UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

JOCELEIDE DALLA COSTA MUMBELLI

APLICAÇÃO DE REDES ADVERSARIAIS GENERATIVAS PARA MELHORAR A INSPEÇÃO AUTOMÁTICA EM LINHAS DE MONTAGEM AUTOMOTIVA

DISSERTAÇÃO

PATO BRANCO
2022

JOCELEIDE DALLA COSTA MUMBELLI

APLICAÇÃO DE REDES ADVERSARIAIS GENERATIVAS PARA MELHORAR A INSPEÇÃO AUTOMÁTICA EM LINHAS DE MONTAGEM AUTOMOTIVA

AN APPLICATION OF GENERATIVE ADVERSARIAL NETWORKS TO IMPROVE AUTOMATIC INSPECTION IN AUTOMOTIVE ASSEMBLY LINES

Dissertação apresentado(a) como requisito parcial à obtenção do título de Mestre em Engenharia Elétrica, do Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Marcelo Teixeira
Coorientador: Prof. Dr. Dalcimar Casanova

PATO BRANCO

2022

Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Pato Branco

JOCELEIDE DALLA COSTA MUMBELLI

AN APPLICATION OF GENERATIVE ADVERSARIAL NETWORKS TO IMPROVE AUTOMATIC INSPECTION IN AUTOMOTIVE ASSEMBLY LINES

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestra Em Engenharia Elétrica da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Sistemas E Processamento De Energia.

Data de aprovação: 14 de Março de 2022

Prof Marcelo Teixeira, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Dalcimar Casanova, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Giovanni Alfredo Guarneri, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Yuri Kaszubowski Lopes, Doutorado - Fundação Universidade do Estado de Santa Catarina (Udesc)

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 14/03/2022.

À meu namorado, amigos e família, pelo apoio

de sempre.

ACKNOWLEDGEMENTS

Este trabalho não poderia ser terminado sem a ajuda de diversas pessoas e/ou instituições às quais presto minha homenagem. Certamente esses parágrafos não irão atender a todas as pessoas que fizeram parte dessa importante fase de minha vida. Portanto, peço desculpas àquelas que não estão presentes entre estas palavras, mas elas podem estar certas que fazem parte do meu pensamento e de minha gratidão.

Inicialmente agradeço a Deus, por sempre me guiar e pela proteção durante todo esse período de mestrado e pandemia.

A minha família, pelo carinho, incentivo e total apoio em todos os momentos da minha vida.

A meu namorado, pela compreensão e apoio nos momentos difíceis.

Ao meu orientador e coorientador, que me mostraram os caminhos a serem seguidos.

A todos os professores do departamento, banca avaliadora, e demais que ajudaram de forma direta e indireta na conclusão deste trabalho.

O presente trabalho foi realizado com apoio financeiro do Edital 21/2019 do Programa de Bolsas Fundação Araucária e Renault do Brasil, chamada pública 6/2019. Em parceria com a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, aos quais deixo aqui meu agradecimento.

Pois eu, o senhor teu Deus, eu te seguro pela mão e te digo: "Nada temas, eu venho em teu auxílio." Isaías 41:13

ABSTRACT

MUMBELLI, Joceleide Dalla Costa. **Aplicação de Redes Adversariais Generativas para melhorar a inspeção automática em linhas de montagem automotiva**. 2022. 42 p. Dissertation (Master's Degree in Electrical Engineering) – Universidade Tecnológica Federal do Paraná. Pato Branco, 2022.

In manufacturing systems, quality inspection is a critical issue. This can be performed by humans, or by means of *Computer Vision Systems* (CVS), which are trained using representative sets of images, modeling classes of defects that may possibly occur. In practice, the construction of such datasets strongly limits the use of most CVS methods, as the variety of defects is of combinatorial nature. Alternatively, instead of recognizing defects, a system can be trained to detect non-defective cases, becoming appropriate for some application profiles. In flexible automotive manufacturing, for example, parts are assembled within a reduced set of correct combinations, while the number of possible incorrect assembling is enormous. In this paper, we show how a CVS can be extended with a *Deep Learning*-based approach that exploits a *Generative Adversarial Network* (GAN) to detect non-defective production eliminating the need for constructing expensive defect image datasets. The proposal is tested over the assembly line of Renault, in Brazil. Results show that our approach has better accuracy in inspection, compared with the currently used CVS. We also show that the same method can be used in different components inspection, without any modification.

Keywords: Automatic inspection. Deep learning. Generative Adversarial Networks. Automotive manufacturing.

RESUMO

MUMBELLI, Joceleide Dalla Costa. **Aplicação de Redes Adversariais Generativas para melhorar a inspeção automática em linhas de montagem automotiva**. 2022. 42 f. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Tecnológica Federal do Paraná. Pato Branco, 2022.

Em sistemas de manufatura, a inspeção de qualidade é uma questão crítica. Isso pode ser feito por humanos, ou por meio de *Sistemas de Visão Computacional* (CVS), que são treinados usando conjuntos representativos de imagens, modelando classes de defeitos que eventualmente possam ocorrer. Na prática, a construção de tais conjuntos de dados limita fortemente o uso da maioria dos métodos CVS, pois a variedade de defeitos é de natureza combinatória. Alternativamente, ao invés de reconhecer defeitos, um sistema pode ser treinado para detectar casos não defeituosos, tornando-se apropriado para alguns perfis de aplicação. Na fabricação automotiva flexível, por exemplo, as peças são montadas dentro de um conjunto reduzido de combinações corretas, enquanto o número de possíveis montagens incorretas é enorme. Neste artigo, mostramos como um CVS pode ser estendido com uma abordagem baseada em *Deep Learning* que explora uma *Generative Adversarial Network* (GAN) para detectar produção não defeituosa, eliminando a necessidade de construir conjuntos de dados de imagem de defeito. A proposta é testada na linha de montagem da Renault, no Brasil. Os resultados mostram que nossa abordagem melhor precisão na inspeção, em comparação com o CVS atualmente usado. Mostramos também que o mesmo método pode ser utilizado em diferentes inspeções de componentes, sem nenhuma modificação.

Palavras-chave: Inspeção automática. Aprendizagem profunda. Redes Adversárias Geradoras. Fabricação automotiva .

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

As *Manufacturing Systems* (MSs) (KIRAN, 2019; YIN *et al.*, 2018; ESMAEILIAN *et al.*, 2016) progress towards digitalization, they require more flexibility to process multiple types of products over the same plant. An example emerges from the automotive domain, where different models of vehicles are manufactured, each using its own set of components. Environments like this create new challenges and requirements for classic control and automation practices, such as the need for visual inspections to be integrated at many steps of manufacturing, over many types of assembly sets (CAPUTO *et al.*, 2015). Inspections aim to guarantee that manufactured products are zero-defect and meet certain security and quality requirements.

In practice, several types of defects can be inspected, e.g., component placement, position, soldering, type consistency, etc. The inspection task is in general conducted by a human agent, so that faulty components are sent back to rework, while the others progress through the process. As this is error-prone and unproductive, there are many attempts to replace human inspection by automatic decisions taken by CVS (BOUDELLA *et al.*, 2018). A CVS includes a capture camera and embedded computing that allow to infer about visual details into an image.

Unfortunately, the accuracy of a CVS still fails to match visual inspections for most applications, despite so many efforts in the literature (FENG *et al.*, 2019). The main barrier is that the result of a CVS inspection varies according to the problem nature, suffering from environmental influences, such as noise, lighting, and angle of image capturing (QUINTANA *et al.*, 2015). In the classic view of CVS inspection, each type of defect requires a specialized computational method to be detected, implying substantial efforts in research, engineering, and implementation.

Recent novelties in DL have fulfilled many gaps of CVS, making them more applicable and accurate to industry (LI *et al.*, 2018; LUCKOW *et al.*, 2016; MAZZETTO *et al.*, 2019), with promising impact on the automotive domain (HUVAL *et al.*, 2015; POMERLEAU, 1991; MAZZETTO *et al.*, 2020). Yet, classical DL approaches (i.e., based on supervised training) still rely on the dataset quality and representativeness. In fact, most DL methods are trained over datasets that include a significant number of images, that represent each type/class of defect that may possibly occur in manufacturing. When therefore the diversity of defects is wide and combinatorial, the construction of such a dataset imposes a strong limitation for the use of most DL CVS in a supervised way, making them weakly generalizable to industry.

Alternatively, instead of training the DL method on a supervised way, with a reasonable set of image samples for each possible defect, a system could be trained to detect only non-defective cases, so that any situation otherwise is considered a defect. This becomes appropriate for applications where the number of possible defects is (much) greater than the non defective cases. In flexible automotive manufacturing, for example, parts are presumably assembled within a reduced set of correct combinations, while the number of possible incorrect assembling is enormous. Thus, training a CVS to recognize incorrect cases may be an exhaustive, sometimes unfeasible, task that delays the production flow and allows errors to survive post production.

Computationally, this approach can be seen as an anomaly detection task, and the *Generative Adversarial Networks* (GANs) are nowadays the most promising unsupervised learning method to address it (MATTIA *et al.*, 2019). GANs dispense the construction of image datasets to represent defects, as their training requires only correct images, which are in general easier to define and much more available in factory floors. Therefore, GANs are expected to generalize better for the automatic inspection of several types of defects, evidencing practical appeal to industry. GANs have been applied to agriculture (WANG *et al.*, 2019), optimization problems (MUKHERJEE *et al.*, 2021), general manufacturing (SAIZ *et al.*, 2021; KUSIAK, 2019), machining (DESHPANDE *et al.*, 2020a), surface inspection (PERES *et al.*, 2021), among others. Yet, applications for visual inspection in the automotive assembly lines are still incipient.

This paper extends a CVS currently used by Renault, in Brazil. This system faces limitations related to the number of points that need to be verified, speed required for image processing, and verification angle. Those difficulties, in conjunction with the classical CVS embedded on the capturing system, reduce the quality of inspection, affecting the advantages of the CVS. The approach proposed here is a DL-based inspection system that exploits a GAN to improve the automatic inspection tests along a flexible automotive assembly line at Renault. By dismissing defective images examples in the training phase, our approach recognizes more defective types with a single model, evidencing better performance in comparison with the current CVS system.

It is further shown that the same method can be generalized and applied to several types of components without modifications. Tests suggest an increase of 2,13% in the mean accuracy for detecting defects that, comparably, are also detected by the current CVS, and identification of countless other defects that currently have not yet being captured.

Structurally, the manuscript is organized as follows: a literature is presented in Chapter 2;

Chapter 3 presents the related concepts; Chapter 4 introduces the main results; while conclusions and perspectives are discussed in Chapter 5.

2 RELATED WORK

In industry, automatic inspection is usually made by humans. This has been gradually replaced by robots that embed CVSs to perceive the environment, collect, and process information, taking decisions accordingly. A sketch of this relation, with some options available in the literature, is shown in Figure 1.

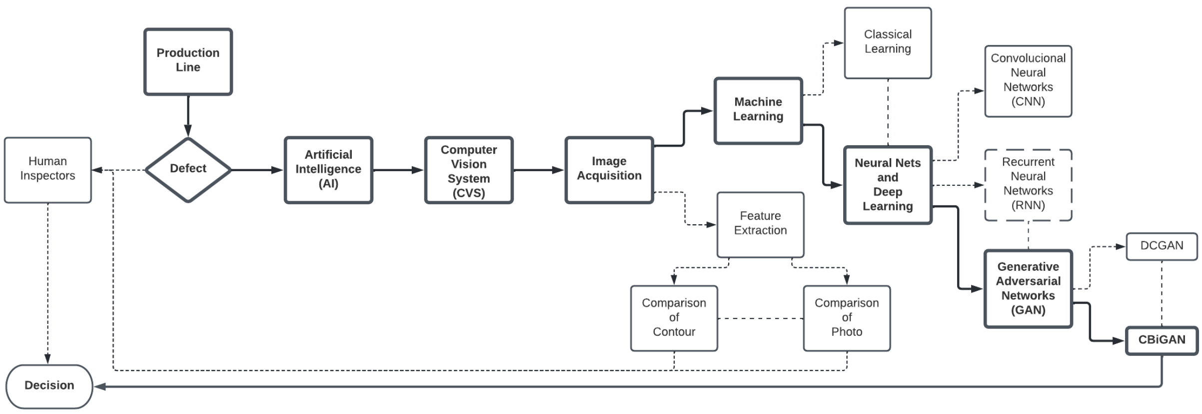

Figure 1 – Overview of the possible manufacturing defects detection pathways based on the two main directions: human-centered and automatic approaches. In bold, we identify the path pursued in this research, while the dashed lines highlight options in the literature that are not exploited here.

A CVS takes pictures of a certain region of interest in objects that include points to be verified. These images may be processed with or without automatic learning. Feature extraction via contour detection (MARTIN *et al.*, 2004), for example, checks the edges of an image to determine classes within it belongs. Its main gap is to be dependent extensively on human effort to complete each test.

Recent advances in and DL have shown potential to eliminate human dependency to some extent, and have been the base for inspecting, for instance, aircraft fuselage (MALEKZADEH *et al.*, 2017), cement cracks (CHA *et al.*, 2017), conformance of automotive assembling (MAZZETTO *et al.*, 2020), besides having been served well for the metallurgical industry (MERY *et al.*, 2005; MERY; FILBERT, 2002). Despite their usefulness, DL approaches require image datasets to be constructed in such a way that every possible defect to be detected is represented by a significant number of images, so that its identification can be learned after training. Computationally, this approach is called supervised learning.

This is simple to be done when defects have minor diversity, i.e., when they belong to a limited set of types, facilitating them to be mapped by images. In some applications, e.g.

automotive manufacturing, defects have high diversity, challenging the construction of a dataset that reasonably reflects this.

One option is to complement the dataset with images generated artificially to represent a defect. DL itself can be used to generate images that improve the defects diversity. *Recurrent Neural Networks* (RNNs) (OORD *et al.*, 2016), *Restricted Boltzmann Machines* (RBMs) (HINTON *et al.*, 2006; HUANG *et al.*, 2017), and *Variational Autoencoders* (VAEs) (KINGMA; WELLING, 2013), are examples of such techniques. They work well for low resolution images, but may be unsuitable when defects are sensitive and require more precise images.

In this case, an alternative is to assume that, instead of recording every possible manufacturing defect, one can record and learn/model only over correct cases, so that any other possible variation is a defect. Computationally this defect is called an anomaly/outlier on the dataset. More specifically, when training data is not polluted by outliers, and we are interested in detecting whether a new observation is an outlier, we call this task as novelty detection. Novelty detection is a semi-supervised task, because it is known that the training step includes only normal (i.e. non-defective) samples. A DL method that implement this idea of novelty detection as semi-supervised task is the GANs. This method simultaneously trains a generator, to produce fake images, and a discriminator, to distinguish between real and fake images. Images generated by GANs have in general good quality and are qualified to enrich image datasets for classification tasks subject to high diversity of defects.

Variations of GANs include: *Deep Convolutional Generative Adversarial Networks* (DCGANs), which uses convolution to ensure stability and convergence during training (RADFORD *et al.*, 2015); *Wasserstein Generative Adversarial Networks* (WGANs) (ARJOVSKY *et al.*, 2017), which uses the Wasserstein distance to effectively solve the problem of gradient disappearance during training; *Cycle Generative Adversarial Networks* (CycleGANs) (ZHU *et al.*, 2017), which allows images from two domains to be generated without paired images, among many others (MATTIA *et al.*, 2019).

In industry, GANs have been tested in steel defect classification (DESHPANDE *et al.*, 2020b), palm print recognition (WANG *et al.*, 2018), people identification (ZHENG *et al.*, 2017), vehicle license plate recognition (WANG *et al.*, 2017), medical image synthesis (FRID-ADAR *et al.*, 2018; SHIN *et al.*, 2018), texturing industry (CARRARA *et al.*, 2021), among others. However, applications in the automotive domain are still emerging. Examples can be found in autonomous vehicles to model longitudinal errors of sensors (ARNELID, 2018), or to generate sensor errors as an attempt to test the safety of advanced driver assistance systems (ZEC *et al.*, 2019). In this paper, GANs are exploited for: (i) increasing the dataset that maps critical manufacturing defects in vehicles production; and (ii) learning how to differentiate defective and non-defective images without having images for all possible defects that may occur.

3 COMPUTER VISION SYSTEMS TO AUTOMATIC INSPECTION

Parallel and multidisciplinary advances in engineering have allowed the industry to produce flexibly, on demand, quickly, and at reduced costs (SALDIVAR *et al.*, 2015). Attempts to adapt the factory floor to these concepts are non-trivial and require the integration of multiple and complex technologies. Among them, the cognitive approaches, which aim to enrich automatic production systems with abilities hitherto only perceived by humans, such as touch, hearing, and vision. The latter is of particular interest in this article.

3.1 COMPUTER VISION CHALLENGES

Specially in manufacturing systems, it is usual to face multiple types of components and to handle dynamic assemblies that constantly move throughout the factory. They need to be inspected to avoid defects and to meet the expected security and quality requirements.

However, automatic inspection results are usually not-trivial to be obtained, mainly because objects used in manufacturing may have a quite similar visual appearance, and defects may be caused by minor physical details. Besides, it is possible that the same object appears differently based on the light, observation angle, distortion and occlusion of the environment, etc. (GAD *et al.*, 2018).

This diverse mix of features makes visual inspection difficult to be fully automated, and it remains a human-centered task. As such, it is monotonous, not rarely impacting on the human health, causing stress, burnout, and also ergonomic problems (GENAIDY *et al.*, 1993), besides tending to be imprecise. The essence of a CVS is to reproduce the visual ability of humans, aiming at reducing dependency and improving precision (SZELISKI, 2010).

Usually, a CVS includes the orderly steps of acquisition, pre-processing, feature extraction, segmentation, and classification. The first two steps extract and process the images obtained from the manufacturing environment, making them suitable for identifying features, after mathematical processing, and isolating segments containing such features. Finally, classifier algorithms identify patterns in those segments, usually by evaluating similarities between regions of interest and predefined templates, comparing the values of *pixels* within an acceptance threshold.

Commercially, image acquisition systems are composed of modern equipment that seek to improve environmental interference, such as lighting and sharpness. An example is shown in Figure 2, which is a hardware from Keyence[1]. It has a quadrant lighting system, called *Multi-spectrum*, with LEDs of eight different colors and a monochrome camera. In theory, system like this should solve most of the problems related to environment instability. Yet, they evidence several limitations, such as the short distance required between the the camera and the object under observation. In practice, this simple detail makes unfeasible the application of image treatment approaches that require more general images of the problem, as for example images with several checkpoints within.

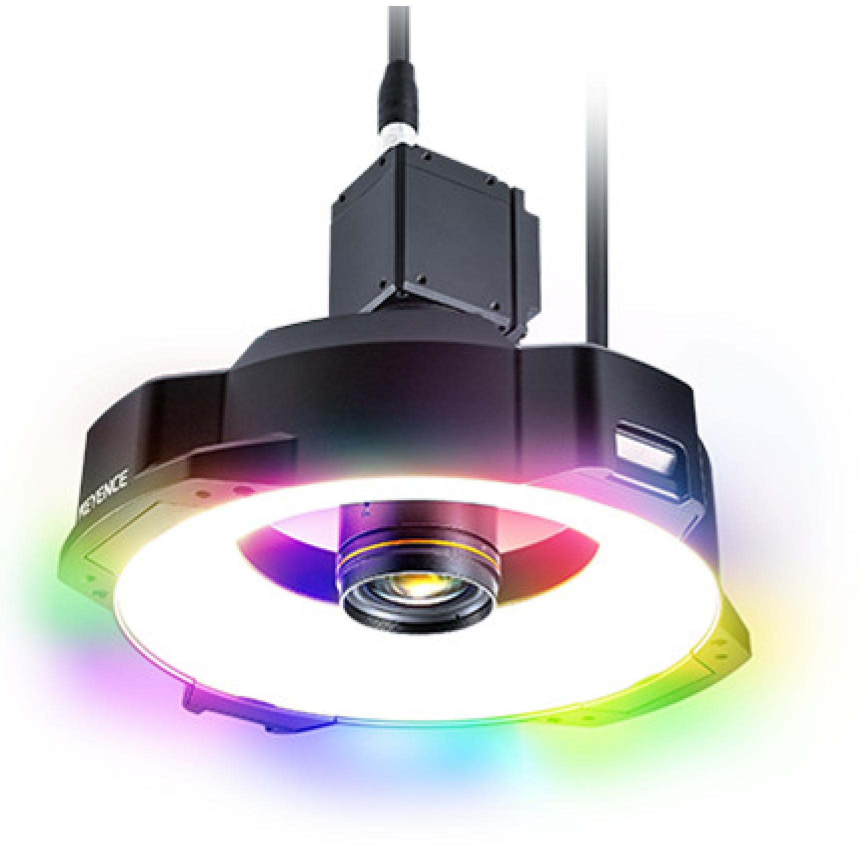

Figure 2 – Example of industrial device for CVS.

Commercially, companies that work with this type of solution usually apply photo-comparison and contour-detection methods (KEYENCE, 2020; WENGLOR, 2020; COGNEX, 2020) to do automatic inspection.

[1] Keyence manufactures sensors, machine vision systems, measuring instruments, barcode scanners, and other factory automation products (KEYENCE, 2020).

3.2 AUTOMATIC INSPECTION CHECK ON AUTOMOTIVE ASSEMBLY LINE

A domain application for which automatic inspection is clear is automotive manufacturing. Here, the quality of the manufactured product is associated with aesthetics, safety, and human life risks (PACANA; CZERWIŃSKA, 2020), besides to be linked with the high added value of the product. Thus, any anomaly in manufacturing implies restarting a long, complex, and expensive chain of rework, sometimes after the product is already in use.

Ensuring the quality of a vehicle depends directly on the quality of the parts, their compatibility of type, and how they are assembled. Thus, defects can be physical (e.g., the morphology of a part), or logical (e.g., the use of a non-defective part in the wrong vehicle). Therefore, detect and preventing possible errors is decisive (CZERWIŃSKA *et al.*, 2019), and improving the CVSs for this job becomes strategic (APOSTOLOPOULOS; TZANI, 2020; WANG *et al.*, 1998).

In contrast, most commercial CVS do not generalize well for real problems in manufacturing. For instance, the need to manually define the template for classification, both for pre-processing methods and for lighting adjustments, are strong limitations of a CVS. In fact, this depends on experts to be set up, and may vary from one application to another. Also the manual definition of thresholds for classification requires experienced knowledge.

The use of DL-based techniques has shown potential to reduce these limitations by automating some of the expert-dependent tasks (APOSTOLOPOULOS; TZANI, 2020), (HAN *et al.*, 2018). Yet, this alternative still faces a crucial problem: the number of possibilities in which a defect can be created and interpreted. When a manufacturing system requires this type of multiple-objects flexibly-assembled inspection, the use of CVSs becomes practically unfeasible due to their weak generalization capabilities to recognize multiple and unexpected defects.

The proposal in this article emerges from this reasoning and hypothesizes that those limitations can be overcome, to some extent, by applying modern learning systems based on *Generative Adversarial Networks* (GANs) (CARRARA *et al.*, 2021).

3.3 GENERATIVE ADVERSARIAL NETWORKS

A *Generative Adversarial Network* (GAN) is part of a deep neural network architecture that consists in training two models (players) to take decision by competing. One player, called *generator* (G), is a neural network that generates new (fake) data instances, while the other, called *discriminator* (D), evaluates their authenticity. By analogy, G acts like an intruder who tries to create and spread false leads to make its identification harder, while D is a detective who tries to sort out the leads that make sense.

In order to make the false leads, G applies random noises to generate data as real as possible to those in a training dataset. This aims to fool D, which then has to decide whether or not each data instance belongs to the training dataset, or it has been created by G. The Figure 3 illustrates a GAN scheme.

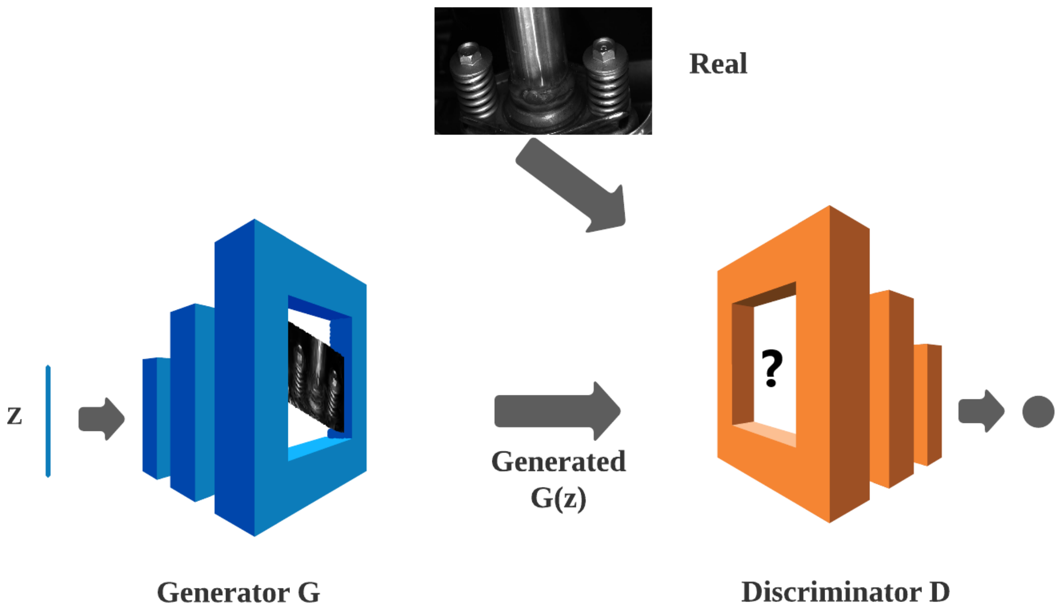

Figure 3 – GAN **architecture.**

The game challenge can then be summarized as follows: given a dataset with training data samples, a generator of fake samples G, and a discriminator of instances D, consider that D is trained to maximize the chances of assigning a correct label for samples coming indistinctly from the dataset or from G; inversely, G is trained to minimize the hits of D.

Remark that the rules for G and D to compete can be intuitively seen as a two-player *minimax* game, inherited from the Games Theory (OWEN, 2013). It assumes that there is always a rational solution to a well-defined conflict between two individuals whose interests are opposite. By following this reasoning, G and D can be seen as two opposite players that take optimized decisions though a value function $V(G,D)$ (GOODFELLOW *et al.*, 2014), according to the Eq. (1):

$$\min_{G} \max_{D} V(D,G) = \mathbb{E}_{\mathbf{x} \sim \mathrm{p}_d(\mathbf{x})}[\log(D(\mathbf{x}))] + \mathbb{E}_{\mathbf{z} \sim \mathrm{p}_{\mathbf{z}}(\mathbf{z})}[\log(1\text{-}D(G(\mathbf{z}))] \quad (1)$$

where: $\mathbf{x}$ is the real data sample, with distribution $\mathrm{p}_d(\mathbf{x})$, that approximates the expected value $\mathbb{E}_{\mathbf{x}}$; and $\mathrm{p}_{\mathbf{z}}(\mathbf{z})$ is a latent variable associated with an input noise vector with random uniform distribution $z \sim U(-1,1)$, that approximates the expected value $\mathbb{E}_{\mathbf{z}}$. Then, $G(\mathbf{z})$ generates data from the input noise $\mathrm{p}_{\mathbf{z}}(\mathbf{z})$, while $D(\mathbf{x}) \in [0,1]$ discerns how likely its input is to be true, or inversely fake.

The GAN training converges when *Nash-equilibrium* is reached in the minimax (zero-sum) game (KREPS, 1989). In Game Theory, the Nash-Equilibrium is reached when the actions of one player do not change depending on the opponent's actions. Here, this means that the GAN generator $G(\mathbf{z})$ produces realistic images, and the discriminator $D(\mathbf{x})$ outputs random predictions (probabilities close to 0.5) (YEH *et al.*, 2017).

However, GANs are typically trained using gradient descent techniques that are designed to find a minimum for a cost function, instead of finding the Nash-Equilibrium, as it may lead the search not to converge (GOODFELLOW, 2014). In other words, achieving Nash-equilibrium often proves difficult due to training instability (SALIMANS *et al.*, 2016), and approaches such as Wasserstein GANs arise.

3.3.1 WGAN

The Wasserstein GANs (WGANs) are alternatives for training conventional GANs that tend to improve the learning stability. This also prevents problems like the mode collapse, and provide meaningful learning curves that are useful for debugging and hyperparameter searches (ARJOVSKY *et al.*, 2017).

Here, $D(\mathbf{h}) \in \mathbb{R}$ is an auxiliary scalar function, which is used to calculate the Wasserstein distance that replaces $D(\mathbf{x})$, in the minimax game, so that:

$$\min_{G} \max_{D} V(D,G) = \mathbb{E}_{\mathbf{h} \sim p_d(\mathbf{h})}[(D(\mathbf{h}))] - \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})}[D(G(\mathbf{z}))]\cdot \quad (2)$$

In this way, the D function moves from a classifier to a critic, producing an authenticity score. This tends to assign high scores to real samples, and low scores to simulated samples. More details on WGANs can be found in (GULRAJANI *et al.*, 2017; PETZKA *et al.*, 2017).

In addition to the WGANs, other variations of GANs have emerged to optimize the original theory, as it is the case of the BiGANs and CBiGANs.

3.3.2 BiGAN

These are GANs improved to optimize the latent space, exposing this space to the discriminator along with the images generated from it (BERG *et al.*, 2019). A codifier module, denoted $E(\mathbf{h})$, is introduced and trained in conjunction with G, in a way to map the real samples to their respective latent spaces. The discriminator $D(\mathbf{h}, \mathbf{z}) \in [0, 1]$ is trained to discern whether the couple $(\mathbf{h}, \mathbf{z})$ comes from a real or generated image. The minimax problem for BiGANs can then be introduced as follows:

$$\begin{aligned} \min_{G,E} \max_{D} V(D, E, G) &= \mathbb{E}_{\mathbf{h}\sim p_d(\mathbf{h})}[\log D((\mathbf{h}), E(\mathbf{h}))] \\ &+ \mathbb{E}_{\mathbf{z}\sim p_z(\mathbf{z})}[\log(1\text{-}D(G(\mathbf{z})), \mathbf{z})))]\cdot \end{aligned} \tag{3}$$

Thus, fooling D causes G and E to minimize the difference between the ordered pairs $(G(\mathbf{z}), \mathbf{z})$ and $(\mathbf{h}, E(\mathbf{h}))$.

3.4 CBIGAN

The *Consistency Bi*GAN (CBiGAN) (CARRARA *et al.*, 2021) aims to combine features from WGAN and BiGAN. This approach improves the modeling of the latent space by exposing it to the discriminator (BiGAN) and also produces authenticity scores (not classification) from WGAN. In this way the CBiGAN tackles anomaly detection as a one-class classification problem, assuming for the training only non-anomalous samples. Given a test sample, the CBiGAN labels it as normal/nonanomalous/defect-free (considering a negative class) or anomalous (otherwise).

CBiGAN can be seen as a GAN that captures the distribution of latent space $Z = \mathbb{R}^n$. The generative model is a BiGAN and the loss function is modeled as Wasserstein distance.

Then, the new minimax problem can be stated as:

$$\min_{G,E} \max_{D} V(G,E,D) = \mathbb{E}_{\mathbf{h}\sim p_d(\mathbf{h})}[D(\mathbf{h},E(\mathbf{h})] - \mathbb{E}_{\mathbf{z}\sim p_z(\mathbf{z})}[D(G(\mathbf{z}),\mathbf{z}))] \tag{4}$$

where:

- $G : Z \longmapsto h$ is the generator that produces false images of latent variable associated with an input noise vector.
- $E : h \longmapsto Z$ is an encoder that map real samples to the corresponding latent space and trained together with G.
- $D : h \times Z \longmapsto \mathbb{R}$ is the discriminator/critic that produces authenticity scores to each samples using Wasserstein distance.

This model is of particular interest in this work, since our training is performed only on non-anomalous/defect-free samples. They are easily acquired in the industrial environment, while samples with defects are more rare and variable.

4 APPLICATION

This section introduces our approach to detect defects that currently results from the experimental CVS at the Renault manufacturing.

A condensed version of the real process is shown in Figure 4. This environment is called *image island*. It was created by Renault to carry out the final inspection of the underside of the vehicles, just before it is attached to the painted body. It consists of a CVS that includes robotic arms, two high resolution cameras, and a computing system. The production line aims to deliver 60 cars per hour, and the cycle time is 54 seconds. Therefore, at each point on the line, employees or equipment have 54 seconds to complete their work, while the remaining 6 seconds are spent for a vehicle to pass from one checkpoint to another.

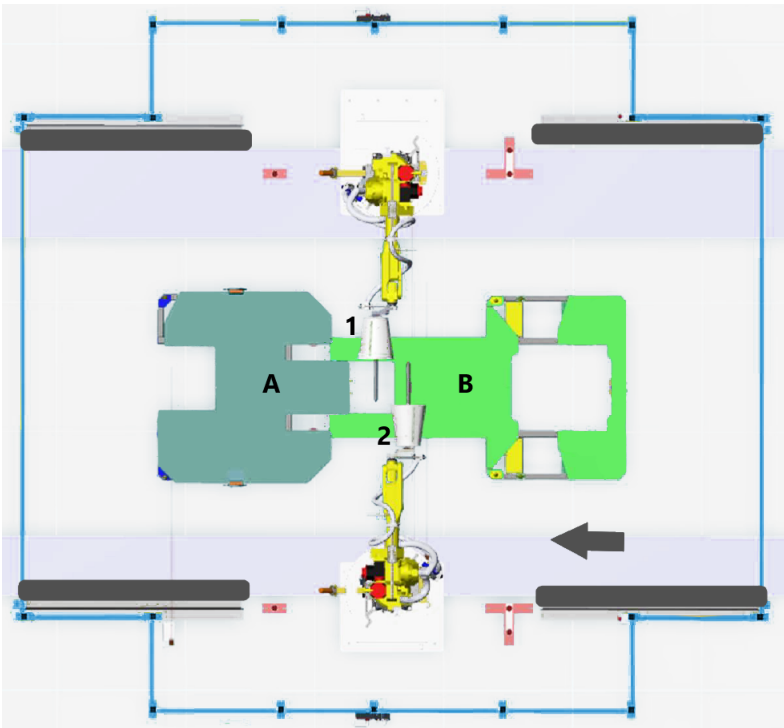

Figure 4 – Representation of an image island at Renault, with a car positioned under the two cameras that capture images from points of interest.

Each car that passes though the island is seen as two parts, A and B, that are positioned in front of two cameras, 1 and 2. They both collect images of certain regions of interest, first from A and then from B. Remark that cars belong to different models, so that the image processing consists in automatic inspection of many components of several car models.

For the automated checking, A and B are further conceptually split in 4 sections. Each section has a number of points to be verified, which varies from one section to another. In total, at most 11 points are checked through the 4 sections, which may vary according to the model. The line does not stop during the verification, so that a sensor detects the presence of a new vehicle and starts the verification process automatically. When a car enters the inspection cell, a PLC controller receives an “ok" signal and then it informs to the robot controller the coordinates of the point to be checked. When the arm reaches that correct position, its controller reports this to the PLC, which finally authorizes the camera controller to collect an image of the item to be verified. For each image capture, the lighting ring switches on and off, in order to achieve best shooting condition possible after the image is captured, the native verification algorithm is triggered for that particular point and it returns a Boolean value associated with the anomalous/defect or non-anomalous/defect-free assembled component. The test result is informed to the main PLC, which interfaces it though a man-machine interface. The same procedure repeats for the other points.

The main problems with the aforementioned system is that it uses classic methods, that are not suitable for the problem to be addressed, mainly because its training is limited by the lack of images representing defects. As a result, the equipment returns a substantial rate of false negatives (FN) (i.e., images with correct components, defect-free, correctly assembled, but detected as defective), and some false positives (FP) (i.e., images for which it has not been possible to be sure about correctness, but the equipment detected as correct).

Implications of mistaken inspections can be severe: first, all cars detected as defective need to leave the line and go through a thorough manual inspection. Cars detected as defective that are really defective, would have to be reworked anyway, so they are an inevitable problem. Cars that are not really defective, but have been classified as such, i.e., that lead to FN (more rare events in this case study, but still) are also a problem because they are removed from the line unnecessarily, consuming time and resources. However, the main problem with the current CVS appears from FP classification, i.e., with defective cars that are detected as normal by the CVS. This type of verification error is quite unacceptable in practice, as it takes effort and time, increases production costs, besides to jeopardize the final integrity of the product and human lives. The next section quantifies the error rate of the existing CVS.

4.1 PROBLEM QUANTIFICATION

The production line was observed for a time window during which 10 cars were inspected. Among these cars, 7 were of the same model and included 11 verification points in their underbase. The current automatic inspection system resulted in 77 verified points, from which 62 were evaluated as non defective, and 15 as defective.

After manual inspection, it was confirmed that the 15 points initially classified as defective were actually correctly assembled, that is, they were FN points. This outcome corresponds to an error rate of 19.5%, which is impractical and makes the equipment unfeasible to be used, as it implied in 15 completely unnecessary rechecks.

Tables 1 and 2 show respectively the confusion matrix and the accuracy for some points that was checked for a time window. A defect-free inspection is represented as positive classifications (P), while a defect is expressed as negative classifications (N), which are further associated with a Boolean value to form the following assertions:

- TP = Positive evaluated as positive (true defect-free).
- TN = Negative evaluated as negative (true defect).
- FP = Negative evaluated as positive (false defect-free).
- FN = Positive evaluated as negative (false defect).

Component	TP	TN	FN	FP
C20	102	0	1	1
C60	105	0	1	0
C100	270	0	0	0
C147	256	1	5	3
C231	96	0	28	0
C259	238	0	34	2
C267	232	10	32	0
C287	101	5	8	17
C329	40	0	0	1
C369	247	0	0	0
C492	254	4	14	0

Table 1 – Confusion matrix of the current CVS model. It considers each component of a given car model that was checked over a certain time window.

Table 1 reveals a larger amount of classification errors (FN and FP). For example, the amount of FN reaches 34 cases for the component number C259, while FP reaches 17 for the component C287. These were the most extreme cases of errors observed on the dataset.

For testing, we selected the sets of components that included reliable amount of TN images, which were the components C147 and C287. Some components have many TN images, but they were detected as unreliable, as they contain noises like the hand of operators or other objects captured in from of the camera. The component C369 was also selected for testing, as it refers to a part similar to others available outside the factory environment, which allowed us to better compare our results. Table 2 shows the accuracy of the tested items.

Component	Accuracy
C147	96.88
C287	76.41
C369	100.00

Table 2 – Accuracy of the current CVS **at Renault manufacturing system checking the components of a given car model over a certain time window.**

Components C147 and C287 have low accuracy when it comes to the industrial environment. Although C369 has acceptable accuracy, it is a component that has not been exposed to any possibility of failure, and the result when exposed to defects is unclear. Therefore, the practical use of the equipment is compromised by the lack of classification confidence.

4.2 IMAGE DATASET AND EXPERIMENTAL SETUP

Our experiments exploit two distinct setups: one including images collected directly from the assembly line; and another, called controlled setup, constructed with real workpiece images that were collected outside the factory floor. In this second setup, the workpiece was physically modified to simulate a number of possible anomalies to be detected by the inspection system. The two setups are described as follows.

4.2.1 DATASET 1 - real images

The first dataset (DS1) is composed of images collected from the real production environment. With the factory in its full production, the images were captured from the island by cameras attached to robotic arms. This dataset includes 3 distinct partitions, each one referring to a distinct type of component, each type is further divided into train and test data, as follows:

- Component C147: Steering Case Pin Rubber.
- Component C287: Gearbox Sensor.
- Component C369: Exhaust Pipe Screw.

Samples of each type of component are shown in Figures 5a, 5b, and 5c. The configuration of training and testing datasets is presented in Table 3. All the images into the training set are TP images, i.e., without defects, while the testing dataset include indistinctly TP and TN images.

Partition	Profile	C147	C287	C369
Train dataset		165	48	176
	Total	165	48	176
Test dataset		35	20	25
		65	63	46
	Total	100	83	71
Total dataset		265	131	247

Table 3 – DS1 partition settings for training and testing.

As the image capturing system is positioned considerably distant from the components (approximately 40 cm) the resulting images include fragments do not associated with the defects under analysis (e.g., background and border mechanical sets). Thus, a preliminary region of interest (ROI) was defined to allow focusing on particular objects under consideration. The result of this pre-processing phase can be seen in Figures 5d, 5e, and 5f.

(a) C147. (b) C287. (c) C369.

(d) C147. (e) C287. (f) C369.

Figure 5 – Original images (a), (b), and (c), for the components C147, C287, and C369, captured directly from the assembly line, and respective ROIs version (d), (e), and (f).

For the training dataset, a DA technique was applied in order to increase the number of training images. We applied variations of rotations, width and height shifts, zoom scale, and filling mode. The final number of images of each component after augmentation can be seen in Table 4.

Remark that training our method requires only non-defective images to be available, which is a huge advantage in comparison with other approaches in the literature. In fact, TP images are usually more frequent and easier to obtain, in contrast with DL-based approaches that require to collect a reasonably large set of TN images, including all possible types of defects.

Partition	Profile	C147	C287	C369
Train	Defect − free Real	165	48	176
	Defect − free DA	1650	1344	1760
Total dataset		1815	1392	1936

Table 4 – Training DS1 **after** DA**.**

After the training phase, it is necessary to evaluate the accuracy and stability of the trained model when subject to an independent dataset that may include both TP and TN images. As the images are captured in a real environment, real defects are less frequent events. Therefore, to fairly test the model face to more complex classification tasks, we conducted a manual DA over the test images so that more instances of TN parts are produced. The imposed defects are similar to real defects, observed in real TN images, extended with new features of possible defects. The defects have been separated into small and large size defects, as listed in Table 5. Examples of TN images in the testing dataset with real defects (when they exist) and manually added defects, can be seen in Figure 6.

Partition	Class	Profile	C147	C287	C369
Train dataset	Defect − free		1815	1392	1936
		Total	1815	1392	1936
Test dataset	Defect − free		35	20	25
		Total	35	20	25
	Defect	Real	04	22	00
		Manual small	32	19	23
		Manual huge	29	22	23
		Total	65	63	46
Total dataset			1925	1475	2007

Table 5 – Number of images in training and test DS1 **after** DA**.**

4.2.2 DATASET 2 - controlled images

The second dataset (DS2) was composed with images of parts collected by a photograph camera apart from the island in the assembly line. This controlled images generation allows to represent a large number of possible defects and test whether or not the model can recognize them.

The Table 6 summarizes the number of images in DS2. As DS2 consists of only 1 part, and this part is the same as C369 of DS1, the total of images contained in DS2 is equal to DS1 for the part C369. Analogous to DS1, DA was also applied to the training images in DS2 in order to make them comparable.

From the controlled dataset DS2, we can simulate a wide diversity of defects. The test

(a) Big defect — C147.

(b) Small defect — C147.

(c) Real defect — C147.

(d) Big defect — C287.

(e) Small defect — C287.

(f) Real defect — C287.

(g) Big defect — C369.

(h) Small defect — C369.

Figure 6 – Production line DS1 **parts defects**

Partition	Profile	CR
Training dataset	Defect − free	1936
	Total	1936
Test dataset	Defect − free	25
	Defect	46
	Total	71
Total dataset		2007

Table 6 – DS2 **partition settings for training and testing datasets.**

dataset did not undergo any computational manipulation, so the defects present in this DS2 partition are defects that would actually pass through the assembly line, that is, the photographed part was defective, with no need to create defects via software. Examples of images used for training and testing are shown in Figure 7.

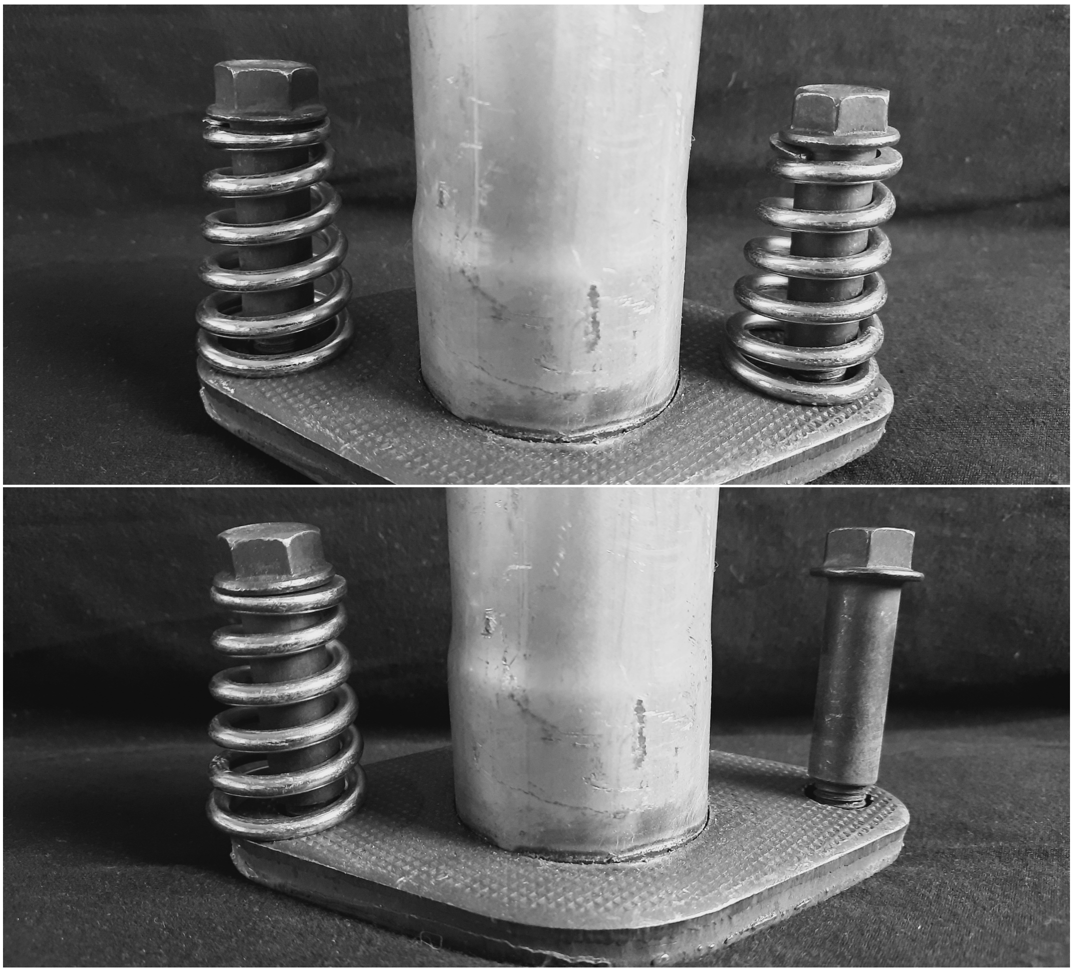

Figure 7 – Images with (a) and without (b) defect from the controlled dataset DS2.

4.2.3 Experimental analysis

In order to assess the performance of both experiments, we used the 2-way holdout (training/test split) method, with confidence interval via normal approximation. We feed the training data to the method to learn from, and then we estimated the performance over unseen data, i.e., the test was entirely conducted over images not used for training. No cross-validation or hypermeter adjustment scheme was applied to perform the model selection, since this is already known in the literature (CARRARA *et al.*, 2021), and also due to the computational costs that this would require.

As our primary evaluation, we measured the accuracy of correct classifications between TP and TN assembly parts. Accuracy is a simple metric that represents the number of correct predictions of the model. It can be defined as the number of correctly classified test cases, divided by the total number of test cases. We then compare these results with those returned by the CVS installed on the real manufacturing line.

Figures 8, 9 and 10 show examples of the images created by CBiGAN. The network was trained with images of parts from the industrial environment, that is, the generated images are very similar to factory floor images.

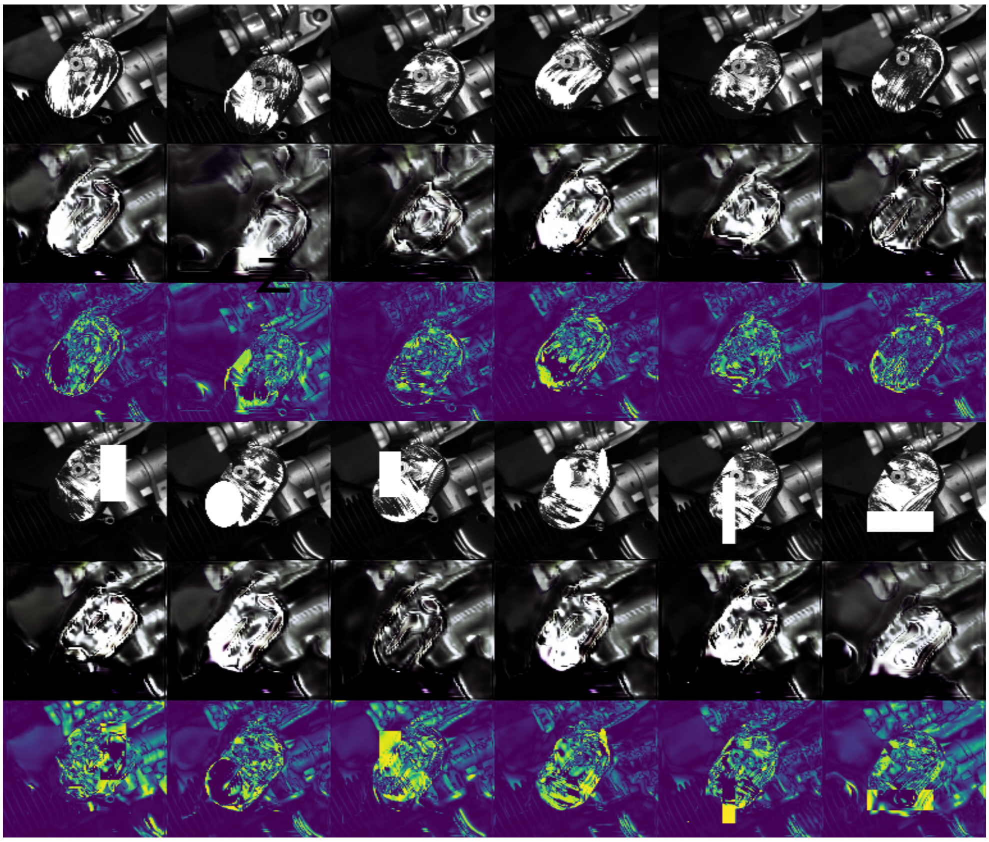

Figure 8 – Image produced by CBiGAN for part C147, which shows in its lines the training and test images, in addition to the images generated by CBiGAN and the difference between the input and generated images.

The Figures 8, 9 and 10 são divididas ao meio de forma horizontal, cada uma das colunas da parte superior representa uma peça diferente contida no banco de testes, o mesmo acontece na parte inferior. Já no caso das linhas, considerando a primeira metade das Figuras, iniciando pela linha superior, elas representam respectivamente, imagens de peças contidas no banco de testes, imagens dessas peças recriadas pela rede, e a diferença entre ambas. As linhas da segunda segunda metade da imagem são divididas da mesma maneira.

Tables 7 and 2 show the accuracy obtained by our tests, compared with the performance

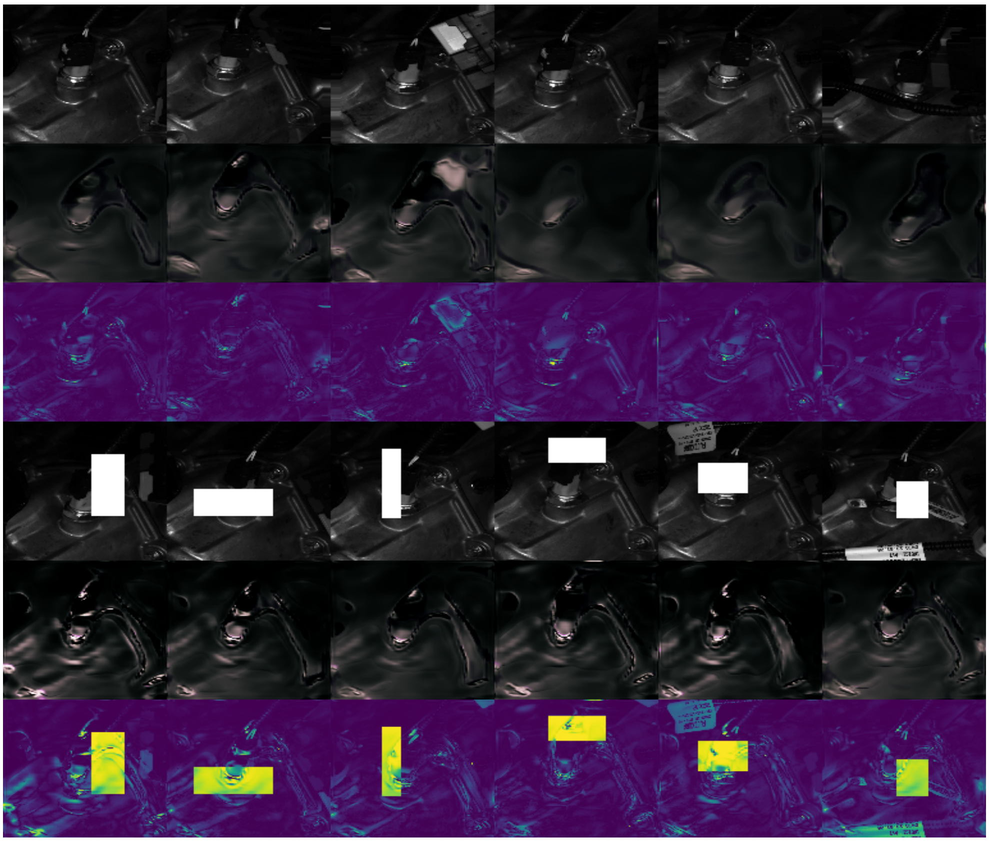

Figure 9 – Image produced by CBiGAN for part C287, which shows in its lines the training and test images, in addition to the images generated by CBiGAN and the difference between the input and generated images.

of the Renault equipment to classify each type of parts.

Component	Accuracy %
C147	98.4
C287	85.9
C369	100.0

Table 7 – Accuracy achieved in the test with CBiGAN for each set of components.

One observes an accuracy increase from 96.88% to 98.4% to classify the component C147. The improvement rate, in this case, seems to be low, which is explained by the low amount of images of defects of this component. Differently, when classifying the component C287, the accuracy increase was from 76.41% to 85,9%, which is more than 10%. This could be further improved by a more robust training step, including more images of real defects to feed the model. Finally, for the component C369 the obtained accuracy was on the order of 100%, both for DS1 and DS2.

Upon comparison, we conclude that our model was capable of improving, or at least maintaining, the accuracy of the real vision system for all the image profiles we tested. We highlight, however, that our model is further advantageous in the sense that it also identifies

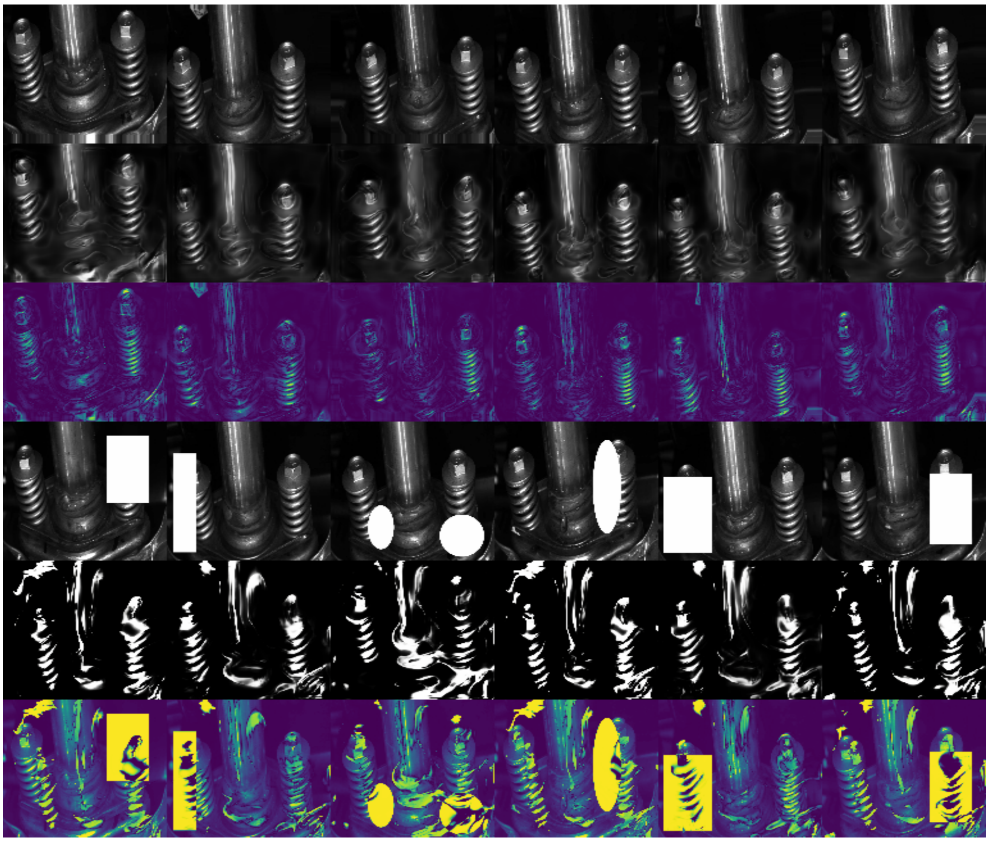

Figure 10 – Image produced by CBiGAN for part C369, which shows in its lines the training and test images, in addition to the images generated by CBiGAN and the difference between the input and generated images.

abnormal components and is able to identify all possible defective components, as it considers defective anything different from a good component used for training.

To verify that the proposed method is able to identify all possible defective components the DS2 was created, containing a wide diversity of defects. With the aim of simulate a wide diversity of defects. The Figure 11 shows examples of images generated by the CBiGAN using DS2.

As DS2 is composed of the same C369 component as DS1, we were able to compare how CBiGAN behaves with controlled images. In our tests, the result was considered excellent, keeping an accuracy of 100%. Although more real tests on the floor of the industry are needed, it is a good indication that the method can bring real benefits in the identification of defects.

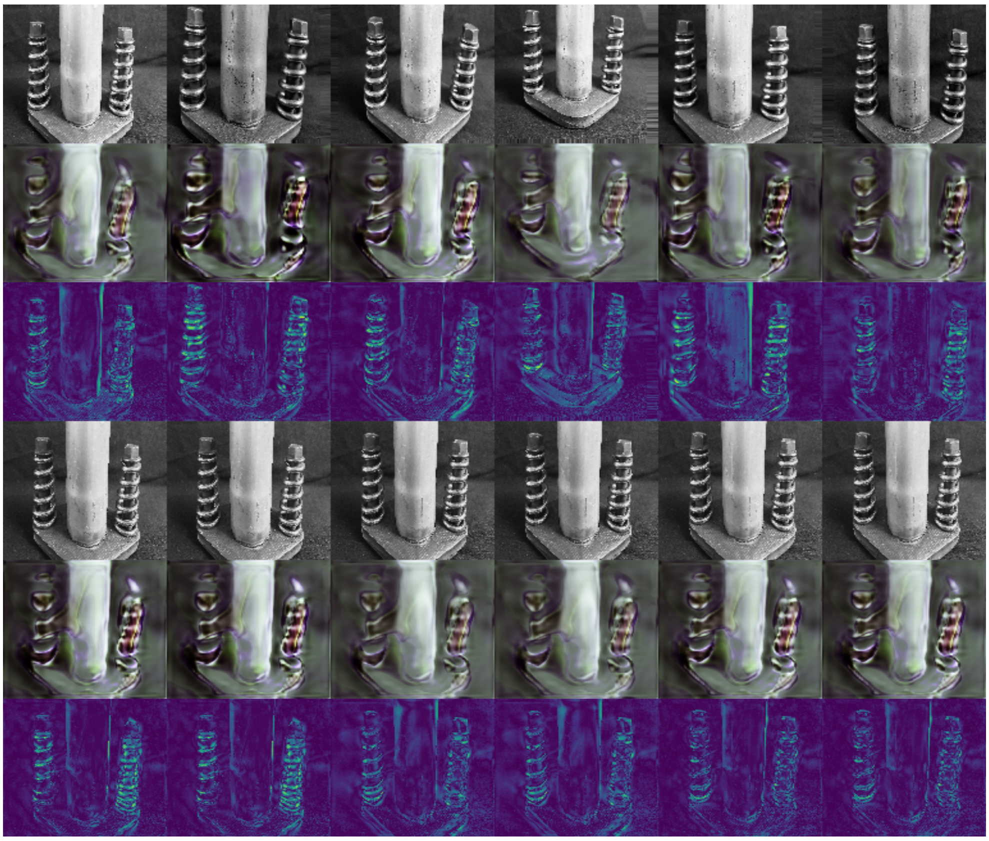

Figure 11 – In the picture we can see some real images, some produced by CBiGAN, and the different between them for the $\mathrm{DS2}$ set of components.

5 CONCLUSIONS AND PERSPECTIVES

In industry, ensuring product quality is increasingly important and challenging, especially in flexible manufacturing plants. In the automotive industry, for example, a single production line is usually responsible for manufacturing several models of cars, each one with multiple components. When a car has problems after sold, it denigrates the brand image and increases costs due to repairs. In this context, applying manual inspection as a quality control mechanism is usually not the most efficient; therefore, approach and and automated visual inspections emerge as feasible alternatives.

This type of technological solution, however, has two strong limitations: (1) the need for a specialized solution for each type of defect in each product; and (2) the need for examples (i.e., images) of any and all possible types of defect, for all components to be inspected. This work presented a method that allows to tackle both limitations in parallel. The proposed GAN model only needs defect-free images to be trained, thus solving limitation (2) and, additionally, it also solves limitation (1) as the same model can be used for different components, requiring only one retraining with the images of the component to be inspected.

The proposal was validated using 2 different scenarios. The first exploited a real industrial manufacturing environment with multiple components. The second, used images with acquisition and simulation of controlled defects. In scenario 1, the same method was used to inspect 3 different components, without the need for adaptations in the GAN. The result evidenced an increase in accuracy of inspection and identification of defects, compared with the current system used in the factory floor. In scenario 2, it was concluded that a wide variety of defects can be identified without the need for them to be part of the training step.

The obtained increase in the quality of inspection results in substantial gains along the manufacturing process, as each anticipated error implies less manual re-checking and avoids having to stop of the production line. In future researches we aim to extend the GAN-based approach to cover other types of inspection, such as conformance tests, which is the kernel for successful applications in flexible manufacturing.

REFERENCES

APOSTOLOPOULOS, Ioannis D; TZANI, Mpesiana. Industrial object, machine part and defect recognition towards fully automated industrial monitoring employing deep learning. the case of multilevel vgg19. **arXiv preprint arXiv:2011.11305**, 2020.

ARJOVSKY, Martin; CHINTALA, Soumith; BOTTOU, Léon. Wasserstein generative adversarial networks. *In*: PMLR. **International conference on machine learning**. *[S.l.]*, 2017. p. 214–223.

ARNELID, Henrik. **Sensor modelling with recurrent conditional gans**. 2018. Master's Thesis (Master's Thesis), 2018.

BERG, Amanda; AHLBERG, Jörgen; FELSBERG, Michael. Unsupervised learning of anomaly detection from contaminated image data using simultaneous encoder training. **arXiv preprint arXiv:1905.11034**, 2019.

BOUDELLA, Mohamed El Amine; SAHIN, Evren; DALLERY, Yves. Kitting optimisation in just-in-time mixed-model assembly lines: assigning parts to pickers in a hybrid robot–operator kitting system. **International Journal of Production Research**, Taylor & Francis, v. 56, n. 16, p. 5475–5494, 2018.

CAPUTO, Antonio C; PELAGAGGE, Pacifico M; SALINI, Paolo. Modeling errors in kitting processes for assembly lines feeding. **IFAC-PapersOnLine**, Elsevier, v. 48, n. 3, p. 338–344, 2015.

CARRARA, Fabio; AMATO, Giuseppe; BROMBIN, Luca; FALCHI, Fabrizio; GENNARO, Claudio. Combining gans and autoencoders for efficient anomaly detection. *In*: IEEE. **2020 25th International Conference on Pattern Recognition (ICPR)**. *[S.l.]*, 2021. p. 3939–3946.

CHA, Young-Jin; CHOI, Wooram; BÜYÜKÖZTÜRK, Oral. Deep learning-based crack damage detection using convolutional neural networks. **Computer-Aided Civil and Infrastructure Engineering**, Wiley Online Library, v. 32, n. 5, p. 361–378, 2017.

COGNEX. **COGNEX**. 2020. https://www.cognex.com/.

CZERWIŃSKA, Karolina; SIWIEC, Dominika; PACANA, Andrzej; MALINDŽÁK, Dušan. Analysis of non-compliance of industrial robot arm parts. **Zeszyty Naukowe. Organizacja i Zarządzanie/Politechnika Śląska**, 2019.

DESHPANDE, Aditya; MINAI, Ali; KUMAR, Manish. One-shot recognition of manufacturing defects in steel surfaces. **Procedia Manufacturing**, v. 48, p. 1064–1071, 01 2020.

DESHPANDE, Aditya M; MINAI, Ali A; KUMAR, Manish. One-shot recognition of manufacturing defects in steel surfaces. **Procedia Manufacturing**, Elsevier, v. 48, p. 1064–1071, 2020.

ESMAEILIAN, Behzad; BEHDAD, Sara; WANG, Ben. The evolution and future of manufacturing: A review. **Journal of Manufacturing Systems**, Elsevier, v. 39, p. 79–100, 2016.

FENG, Xin; JIANG, Youni; YANG, Xuejiao; DU, Ming; LI, Xin. Computer vision algorithms and hardware implementations: A survey. **Integration**, Elsevier, v. 69, p. 309–320, 2019.

FRID-ADAR, Maayan; KLANG, Eyal; AMITAI, Michal; GOLDBERGER, Jacob; GREENSPAN, Hayit. Synthetic data augmentation using gan for improved liver lesion classification. *In*: IEEE. **2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)**. *[S.l.]*, 2018. p. 289–293.

GAD, Ahmed Fawzy; GAD, Ahmed Fawzy; JOHN, Suresh. **Practical computer vision applications using deep learning with CNNs**. *[S.l.]*: Springer, 2018.

GENAIDY, AM; AL-SHEDI, AA; SHELL, RL. Ergonomic risk assessment: preliminary guidelines for analysis of repetition, force and posture. **Journal of human ergology**, Human Ergology Society, v. 22, n. 1, p. 45–55, 1993.

GOODFELLOW, Ian; POUGET-ABADIE, Jean; MIRZA, Mehdi; XU, Bing; WARDE-FARLEY, David; OZAIR, Sherjil; COURVILLE, Aaron; BENGIO, Yoshua. Generative adversarial nets. **Advances in neural information processing systems**, v. 27, 2014.

GOODFELLOW, Ian J. On distinguishability criteria for estimating generative models. **arXiv preprint arXiv:1412.6515**, 2014.

GULRAJANI, I; AHMED, F; ARJOVSKY, M; DUMOULIN, V; COURVILLE, AC. Improved training of wasserstein gans. in advances in neural information processing systems. Springer, 2017.

HAN, Junwei; ZHANG, Dingwen; CHENG, Gong; LIU, Nian; XU, Dong. Advanced deep-learning techniques for salient and category-specific object detection: a survey. **IEEE Signal Processing Magazine**, IEEE, v. 35, n. 1, p. 84–100, 2018.

HINTON, Geoffrey E; OSINDERO, Simon; TEH, Yee-Whye. A fast learning algorithm for deep belief nets. **Neural computation**, MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , v. 18, n. 7, p. 1527–1554, 2006.

HUANG, Gao; LIU, Zhuang; MAATEN, Laurens Van Der; WEINBERGER, Kilian Q. Densely connected convolutional networks. *In*: **Proceedings of the IEEE conference on computer vision and pattern recognition**. *[S.l.: s.n.]*, 2017. p. 4700–4708.

HUVAL, Brody; WANG, Tao; TANDON, Sameep; KISKE, Jeff; SONG, Will; PAZHAYAMPALLIL, Joel; ANDRILUKA, Mykhaylo; RAJPURKAR, Pranav; MIGIMATSU, Toki; CHENG-YUE, Royce *et al.* An empirical evaluation of deep learning on highway driving. **arXiv preprint arXiv:1504.01716**, 2015.

KEYENCE. **KEYENCE**. 2020. https://www.keyence.com.br/.

KINGMA, Diederik P; WELLING, Max. Auto-encoding variational bayes. **arXiv preprint arXiv:1312.6114**, 2013.

KIRAN, DR. **Production Planning and Control: A Comprehensive Approach**. *[S.l.]*: Butterworth-Heinemann, 2019.

KREPS, David M. Nash equilibrium. *In*: **Game Theory**. *[S.l.]*: Springer, 1989. p. 167–177.

KUSIAK, Andrew. Convolutional and generative adversarial neural networks in manufacturing. **International Journal of Production Research**, v. 58, p. 1–11, 09 2019.

LI, Liangzhi; OTA, Kaoru; DONG, Mianxiong. Deep learning for smart industry: Efficient manufacture inspection system with fog computing. **IEEE Transactions on Industrial Informatics**, IEEE, v. 14, n. 10, p. 4665–4673, 2018.

LUCKOW, Andre; COOK, Matthew; ASHCRAFT, Nathan; WEILL, Edwin; DJEREKAROV, Emil; VORSTER, Bennie. Deep learning in the automotive industry: Applications and tools. *In*: IEEE. **2016 IEEE International Conference on Big Data (Big Data)**. *[S.l.]*, 2016. p. 3759–3768.

MALEKZADEH, Touba; ABDOLLAHZADEH, Milad; NEJATI, Hossein; CHEUNG, Ngai-Man. Aircraft fuselage defect detection using deep neural networks. **arXiv preprint arXiv:1712.09213**, 2017.

MARTIN, David R; FOWLKES, Charless C; MALIK, Jitendra. Learning to detect natural image boundaries using local brightness, color, and texture cues. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 26, n. 5, p. 530–549, 2004.

MATTIA, Federico Di; GALEONE, Paolo; SIMONI, Michele De; GHELFI, Emanuele. A survey on gans for anomaly detection. **arXiv preprint arXiv:1906.11632**, 2019.

MAZZETTO, Muriel; SOUTHIER, Luiz FP; TEIXEIRA, Marcelo; CASANOVA, Dalcimar. Automatic classification of multiple objects in automotive assembly line. *In*: IEEE. **2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)**. *[S.l.]*, 2019. p. 363–369.

MAZZETTO, Muriel; TEIXEIRA, Marcelo; RODRIGUES Érick Oliveira; CASANOVA, Dalcimar. Deep learning models for visual inspection on automotive assembling line. **INTERNATIONAL JOURNAL OF ADVANCED ENGINEERING RESEARCH AND SCIENCE**, v. 7, n. 1, p. 473–494, 2020.

MERY, Domingo; FILBERT, Dieter. Automated flaw detection in aluminum castings based on the tracking of potential defects in a radioscopic image sequence. **IEEE Transactions on Robotics and Automation**, IEEE, v. 18, n. 6, p. 890–901, 2002.

MERY, D; HAHN, D; HITSCHFELD, N. Simulation of defects in aluminium castings using cad models of flaws and real x-ray images. **Insight-Non-Destructive Testing and Condition Monitoring**, The British Institute of Non-Destructive Testing, v. 47, n. 10, p. 618–624, 2005.

MUKHERJEE, Debarghya; GUHA, Aritra; SOLOMON, Justin M; SUN, Yuekai; YUROCHKIN, Mikhail. Outlier-robust optimal transport. *In*: MEILA, Marina; ZHANG, Tong (Ed.). **Proceedings of the 38th International Conference on Machine Learning**. PMLR, 2021. (Proceedings of Machine Learning Research, v. 139), p. 7850–7860. Available at: http://proceedings.mlr.press/v139/mukherjee21a.html.

OORD, Aaron Van; KALCHBRENNER, Nal; KAVUKCUOGLU, Koray. Pixel recurrent neural networks. *In*: PMLR. **International Conference on Machine Learning**. *[S.l.]*, 2016. p. 1747–1756.

OWEN, Guillermo. **Game theory**. *[S.l.]*: Emerald Group Publishing, 2013.

PACANA, Andrzej; CZERWIŃSKA, Karolina. Improving the quality level in the automotive industry. **Production Engineering Archives**, v. 26, 2020.

PERES, Ricardo Silva; AZEVEDO, Miguel; ARAúJO, Sara Oleiro; GUEDES, Magno; MIRANDA, Fábio; BARATA, José. Generative adversarial networks for data augmentation in structural adhesive inspection. **Applied Sciences**, v. 11, n. 7, 2021.

PETZKA, Henning; FISCHER, Asja; LUKOVNICOV, Denis. On the regularization of wasserstein gans. **arXiv preprint arXiv:1709.08894**, 2017.

POMERLEAU, Dean. Rapidly adapting artificial neural networks for autonomous navigation. *In*: **Advances in neural information processing systems**. *[S.l.: s.n.]*, 1991. p. 429–435.

QUINTANA, Marcos; TORRES, Juan; MENÉNDEZ, José Manuel. A simplified computer vision system for road surface inspection and maintenance. **IEEE Transactions on Intelligent Transportation Systems**, IEEE, v. 17, n. 3, p. 608–619, 2015.

RADFORD, Alec; METZ, Luke; CHINTALA, Soumith. Unsupervised representation learning with deep convolutional generative adversarial networks. **arXiv preprint arXiv:1511.06434**, 2015.

SAIZ, Fátima A.; ALFARO, Garazi; BARANDIARAN, Iñigo; GRAñA, Manuel. Generative adversarial networks to improve the robustness of visual defect segmentation by semantic networks in manufacturing components. **Applied Sciences**, v. 11, n. 14, 2021.

SALDIVAR, Alfredo Alan Flores; LI, Yun; CHEN, Wei-neng; ZHAN, Zhi-hui; ZHANG, Jun; CHEN, Leo Yi. Industry 4.0 with cyber-physical integration: A design and manufacture perspective. *In*: IEEE. **Automation and computing (icac), 2015 21st international conference on**. *[S.l.]*, 2015. p. 1–6.

SALIMANS, T; GOODFELLOW, I; ZAREMBA, W; CHEUNG, V; RADFORD, A; CHEN, X. Improved techniques for training gans, in ‘advances in neural information processing systems’. 2016.

SHIN, Hoo-Chang; TENENHOLTZ, Neil A; ROGERS, Jameson K; SCHWARZ, Christopher G; SENJEM, Matthew L; GUNTER, Jeffrey L; ANDRIOLE, Katherine P; MICHALSKI, Mark. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. *In*: SPRINGER. **International workshop on simulation and synthesis in medical imaging**. *[S.l.]*, 2018. p. 1–11.

SZELISKI, Richard. **Computer vision: algorithms and applications**. *[S.l.]*: Springer Science & Business Media, 2010.

WANG, Dongyi; VINSON, Robert; HOLMES, Maxwell; SEIBEL, Gary; BECHAR, Avital; NOF, Shimon; TAO, Yang. Early detection of tomato spotted wilt virus by hyperspectral imaging and outlier removal auxiliary classifier generative adversarial nets (or-ac-gan). **Scientific reports**, Nature Publishing Group, v. 9, n. 1, p. 1–14, 2019.

WANG, Gengxing; KANG, Wenxiong; WU, Qiuxia; WANG, Zhiyong; GAO, Junbin. Generative adversarial network (gan) based data augmentation for palmprint recognition. *In*: IEEE. **2018 Digital Image Computing: Techniques and Applications (DICTA)**. *[S.l.]*, 2018. p. 1–7.

WANG, Wei; LOH, Robert NK; GU, Edward Y. Passive compliance versus active compliance in robot-based automated assembly systems. **Industrial Robot: An International Journal**, MCB UP Ltd, 1998.

WANG, Xinlong; MAN, Zhipeng; YOU, Mingyu; SHEN, Chunhua. Adversarial generation of training examples: applications to moving vehicle license plate recognition. **arXiv preprint arXiv:1707.03124**, 2017.

WENGLOR. **WENGLOR**. 2020. https://www.wenglor.com/.

YEH, Raymond A; CHEN, Chen; LIM, Teck Yian; SCHWING, Alexander G; HASEGAWA-JOHNSON, Mark; DO, Minh N. Semantic image inpainting with deep generative models. *In*: **Proceedings of the IEEE conference on computer vision and pattern recognition**. *[S.l.: s.n.]*, 2017. p. 5485–5493.

YIN, Yong; STECKE, Kathryn E; LI, Dongni. The evolution of production systems from industry 2.0 through industry 4.0. **International Journal of Production Research**, Taylor & Francis, v. 56, n. 1-2, p. 848–861, 2018.

ZEC, Edvin Listo; ARNELID, Henrik; MOHAMMADIHA, Nasser. Recurrent conditional gans for time series sensor modelling. *In*: **Time Series Workshop at International Conference on Machine Learning,(Long Beach, California)**. *[S.l.: s.n.]*, 2019.

ZHENG, Zhedong; ZHENG, Liang; YANG, Yi. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. *In*: **Proceedings of the IEEE international conference on computer vision**. *[S.l.: s.n.]*, 2017. p. 3754–3762.

ZHU, Jun-Yan; PARK, Taesung; ISOLA, Phillip; EFROS, Alexei A. Unpaired image-to-image translation using cycle-consistent adversarial networks. *In*: **Proceedings of the IEEE international conference on computer vision**. *[S.l.: s.n.]*, 2017. p. 2223–2232.