eISSN 2783-6851

NEW TRENDS in COMPUTER SCIENCES

2024
Volume 2
Issue 1
Pages 31–45
https://doi.org/10.3846/ntcs.2024.20515

LEVERAGING GENERATIVE ADVERSARIAL NETWORKS TO IMPROVE TRAINING IMAGE DATASET

Henrikas GIEDRA ✉, Gabriela VDOVIAK

Deparment of Electronic Systems, Faculty of Electronics, Vilnius Gediminas Technical University, Vilnius, Lithuania

Article History:
- received 3 December 2023
- accepted 17 April 2024

Abstract. Convolutional neural networks (CNNs) are powerful models of deep learning that are widely used in computer vision classification tasks. The purpose of this study is to investigate the impact of datasets on CNN performance, employing original datasets and expanded datasets with synthetically generated images. The Generative Adversarial Network (GAN) is an unsupervised deep learning method used for synthetic data generation and can address the limitations of image augmentations. In this study, a new GAN architecture is used to synthesize high-resolution images when dealing with limited training data. The StyleGAN2-ADA model is specifically designed to generate high-quality images using limited datasets. Adaptive Discriminator Augmentation (ADA) dynamically adjusts data augmentation, enhancing discriminator efficiency and stability. The findings indicate a reduction in the likelihood of overfitting, enhancement in network generalization, mitigation of class imbalance concerns, and a concurrent increase in the accuracy and stability of network classification.

Keywords: computer vision, convolutional neural networks, deep learning, generative adversarial networks, image classification, image synthesis.

✉Corresponding author. E-mail: *henrikas.giedra@vilniustech.lt*

1. Introduction

The use of artificial Convolutional Neural Networks (CNNs) for image recognition and classification tasks is rapidly increasing in industrial automation. However, the effectiveness of CNN classifiers is directly dependent on the size and quality of the training dataset; this is especially true in the case of images (Sarker, 2021). There are domains where acquiring a sufficient number of images is prohibitively expensive, too challenging, or they simply do not exist. Data augmentation techniques are commonly used to address this issue. However, traditional distortion operations may not provide sufficient quality or diversity in the data (Shorten & Khoshgoftaar, 2019). An excellent solution to this problem can be a Generative Adversarial Network (GAN). This network is capable of producing high-quality images that can be used to improve the training of CNN classifiers.

This research aims to explore the potential of Generative Adversarial Networks (GANs) in enhancing training image datasets for manufacturing applications. The main goals of this study are to assess the adaptability of GANs in generating realistic images of manufactured parts and their defective elements, with a particular focus on human perception. Additionally, we aim to evaluate the impact of advanced GAN architectures that use weight transfer techniques and examine the contribution of synthetic images to the effectiveness of CNNs.

Through these objectives, the research seeks to provide valuable insights into optimizing the process of training dataset expansion with synthetic images.

Similar experiments were conducted in Chan et al. (2023), where class imbalance and data scarcity issues were addressed using StyleGAN2-ADA, and classification performance was assessed on the MobileNetV3 model. These experiments focused on microscopic multi-class algae images, resulting in an 8.8% improvement in the F1-score for all-class classification. Our goal is to replicate these experiments on an even smaller dataset of industrial images and additionally apply domain knowledge during image preprocessing to achieve higher accuracy and investigate GANs capability to generate a realistic representation of defects.

The rest of this paper is structured as follows: Section 2 focuses on challenges associated with dataset collection, data augmentation, and the possible application of GANs for enhancing CNN performance. Section 3 explores two main methodologies used in the experiment: MobileNet network and StyleGAN2-ADA. Section 4 discusses the selection and purpose of the initial dataset, the requirements of the experiment and show cases of the achieved results. The presentation of results includes both original and GAN generated images representing defect-free and defective samples. Section 5 focuses on the presentation and comparative analysis of the results obtained using initial and expanded datasets. The results of CNN performance are assessed using accuracy and F1 metrics, providing a comprehensive evaluation of the experimental outcomes. Section 6 summarizes the main findings discussed throughout the study.

2. Challenges in improving the performance of convolutional neural networks

In machine learning, the quality and characteristics of a dataset have a significant impact on the performance of artificial neural networks. Various issues with datasets can adversely affect the efficiency of these networks, as they require a considerable amount of data to be trained on to learn features and perform accurate classification or other predictions (Alzubaidi et al., 2023). An improvement in the architecture of the model is insufficient to enhance the performance of CNNs. Thus, it is essential to identify and solve various problems that arise in datasets. Data cleansing, processing, and labeling steps are a critical strategy aimed at improving the performance of CNNs.

2.1. Dataset collection challenges

The preparation of a dataset involves several stages: data collection, cleaning, processing, and labeling (Munappy et al., 2022). This process may present various challenges, such as class imbalance, data scarcity, feature selection, data leakage, label inaccuracy, and outliers.

- *Class imbalance* is an issue when the quantity of data in classes is distributed unevenly. Convolutional neural network models trained on such datasets perform poorly on classes with fewer samples because they become biased and adapt to classes with more data (Seliya et al., 2021).
- *Data scarcity* is an expensive and time-consuming process. In some domains, collecting useful data may be impossible for various reasons, such as privacy requirements for personal data, safety, and ethical issues (Adadi, 2021).

- *Feature selection* is an essential procedure done during model training, where relevant features are selected from the available data to enhance the model's effectiveness (Alzubaidi et al., 2021).
- *Data leakage* problems occur when information from testing or validation datasets indirectly influences the network training process.
- *Label inaccuracy* in the dataset can mislead the machine learning model during training and evaluation. Human errors in labeling or biased assessment can significantly degrade model performance (Bernhardt et al., 2021).
- *Outliers* are data samples that may significantly deviate from the general dataset, distorting the model training process, introducing noise, or affecting the model's ability to adapt to new data (Pang et al., 2020).

2.2. CNN classifiers

Convolutional neural networks (CNNs) are powerful models of deep learning that are widely used in computer vision classification tasks. By using convolutional layers, they are able to effectively distinguish and internalize key visual features that reveal local regularities and spatial relationships in images. CNN classification achieves highly remarkable results in various areas, including the scientific, industrial, and commercial domains.

The structure of a CNN classifier typically consists of several layers: convolutional, pooling, and fully connected layers.

The training process of a CNN can be generalized into the following stages (LeCun et al., 2015): initialization, forward propagation, loss computation, backward propagation, parameter update, and iterative training.

Ideally, the accuracy of CNN models during training should be evaluated using different metrics. The most common ones include precision, recall, and F1 scores. The F1 score represents the harmonic mean of the precision and recall metrics. This score is particularly suitable for evaluating models when facing class imbalance issues within a dataset.2.3. Data augmentation and Generative Adversarial Networks.

2.3. Data augmentation and Generative Adversarial Networks

The main challenges to improving the performance of the CNN are class imbalance and the lack of data (Motamed et al., 2021). In these cases, network performance can be improved by expanding the available data set. This can be achieved through two methods: image augmentations and generative adversarial networks.

Data augmentation is the most common method to expand datasets, involving a variety of geometric transformations in images (Alomar et al., 2023). Data augmentations are straightforward to implement and can be applied during the convolutional neural network training step. However, this method has several drawbacks: lack of new information, limited number of new images, and restriction of possible transformations.

The generation of realistic synthetic data is a complex task, as it requires learning to mimic the distribution of the original dataset. GANs overcome traditional augmentation limitations by producing new synthetic information. It enables diverse data generation beyond augmentation's capabilities, which is particularly useful for specific image generation where conventional methods fail.

The issue of data scarcity in training is relevant for both classifiers and GAN models. Transfer learning, which starts by training CNNs with pre-trained weights on large sets of data that are unrelated to the initial task, can be used to overcome this problem. This method reduces the computational resources required during training and significantly shortens the training process, allowing the network to achieve good results without the need for extensive data.

Another method to solve the problem of a small GAN dataset is to use data augmentations. Unfortunately, the conventional data augmentations used in CNN classifiers are unsuitable for training GAN networks. Such augmentations may be identified as essential image features and repeated in synthetic GAN images, introducing unwanted noise, and deviating from real original images. A solution to this problem is adaptive discriminator enhancement, applied in the StyleGAN2-ADA model, specifically designed to generate high-quality images using limited datasets. More details of this network are discussed further.

2.4. Generative adversarial networks architecture

A Generative Adversarial Network (GAN) is a fundamental framework in unsupervised deep learning (Goodfellow et al., 2020). It is comprised of two neural networks: a generator and a discriminator. This dual network structure engages in a competitive learning dynamic, where the generator learns to synthesize realistic data by assimilating information from real samples and previously generated outputs. Leveraging a latent noise vector, the generator creates images while the discriminator, employing a fully connected neural network design, tries to differentiate between authentic and synthetic data. The discriminator improves its ability to differentiate between real and synthetic data through iterative training. On the contrary, the generator employs an inverse architecture to reconstruct images, contributing to continually refining its output. GANs iteratively train until convergence, a critical point at which the discriminator accurately identifies synthetic data as indistinguishable from real samples. The GAN architecture is illustrated in Figure 1.

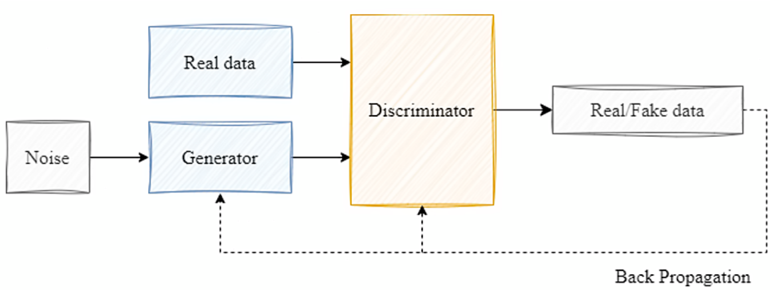


Figure 1. GAN architecture

3. Methodologies applied during the experiment

3.1. MobileNet network

To explore the effect of training images on the effectiveness of a CNN, it is imperative to perform fast and reliable training of the classifier network. Due to the experimental nature of this study, it is necessary to explore various combinations of images and network hyperparameters. To achieve this, the MobileNet model was selected as a binary classifier.

MobileNet is a CNN architecture specifically designed for application on mobile and embedded devices. This model maintains operational efficiency even when deployed on constrained resource platforms. MobileNet employs deeply separable convolutional blocks, each comprising a 3×3 depthwise convolutional layer followed by a 1×1 pointwise convolutional layer. This approach reduces computational costs by decreasing parameters and the number of operations while retaining the network's ability to capture spatial and channel-wise characteristics. Due to its straightforward architecture, the model achieves high accuracy and operational efficiency on resource-limited platforms, making it a commonly utilized choice in real-time applications on such devices.

3.2. StyleGAN2-ADA architecture

Numerous advancements in the architecture of GANs have been achieved through the development of the discriminator model. These modifications are based on the idea that an enhanced discriminator enables the generation of more realistic synthetic images. A shift in this approach occurred in 2018 when researchers at Nvidia introduced the StyleGAN generative network (Karras et al., 2019), a progressive extension of the GAN framework outlined by (Karras et al., 2017). The aim of this method is to enhance generator models for synthesizing high-resolution, high-quality images. Achieving this goal involves a gradual refinement of both the generator and the discriminator.

When StyleGAN is compared to previous GAN iterations, these distinctive features can be highlighted:

- Adoption of two new randomness sources (representation modulation channels and noise layers) in the generator's input, replacing the latent space's potential point.
- Augmentation of the generator with an Adaptive Instance Normalization layer housing a mapping network, empowering the model to control image characteristics.
- Introduction of stochastic variation via noise embedded in feature maps, allowing fine-grained style interpretation at the pixel level.

StyleGAN2, introduced by Karras et al. (2020b), is an advanced generative model derived from StyleGAN. StyleGAN2 integrates several new techniques and enhancements, allowing the generation of even higher-resolution images.

- A new convolutional architecture replaces StyleGAN's progressive growing method, employing a fixed-resolution generator network and multiscale discriminator, enabling higher-resolution image generation.
- Improved regularization methods, such as noise and path-length regularization, mitigate model overfitting, enhance image diversity, and reduce undesired artifacts.

StyleGAN2-ADA serves as an extension of the StyleGAN2 architecture, primarily to improve the quality, diversity, and stability of generated images when dealing with limited training data (Karras et al., 2020a). A key challenge when working with a small dataset is that the discriminator quickly adapts to the available images, rendering its response to the generator insignificant, leading to divergence in learning.

StyleGAN2-ADA technology applies a broad spectrum of augmentations to mitigate discriminator overfitting and ensure that distortions in images do not transfer to generated visuals. StyleGAN2-ADA exclusively involves augmented images when assessing the discriminator, employing a similar approach during the generator's training. This process eliminates distortions in the final images and maintains a correct distribution (Karras et al., 2020a).

The Adaptive Discriminator Augmentation (ADA) method uses 18 transformations grouped into six categories. These transformations are applied during generator training and must be differentiable. This adaptive process enhances the discriminator's capacity to comprehend complex decision boundaries, improving its ability to discern real from synthetic images. At the same time, it allows the generator to learn from a broader and higher-quality range of visuals.

4. Synthetic image generation with StyleGAN2-ADA

The research aims to examine the ability of the GAN to increase CNN classification accuracy. To address this issue, choosing an appropriate dataset with a limited amount of data was necessary to. Ideally, such a dataset should showcase data scarcity and class imbalance problems and consist of images that do not portray objects requiring specialized expertise.

In experimental research, a limited cable defect dataset was selected from the MVTEC AD dataset, which consists of 5069 high-resolution images and contains 15 object classes. The cable category was selected due to its diverse range of defects. The chosen cable subset contains 316 RGB (Red-Green-Blue) images, 92 of which have defects and 224 of which do not. The resolution of images is 1024×1024 pixels, and the dataset was split into training and testing subsets, with the latter containing 15% of the original images. It is essential to note the arrangement of cable strands in images. Each cable consists of 3 colored wires: blue, green and gray. The correct arrangement of these wires is crucial for non-defective cables, with a green strand positioned at the top, followed by blue and gray strands from left to right. Otherwise, a different arrangement of 3 wires is considered as a defect. Defects in such images can be characterized by various breaks in the cable and protruding or missing wires.

Note that the primary objective of this dataset is to detect cable images with defects. The assessment of CNN accuracy does not depend on the type of classification; it only needs to remain constant throughout all experiments. Therefore, a binary classification (defect and non-defect classes) was chosen for simplicity.

The StyleGAN2-ADA model was selected for its advanced architectural solutions, high image quality, control over generated images, and network stability during the training process for limited dataset expansion. The model was adapted for 512×512 image resolution, chosen for robust visual assessment, greater information transfer, and efficient processing within the Colab environment.

In addition, the experiment used a transfer learning technique due to its efficiency in reducing computational resources and achieving faster results. Pre-trained networks, particularly on 512×512 resolution animal faces (AFHQWild), were selected for their ability to discern useful image features, enhance model performance, and generalize better with new data.

The network was trained with a kimg parameter set to 300. Kimg parameter represents the number of thousands of real images presented to the discriminator during training. Using a single Nvidia Tesla V100 graphics processing unit (GPU), the training took approximately 5–8 hours, depending on factors such as dataset size, image resolution, GPU quantity, desired image quality, and hyperparameter selection. The results showed that transfer learning led to faster network convergence.

The images generated by the StyleGAN2-ADA network are shown in Figures 2 and 3. The illustrations show two sections: A presents original dataset images, while B showcases images generated by a generative adversarial network. Synthetic defect-free images were generated to address the issue of class imbalance. These images intentionally lack distortions and significant alterations. As shown, the synthetic defect-free class images closely resemble the original ones, avoiding any unwanted artifacts. Only minor changes in perspective are captured to simulate images taken from different angles. This outcome was the desired result when generating defect-free images.

The synthetic images with defects show the mixture of randomized features of defected wires from the training dataset. As illustrated, the synthetically generated images of the defect class appear realistic and meet the requirements of the experiment. However, due to the lack of training images, there is also a lack of variance in their defect features. Therefore, to introduce more variance, image processing with external software was attempted.

a)
b)

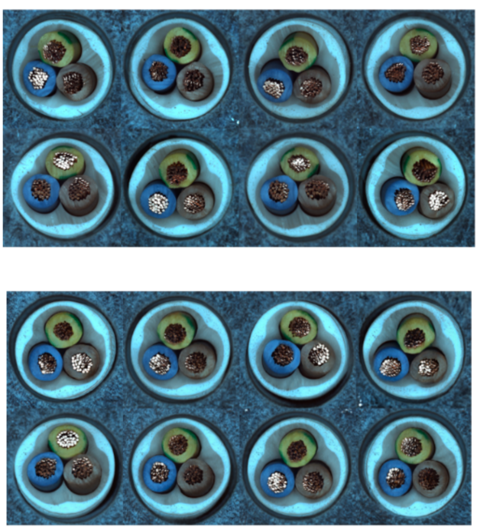


Figure 2. StyleGAN2-ADA generated images for non-defect class: a – original images; b – synthetic images

a)
b)

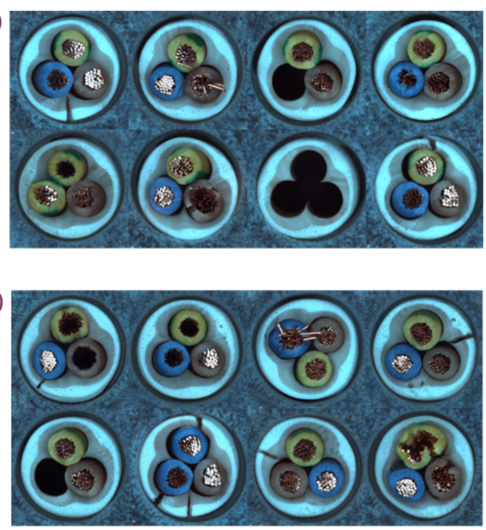


Figure 3. StyleGAN2-ADA generated images for defect class: a – original images; b – synthetic images

a)

b)

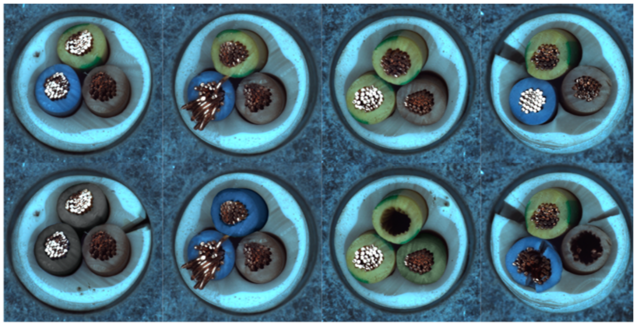
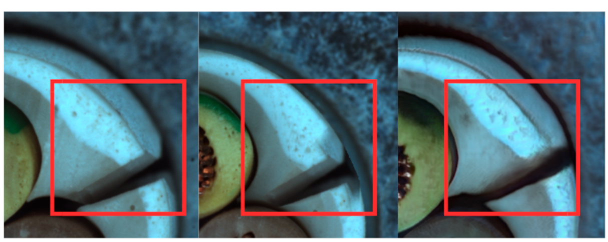


Figure 4. Images altered with GIMP software: a – original images; b – altered images

Fifty images were selected from the original defect class dataset and edited using GIMP software to increase the diversity of defects in cables (refer to Figure 4). The processed images introduced various new combinations of defects, such as missing wires, swapped cable positions, and cables of matching color, which were absent in the original dataset. The second experiment was then performed by expanding the original dataset with edited images.

When comparing the results obtained from the second experiment, it was found that the use of processed images resulted in generating images with a more diverse range of defect combinations than in the first experiment. Furthermore, upon comparison of the edited images with those generated during the second experiment, it is evident that the defects produced by the GAN appear more realistic. This indicates that the network can accurately interpret the modifications made to edited images and replicate them more realistically (see Figure 5). For instance, upon closer inspection of edited pictures, one can observe the remaining original color of the cable around the edges instead of the modified color, or there may be other visual discrepancies in the added defects. This makes the edited images appear less realistic than the generated ones.

Also, when generating images with defects, there are instances where the newly generated images portray a defect-free appearance due to the blending of styles. Some cases can occur where the image is blended or unwanted artifacts are captured in the background (see Figure 6). Therefore, the generated defective class images required further visual evaluation and removal before being compiled into the training set for the classifier.



a) b) c)

Figure 5. Comparison of cable defect generation: a – original defect; b – altered defect; c – GAN generated defect

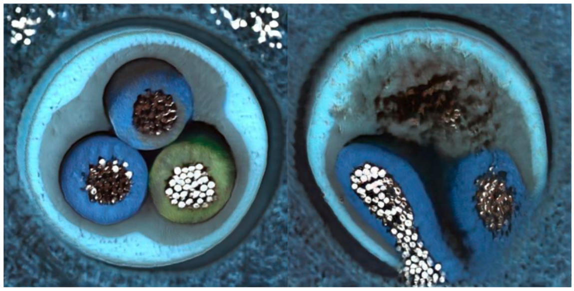


Figure 6. GAN generated images with artifacts

5. The results of the convolutional neural network performance enhancement

The aim of this study is to investigate the impact of images used in CNN training. We will compare the network classification performance when trained with the original dataset and the dataset expanded with GAN to achieve this.

The experiment uses a binary classifier consisting of a MobileNetV2 deep learning model that was trained on a large-scale dataset (ImageNet) available in the Keras library. These pre-trained models serve as robust transfer learning tools, leveraging pre-trained network representations to tackle novel tasks within specific datasets. Adapted for cable defect classification, this model involves the removal of specific pre-trained top layers, replaced by output layers responsible for the final stage of image classification. The new architecture of the model's top section includes the following.

- Global Average Pooling Layer
- Dropout Layer
- Fully connected layer with kernel and bias L2 regularizers
- Batch Normalization
- Swish activation layer
- Fully connected Softmax layer with kernel and bias L2 regularizers.

5.1. The results of the experiment using an initial dataset

The initial training dataset comprises only real images: 190 in the defect-free class and 78 in the defective class. Due to the small size of the dataset, it remains undivided into training and validation sets, because each image is crucial for training. The goal is to evaluate how the model learns using both real and synthetic images, so testing is performed after each epoch using the testing data for validation. 15% of the real images, 14 defective and 34 non-defective, were reserved for testing.

To accurately assess the network, a larger testing data set was needed. Therefore, additional image augmentations were carefully selected to expand the testing dataset to 226 images. This dataset was used for both experiments (using the initial dataset and using the expanded dataset).

The network was trained using the Adam optimizer with a fixed learning rate of 0,00001 and the Binary Focal Crossentropy loss function to maintain a consistent training duration of 40 epochs across all experiments. Using the testing dataset, the network performance was assessed, and the model was evaluated after each training epoch. During the experiment, the model underwent an average of five training cycles without structural modifications. Each training iteration showed notable fluctuations in metrics and convergence rates. Consequently, the averaged metric values were calculated and presented along with a 95% confidence interval to provide a comprehensive overview of the findings. Figure 7 displays the resulting metrics, providing a clear and informative visualization.

The visual representation of loss values on the testing dataset indicates that the model starts to overfit rapidly. In particular, from the 18th training epoch onward, there is a discernible increase in the network error values. Standard practice involves employing separate validation and testing datasets during model training, ones that remain untouched throughout the learning process. In practice, when selecting the best-performing model, emphasis is often placed on the validation error or, in this experimental context, the testing error. To identify the best performing model, it is recommended that early stopping techniques be implemented.

Based on the graphical results of the model accuracy, it is clear that the network achieves a maximum accuracy of only 81.81% during the best training cycle with the initial dataset. In this scenario, the model overfits and adapts to the training set images, leading to sub-optimal performance with the testing images. The primary cause appears to be the insufficient data in the training set. The limited diversity of cable defects in the images is also considered to contribute to this issue.

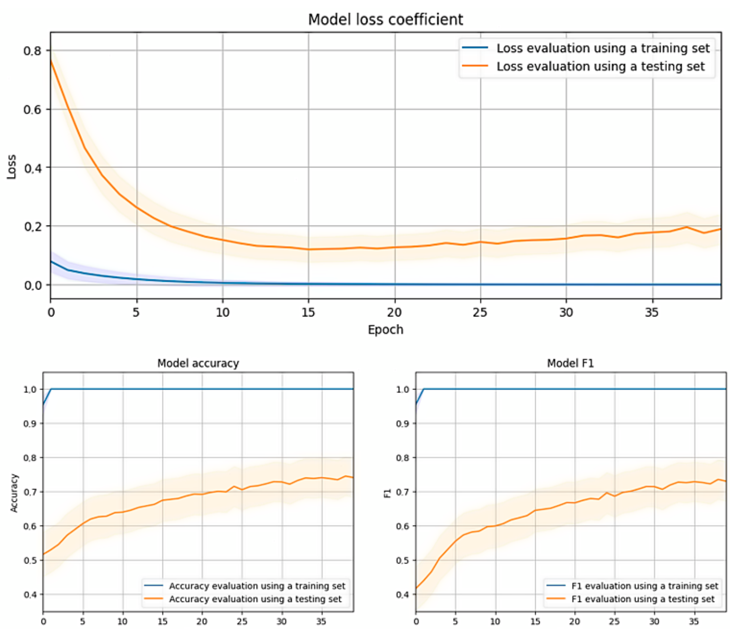


Figure 7. MobileNetV2 model trained using the initial dataset (real images), evaluation metrics

As a result of overfitting during training, the F1 score appears to approach 1 at a slow rate. A lower F1 score suggests that the classifier is less effective at classifying cable images and makes a significant number of errors in the process. The maximum F1 score achieved (during the best training cycle) is 81.59%.

5.2. The results of the experiment using an expanded dataset

The expanded training dataset consists of original real images and synthetic images generated by a StyleGAN2-ADA network. Prior to expanding the original training dataset, synthetically generated images underwent visual assessment, and any inappropriate data was removed. This ensured the accuracy of the data, preventing the presence of unrelated images in the cable defects class. Also, the dataset was expanded with images that provided new information to the network during the training process. To achieve this, a classification of the generated synthetic images was performed using a network previously trained with real data. In this way, synthetic images that were misclassified and had a classification probability below the 0.9 threshold were selected. Approximately 10% of the synthetic images were rejected during this step. By adding additional generated images, the training dataset of the classifier was increased approximately 28 times, from 268 images to 7 436. Each class was increased to 3 584 images, adding a total of 7 168 synthetic images, to maintain class data balance. The same test dataset as in the first experiment, which contained 226 images, was used to assess the network accurately.

The network was trained using the Adam optimizer with a learning rate of 0.00001 and the Binary Focal Crossentropy loss function for 40 epochs. This approach aimed to examine how classifier results change when the initial dataset is increased with synthetic images. The results of the metrics obtained are presented with a 95% confidence interval in Figure 8.

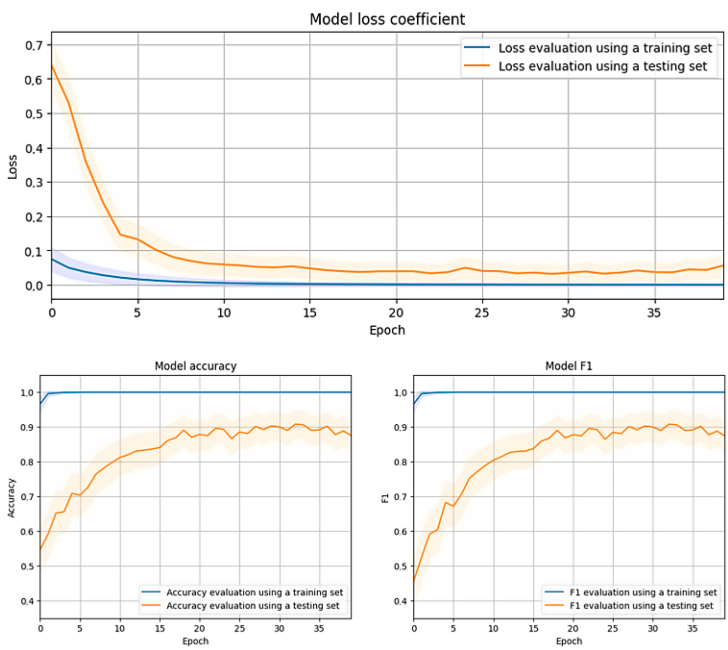


Figure 8. MobileNetV2 model trained using the expanded dataset (real + synthetic images), evaluation metric

Figure 8 shows that the model did not overfit as quickly as in the initial experiment. The error value of the network trained on the initial dataset started to increase noticeably from the 18th epoch. With the addition of synthetic images, this value remained relatively stable until the last 40th epoch and then slowly decreased. Early stopping is not necessary in this case, as the probability of selecting a model from an inappropriate epoch decreases.

The comparison of model accuracy results indicates that the convolutional neural network (CNN) accuracy reaches a maximum of 95.05% during the best training cycle when the original dataset is expanded with synthetic images generated by the generative adversarial network. It is considered that the accuracy of the model with an expanded dataset achieved significantly better results, not only due to the increased amount of data but also due to the high quality of synthetic images.

Another important metric to consider when evaluating CNN models is the F1 score. The higher the value of this score, the more efficient the classifier performance. As shown in the F1 score figure, its value gradually approaches 1, indicating that the classifier trained on the expanded dataset can more accurately distinguish and classify images with and without defects than the one trained only with original data. The maximum F1 score achieved is 94.99% (during the best training cycle).

5.3. The comparison of the experimental results

When developing a CNN classifier, it is beneficial to explore how it makes classification decisions during testing. The Gradient-weighted Class Activation Maps (Grad-CAM) can be used as an informative tool for this purpose. These maps help to evaluate the parts of an image that the model utilizes for its predictions. They are constructed using gradient information that is fed into the final convolutional block of the CNN classifier. Additionally, these maps assist in generating bounding boxes, which can be interpreted as the location of the object defect. Figure 9 showcases a map of images generated using the initial (unexpanded) training set.

The illustration shows that the model, when trained on the initial dataset, highlights image portions not indicative of defects in misclassified images. This includes using inconsequential background areas and similar features for predictions.

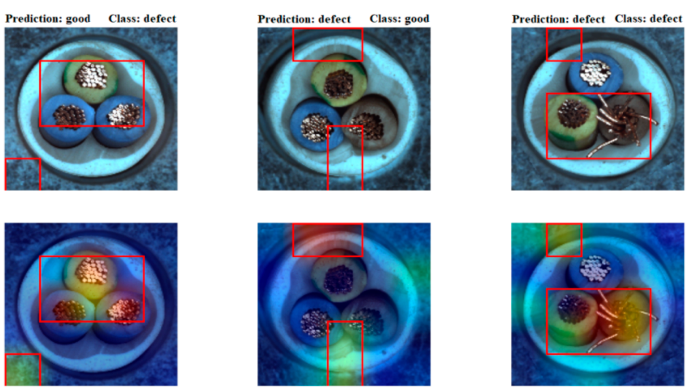


Figure 9. Grad-CAM results of MobileNet V2 network trained on initial dataset

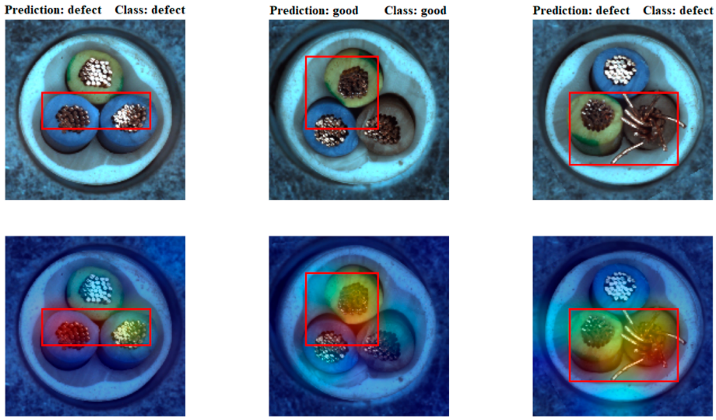


Figure 10. Grad-CAM results of MobileNet V2 network trained on expanded dataset

Figure 10 shows the gradient-weighted class activation maps for images generated using the expanded dataset.

It can be seen that the model trained on the expanded dataset captures areas of defects in the images and makes predictions based on this information. A clear improvement in classification is visible compared to the model trained on the original dataset, i.e., the same images are classified correctly, and cable defect areas are accurately identified.

Table 1 compares the performance measurements of the MobileNet V2 classifier using the original and expanded datasets. The measured values were obtained during the best training cycles in the networks.

Table 1. Comparison of MobileNet V2 classifier results

	F1, %	Accuracy, %
CNN (original dataset)	81.59	81.88
CNN (expanded dataset)	94.99	95.05

Summing up, based on the results obtained with the expanded dataset, the application of GANs improves the performance of CNNs. This is evidenced by a decrease in overfitting, an improvement in network generalization, a reduction in the problem of class imbalance, and an increase in the accuracy and stability of network classification. It is worth noting that using GANs to generate synthetic images can save time and financial resources when creating training datasets.

6. Conclusions

The research addresses the challenge of limited training datasets for Convolutional Neural Networks (CNNs) training in the manufacturing industry, particularly in defect detection. It proposes using Generative Adversarial Networks (GANs) to generate synthetic images that resemble real manufacturing parts, focusing on defect representation. The objectives include testing the ability of GANs to produce realistic images, evaluating advanced GAN architectures with a weight transfer technique, and assessing the impact of synthetic images on CNN training. This study expands on previous work by applying GANs to industrial images, using an insufficient dataset, and exploiting domain knowledge to improve accuracy.

The experiments on cable classification demonstrated that the efficiency of CNNs was significantly improved by adding synthetic images generated by GANs. The results were analyzed using metrics with a 95% confidence interval. The results could be evaluated more quickly by using the MobileNet V2 model, which is known for its faster performance compared to standard CNN models. The network, trained on a small dataset, quickly adapted to the training data, and struggled with accuracy, achieving 81.81% accuracy and an F1 score of 81.59%, particularly on unseen test images. However, the addition of synthetic images to the dataset, which increased its size by approximately 28 times, significantly improved the evaluation metrics. In this scenario, overfitting during training was not observed, and the classifier achieved a maximum accuracy of 95.05% and an F1 score of 94.99%. These results demonstrate improved generalization.

Future work could explore the scalability and adaptability of GAN-generated synthetic images in different manufacturing domains and datasets to assess their broader applicability and effectiveness. In addition, alternative GAN architectures and training methods could be investigated to refine the generation of synthetic images tailored to specific manufacturing contexts, potentially improving the performance and efficiency of the dataset augmentation process.

Author contributions

All authors contributed to this work equally.

Disclosure statement

The authors declare that they have no competing financial, professional, or personal interests that could influence the conduct or reporting of this research.

References

Adadi, A. (2021). A survey on data-efficient algorithms in big data era. *Journal of Big Data*, *8*(1), Article 24. https://doi.org/10.1186/s40537-021-00419-9

Alomar, K., Aysel, H. I., & Cai, X. (2023). Data augmentation in classification and segmentation: A survey and new strategies. *Journal of Imaging*, *9*(2), Article 46. https://doi.org/10.3390/jimaging9020046

Alzubaidi, L., Bai, J., Al-Sabaawi, A., Santamaría, J., Albahri, A. S., Al-dabbagh, B. S. N., Fadhel, M. A., Manoufali, M., Zhang, J., Al-Timemy, A. H., Duan, Y., Abdullah, A., Farhan, L., Lu, Y., Gupta, A., Albu, F., Abbosh, A., & Gu, Y. (2023). A survey on deep learning tools dealing with data scarcity: Definitions, challenges, solutions, tips, and applications. *Journal of Big Data*, *10*(1), Article 46. https://doi.org/10.1186/s40537-023-00727-2

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1), Article 53. https://doi.org/10.1186/s40537-021-00444-8

Bernhardt, M., Castro, D. C., Tanno, R., Schwaighofer, A., Tezcan, K. C., Monteiro, M., Bannur, S., Lungren, M. P., Nori, A. V, Glocker, B., Alvarez-Valle, J., & Oktay, O. (2021). *Active label cleaning: Improving dataset quality under resource constraints.* ArXiv. https://doi.org/10.48550/arXiv.2109.00574

Chan, W. H., Fung, B. S. B., Tsang, D. H. K., & Lo, I. M. C. (2023). A freshwater algae classification system based on machine learning with StyleGAN2-ADA augmentation for limited and imbalanced datasets. *Water Research*, *243*, Article 120409. https://doi.org/10.1016/j.watres.2023.120409

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139–144. https://doi.org/10.1145/3422622

Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). *Progressive growing of gans for improved quality, stability, and variation.* ArXiv. https://doi.org/10.48550/arXiv.1710.10196

Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., & Aila, T. (2020a). Training generative adversarial networks with limited data. *Advances in Neural Information Processing Systems*, *33*, 12104–12114.

Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4401–4410). IEEE. https://doi.org/10.1109/CVPR.2019.00453

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020b). Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 8110–8119). IEEE. https://doi.org/10.1109/CVPR42600.2020.00813

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Motamed, S., Rogalla, P., & Khalvati, F. (2021). Data augmentation using Generative Adversarial Networks (GANs) for GAN-based detection of Pneumonia and COVID-19 in chest X-ray images. *Informatics in Medicine Unlocked*, *27*, Article 100779. https://doi.org/10.1016/j.imu.2021.100779

Munappy, A. R., Bosch, J., Olsson, H. H., Arpteg, A., & Brinne, B. (2022). Data management for production quality deep learning models: Challenges and solutions. *Journal of Systems and Software*, *191*, Article 111359. https://doi.org/10.1016/j.jss.2022.111359

Pang, G., Shen, C., Cao, L., & van den Hengel, A. (2020). *Deep learning for anomaly detection: A review*. ArXiv. https://doi.org/10.48550/arXiv.2007.02500

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, *2*(3), Article 160. https://doi.org/10.1007/s42979-021-00592-x

Seliya, N., Abdollah Zadeh, A., & Khoshgoftaar, T. M. (2021). A literature review on one-class classification and its potential applications in big data. *Journal of Big Data*, *8*(1), Article 122. https://doi.org/10.1186/s40537-021-00514-x

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), Article 60. https://doi.org/10.1186/s40537-019-0197-0

APPENDIX

Notations

Abbreviations

CNN – Convolutional Neural Network.

GAN – Generative Adversarial Network.

ADA – Adaptive Discriminator Augmentation.

GPU – Graphics Processing Unit.