remains unclear. The lack of comprehensive real-world datasets for such scenarios hampers the development of suitable machine-learning models. Addressing this gap, large-scale datasets like MNIST (LeCun, 1998), CIFAR10 (Hinton), or ImageNet (Krizhevsky, 2012) have significantly advanced computer vision in recent years.

For object detection, the advanced model of the YOLO family of object detection models, which are renowned for their accuracy and speed, is called YOLOv8 (You Only Look Once version 8). (Zhang, Ren, et al. , 2016) YOLO models use a single neural network to recognize and classify items in real time. It is known for its better design compared to earlier iterations, and YOLOv8 boasts an improved design that increases accuracy and performance. This could

Hudnurkar, M., Sahu, G., Ambekar, S., Vadke, J. and Singh, K. K.
Screw Anomaly Detection Comparison of YoloV8 with Variational Auto Encoders and Generative Adversarial Networks.
DOI: 10.5220/0013598100004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 2*, pages 598-608
ISBN: 978-989-758-763-4

involve enhanced detecting heads, enhanced backbone networks, and enhanced feature extraction. (Sohan, SaiRam, et al. , 2024), (Redmon, Divvala, et al. , 2016). The goal of YOLOv8 is to provide increased speed and accuracy. (Ren, Girshick, et al. , 2017). This implies that it has faster and more accurate object detection. (Hussain, 2023).

(Salimans, 2018)In recent years, the integration of advanced generative models has propelled the field of computer vision, offering unprecedented capabilities in image synthesis and manipulation. The decoder then uses the sampled points to rebuild the images, guaranteeing various realistic outputs. Adversarial learning is incorporated into the model using a specialized discriminator. By assessing the created images' authenticity, this discriminator develops a dynamic learning process that improves the generator's capacity to produce realistic content.

Custom loss functions like correlation loss are introduced to improve training. (Lee, 2021)The codebase has smooth integration of visualization tools that facilitate the exploration of latent space and evaluation of generated image quality. The comprehensive assessment of training and test sets at the end of the study paper shows how well the model can reconstruct and produce high-quality images.

Outstanding generative models are critical in image synthesis, manipulation, and other fields where our novel combined VAE-GAN system has potential applications. The following sections explore the specifics of our new method's design, training procedure, and outcomes, adding to the current discussion on how adversarial training and probabilistic modelling in deep generative models are coming together. The deficiencies mentioned above led to the formulation of the following research questions:

RQ1: To what extent can image reconstruction be used to detect irregularities in screws? The code creates a single framework for sophisticated picture production and modification that integrates Generative Adversarial Networks (GAN) with Variational Auto encoders (VAE).

Leveraging Tensor Flow and Keras, the model employs VAE for latent space exploration and GAN for adversarial training. The process involves data loading, augmentation, and the construction of an encoder-decoder architecture for VAE, complemented by a discriminator for GAN training.

The study makes use of a large screw dataset that is accessible on Kaggle and includes pictures of screws with various characteristics, both good and bad. The basis for identifying abnormalities in screws is the test dataset, which contains screws with altered features, including front, scratch head, scratch neck, thread side, and thread top. To create efficient methods for image reconstruction and precise screw anomaly identification, the study technique systematically examines abnormalities in various screw images.

Building an advanced encoder-decoder framework for VAE and a discriminator specifically designed for GAN training constitute the fundamental components of the architecture.

Custom loss functions are introduced to optimize the training process, most notably the addition of correlation loss. These loss functions are essential for improving the model's generative and discriminative skills. Rigorous training and test set evaluation systematically demonstrate the model's competency. This assessment verifies that the combined VAE-GAN framework produces outputs of excellent quality and can rebuild images.

In summary, this research endeavours to push the boundaries of generative models by integrating VAE and GAN, capitalizing on the diverse screw dataset from Kaggle. The comprehensive approach, from data pre-processing to model evaluation, showcases the potential of this integrated framework for real-world applications where advanced image generation is paramount.

The whole paper structure starts with an introduction of the research study found in section 1, where the document is divided into sections. Section 2 talks about reviews of the literature. The problems and inadequacies in the research are discussed in Section 3. Section 4 talks about the study's goals. The application of the research approach is covered in Section 5. Section 6 discusses the analysis and findings of the experiment. The debate based on the YoloV8's comparison with VAE and GAN is covered in Section 7. The conclusion is covered in Section 8. Section 9 concludes with a discussion of the research's future scope.

2 LITERATURE REVIEW

To extract information from input photos, YOLOv8 uses an advanced backbone network. This backbone network is probably built on top-tier convolutional neural networks (CNNs), incorporating newer techniques to improve feature extraction. YOLOv8 uses sophisticated convolutional layers such as wise separable or re-parameterized convolutions to minimize computational complexity while retaining high accuracy. The model may use enhanced feature fusion approaches to better mix features from various network scales and levels and improve the detection of small and large objects (ALRUWAILI, ATTA, et al. , 2023).

The methodical technique for applying YOLOv8 for defect identification involves preparing the dataset and gathering information about faults pertinent to your application, such as dents, scrapes, cracks, and missing pieces. (Redmon, Divvala, et al. , 2016). To make the model robust, ensure the dataset has photographs with various backgrounds, lighting situations, and perspectives. Bounding Boxes: Use bounding boxes to indicate any image flaws. If there are multiple flaws, a class should identify each one. To expand the training data's diversity and strengthen the model's resilience, apply data augmentation techniques such as rotation, scaling, flipping, and colour tweaks (Carrera, 2017).

Introduce NanoTWICE, a dataset featuring 45 grey-scale images showcasing Nano fibrous material captured by a scanning electron microscope. Forty photographs with anomalous regions (dust particles, flattened areas, etc.) are left for training, with five defect-free images provided. Nevertheless, as the dataset only provides one kind of texture, it is unknown whether the techniques tested on this dataset may be applied to other textures from other domains.

In a 2007 DAGM workshop, Wieler and Hahn (Bergmann, 2019) They presented a dataset designed explicitly for the optical inspection of textured surfaces. This dataset includes ten classes of artificially generated grey-scale textures with defects weakly annotated as ellipses. A total of 1000 flawless texture patches for training and 150 flawed patches for testing are included in each lesson. However, the annotations are rough, and there is little difference in appearance between different textures because relatively similar texture models are used. Additionally, artificially generated datasets serve as approximations to real-world scenarios.

The notable work by (Bergmann, 2019) Stands out. They have used the MVTec Anomaly Detection dataset. This dataset encompasses 5354 high-resolution colour images across various texture categories. It includes standard (defect-free) images for training and images with anomalies for testing. The anomalies span over 70 defects, including scratches, dents, contaminations, and structural changes. Notably, pixel-precise ground truth regions for all anomalies are provided. The study thoroughly assesses state-of-the-art unsupervised anomaly detection techniques, combining traditional computer vision techniques with deep architectures such as convolutional auto encoders, generative adversarial networks, and feature descriptors using pre-trained convolutional neural networks. In our research, we leverage this dataset with a specific focus on screws, aiming to achieve accurate image reconstruction for precise depiction of anomalies in screws.

(Wang, 2021)Repurposing existing classification datasets with available class labels, such as MNIST, is common in evaluating outlier detection methods within multi-class classification scenarios. (LeCun, 1998), CIFAR10 (Hinton, 2009), and ImageNet (Krizhevsky, 2012). This approach, widely adopted (Cho, 2015), (Chalapathy, 2018), (Ruff, 2018), involves selecting a subset of classes and relabelling them as outliers. The novelty detection system is then trained exclusively on the remaining inlier classes, with the testing phase assessing the model's ability to correctly predict whether a test sample belongs to an inlier class. (Buterin, 2013). While this yields substantial training and testing data, the anomalous samples significantly differ from those in the training distribution. Consequently, evaluating how a proposed method generalizes to anomalies with less pronounced differences from the training data manifold remains uncertain.

To address this challenge, (Saleh, 2015) Introduce a dataset from internet search engines, which includes six categories of abnormally shaped objects (e.g., oddly shaped cars, aeroplanes, and boats). The purpose of these objects in the PASCAL VOC dataset is to set them apart from regular samples of the same class.

The purpose of these objects in the PASCAL VOC dataset is to set them apart from regular samples of the same class. (Everingham, 2015). Even though their data may be more similar to the training data set, the choice is based on the complete image rather than identifying the unique or unusual elements. This technique finds widespread application across various domains, including cybersecurity, fraud detection, fault monitoring in industrial processes, and healthcare. In cybersecurity, anomaly detection can help identify unusual network activities indicative of potential security breaches.

(Karame, 2018) Fraud detection aids in spotting atypical transaction patterns that may suggest fraudulent activities. In industrial settings, anomaly detection is valuable for detecting early equipment failures or deviations from standard operational behaviour. Anomaly detection is pivotal in enhancing data-driven decision-making by highlighting irregularities that might go unnoticed (Welling, 2019)

We offer a summary of popular datasets for natural image anomaly identification. We highlight the need for a new dataset, distinguishing between those that necessitate a binary choice between images with and without defects and those that allow anomalous region segmentation.

When evaluating methods for segmenting anomalies in images, the availability of public datasets is limited, with a focus on textured surfaces. There is a notable absence of a comprehensive dataset allowing for segmenting abnormal regions in natural images.

3 RESEARCH GAPS AND CHALLENGES

- Given the limited number of defective samples, unsupervised algorithms used for anomaly detection, particularly in applications such as industrial optical inspection, encounter challenges.
- In cases where the distributions of inliers and outliers exhibit significant divergence, existing methods often prioritize outlier detection in classification scenarios.
- Reliably determining the anomalies in screws through image reconstruction is a significant issue.
- The assessment includes categorizing classes in object classification datasets as outliers and the rest as inliers, which is a bit difficult.

Uncertainty surrounds evaluating cutting-edge methods for anomaly detection tasks and detecting subtle deviations in specific areas.

The absence of extensive real-world datasets for such situations hinders the development of appropriate machine-learning models.

4 OBJECTIVES OF THE RESEARCH

To systematically investigate anomalies in different screw images, aiming to develop effective strategies for image reconstruction and accurate identification of screw anomalies.

To introduce object detection, segmentation & classification using the YoloV8 model, which can detect the object with & without anomalies.

To implement a unified framework integrating Variational Auto encoders (VAE) and Generative Adversarial Networks (GAN), i.e. VAE-GAN, for advanced image generation, synthesis and manipulation.

To compare the results of YoloV8 with VAE & GAN.

Custom loss functions and correlation loss are introduced to optimize and maximize the model's performance of YoloV8 against VAE & GAN.

5 RESEARCH METHODOLOGY

The suggested model, YoloV8, provides a thorough approach to visual anomaly identification.

5.1 Approach

First, picture datasets are loaded and enhanced. The model incorporates loss functions like correlation loss to maximise the training process. Figure 1 represents the research approach used in the study, as shown below.

Figure 1: The research approach employed for the study

5.1.1 Identify the Problem and Define the Objective

One of the main goals of this research project is to create an efficient model for picture anomaly detection that specifically addresses screw inspection. This project aims to create a model that can precisely detect irregularities in screw images and distinguish between screws that are devoid of defects and those that have irregularities like scratches, manipulation, and unevenness. We aim to build a strong framework that combines image reconstruction, anomaly recognition, and image production while utilizing deep learning approaches.

5.1.2 Data Collection

The dataset used in this study is sourced from Kaggle, with the underlying data originating from mvtech.com. Comprising three files: train, test, and ground_truth-the dataset includes 320 high-resolution (1024x1024) images of defect-free screws for training and 160 test images categorized into classes such as sound, manipulated_front, scratch_head, scratch_neck, thread_side, and thread_top. An accurate assessment and validation of the anomaly detection capabilities of the model are guaranteed by the availability of ground_truth annotations.

5.1.3 Exploratory Data Analysis (E.D.A.) and Feature Selection

To understand the dataset's features, EDA entails carefully reviewing it. The photos are enhanced, converted, and pre-processed for functional model training. The selected characteristics consist of picture pixels; it is rotated and flipped to increase the dataset's diversity.

5.1.4 Model Creation

To create realistic images through adversarial training, the architecture learns the latent representations of the input images. Custom loss functions like correlation loss are integrated, and the model optimizes training. By combining reconstruction loss, KL divergence, and GAN loss, the VAE-GAN framework is trained. By precisely re-constructing flawless photos and recognising irregularities in screws, the model is refined to demonstrate its ability to detect anomalies in the image dataset. KL divergence, GAN loss, and reconstruction loss are combined to train the integrated VAE-GAN framework.

5.1.5 Discussion

The methodology evaluates training and test sets to show the model's effectiveness. The code also makes it easier to explore rebuilt images and anomalies, demonstrating the model's accuracy in reconstructing and identifying anomalies in screw images. A comparison of Yolo8 with VAE & GAN is also introduced.

5.2 Tools & Dataset Utilized

More significantly, the Python programming language, Jupyter Notebook, will be utilized in this study's data cleaning, preparation, and analysis stages. Several libraries, including Seaborn, Numpy, Matplotlib, and Pandas, will be used for data analysis and visualisation.

The primary data for this study originated from mvtech.com, and the dataset used was obtained from Kaggle. The dataset, which consists of three separate files (train, test, and ground_truth), is essential to our research. The training set consists of 320 high-resolution (1024x1024) pictures of regular screws that serve as a baseline for training the model. In contrast, the test file has 160 photos divided into thread_top, thread_side, good, manipulated_front, scratch_head, and scratch_neck categories. The test photos have corresponding ground_truth annotations that help properly assess and validate the model's anomaly detection skills.

5.3 Techniques Utilized

- YoloV8: A well-liked object detection model created by Ultralytics is called YoloV8 (You Only Look Once version 5). It is an improvement on the deep learning models of the YOLO series, which are intended to detect objects in real-time. YoloV8 is well-suited for quickly detecting objects in applications like robotics, autonomous cars, and surveillance since it can process photos and videos at a high frame rate. It successfully balances precision and speed.
- VAE: Variational Auto encoders (VAEs) are an advanced version of traditional auto encoders that map input data into a probabilistic latent space using an encoder-decoder architecture. The Variational Auto Encoder (VAE) draws inspiration from the Helmholtz Machine. (Dayan, 1995), which introduced the concept of a recognition model. Its lack of optimization for a single objective paved the way for the development of V.A.E.s. Nevertheless, using the reparameterization approach, it back propagates via the numerous layers of the deep neural networks that are nested within it. (Welling, 2019).
- Since its inception, the VAE framework has undergone various extensions, including applications to dynamic models. (Johnson, 2016), models with attention (Gregor, 2014), and those incorporating multiple levels of stochastic latent variables (Kingma, 2016). VAEs have proven to be a fertile ground for building diverse generative models. The Generative Adversarial Network (GAN) model has also garnered noteworthy interest (Goodfellow, 2014). Recognising these complementary strengths has led to the proposal of hybrid models to leverage both approaches' advantages. (Dumoulin, 2017), (Grover, 2018), (Rosca, 2018)

- GAN: Generative Adversarial Networks (GANs) are prominent in Machine Learning (ML) frameworks. (Grnarova, 2019). The practical application of GANs gained momentum in 2017, with an initial focus on refining the generation of human faces, showcasing the technology's capability for image enhancement and producing more compelling illustrations at high-intensity levels. This historical background highlights the development of GANs and their revolutionary influence on several machine-learning domains. (Aggarwal, 2021). Two neural networks comprise the GAN architecture: a discriminator and a generator. The generator attempts to mimic the properties of accurate training data by creating synthetic data out of random noise. In addition, the discriminator serves as a binary classifier that discerns between real and fake data.

6 EXPERIMENTAL RESULTS & ANALYSIS

6.1 Yolo8

This section has demonstrated the practical significance of the suggested methods by discussing their outcomes in several performance metrics, including accuracy, mAP, precision, Confusion matrix, etc. According to an analysis of the trained models ' observations and results, models trained on segmented images perform better than colour and grayscale images. The minimal noise in the photos adds to the incredible accuracy of the models.

A new Transfer-learning model is called YoloV8. This model is trained on the available dataset for 100 epochs. VAE & GAN model achieved an accuracy of approximately 90%. But the accuracy achieved in YoloV8 is highest, around 97%, with a confusion matrix of 0.93 in the correct prediction of the screws with & without anomalies. Figure 2 below shows object detection & classification from the images dataset of screws. Table 1 shows the classification for anomaly detection. Table 2 gives the experimental setup utilised in YoloV8.

Table 1: Classification of Anomaly Detection

Sr. No.	Class
1	manipulated_front
2	scratch_head
3	scratch_neck
4	thread_side
5	thread_top

Table 2: Experimental Setup Utilized in Yolov8

Sr. No.	Parameters	Experimental Values
1	Epochs	100
2	Batch_Size	16
3	Image_size	640
4	Optimizer	Stochastic Gradient Descent (SGD)

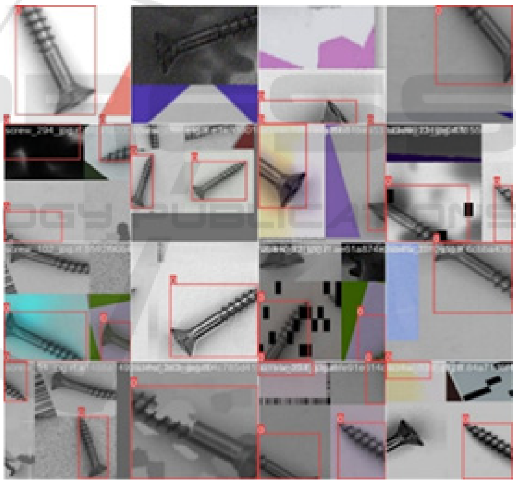


Figure 2: Training the model using Yolo8 object detection & classification from the image dataset.

The results obtained in Figure 3 below indicate the rate of accuracy in terms of performance when the YoloV8 is used for detecting the screws with and without anomalies. The value '1.0' indicates the screws are perfectly alright without any anomaly. The values' 0.6, 0.7' show the screws are slightly distorted and in the anomaly detection category.

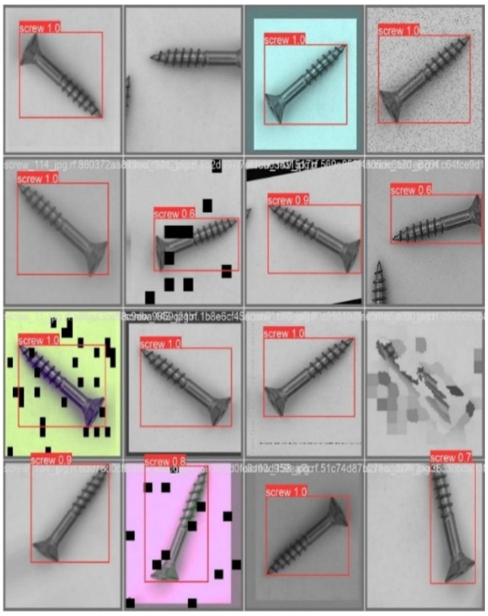


Figure 3: Accuracy achieved when Yolo8 object detection is utilized with & without anomaly.

Utilizing YoloV8, the overall loss decreased from 1.29 to almost 0.230 as the number of iterations increased from 0 to 99. The loss while training the dataset with & without anomaly images at the 99th iteration is 0.214, which is very low compared to other conventional methods. The metric-mAP50 values gradually increased to 0.897. Figure 4 below is the graphs displaying YoloV8's various parameter values at multiple epochs.

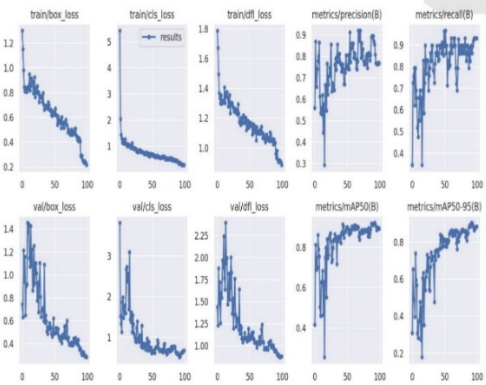


Figure 4: Graphs displaying YoloV8's various parameter values at various epochs

For testing purposes, the model has undergone 100 epochs; at the initial stage of the first 20 epochs, the precision obtained is 0.76. In subsequent iterations between 80-100 epochs, the precision obtained is 0.89 at the 95th iteration, which is higher than other conventional methods & in comparison to VAE-GAN. Similarly, the recall value at the 97th iteration is 0.9310.

Figure 5 shows the confusion matrix for the training dataset, which obtained a value of 0.93 and achieved good accuracy when measured against VAE & GAN.

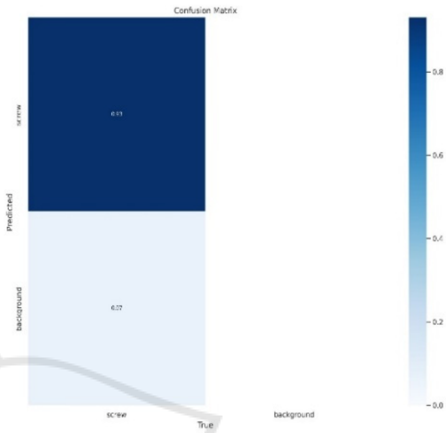


Figure 5: Confusion Matrix

Figures 6,7,8 show the positive metrics calculated: precision-confidence, precision-recall, and recall-confidence curves, respectively.

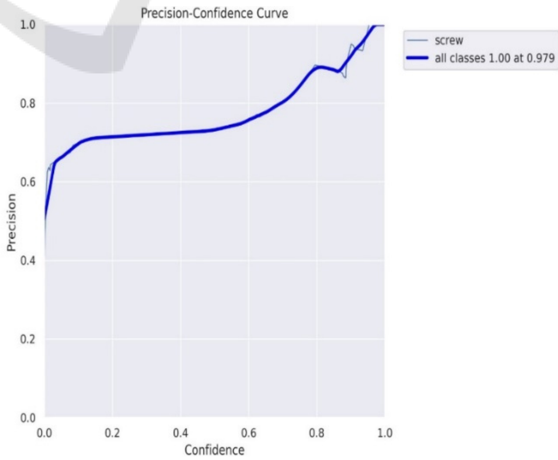


Figure 6: Precision-Confidence Curve

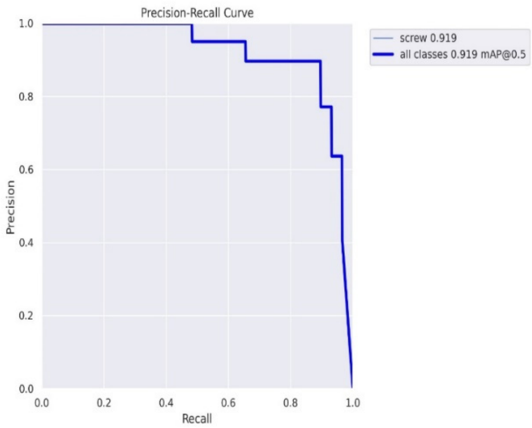


Figure 7: Precision-Recall Curve

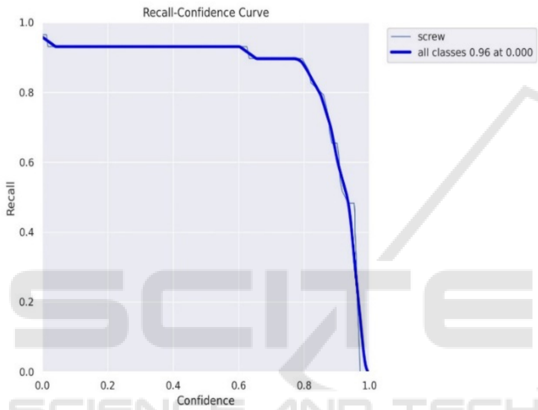
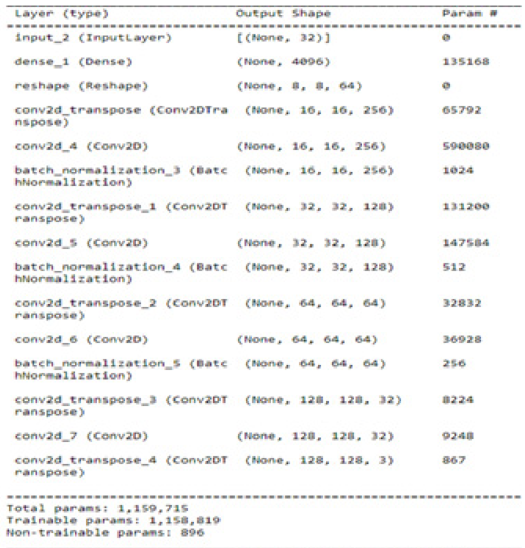


Figure 8: Recall-Confidence Curve



```
Layer (type)                 Output Shape              Param #
=================================================================
input_2 (InputLayer)         [(None, 32)]              0

dense_1 (Dense)              (None, 4096)              135168

reshape (Reshape)            (None, 8, 8, 64)          0

conv2d_transpose (Conv2DTra  (None, 16, 16, 256)       65792
nspose)

conv2d_4 (Conv2D)            (None, 16, 16, 256)       590080

batch_normalization_3 (Batc  (None, 16, 16, 256)       1024
hNormalization)

conv2d_transpose_1 (Conv2DT  (None, 32, 32, 128)       131200
ranspose)

conv2d_5 (Conv2D)            (None, 32, 32, 128)       147584

batch_normalization_4 (Batc  (None, 32, 32, 128)       512
hNormalization)

conv2d_transpose_2 (Conv2DT  (None, 64, 64, 64)        32832
ranspose)

conv2d_6 (Conv2D)            (None, 64, 64, 64)        36928

batch_normalization_5 (Batc  (None, 64, 64, 64)        256
hNormalization)

conv2d_transpose_3 (Conv2DT  (None, 128, 128, 32)      8224
ranspose)

conv2d_7 (Conv2D)            (None, 128, 128, 32)      9248

conv2d_transpose_4 (Conv2DT  (None, 128, 128, 3)       867
ranspose)

=================================================================
Total params: 1,159,715
Trainable params: 1,158,819
Non-trainable params: 896
```

Figure 9: Parameters applied in VAE & GAN

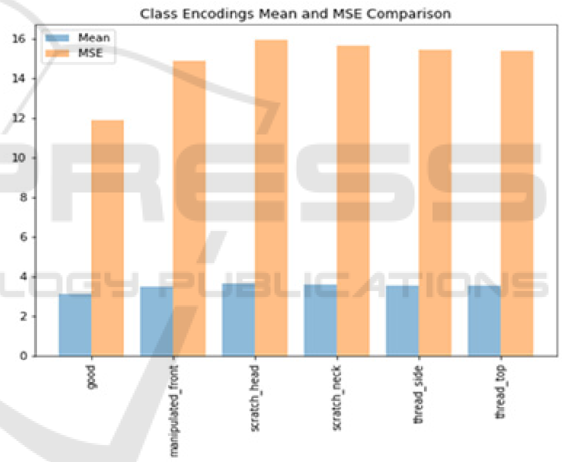


Figure 10: Class Encoding and Mean Comparison values obtained in class encoding mean and use comparison

6.2 VAE & GAN

Reconstruction error levels (criteria) are varied to perform more sophisticated analysis, and the sci-kit-learn algorithm is used to compute extensive classification metrics, including precision, recall, and F1-score. Furthermore, our approach includes a class-by-class analysis of the test data for every class. Figure 9 indicates the parameters applied while using VAE & GAN architecture. Figure 10 is the class encoding, and mean comparison is calculated to show the performance in terms of the classification of screws with and without anomalies detection. Table 3 shows the values obtained in the class encoding mean and MSE comparison.

Table 3: shows the values obtained in the class encoding mean and MSE comparison.

Screw Class	Class Encoding Mean	MSE Comparison
Good	3.21	11.99
Manipulated_front	3.62	14.82
Scratch_head	3.72	14.82
Scratch_neck	3.72	15.51
Thread_side	3.72	14.90
Thread_top	3.72	14.8

7 DISCUSSION

A thorough examination of the techniques used, centred on the use of YoloV8, Generative Adversarial Networks (GAN) and Variational Autoencoders (VAE) in the complex field of anomaly detection for screw pictures, opens the debate. These cutting-edge deep learning approaches solve problems frequently encountered in anomaly detection jobs, such as imbalanced datasets and limited data availability for particular classes. The autoencoder, VAE, and GAN work to identify abnormalities in later testing stages by comparing reconstruction errors to predetermined criteria.

A thorough quantitative assessment of the overall framework's effectiveness in differentiating between normal and abnormal cases is obtained by incorporating a classification report with precision, recall, and F1-score metrics. The results discussed indicate the accuracy achieved while using YoloV8 is between 95-97%, and when VAE-GAN is applied, the accuracy achieved is around 90%. Thus, the performance of YoloV8 is high and has faster processing than other conventional techniques

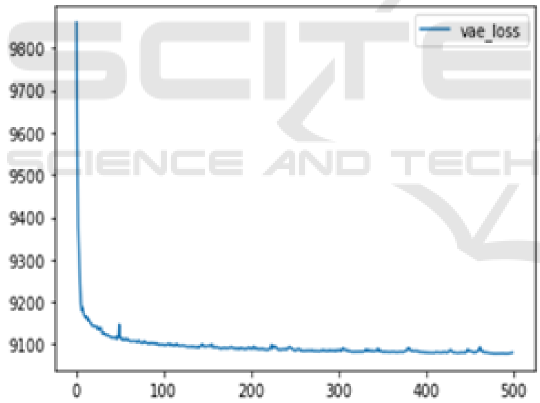


Figure 11: Graph showing the vae_loss

The figure 11 a graph on 'vae_loss creates a plot of the 'vae_loss' column over epochs. Figure 11 shows the graph of the VAE loss calculated. The loss calculated in YoloV8 is 0.2141, again less than VAE & GAN.

8 CONCLUSION

As a result, this study's one-class classification paradigm for anomaly detection in screw images incorporates robust and synergistic integrations of sophisticated deep learning approaches, namely YoloV8, Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN). The empirical investigation, characterised by thorough testing and assessment, continually shows how effective the integrated method is. It continuously demonstrates its skill in recreating images and detecting anomalies, guided by Mean Squared Error (MSE) loss. The combined use of GANs' adversarial training dynamics and VAE's probabilistic nature improves anomaly identification of overall screw and gives users a deeper comprehension of the dataset.

The effectiveness of the integrated methodology in differentiating between normal and abnormal cases is confirmed by quantitative validation using classification reports with precision, recall, and F1-score metrics. Moreover, the examination of variations in encoded representations between classes highlights how flexible the models are for a variety of screw types. However, YoloV8 yields better accuracy at 97% and 96% than VAE & GAN. Visual assessments, comparing original and reconstructed images, highlight the collective effectiveness of these advanced deep learning techniques, especially in the challenging task of accurately reconstructing images with anomalies.

9 FUTURE SCOPE OF RESEARCH

Looking ahead, future work can explore enhancements to the proposed methodology. Techniques for transfer learning could be investigated to adapt the model to new datasets or novel anomaly types. Additionally, the scalability and generalization of the approach to more extensive and diverse datasets could be a focus for further research. Integration with real-time monitoring systems and deployment in practical industrial environments could pave the way for effective implementation.

This research advances anomaly detection methodologies and emphasizes the practical relevance of integrating sophisticated deep learning techniques in addressing industrial challenges. The findings contribute to the ongoing dialogue on anomaly detection, showcasing the potential for transformative applications in quality control processes within industrial settings. In essence, our research bridges the theoretical foundations of deep learning with the pragmatic demands of industrial quality control. The findings underscore the viability and relevance of one-class classification methodologies, especially auto-encoder architectures, in addressing the challenges posed by image-based anomaly detection. As industries increasingly adopt automated systems for quality assurance, our work contributes to the evolving landscape of artificial intelligence applications to enhance precision and reliability in industrial processes.

ACKNOWLEDGEMENTS

My deepest gratitude goes out to Dr. Manoj Hudnurkar and Dr. Suhas Ambekar from Symbiosis Centre for Management and Human Resource Development, (SCMHRD), Symbiosis International (Deemed University) and to Geeta Sahu, Assistant Professor in the Department of Information Technology & Data Science at Vidyalankar School of Information Technology (VSIT), Mumbai, for their unwavering support. They continuously offered support and direction while preparing this research paper. I also want to thank everyone who assisted us directly and indirectly with the study documentation.

REFERENCES

S. Dlamini, C.-Y. Kao, S.-L. Su and C.-F. J. Kuo, “Development of a real-time machine vision system for functional textile fabric defect detection using a deep YOLOv4 model,” Textile Research Journal, vol. 92, no. 5-6, 6 September 2021.

L. B. Y. B. a. P. H. Y. LeCun, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE,vol. 86, pp. 2278-2324, 1998.

I. S. a. G. E. H. A. Krizhevsky, “ImageNet Classification with Deep Convolutional Neural Networks,” Proceedings of the 25th International Conference on Neural Information Processing Systems, pp. 1097-1105, 2012.

K. He, X. Zhang, S. Ren and J. Sun, “Deep residual learning for image recognition,” in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016.

M. Sohan, T. SaiRam and C. V. RamiReddy, “A Review on YOLOv8 and its Advancements,” in Data Intelligence and Cognitive Informatics, ResearchGate, 2024, pp. 529-545.

J. Redmon, S. Divvala, R. Girshick and A. Farhadi, “You Only Look Once:Unified, Real-Time Object Detection,” in 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016.

S. Ren, K. He, R. Girshick and J. Sun, “Networks, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, 2017.

M. Hussain, “YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection,” Machines, 2023.

C. D. P. T. &. A. M. Salimans T, “Ethereum Fraud Detection using Machine Learning,” 2018.

T. B. P. Z. L.-H. Lee, “All one needs to know about metaverse: a complete survey on technological singularity, virtual ecosystem, and research agenda,” Journal of Latex Class FIles, 2021.

M. ALRUWAILI, M. N. ATTA, M. H. SIDDIQI, A. KHAN, A. KHAN, Y. ALHWAITI and S. ALANAZ, “Deep Learning-Based YOLO Models for the Detection of People With Disabilities,” IEEE Access, 25 December 2023.

F. M. G. B. a. E. L. D. Carrera, “Defect Detection in SEM Images of Nanofibrous Materials,” IEEE Transactions on Industrial Informatics, pp. 551-561, 2017.

M. F. D. S. a. C. S. P. Bergmann, “MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection,” IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9584-9592, 2019.

S. X. Z. Z. &. L. H. Wang Y, “Ethereum Fraud Detection using Transaction Network Analysis,” 2021 IEEE International Conference on High Performance Computing and Communications, pp. 1920-1927, 2021.

A. K. a. G. Hinton, “Learning multiple layers of features from tiny images,” Technical report, University of Toronto, p. 2009.

J. A. a. S. Cho, “Variational Autoencoder based Anomaly Detection using Reconstruction Probability,” Technical report SNU Data Mining Center, 2015.

A. K. M. a. S. C. R. Chalapathy, “Anomaly Detection using One-Class Neural Networks,” 2018.

R. V. N. G. L. D. S. A. S.-d. A. B. L. Ruff, “Deep One-Class Classification,” Proceedings of the 35th International Conference on Machine Learning volume 80 of Proceedings of Machine Learning Research,, pp. 4393-4402, 2018.

V.Buterin, “Ethereum White Paper: a next generation smart contract & decentralized application platform,” 2013.

A. F. a. A. E. B. Saleh, “Object-Centric Anomaly Detection by Attribute-Based Reasoning,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 787-794, 2013.

S. M. A. E. L. V. G. C. K. I. W. J. W. a. A. Z. M. Everingham, “The Pascal Visual Object Classes Challenge: A Retrospective,” International Journal of Computer Vision, pp. 98-136, 2015.

A. E. &. C. S. Karame G.O, “Identifying Ethereum Fraud Detection using Graph Based Analysis,” In Proceeding of the 2018 ACM SIGSAC Conference on Compute and Communications Security, pp. 177-194, 2018.

P. G. E. H. R. M. N. a. R. S. Z. Dayan, “The Helmholtz Machine,” Neural Computation, pp. 889-904, 1995.

D. P. K. a. M. Welling, “An Introduction to Variational Autoencoders,” Foundations and Trends in Machine Learning: Vol xx, No. xx, pp. 1-18, 2019.

M. D. K. D. A. W. R. P. A. a. S. R. D. Johnson, “Composing graphical models with neural networks for structured representation and fast inference,” Advances in Neural Information Processing Systems, pp. 2946-2954, 2016.
K. I. D. A. M. C. B. a. D. W. Gregor, “Deep AutoRegressive Networks,” International Conference on Machine Learning, pp. 1242-1250, 2014.
D. P. T. S. R. J. X. C. I. S. a. M. W. Kingma, “Improved Variational inference with inverse autoregressive flow,” Advances in Neural Information Processing Systems, pp. 4743-4751, 2016.
I. J. P.-A. M. M. B. X. D. W.-F. S. A. a. Y. B. Goodfellow, “Generative adversarial nets,” Advances in Neutral Information Processing Systems, pp. 2672-2680, 2014.
V. I. B. B. P. A. L. M. A. O. M. a. A. C. Dumoulin, “Adversarially learned inference,” International Conference on Learning Representations, 2017.
A. M. D. a. S. E. Grover, “Flow-GAN: Combining maximum likelihood and adverserial learning in generative models,” AAAI Conference on Artificial Intelligence., 2018.
M. B. L. a. S. M. Rosca, “Distribution matching in variational inference,” arXiv preprint arXiv:1802.06847, 2018.
K. Y. L. A. L. N. P. I. G. T. H. A. K. Paulina Grnarova, “A Domain Agnostic Measure for Monitoring and Evaluating GANs,” Advances in Neural Information Processing Systems 32 (NeurIPS 2019), 2019.
M. M. G. B. Alankrita Aggarwal, “Generative adversarial network: An overview of theory and applications,” International Journal of Information Management Data Insights, 2021.