الجمهورية الجزائرية الديمقراطية الشعبية
People's Democratic Republic of Algeria
وزارة التعليم العالي والبحث العلمي
Ministry of Higher Education and scientific research
المدرسة الوطنية العليا للتكنولوجيا
National Higher School of Technology
Department of Electrical Engineering And Industrial Computing

Final Year Project to Obtain the Diploma of

Master

Field:

Automation

Speciality:

Automation and Industrial Computing

Deep Learning for product quality inspection: State of the art

Realised by
FEZZAI Oussama Chams Eddine
TAIBECHE Ahmed

Examination Committee:

Mr. REZOUG Amar	Chair
Mme. REBAI Karima	Supervisor
Mr. CHOUTRI Kheireddine	Examiner
Mme. CHOUAF Seloua	Examiner

Algiers, the 7/ 4/2023

Academic year 2022 –2023

Deep Learning for product quality inspection: State of the art

TAIBECHE Ahmed
National Higher School of Technology
(Algiers - Algeria)
a_taibeche@enst.dz

FEZZAI Oussama Chams Eddine
National Higher School of Technology
(Algiers - Algeria)
o_fezzai@enst.dz

REBAI Karima
National Higher School of Technology
(Algiers - Algeria)
karima.rebai@enst.dz

Abstract: **The examination of visual components holds great importance in diverse industrial sectors, where conventional methods have historically played a crucial role in producing valuable outcomes. However, recent advancements have introduced a novel approach that leverages the capabilities of deep learning, resulting in a remarkable transformation of anomaly detection techniques. The objective of this paper is to provide a thorough overview of cutting-edge deep learning methodologies implemented in the realm of visual inspection, with a specific emphasis on their effectiveness in detecting defects. By harnessing the power of deep learning algorithms, this study explores the progress and breakthroughs made in this field, highlighting the notable enhancements achieved in industrial processes. Through an extensive analysis of various methodologies and approaches, this paper presents a comprehensive evaluation of the present state of deep learning in visual inspection, shedding light on its potential to revolutionize quality control measures and ensure efficient defect detection within industrial settings.**

Keywords: **Visual inspection, Deep learning, Transfer Learning, Supervised Learning, Unsupervised Learning.**

I. INTRODUCTION

During the late 1960s and early 1970s, computer vision methods emerged in order to imitate human vision using camera schematics, projections, and photogrammetry. At present, computer vision has advanced to performing complex tasks, including object recognition, autonomous vehicle navigation, face detection and fast image processing.

Industries that rely on human experts for visual inspection and quality control can benefit from computer vision's speed, repeatability, and accuracy. These systems utilize a light source, image capture device, and computational systems for feature extraction and processing, making them ideal for repetitive, high-speed production environments where fast and precise measurements are crucial. By eliminating the drawbacks of human inspection, such as subjectivity, fatigue, slowness, and expenses, computer vision can be an effective tool for industries that face challenging, monotonous, or dangerous inspection tasks.

Visual inspection (VI) has been the oldest Non-Destructive Testing (NDT) method used for quality enhancement. It involves observing the surface of an object to look for physical defects like cracks, scratches, corrosion, and misalignments [1]. Visual inspection can detect defects, anomalies, or inconsistencies in products, materials, or processes, Therefore, It plays a crucial role in quality enhancement, by identifying defects early in the manufacturing process, companies can take corrective actions and prevent costly recalls or rework [2].

VI is essential and important for manufacturing companies, especially for big brands that produce expensive products as they have a lot to lose if defective products are sent out. For example, car manufacturers like Rolls-Royce use visual inspection to ensure the fit and finish of each component as 70% of all faulty components are rejected due to discrepancies found during visual inspection [3]. It is essential to note that visual inspections are not only reserved for high-end products but also used in everyday consumer goods like food and beverages. In the food and beverage industry, quality control and inspection are crucial to ensure the safety of consumers and prevent costly errors. Visual inspections play a significant role in enhancing the food quality by identifying defects or contaminants that may cause harm [4].

While human inspectors are often used in many industries to perform visual inspections, this approach can be slow, labor-intensive, and susceptible to errors. Consequently, automated systems that incorporate image processing, computer vision, and machine learning have become a more efficient and trustworthy option for identifying product defects. Automated VI systems offer the advantages of reducing the risks of inconsistency and subjectivity in inspection results, ultimately leading to improved product quality and cost savings.

Advancements in deep learning have revolutionized the field of visual inspection in recent years. One recent advancement in deep learning for visual inspection is Generative Adversarial Networks (GANs) that can enhance deep learning models by generating realistic images for training datasets [5]. This augmentation provides more diverse examples for deep learning models to learn from, ultimately improving model accuracy [5]. Another promising development is the use of unsupervised learning techniques such as Autoencoders for anomaly detection in images. These techniques can identify abnormal patterns in data without the need for labeled examples, making them useful for applications where labeled data is scarce [6].

Transfer learning is another technique used to enhance the performance of deep learning models in visual inspection tasks. Specifically, pre-trained models such as ResNet, VGG,

and Inception are fine-tuned on specific visual inspection datasets to improve their accuracy and reduce the amount of training required. This approach has demonstrated success in a variety of visual inspection applications [7][8][9][10].

Finally, robotics integration with deep learning represents an exciting frontier in the field. Autonomous visual inspection systems use cameras and sensors to capture images of objects and analyze them in real-time using deep learning algorithms to detect defects and anomalies [11].

This paper aims to provide an exhaustive overview of the current state of visual inspection and the use of advanced techniques such as deep learning to enhance quality and efficiency. To this end, we examine the importance of VI and how advanced techniques, particularly Deep learning, can improve its efficiency and accuracy. The remainder of this paper is divided into five sections. The first one presents the structure and the components of visual inspection systems. The second section discusses the traditional techniques used for VI and their limitations. The third section explains how deep learning techniques like transfer learning and CNNs can be used in VI with examples from the industry. Lastly, the fourth part briefly discusses commercially available products for VI and their advantages and disadvantages and the five section concludes this paper.

II. Visual Inspection Systems

Visual inspection system is composed principally from camera and lighting sources. It can be installed at different position in production line for product quality inspection.

A. Cameras

According to the key characteristics to be consider for product quality inspection, there are different industrial video camera types. For example:

Line scan cameras that process very narrow bands, one at a time. They capture single rows of pixels in quick succession (Figure 1.a), while the camera is moved over an object or objects pass under the camera. Line cameras are used when there is a need to analyze many items in a row without stopping, or when trying to capture a very long image.

Since line scan systems require parts in motion to build the image, they are often well-suited for products in continuous motion, or high-speed applications [12].

Area scan cameras use a matrix of pixels to capture a scene (Figure 1.b), similar to how the camera on smartphone works. They have a fixed resolution and are easier to install and set up. Area scan cameras are more general purpose and are often used to capture a defined area quickly, where an image can be segmented into regions to look for specific objects [12].

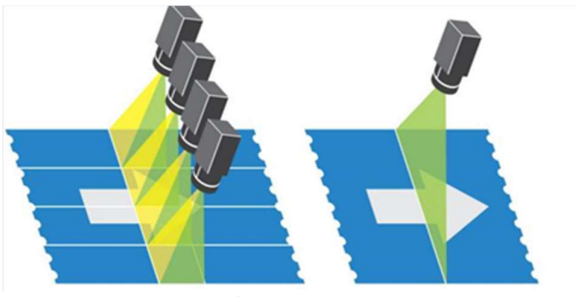


Fig. 1. Structure of Area Scan and Line Scan [13]

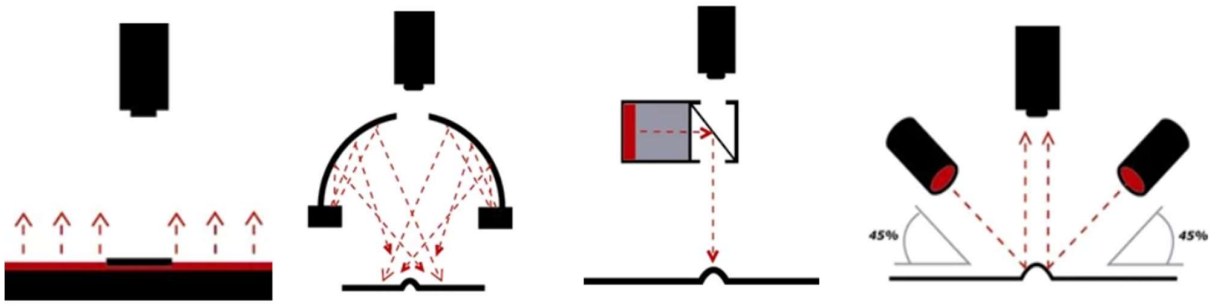
This type of camera is best suited towards applications where the object is stationary, even if only shortly, or moving at a slow pace. It is also the most suitable techniques when inspecting individual objects, where a single image is sufficient to capture the whole area of interest [12].

B. Camera lenses

The choice of camera lenses has various effects on the image quality and the image processing performances and costs. More expensive lenses that produce high resolution and high-quality images can reduce processing times, while lenses that produce more basic imaging resolutions are more cost effective [12]. Below are some factors to consider when choosing lenses:

- Sensor and image circle sizes: It is important to choose a lens that is large enough not to create any vignette at the edge of the sensor area, but not so large that there is "extra" area in the image circle that is not used [14].
- Resolution: for high resolution image, lens with a high megapixel count that can resolve the pixel size is sufficient [14].
- Focal length: it determines the angle of view, or how large of an area a lens can capture. The focal length must match the sensor size [14].
- Aperture and lighting conditions: It is important to choose a lens that has sufficient f-stops (settings that allow the aperture to be more open or closed) according to the environment's lighting conditions. The lower the f-stop number, the wider the aperture will open, but this will also reduce the depth of field [12].
- Filters: some filters can be attached to the lens, such as a polarizing filter to help reduce glare on reflective objects. Some colored filters can also help highlight important areas on inspected objects [12][14].

Another important aspect to be consider for selecting the used camera is the sensor type. A camera can be built either around charge-coupled device (CCD) or complementary metal oxide semiconductor (CMOS) image sensor technology. Both have the same task of transforming light (photons) into electrical signals (electrons). This information is, however, transmitted by both sensor types using different ways and means and the design of each is also fundamentally different. In CCD sensors, the charges of the light-sensitive pixels are shifted and converted into signals. The charges of the pixels, which are created by exposure to a semiconductor, are transported to a central A/D converter. In the CMOS sensors, a capacitor as a charge storage is put in parallel to each individual pixel. This capacitor is charged with the exposure of each pixel by its photoelectric current. The voltage created in the capacitor is proportional to the brightness and the exposure time. Currently, the sensor market is increasingly pointed toward CMOS technology [12].



(a) Back lighting (b) Dome diffuse (c) On-axis diffuse (d) Dark field Lighting

Fig. 2: Emplacement and types of lighting source [15]

C. Lighting sources

Several types of light sources are commonly used in machine vision [12]. Quartz-halogen lamps, fluorescent lamps, and Light-Emitting Diodes (LEDs) are by far the most widely used lighting sources in visual inspection systems [15]. Form the lighting source emplacement and types, different configurations are possible as shown in Figure 2.

- *Back lighting*

In this method, the subject is illuminated from behind, which creates a contrast between the subject and the bright background. This technique is useful for detecting holes and gaps, positioning or orienting parts, or measuring objects.

- *Diffuse lighting*

It is the most commonly used on shiny specular or mixed reflectivity samples where even but multidirectional light is needed. Several implementations of diffuse lighting are generally available, but there are different types (Figure 2.b and Figure 2.c), with hemispherical dome/cylinder or on-axis being the most common.

- *Dark field lighting*

It is perhaps the least well understood of all the techniques, although these techniques can be used in everyday life. For example, the use of automobile headlights relies on light incident at low angles on the road surface, reflecting back from the small surface imperfections, and also nearby objects (Figure 2.d).

D. Placement of visual inspection system in production line

Vision-based inspection systems are placed differently on manufacturing lines based on the specific defects they inspect. They are only used during production when the design is locked as configuring them is expensive and complicated due to the custom setup of cameras, lighting, and alignment materials [16].

The placement of a visual inspection system in a production line is critical to its performance. Factors such as the type of inspection needed, the product being produced, and the layout of the production line should be taken into consideration when determining the location of the system. In addition, ensuring that the system is placed in an area with sufficient lighting and minimal vibration can help to ensure accurate inspections [17].

a) ***End of line Visual inspection systems*** are typically placed at the end of the production line, where they can inspect the final product after all the manufacturing processes have been completed.

b) ***in-line inspection****,* visual inspection systems can be placed in between the manufacturing processes to inspect the product at each stage of production. Placement of the visual inspection system should consider the speed of the production line, the type of inspection needed, and the product being produced. The table 1 resumes the advantages and disadvantages of each system [12][14].

E. Visual inspection system architecture

Mono-image inspection is a simple and straightforward approach to visual inspection that can be easily integrated into existing production processes. However, it may have limitations for detecting defects in complex geometries or textures or for detecting changes over time. Multi-image inspection is a more complex approach that can capture more information from different perspectives, making it more accurate for detecting defects in complex geometries or textures or for detecting changes over time. However, it may be more difficult to integrate into existing production processes and may require more sophisticated software algorithms. The choice between mono-image and multi-image inspection will depend on the specific requirements of the manufacturing process and the types of defects that need to be detected.

TABLE 1. ADVANTAGES AND DISADVANTAGES OF END-OF-LINE AND IN-LINE INSPECTION

	End-of-line inspection	In-line inspection
Advantages	Allows inspection of the final product Faster inspection speeds easy integration into existing production lines	Real-time defect detection Ability to correct defects in real-time Reduction in scrap and rework costs, can identify which manufacturing process caused the defect The goal is zero-defect manufacturing, with an emphasis on in-process inspection and quality control at the source rather than the end of the process [18]
Disadvantages	Inability to correct defects in real-time May result in wasted materials and production time if defects are found after production	Slower inspection speeds May require a dedicated production line for inspection Can be more complicated to set up compared to end-of-line inspection

III. CONVENTIONAL METHODS FOR VISUAL INSPECTION

In traditional image processing techniques, a primitive image matching technique is used to compare the pixel intensities of the two images directly by using some correlation measure. This can be a good method for the images generated using software, but it is extremely sensitive to even slightest noise, alignment, intensity and other transformations due to image capturing conditions, and therefore it is rarely used. A more common approach is to compare feature vectors of two images, which are extracted by applying some algorithms on the images. Many feature extraction techniques have been developed.

In [19], the Automatic Optical Inspection (AOI) methods are classified as statistical, structural, filter-based, and model-based. Statistical methods are used to compute the spatial distribution of pixels of the image, and most of the methods are based on first order statistical computations such as mean, variance, range, and other histogram-based measurements, along with the higher order statistics. Devi et al. [20] contrived a Machine Vision (MV) system to inspect rice grains for size and quality. The system proposed uses Canny Edge Detection to obtain rice grain contours, and then develops a region of interest around the contour. The extracted features using the system are area, length, breadth, and diagonal, which were used to classify rice types at a 90% accuracy. Further contributing to quality inspection of rice was proposed in Chen et al. [20], it identified broken, chalky, and damaged or spotted rice grains. A unique algorithm was developed for each defect with the former two using support vector machines and the latter using Sobel gradient operator and Otsu segmentation. The reported defect identification accuracies were 99.3%, 96.3%, and 93.6%, respectively. In [21], an algorithm for image matching based on structural similarity using statistical parameters is presented. Filter-based methods employ filter banks, such as Fourier transform, Gabor filters, and gradient filters, to generate feature vectors which consist of the filter responses.

In [22], a Phase-Only Correlation based image matching technique was developed, which uses the phase information in the Discrete Fourier Transforms for computing similarity. If the knowledge of the underlying process that generates the texture pattern is available, and it can be defined by a stochastic model, then the model-based approaches can yield very accurate results in texture defect detection.

In [23, 24], there are two widely used techniques for feature extraction, which have demonstrated good performances in many pattern matching applications. Soldering defects in PCB is a highly specialized domain, which has its own set of defect detection techniques, and most of the research works for finding such defects employ traditional image processing techniques [18–20]. The work presented in [21] used the HAAR wavelet transform to detect defects in bare PCBs (with no parts mounted on them).

Huang et al. [25] improved inspections on glass bottle breakage and contamination detection with no loss in observation time. The developed system uses a radial scanning method followed by contour fitting to identify breakage on bottles mouths. Bottles with damaged bottoms were detected using a Blob algorithm. The system accomplished an inspection rate of 72,000 bottles per hour while detecting 100% of defective bottles and mislabeling 0.297% of bottles

In some fields like the field of additive manufacturing, the detection of defects over the manufacturing and loading processes of finished products which needs to develop pertinent Non-Destructive Testing (NDT) techniques to detect these defects. Ultrasonic testing (UT) is one of the most powerful non-destructive evaluation (NDE) methods employed to detect and evaluate surface, subsurface and internal defects. In most metals, flaws as deep as several meters are detectable through UT while the material thickness is a limiting factor for most NDT methods. The sensitivity of UT in detecting flaws highly increases with the testing frequency.

Kazemian et al [26] designed a machine vision system to inspect Large-Scale Additive Manufacturing (LSAM) extrusion in real time. LSAM uses building materials, such as cement, to print buildings and other types of structures. The material quality is paramount to ensuring safe and long-standing shelters are created. MV is used here to create a closed-loop inspection system that determines a LSAM's extrusion rate and make adjustments if needed. Since the extruded material in this system is seen from atop, the system perceives a thick line whose width is used to identify extrusion rate deviations. Test results demonstrated correction capabilities within ±10% of the desired rate in less than three seconds.

Hongbin et al [27] proposed a real-time quality control architecture for rolled steel manufacturing using Support Vector Machines, achieving 90.5% accuracy in defect detection and classification. Yıldız et al. [28] preprocessed images with wavelet, threshold, and pathological operations and then used the International Journal of Precision Engineering and Manufacturing-Green Technology GLCM

method to extract features. Finally, defect images were classified based on a KNN algorithm with an average accuracy rate of 96%

Wu et al. [29] use Feature Extraction Method and its application to Surface Defect Recognition by LQV of Hot Rolled Strips training average accuracy: recognition rate under frequency-domain feature set is 84.56% and recognition rate under space-domain feature set is 83.40% in test result

Kumar et al. [30] developed a vision system that utilizes a CCD camera to capture images of welding surfaces for defect detection. The system employs a back-propagation neural network to classify welded joints into four predefined categories. The authors achieved an impressive overall accuracy rate of 95% in testing with 80 real samples.

IV. Deep Learning For Visual Inspection

Deep learning has revolutionized visual inspection by automating and enhancing the accuracy of the process. In supervised learning, Convolutional Neural Networks (CNN) are used for defect identification and classification, while transfer learning enables the transfer of knowledge from pre-trained models to new inspection tasks. In unsupervised learning, deep learning methods like autoencoders and ANOGAN provide effective means for anomaly detection and outlier identification. Deep learning has significantly contributed to improving accuracy, efficiency, and scalability in visual inspection and is widely used across industries.

A. Supervised learning

- *Convolutional Neural Network (CNN)*

CNNs are mostly useful for identifying patterns in images to recognize objects, faces, and scenes. They learn directly from image data, using patterns to classify images and eliminating the need for handcrafted feature extraction. Applications that call for object recognition and computer vision such as self-driving vehicles and face-recognition applications are mostly based on CNNs. Using CNNs for deep learning has become increasingly popular due to three important factors:

a) CNNs eliminate the need for handcrafted feature extraction by learning the features;

b) CNNs yield state-of-the-art recognition results

c) CNNs can be retrained for new recognition tasks, allowing us to develop on pre-existing networks. CNNs offer a best architecture for pattern detection and image recognition [31].

The works addressed the problem of visual inspection can be classified in two categories: pre-trained model adapted to this problem by transfer learning and specific models developed for VI.

Transfer learning of pre-trained models

d) Architectures of pre-rained models

- AlexNET[32]

AlexNet is a classic convolutional neural network architecture that achieved state-of-the-art performance on the ImageNet dataset in 2012. It is characterized by the following properties:

Deep Architecture: AlexNet was one of the first deep convolutional neural networks consists of 8 layers in total, including five convolutional layers and three fully connected layers.

ReLU Activation: AlexNet used the rectified linear unit (ReLU) activation function instead of the more traditional sigmoid function

Dropout Regularization: AlexNet used a technique called dropout regularization to prevent overfitting during training

Local Response Normalization: AlexNet used local response normalization (LRN) as a form of regularization to reduce overfitting

- VGGNET[7]

VGGNet is a convolutional neural network architecture that was introduced in 2014 by the Visual Geometry Group (VGG) at the University of Oxford. It achieved state-of-the-art results on the ImageNet dataset. It is characterized by:

Uniform Architecture: VGGNet has a uniform architecture, with all convolutional layers having a 3x3 filter size and all max-pooling layers having a 2x2 filter size with stride 2.

Deep Architecture: VGGNet is a deep neural network with up to 19 layers, including 16 convolutional layers and 3 fully connected layers. This depth allows it to learn more complex features from images, leading to better performance.

High Accuracy: VGGNet achieved state-of-the-art performance on the ImageNet dataset, achieving a top-5 error rate of 7.3%, which was significantly better than previous state-of-the-art results.

Small Filter Size: VGGNet uses small filter sizes of 3x3, which allows for a larger receptive field with fewer parameters compared to larger filter sizes. This reduces overfitting and makes the network more efficient.

- GOOGLENET[8]

GoogleNet, also known as Inception v1, is a convolutional neural network architecture that was introduced by Google researchers in 2014.

Inception Modules: GoogleNet uses inception modules, which are small networks that perform parallel convolutions at multiple scales and aggregate their outputs. This allows the network to learn features at multiple scales, leading to better performance. The properties of this networks are:

Deep Architecture: GoogleNet is a deep neural network with 22 layers, including nine inception modules. This depth allows it to learn more complex features from images, leading to better performance.

Efficient Use of Parameters: GoogleNet makes efficient use of its parameters by using 1x1 convolutions to reduce the dimensionality of the input feature maps before applying larger convolutions. This reduces the number of parameters and helps to prevent overfitting.

Global Average Pooling: GoogleNet uses global average pooling instead of fully connected layers to produce its output. This reduces the number of parameters and helps to prevent overfitting.

- RESNET[9]

ResNet, short for Residual Network, is a convolutional neural network architecture that was introduced by Microsoft researchers in 2015. It achieved state-of-the-art results on the ImageNet dataset and won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2015, 2016, and 2017. The properties of this networks are:

Residual Blocks: ResNet uses residual blocks, which are small networks that learn residual mappings between layers. This allows the network to learn the residual information rather than trying to learn the entire transformation from one layer to the next. This reduces the vanishing gradient problem and makes it easier to train deeper networks

Deep Architecture: ResNet is a very deep neural network, with up to 152 layers in its largest version. This depth allows it to learn more complex features from images, leading to better performance

Shortcut Connections: ResNet uses shortcut connections that bypass one or more layers, allowing information to flow more easily through the network. These shortcut connections are used to add the output of one or more previous layers to the output of the current layer, which helps to prevent the vanishing gradient problem.

Versatility: ResNet has been used for a wide variety of computer vision tasks beyond image classification, including object detection, semantic segmentation, and image captioning.

e) State of art achievements and works based on Transfer learning

In several works, these pre-trained networks are used and affined for solving the problem of visual inspection and fault detection for different applications. This process is called transfer learning. In the following of this subsection, some works are reviewed.

In agricultural applications, Mohanty, S.P. et al. trained two well-known deep convolutional neural networks, AlexNet and GoogLeNet, to identify 14 crop species and 26 diseases in PlantVillage; GoogLeNet achieved the top accuracy of 99.35%. However, when tested with two other validation datasets (IPM and Bing), the overall accuracy reduced substantially, to just over 31% [33].

For surface quality inspection, structural crack detection, Yang et al gathered their own data and utilized CNN models trained via Transfer Learning. They achieved an impressive accuracy of 98.8% for automatic tracking and detection of cracks by utilizing YOLOv3 [34].

In shape detection and classification, Dharmawan et al. [10] used Transfer learning of CNN models with 752 images are compared. The five considered models in this study are (AlexNet GoogLeNet VGG16 ResNet50, and InceptionV3). From the five transfer learning models, the VGG16 model obtained the best results for the raw cut dataset with a validation accuracy of 95.37%, whereas GoogLeNet showed the best results for the contrast-limited adaptive histogram equalization (CLAHE) and extraction datasets with 96.32% and 93.17% validation accuracies, respectively.

In the manufacturing industry, Wang et al. [35] employed vision intelligence to inspect product defects (specifically water bottles) using deep learning and Hough transform techniques. As a result, they achieved an exceptional overall accuracy rate of 99.60% on their testing dataset.

In water bottle defect detection using deferent models are summarized in Table 2.

B. Unsupervised defect detection

The authors of [37] present an optical inspection application in quality control for printer with two approaches using supervised and unsupervised learning. In the supervised approach, the Mask-R-CNN shows good results with an accuracy of 78%. While, the detection rate for defect-free images using an unsupervised learning is also very low at 65 %.

The unsupervised methods do not require annotations reduces the effort for data collection by almost 97 %. Table 3 shows the time required for data acquisition and data annotation for each method. Furthermore, the recognition accuracies are also presented in this table.

In [39], a novel Stacked AutoEncoder – Long Short Term Mmory (SAE-LSTM) approach is proposed for anomaly detection in mechanical equipment. The SAE is a type of deep learning neural network architecture used for unsupervised learning that used Autoencoder. This last is a type of neural network that learns to compress and reconstruct data. The "stacked" aspect used in [39] refers to the use of multiple layers of these autoencoders. LSTM is a type of recurrent neural network (RNN) architecture commonly used in deep learning. LSTMs are designed to solve the problem of vanishing gradients in traditional RNNs, where gradients can become very small or disappear over time, making it difficult to train the network effectively.

The method proposed in [40] aims to obtain effective and sufficient features for Mura defects by fusing handcrafted and unsupervised-learned features in a complementary manner. Mura defects typically appear in localized areas with low contrast and uneven brightness present especially on LCD and TFT displays.

The proposed Mura defect recognition process using JFR-DRF (Joint Fourier Transform Radial Profile Filtering) is shown in Figure 3. As can be seen, there are two streams, the unsupervised-learning-based stream and the handcrafted stream. The former uses the size-normalized image (after grayscale normalization), whereas the latter employs the original image (after grayscale normalization) and a sparse autoencoder (SAE). Then, the features extracted by these two data streams will be merged into a single feature vector, which acts as the representation for the Mura defect image. After subsequent processing, it will be classified using an one-against-all (OAA) SVM [41] classifier. It exhibits high robustness and improves the recognition performance by nearly 20% compared with the traditional handcrafted feature descriptor.

TABLE 2. PERFORMANCES OF SOME MODELS USED FOR DEFECT INSPECTION OF WATER BOTTLE [35]

Models	Number of layers	Accuracy (%)		Loss	
		Training	Validation	Training	Validation
VGG16	16	95	97	0.15	0.03
ResNet50	50	96.6	17.6	0.58	9.5
Model proposed in [36]	10	96	99	0.11	0.01

TABLE 3. PERFORMANCE AND TIME CONSUMPTION FOR SUPERVISED AND UNSUPERVISED LEARNING [38]

Approach	Supervised	Unsupervised
Technique	R-CNN	CAE
Accuracy	0.78	0.65
Training acquisition	3	1.5
Annotation [hours]	43.1	0

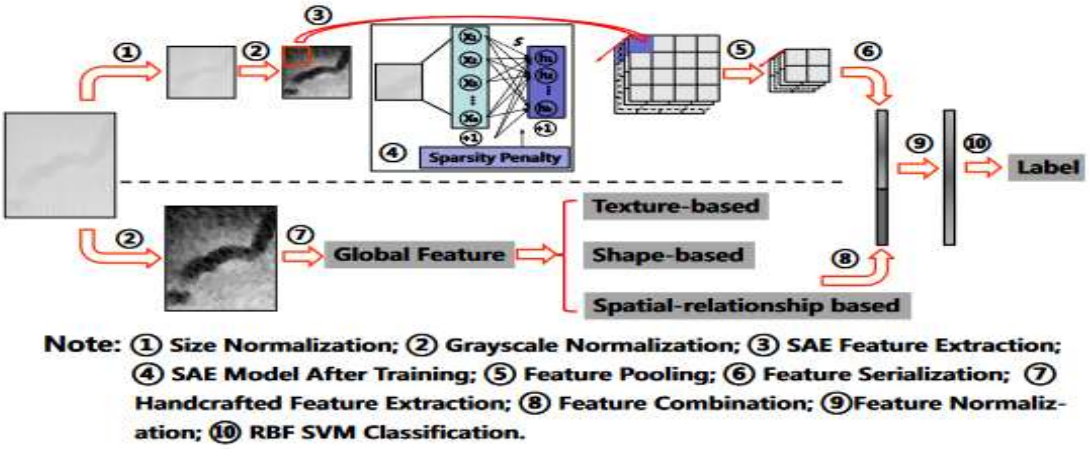


Fig. 3. JFR-DRF Mura defect detection architecture proposed in [40]

In [42], an Anomaly detection with Generative Adversarial Networks approach (AnoGAN) allows for the identification of anomalies in unseen data through unsupervised training of a model on healthy data disease markers in imaging data. This approach is particularly critical in the diagnosis, monitoring of disease progression, and evaluation of treatment response, where the ability to identify anomalies can provide valuable insights and improve patient outcomes. This approach has achieved an accuracy of 88.34%. To learn the distribution pattern of normal images, the core idea is to conduct encoding and decoding on the input normal images and train the neural network with the aim of reconstruction. The reconstruction-based method, for anomaly detection in images using deep learning, involves encoding and decoding input images, analyzing differences between original and reconstructed images, and using anomaly scores to identify anomalies. Classical models like autoencoders, VAE, and GAN are used for reconstruction. Only normal data is modeled during training, and higher anomaly scores in the testing phase indicate the presence of anomalies. The basic flow of the reconstruction-based method is illustrated in Figure 4.

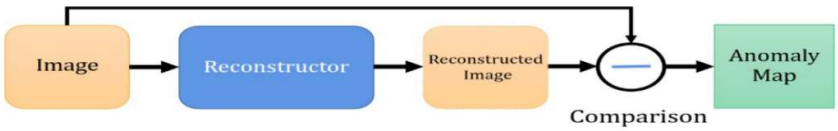


Fig. 4. Architecture of AnoGAN for defect detection [43]

V. AI-BASED COMMERCIAL SOLUTIONS FOR VI

Visual inspection using Artificial Intelligence for anomaly detection in industrial process is an attractive field for investments and business. It is an important element for smart factories and Industry 4.0. Below, some actually commercialized products will be presented.

A. CAD2SCAN

The software of Kitov CAD2SCAN system significantly simplifies setup of robotic vision inspection

systems using CAD-based automated inspection planning. By combining CAD2SCAN (Figure 5) with the Kitov Smart Planner and Deep-learning solutions, customers can



Fig. 5. CAD2SCAN system for VI [44]

automate inspection solutions that are too difficult and time consuming to manually program and deploy. Particularly useful for parts with complex geometries. CAD2SCAN allows quality managers to define their inspection requirements directly on the part CAD model [44].

B. *Inspekto S70*

Combing cutting-edge hardware and software. The INSPEKTO S70 illustrated in Figure 6 is an AI-driven industrial visual quality inspection solution with an interesting quality for use-case versatility and simplicity of use [45].



Fig. 6. INSPEKTO S70 system for VI [45]

The first AI engine of this system focuses on obtaining the image through self-set-up of all optics parameters per specific use case and scenario. It has a self-adjust mechanism during runtime to dynamically mitigate changes in production and smartly selects the optimized image from the live stream for the recognition and inspection engines. The second AI engine focuses on identifying the part by recognizing and detecting rigid objects of nearly any shape and surface. It only requires one reference image for detection and supports variations in location and rotation in the field of view (FOV).

It can be self-triggered or triggered externally. The third AI engine is responsible for finding defects through semi-supervised, non-specific AI technology that overcomes specific data scarcity and leverages generalization. It only requires a few good parts (OK), and defected (NOK) parts can be added for performance fine-tuning. It understands product tolerances and physical attributes and can differentiate between defects and permissible defects without any prior defect references or training. The system's ongoing improvement is via continuous deep-learning.

C. *Cognex tech*

In-Sight D900 Series:The In-Sight D900 vision system leverages advanced AI-based tools to solve challenging Optical Character Recognition OCR, assembly verification, and defect detection tasks. Processing takes place on-device, which eliminates the need for a PC, simplifying application deployment and bringing the power of deep learning to non-programmers [46].

D. *Robotic inspection system VINSPEC*

The Machine Vision System, manufactured by VITRONIC, generally functions by inspecting surfaces and contours of visible, sealing and functional surfaces. Even if the parts to be inspected are curved / planed, this system has also the ability to inspect it. Vitronic proposed AI-based solution for inspection of photovoltaic solar cells, automotive and healthcare [47].

E. *Products of Jidoka*

Jidoka company proposed various AI-based solution for visual inspection in different industrial fields such as automotive, pharmaceutical industries [48].

VI. Conclusion

In conclusion, visual inspection is a crucial aspect of manufacturing, quality control, and many other industries. However, conventional methods like human inspection and manual image analysis are subject to limitations stemming from subjectivity and inconsistency. Advanced techniques like machine learning and computer vision hold tremendous potential for surpassing these constraints and improving the precision and efficiency of visual inspection.

Notwithstanding the revolutionary impact of advanced visual inspection techniques, several challenges must be addressed. The primary hurdle is the need for abundant and high-quality data to train deep learning models, which can be problematic in industries with scarce or sensitive data. Another obstacle is the limited interpretability and explain ability of deep learning models, rendering it difficult to comprehend their decision-making process. This can pose particular difficulties in industries where transparent and understandable decisions are critical, such as medical or safety- critical applications. Integrating advanced visual inspection techniques into established manufacturing processes can also present difficulties, as it can be costly to implement new technologies and equipment. Moreover,training staff to use these tools may require significant time and effort. To deal with these problems and develop solutions that can contribute in the evolution of Industry 4.0, researchers, developer and companies continue their development of smart systems.

References

[1] Christiansen, B., “The What, Why, and How of Visual Inspections: Processes and Benefits,” Limble CMMS, https://limblecmms.com/blog/visual-inspection/. Accessed 1 May 2023.

[2] See, J. E., Drury, C. G., Speed, A., Williams, A., & Khalandi, N. (2017). “The Role of Visual Inspection in the 21st Century,”

Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61(1), 262–266. doi:10.1177/1541931213601548

[3] "A Guide to Visual Inspection." Rolls-Royce, https://suppliers.rolls-royce.com/GSPWeb/ShowProperty. Accessed 9 May 2023.

[4] Callipm. "Vision Inspection in the Food & Beverage Industry," Callipm, 2023, https://callipm.com/blog/vision-inspection-in-the-food-and-beverage-industry/. Accessed 12 May 2023.

[5] Kimura, M., Yanagihara, T., (2019). "Anomaly Detection Using GANs for Visual Inspection in Noisy Training Data," In: Carneiro, G., You, S. (eds) Computer Vision – ACCV 2018 Workshops. ACCV 2018. Lecture Notes in Computer Science, Vol 11367. Springer, Cham. https://doi.org/10.1007/978-3-030-21074-8_31

[6] Bauer, A. (2022). Self-Supervised Training with Autoencoders for Visual Anomaly Detection. arXiv preprint doi.org/10.48550/arXiv.2206.11723

[7] Simonyan, K., & Zisserman, A. (2015). "Very deep convolutional networks for large-scale image recognition," The 3rd International Conference on Learning Representations (ICLR2015)

[8] Szegedy, C., et al. (2015). "Going deeper with convolutions," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-9).

[9] He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep residual learning for image recognition," In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[10] Dharmawan, I.D.M.O.; Lee, J.; Sim, S. "Defect Shape Classification Using Transfer Learning in Deep Convolutional Neural Network on Magneto-Optical Nondestructive Inspection," Applied Sciences. 2022, 12, 7613.

[11] Cho, B.-H., Byun, S.-H., Shin, C.-H., Yang, J.-B., et al., (2004). "KeproVt: underwater robotic system for visual inspection of nuclear reactor internals," Nuclear engineering and design, vol. 231, no 3, p. 327-335.

[12] Tout, K., (20018). "Automatic vision system for surface inspection and monitoring: Application to wheel inspection," PhD Thesis Université de Technologie de Troyes, French 2018, https://www.theses.fr/2018TROY0008. Accessed 12 May 2023.

[13] ML2Grow. "Improve Your Manufacturing with Computer Vision," https://ml2grow.com/improve-your-manufacturing-with-computer-vision/. March 8, 2023

[14] Intel Corporation. "Your Industrial Physical Environment," Intel Corporation,2023, https://www.intel.com/content/www/us/en/edge-computing/resources/eii-and-your-industrial-physical-environment.html. Accessed 12 May 2023.

[15] National Instruments. "A Practical Guide to Machine Vision Lighting," National Instruments Corporation,2019, https://www.ni.com/en-lb/shop/choosing-the-right-hardware-for-your-vision-applications/a-practical-guide-to-machine-vision-lighting.html. Accessed 15 May 2023

[16] Bigas, M., Cabruja, E., Forest, J., and Salvi,J., (2006) "Review of CMOS image sensors," Microelectronics Journal, vol. 37, no. 5, pp. 433 451, 2006.

http://www.sciencedirect.com/science/article/pii/S0026269205002764

[17] Harding, K. G., (2003), "Machine vision lighting," In Encyclopedia of Optical Engineering. CRC Press, 2003, pp.12271336

[18] Kempainen, S., (1997) "CMOS image sensors: eclipsing ccds in visual information?", EDN, vol. 42, no. 21, pp.101120, 1997.

[19] Tikhe, C., and Chitode, J., S., (2014), "Metal surface inspection for defect detection and classification using Gabor Filter," International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, No 6, 2014

[20] Devi, T. Gayathri, P. Neelamegam, and S. Sudha. (2017) "Machine vision based quality analysis of rice grains," 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), IEEE.

[21] Wang, Z., and Bovik, A., (2004), "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Processing. Vol. 13, No 4, 2004

[22] Miura, M., Sakai, S., Aoyama, S., Ishii, J., et al., (2012) "High-accuracy image matching using phase-only correlation and its application," in: SICE Annual Conference, 2012, pp. 307–312

[23] Suvdaa1, B., Ahn, J., and Ko, J., (2012) "Steel surface defects detection and classification using SIFT and voting strategy," International Journal of Software Engineering and Its Applications. Vol. 6, No 2, 2012.

[24] Lowe, D., (2004), "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, Vol. 60, No 2, 2004, pp 91–110

[25] Huang, B., et al. (2018) "Research and implementation of machine vision technologies for empty bottle inspection systems," International journal of Engineering science and technology, Vol. 21, pp 159-169.

[26] Kazemian, A., et al. (2019) "Computer vision for real-time extrusion quality monitoring and control in robotic construction," Automation in Construction, Vol. 101, pp. 92-98

[27] Zhang, X., and Wang, L. (2004). "An intelligent real-time vision system for surface defect detection," In Proceedings of the 2004 IEEE International Conference on Pattern Recognition, pp. 636-640.

[28] Yıldız, K., Buldu, A., and Demetgul, M. (2016). "A thermal-based defect classifcation method in textile fabrics with Knearest neighbor algorithm," Journal of Industrial Textiles, Vol. 45, No 5, pp. 780–795

[29] Wu, G., Zhang, H., Sun, X., Xu, J., and Xu, K. (2007). "A Bran-new Feature Extraction Method and its application to Surface Defect Recognition of Hot Rolled Strips," In 2007 IEEE International Conference on Automation.

[30] Kumar, A., and Natarajan, K. (2011). "Vision inspection system for the identification and classification of defects in MIG welding joints," International Journal of Advanced Manufacturing Technology, Vol. 57, pp. 1849-1857

[31] Asfaw, A., Mohammed, A., and Alemu, A. (2022). "An automatic water bottle defect detection system using deep learning approach," Master thetis, Bahir Dar University, Ethiopia

[32] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "ImageNet classification with deep convolutional neural networks," In Advances in neural information processing systems, pp. 1097-1105.

[33] Mohanty, S. P., Hughes, D. P., and Salathé, M., (2016) "Using deep learning for image-based plant disease detection," In Frontiers in Plant Science, Vol. 7, 2016, doi: 10.3389/fpls.2016.01419

[34] Yang, C., Chen, J., Li, Z., and Huang, Y. (2021) "Structural Crack Detection and Recognition Based on Deep Learning," Applied Sciences. Vol. 11, 2021.

[35] Wang, J., Fu, P., and Gao, R. X. (2019). "Machine vision intelligence for product defect inspection based on deep learning and Hough transform," Journal of Manufacturing Systems, Vol. 51, pp. 52–60. doi:10.1016/j.jmsy.2019.03.002

[36] Gizaw, A., (2022), "Water Bottle Defect Detection Using Computer Vision" Master's thesis, Bahir Dar University, 2022.

[37] Schlüter, M., and Krüger, A. (2019). "Supervised learning vs. unsupervised learning: A comparison for optical inspection applications in quality control," In 2019 International Conference on Image Analysis and Processing (ICIAP) (pp. 1014-1021).

[38] Li, Z., Li, J., Wang, Y., and Wang, K. (2019). "A deep learning approach for anomaly detection based on SAE and LSTM in mechanical equipment". Mechanical Systems and Signal Processing, Vol. 133, 2019.

[39] Mei, S., Yang, H., and Yin, Z. (2017). "Unsupervised-learning-based feature-level fusion method for mura defect recognition," IEEE Transactions on Semiconductor Manufacturing, Vol. 30, No 3, pp. 369-378, 2017

[40] Weston, J., and Watkins, C. (1998). "Multi-class support vector machines," (Technical Report CSD-TR-98-04). Department of Computer Science, Royal Holloway, University of London, Egham, UK.

[41] Schlegl, T., Seeböck, P., et al., (2017). "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," arXiv preprint arXiv:1703.05921.

[42] Cui, Y., Zhang, X., Adak, C., Gong, X., and Yan, S. (2022). "A survey on unsupervised visual industrial anomaly detection algorithms," arXiv preprint arXiv:2204.11161.

[43] Kitov AI. "CAD2SCAN" Kitov AI, https://www.kitov.ai/products/cad2scan/, accessed May 12, 2023

[44] Inspekto. "Inspecto S70 Industrial," Inspekto, https://inspekto.com/inspekto-s70-industrial/, accessed May 12, 2023

[45] Cognex. “In-Sight 2800 Vision System,” Cognex, https://www.cognex.com/products/machine-vision/2d-machine-vision-systems/in-sight-2800, accessed May 12, 2023.

[46] Vitronic. “V-5000.” Vitronic, https://www.directindustry.com/prod/vitronic-machine-vision/product-64309-1443053.html, accessed May 12, 2023

[47] Jidoka Technologies. « Products.” Jidoka Technologies, www.jidoka-tech.ai/products/. Accessed 12 May 2023.