



Article

Automated Fillet Weld Inspection Based on Deep Learning from 2D Images

Ignacio Diaz-Cano ^{1,*}, Arturo Morgado-Estevez ^{2,†}, José María Rodríguez Corral ^{1,†}, Pablo Medina-Coello ^{3,†}, Blas Salvador-Dominguez ^{2,†} and Miguel Alvarez-Alcon ^{3,†}

¹ Department of Computers Engineering, 11519 Cadiz, Spain; josemaria.rodriguez@uca.es

² Department of Automatic, Electronic, Computer Architecture & Communication Networks Engineering, 11519 Cadiz, Spain; arturo.morgado@uca.es (A.M.-E.); blas.salvador@uca.es (B.S.-D.)

³ Department of Mechanical Engineering and Industrial Design, 11519 Cadiz, Spain; pablo.medina@gm.uca.es (P.M.-C.); miguel.alcon@uca.es (M.A.-A.)

* Correspondence: ignacio.diaz@uca.es

† Current address: School of Engineering, University of Cadiz, University Avenue s/n, 11519 Cadiz, Spain.

Abstract: This work presents an automated welding inspection system based on a neural network trained through a series of 2D images of welding seams obtained in the same study. The object detection method follows a geometric deep learning model based on convolutional neural networks. Following an extensive review of available solutions, algorithms, and networks based on this convolutional strategy, it was determined that the You Only Look Once algorithm in its version 8 (YOLOv8) would be the most suitable for object detection due to its performance and features. Consequently, several models have been trained to enable the system to predict specific characteristics of weld beads. Firstly, the welding strategy used to manufacture the weld bead was predicted, distinguishing between two of them (Flux-Cored Arc Welding (FCAW)/Gas Metal Arc Welding (GMAW)), two of the predominant welding processes used in many industries, including shipbuilding, automotive, and aeronautics. In a subsequent experiment, the distinction between a well-manufactured weld bead and a defective one was predicted. In a final experiment, it was possible to predict whether a weld seam was well-manufactured or not, distinguishing between three possible welding defects. The study demonstrated high performance in three experiments, achieving top results in both binary classification (in the first two experiments) and multiclass classification (in the third experiment). The average prediction success rate exceeded 97% in all three experiments.

Keywords: CNN; surface inspection welding; shipbuilding; FCAW; GMAW; welding defects; deep learning; geometric deep learning; yolo



Academic Editors: Lorenzo Seidenari and Alessandro Rozza

Received: 23 November 2024

Revised: 31 December 2024

Accepted: 7 January 2025

Published: 17 January 2025

Citation: Diaz-Cano, I.; Morgado-Estevez, A.; Rodríguez Corral, J.M.; Medina-Coello, P.; Salvador-Dominguez, B.; Alvarez-Alcon, M. Automated Fillet Weld Inspection Based on Deep Learning from 2D Images. *Appl. Sci.* **2025**, *15*, 899. <https://doi.org/10.3390/app15020899>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the context of welding production, a crucial aspect that demands significant attention is the subsequent inspection of the work performed. This activity necessitates a duration that is longer than that of the welding task itself.

Nowadays, the welding process is established in many sectors of the industry. For each of these industries, there are a number of ISO standards rules that must be met for welding work to be accepted. In certain industries, such as the naval sector, welding is of paramount importance. Consequently, the significance of welding testing tasks must be commensurate with that of the welding task itself. The financial burden of welding inspection, encompassing both temporary and economic costs, constitutes a substantial percentage of the overall welding expenditure [1].

1.1. Background

However, the task of visual inspection of the weld begins once the weld seam has already been created, so for a study of this type, what is truly important is to apply some type of technique to check the goodness of the created seam. In this sense, there are several techniques to carry out this inspection work, within what are called non-destructive tests: ultrasonic testing, radiographic testing, magnetic testing, and the simplest technique that involves visually observing the welding carried out by an expert, among others.

Although radiographic testing is the most widely used technique, direct visual inspection by an expert is still common. In any case, due to the number of meters of welding that must be checked and the large amount of time that must be devoted to this task, it is necessary to automate it to the maximum, achieving acceptable automation reliability.

Given the significance and pertinence of the processes, as well as their extensive utilization within the industry, this study concentrates on the examination of the FCAW and GMAW welding processes.

In this way, FCAW is an electric arc welding process in which the arc is established between the continuous tubular electrode and the piece to be welded. The protection is carried out through the decomposition of raw materials contained in the tube, with or without external gas protection and without applying pressure [2], as illustrated in Figure 1. On the other hand, GMAW is an arc welding process that requires automatic feeding of an electrode protected by a gas that is supplied externally [3], as demonstrated in Figure 2.

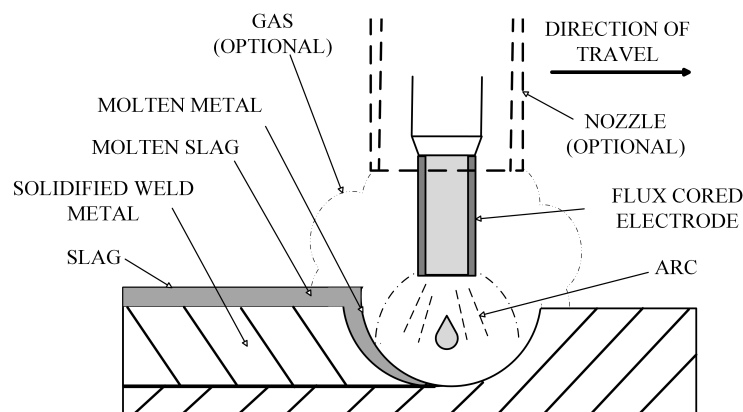


Figure 1. Scheme of the FCAW welding process.

Undoubtedly, the study of the input parameters in robotic welding, as is the case with the works presented in [4,5], the quality of the weld bead is contingent upon the presence of a particular factor, as the authors mention in [4–6]. However, the present study will concentrate on the implementation of a neural network in the analysis of welding seams, with the objective of detecting potential defects and identifying seams that have been executed with a high standard of quality.

The work carried out can be aligned with a geometric deep learning model in accordance with the indications presented by the authors in [7]. It has been demonstrated that these models encompass those dedicated to detecting objects based on a CNN when the objects must also be detected in different positions and angles, as discussed in this article. This process facilitates the identification of works related to the presented topic that have been resolved using a CNN approach.

In order to achieve a correct vision test of the welding parts, it is necessary to prepare the area to be welded beforehand, taking into account the consumable to be used in the welding, as well as the type of steel to be used in it. The proportion of gas is also a

critical factor that must be taken into consideration in each of the welding processes to be carried out.

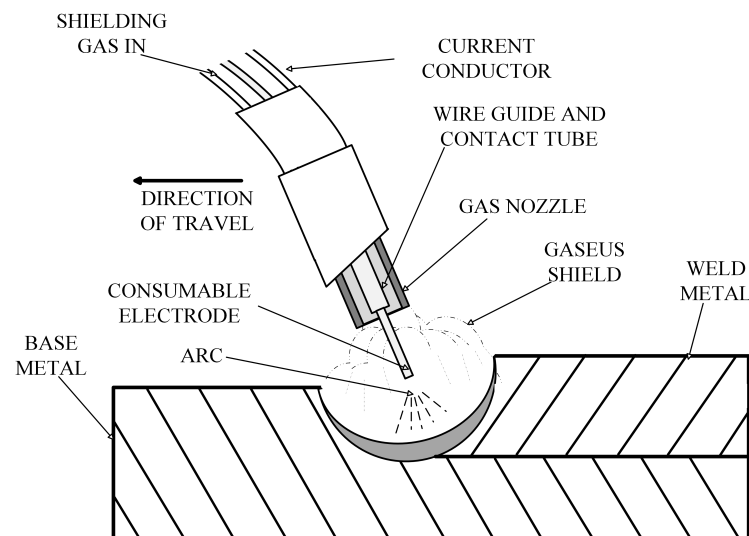


Figure 2. Scheme of the GMAW welding process.

1.2. Related Works

In recent years, a range of methods and models based on the concept of neural networks, in their various variants, have been proposed by other researchers to address the tasks of inspection, monitoring, or diagnosis of a weld. These methods have been employed in various welding processes across multiple sectors. A study presented in Liu [8] employed an artificial neural network (ANN) for the purpose of predicting welding residual stress and deformation in the context of electro-gas welding.

Thus, the authors propose in [9] a real-time monitoring system for laser welding of tailor rolled blanks using CNN. As is the case of [1], where the authors propose the detection of weld defects using a type of CNN called Faster R-CNN, trained with X-ray images of the weld seams. As indicated in reference [10], a welding surface Inspection of Armatures system is demonstrated through the use of CNN and image comparison. In a similar manner, the authors of [11] utilize a deep learning (DL) technique to detect defects in the gas metal arc welding (GMAW) process. The authors conducted ultrasonic, magnetoscopic, and penetrant liquid tests on the welds, though they do not share the dataset used.

The reviewed studies sometimes use their own images to train their artificial neural networks. On other occasions, the works themselves, such as the one presented, have created a dataset of images for the occasion, as is the case of [12], where a database of X-ray images corresponding to five categories is incorporated: castings, welds, baggage, natural objects, and settings. As delineated in [13], the system was developed for the specific purpose of facilitating welding training. It involves the use of a camera-based system that captures a series of images of welding seams. These images are then processed using a CNN to train the system. In the case of [14,15], the authors propose an unsupervised learning model using their own radiographic images and other calculations. In the same way, the authors in [16] create their own dataset of radiographic images to train a DL system based on VGGnet [17].

Traditionally, neural network architectures for object detection are often divided into two categories called one-stage or two-stage. In two-stage detector methods, a Region Proposal Network (RPN) is created. RPN is used for generating high-quality object proposals, which are very vital to improving the performance of detectors [18,19]. This RPN is used to

generate the Regions Of Interest (ROI), proposed in the first stage. These ROI proposals are subsequently used for classification and bounding box regression in the second stage. Two-stage methods have the disadvantage that they are usually slower; however, they have higher precision than one-stage methods. On the other hand, one-stage detectors do not preselect to generate candidates like two-stage methods. They directly attempt to detect the objects by solving a simple regression problem. That is why one-step methods are faster in their resolution but produce results with less precision [20]. In the same way, Single Shot Detection (SSD) is presented by its authors [21] as a method to detect objects in images using a single deep neural network in a single stage. In this case, the output bounding boxes are discretized into a set of already predefined boxes using various aspect ratios and scales, locating the position on the feature map. The prediction of the network yields scores for each category of object in each default box, and it produces adjustments to the box to better match the shape of the object.

Thus, within the one-stage methods studied, the most interesting and traditionally used, which can be highlighted, are YOLO [22]. This method is responsible for solving the object detection problem by applying a regression to bounding boxes that are separated in space, together with the application of class probabilities associated with the boxes. From this point, a single neural network predicts class probabilities and bounding boxes from full images. RetinaNet is a one-stage method; however, it is able to achieve the precision of two-stage methods without losing the speed of its class methods because it addresses the class imbalance between foreground and background [20].

Methods based on a two-stage approach are more common and widespread within the world of DL object detection. Thus, R-CNN [23] is made up of three parts. In the first, proposals for independent candidate regions of each category are generated. The second part is made up of one that extracts features of fixed length. The third part is a set of SVMs. The second module is a large convolutional neural network that extracts a feature of a fixed-length vector of each region. The third module is a set of class-specific linear Support Vector Machines (SVMs). R-FCN, where fully convolutional region-based networks are presented for accurate and efficient object detection. Unlike other models like Fast/Faster R-CNN, which apply a costly subnet per region hundreds of times, in this case the calculations are shared across the entire image, proposing scoremaps relative to position [24]. In the case of Fast R-CNN, this method is based on efficiently classifying object proposals using deep convolutional networks. Compared to previous works, Fast R-CNN brings innovations to improve training and test speed while increasing object detection accuracy [25]. As delineated in the extant literature [26], the Faster R-CNN model is presented with an RPN that exhibits convolutional characteristics analogous to those of the detection network. We already know that an RPN is a fully convolutional network capable of predicting both object limits and object scores at each position. In this case, the RPN generates high-quality region proposals, which Fast R-CNN uses for detection, merging RPN and Fast R-CNN into a single network that shares features, making a substantial improvement on the latter. Alternatively, Mask-RCNN [27] is a refinement of Faster R-CNN, incorporating an additional branch to predict an object mask in parallel with the existing branch that recognizes bounding boxes.

A review of the extant bibliography yielded several articles addressing the creation and provision of image datasets focused on the detection of weld defects. Thus, in the first place, we mention the GDXray database, free to use, that the authors present in [12]. It is a database of images taken with X-rays, intended to perform non-destructive tests, not being explicitly of welding defects.

The authors of [24], while not releasing the image repository to the public, propose the establishment of an image database, in which the frequency is even mentioned, for the

detection of defective spot welds. They focus on studying performance in small datasets and with unbalanced classes, the images being taken with a high-resolution camera. In the same way, taking the images with a high-resolution color camera, in the work presented in [10], the authors propose a welding surface inspection of armatures. They use a CNN-based approach, comparing various CNN models against each other. In both cases, the authors do not make the image dataset public so that it can be studied in other research.

The investigation carried out in [28] shows radiographic imaging to detect welding defects in shipbuilding. Although the authors show us the number of images taken, the percentage composition of the train, validation, and test groups, a reference to the dataset under consideration is noticeably absent.

Definitely, some previous research has been found for automated recognition of surface weld defects based on DL. In few works, new datasets are used, or the data used for training are not offered publicly. The GDXray database is the most frequently used and accessible of the datasets, as referenced in the previous paragraph.

1.3. Contribution

Thus, to our knowledge, there is no study for the detection of fillet weld defects using convolutional neural networks trained with high-quality 2D images. Therefore, the main contributions of the work are the following:

- A set of images depicting weld seams has been developed and made available to the scientific community. This set includes images of seams that have been deemed acceptable as well as seams that exhibit defects, such as a lack of penetration or undercuts. These images have been taken with FCAW and GMAW welding differently. The images are captured with high-precision and high-quality camera instead of X-ray images, as has traditionally been done in other works.
- The development of a methodology with a series of steps that can be used in other research dealing with the detection of objects through 2D images.
- The study and results of three experiments based on the application of neural networks through the analysis of 2D images: the detection of the type of FCAW/GMAW weld, the verification of the goodness of a weld, and the detection of certain defects in a weld seam.

1.4. Organization

Section 2, where we will detail the materials and the process that has been established to define the methodology and the steps that have been followed to carry out this research. Next, in Section 3, the results and conclusions of this work will be presented. Ending with Section 4, include conclusions and future work.

2. Materials and Methods

Throughout this section, we will detail the framework used in the research. In our experimentation we need a framework that allows us to create the welding seams and another one that is capable of allowing us to acquire images to train the neural network. Subsequently, a thorough exposition of the YOLOv8 object detection framework will be furnished. Finally, the methodology used in the research will be explained. These steps will be described appropriately so that they can be used in other projects or research related to object detection.

2.1. Framework

The case study focuses on the manufacture of fillet welds on steel plates, using two welding techniques, FCAW and GMAW. Thus, the FCAW and GMAW welding processes

were configured on a FANUC LR Mate 200 iD 7L robotic arm (Fanuc, Japan), Figure 3a. The apparatus utilized in this experiment is a Lincoln R450 CE welding system (Lincoln Electric, Cleveland, OH, USA), as illustrated in Figure 3b. This system is equipped with the necessary wire and gas to achieve a precise weld, in accordance with the established processes. The shield gas combination used in the manufacture of the weld seams was composed of 80% argon and 20% carbon dioxide for the GMAW welding process and 75% argon and 25% carbon dioxide for the FCAW process. This combination of gases was employed with the objective of combining the optimal performance of each gas in a welding process, namely CO₂ (enhanced penetration) and argon (improved protection). This approach was informed by the specific characteristics of the material to be welded (carbon steel). Furthermore, the reliability of the weld and its final quality are more balanced. This is evidenced by the following observations in [29].



(a) Robotic arm

(b) Multi-Process Welding Machine

Figure 3. On the left side, Fanuc 200i-D 7L robotic arm equipped with a welding torch. In the background of this image, you can see the gas bottles (Argon/Carbon Dioxide) in their conveniently mixed proportions. On the right side is the Lincoln R450 CE Multi-Process Welding Machine, placed under the table of the robotic arm and connected to it.

The SC-420MC titania flux-cored wire is suitable for all-position welding with either 100% CO₂ shielding gas or Ar-20-25% CO₂ shielding gas. The reduced spattering and enhanced slag detachability result in a shorter bead grinding operation [30].

On the other hand, the SM-70 eco wire, a general-purpose AWS ER70S-6 copper-clad solid wire suitable for manual and semi-automatic applications, was utilized in this investigation for GMAW weld seams. Its extensive utilization in a multitude of industrial sectors, including structural fabrication, automotive, heavy machinery, and shipbuilding, attests to its versatility and reliability [31]. Moreover, the steel employed in the experiments was 6 mm thick carbon steel, designated S275JR [32]. Weld beads of approximately 5–7 cm were created by L-shaped fillet welding, as can be seen in Figure 4, where five weld beads can be seen. The welding wires and steel utilized in this process are standard materials commonly employed in the shipbuilding industry.



Figure 4. Steel plate where numbered seams were welded and then treated according to the experiment to be carried out.

Subsequently, each seam was numbered and labeled as acceptable or not acceptable. Furthermore, within the non-acceptable option, its manufacturing defect was labeled. The labels were assigned by two human welding experts, who inspected each seam visually, as would be done in typical surface inspection welding field work.

The subsequent stage of the study will entail the acquisition of a series of images of the aforementioned welds, with the objective of training a neural network. Once this network has been trained, a single-step Deep Learning architecture, YOLOv8, will be used to predict the type of weld performed and the welding defects found. Thus, the pictures have been taken in different tones of light, aided by the luminaire that appears in Figure 5a. The process of image capture and its associated positioning was automated by means of a high-resolution camera placed at the end effector of a robotic arm, as illustrated in Figure 5b.



(a) Luminaire Effilux EFFI-FD-200-200-000



(b) Fanuc LR Mate 200 iD Robot

Figure 5. Equipment utilized for the capture of images in various positions and luminosities included a high-precision camera affixed to the end effector of the robotic arm and a luminaire positioned in different locations, contingent upon the intended image, with the objective of attaining a series of images exhibiting the most diverse range of luminosities feasible, thereby facilitating a more comprehensive training experience.

2.2. YOLOv8 Architecture

A thorough examination of the extant literature on related methods was conducted, leading to the determination that the YOLO algorithm would be employed due to its capacity to integrate a greater number of characteristics than those initially proposed and its exceptional adaptability to the images contained within the datasets. In addition, the authors of [33] demonstrate the efficacy of the method in identifying defects in welds. This task is analogous to the present one, but it is conducted using radiographic images. A comparative analysis of various algorithms is conducted, with YOLO emerging as the most effective option.

In [34], the authors undertake a comparative analysis of all iterations of the selected algorithm. It is evident that, of all the iterations of YOLO, the YOLOv8 version has been the most effective in achieving optimal results, as evidenced by its superior performance in aligning with the aforementioned characteristics. While acknowledging the shortcomings of YOLO in consistently achieving optimal performance in detecting small objects, the YOLOv8 model was meticulously engineered to address these limitations. Its design aims to enhance localization accuracy and optimize the performance-to-computing ratio.

This renders it particularly well-suited to the purposes of this study [35]. This version was released by a company, Ultralytics, that develops tools to build, optimize, and implement deep learning models. In this version, all the characteristics of YOLO—performance, precision, robustness, efficiency, etc.—are far superior to its predecessors.

Consequently, the YOLOv8 architecture has been selected for further investigation. The YOLOv8 model has been developed to incorporate new features and enhancements over its predecessor, YOLOv5 [22], with the objective of achieving an optimal speed-accuracy trade-off. The YOLOv8 architecture comprises five distinct models, designated as “n”, “s”, “m”, “l”, and “x”, which vary in size and thus offer different levels of computational complexity and parameter requirements. The “n” model is the most lightweight, making it particularly well-suited for processes with limited computational resources. The YOLOv8 architecture is composed of three fundamental components: the backbone network, the neck, and the head. The backbone network is responsible for feature extraction, while the neck component integrates these features. Finally, the head component is responsible for object localization and regression.

2.3. Preparation of Fillet Weld Test Piece

To conduct a visual inspection aligned with weld characteristics, the recommendations and suggestions that appear in the ISO 15792-3:2011 [36] standard were taken into account. The standard, entitled “Classification testing of positional capacity and root penetration of welding consumables in a fillet weld”, contains a series of suggestions necessary to be able to carry out an evaluation of the fillet weld in an appropriate manner.

The standard goes on to address general requirements and the type of metal to be used in the weld, as well as the use of different welding wires and their testing. It also covers the material and the thickness that should be used for this type of weld, as discussed in Section 2.1 according to the document of standard, the citation is as follows [36].

With regard to the preparation of the piece for welding and subsequent visual inspection, the standard establishes the following sections:

- Prior to the assembly process, the component must be placed in an L-shape, with two plates of the designated metal forming a 90° angle. This configuration is to be achieved by using a web and a flange, with dimensions of these components complying with those stipulated in section 5.1 of the electrode classification standard [37].
- The subsequent essay will examine the position and conditions of welding. In this instance, the recommendations of the ISO 6947 [38] standard were adhered to with

regard to electrode temperature and single-pass weld deposit. It is recommended that the reader consult Section 5.2 of the standard document ISO 15792-3:2011 [36].

- In order to comply with Section 5.3 of the standard document ISO 15792-3:2011 [36], welding speed recommendations were adhered to in accordance with the consumable used.
- It was determined that welding of the second side was not a prerequisite, thus rendering point 5.4 of the standard document ISO 15792-3:2011 [36] superfluous.

In accordance with the established protocol, the sections pertaining to the examination of the test piece (Section 6) of the document ISO 15792-3:2011 [36] were analyzed. The decisions and actions undertaken in this context are outlined below:

- In relation to the dimensions of the weld throat thickness, the robotic arm's welding parameters were utilized as the reference point. Subsequent to this, a measurement of the throat thickness was conducted, exhibiting a certain degree of flexibility with respect to the standard specifications. This resulted in a range of $\pm 10\text{--}15\%$ for the measurement data.
- In order to comply with the measurements and requirements established in Section 6.2 of the standard document ISO 15792-3:2011 [36], the parameters were established in the robotic arm that was used to perform the welding with maximum precision. The welding beads were then measured, as well as the throat thickness. A convex fillet weld was established.
- It is evident that Section 6.2 of the standard has not been given due consideration, as it was understood that it had already been incorporated into prior preparations in other sections.

Subsequent to the fabrication of the welding beads, the following inspections and measurements were conducted:

- Visual inspection of defects and correct welds.
- Measurement of fillet leg lengths.
- Visualization of the correct convexity of the fillet.
- Verification of the throat, as previously indicated to the welding robot in the parameters.

2.4. Methodology

The experiments in this study were conducted in accordance with a series of steps, as illustrated in Figure 6. These steps include data acquisition (DA), feature extraction (FE), data preprocessing (DP), data augmentation (DAU), hyper-parameter selection (HS), deep learning model (DLM), and finally, performance metrics (PM). Each of these steps is explained in detail below.

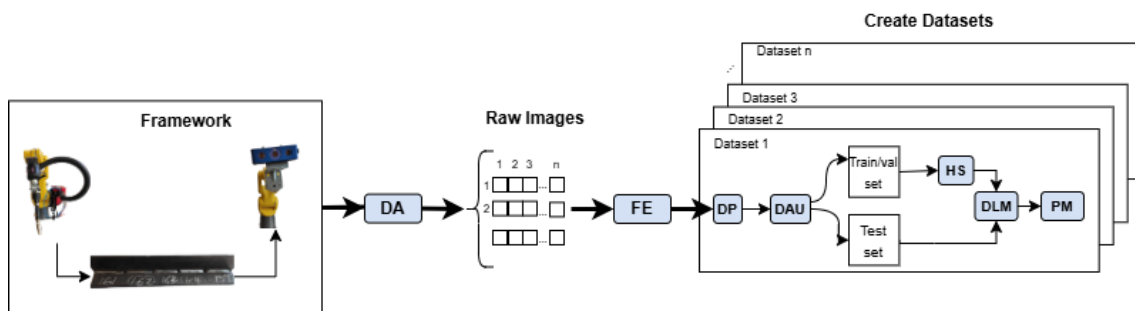


Figure 6. Diagram illustrates the methodological framework employed in this study. The process initiates with the fabrication of the welds necessary for the experimental studies, followed by the acquisition of images of these welds. Subsequently, a series of image transformations are performed to train three models, one for each experiment, capable of detecting the manufactured weld seams.

2.4.1. Data Acquisition (DA)

As is universally acknowledged within the field of training, it is imperative to employ the most accurate data possible. Consequently, the primary focus was directed towards the acquisition of 2D images of the welding seams, characterized by optimal quality and precision, in alignment with the experimental protocol that was to be executed.

The images, which constitute our initial data, represent the most crucial element of an object recognition system based on deep learning techniques. The images were captured using a high-resolution camera, as illustrated in Figure 7. The camera was placed on the end effector of a robotic arm, with the aim of taking a total of seven poses of each weld bead, with different luminosities in order to have more images for the training/validation/test dataset. To do this, a simple program was created in the robot, where it had seven stopping points, taking an image of the weld beads at each of these points.



Figure 7. Industrial camera brand Ensenso model N35 (IDS-IMAGING, Germany), used to take images of weld seams.

The camera was situated on the end effector of a robotic arm, with the objective of capturing seven poses of each weld bead of varying luminosities to augment the images available for the training, validation, and testing datasets. To this end, a rudimentary program was developed within the robot, delineating seven designated stopping points. To accomplish this objective, a rudimentary program was developed within the robot that incorporated seven designated stopping points. At each of these points, the robot captured an image of the weld beads, as illustrated in Figure 5b. The distance at which the images were captured fell within the operational range recommended by the camera manufacturer, with a minimum distance of 270 mm and a maximum distance of 3000 mm [39]. Consequently, images were captured from a variety of positions and distances, ranging from 1200 mm to 1700 mm. The dimensions of the images were consistent across all captures, aligning with the size recommended by the camera for optimal quality.

The three experiments that have been conducted differ in terms of the classes to be recognised and the manner in which the images are treated. In the initial experiment, weld seams were extracted in great detail, with two classes being attempted to be identified. In the second and third experiments, the complete image was utilized, with all welding seams, resulting in the detection of two classes in experiment two and four classes in experiment three.

2.4.2. Feature Extraction (FE)

Subsequent to the acquisition of images, the subsequent task is the extraction of features from each image in order to detect the object under study, a weld bead. In order to accomplish this, it is necessary to pay close attention to the scenario to be proposed, in conjunction with the materials involved. This material stands out among other properties for its high tensile strength, which guarantees the integrity of the structure and serves as an excellent support for creating weld beads, making reliable joints [40]. A complicated scene

is therefore presented, because there is a steel plate or structure, joined by a weld bead, both with similar material and visual characteristics.

Therefore, it is essential to select the scene that the system wants to learn well. For this reason, the boundary box process, or labeling, has been carried out on each image of each dataset created, using an appropriate online tool for this, Roboflow [41]. In this way, it has been possible to label each of the characteristics that have been studied in the research in each dataset. Thus, for each dataset, a series of labels or classes will be specified in each raw image. These labels correspond to the learning and search characteristics you would like to train in the model.

As illustrated in Figure 8, a steel plate that has been welded and appropriately labelled is presented. This plate serves as an example to the system, indicating that welds that have been manufactured to a high standard are labelled 'good label', while welds that exhibit defects are labelled 'bad label'.



Figure 8. Mild steel plate with several welding beads, labeled with the online tool Roboflow, so that the system can detect a type of weld manufactured correctly compared to another weld manufactured with some defect.

2.4.3. Data Preprocessing (DP)

Although it may seem secondary, some image preprocessing techniques can help to detect the objects we intend to find in each scene [42]. Among other techniques, we can find grayscale, image self-orientation, contrast adjustment, and size readjustment.

In our case, the images have already been taken in high-quality grayscale, so it is not necessary to apply this preprocessing. However, it was considered to be one of the most important transformations since the system is able to converge in the same way to detect objects as if the image were taken in color, but investing less computing power due to the absence of color channels.

In the interest of optimizing the system's convergence, it is imperative to implement a uniform size adjustment to each image. To this end, a recalibration of image dimensions has been executed. Concurrently, a self-orientation has been applied to each image, thereby facilitating the system's training with a more robust pattern.

2.4.4. Data Augmentation (DAU)

Data augmentation is a highly effective method for creating useful DL models. This method ensures that the validation error should decrease along with the training error. It achieves this by representing a more complete set so that it is able to minimize the distance

between the training and validation data, including the test set [43]. The augmentation techniques that were applied included the following:

- **Horizontal flips:** this effect will reflect the images horizontally, increasing the variety of vertex orientations.
- **Shear:** add variability to perspective to help the model be more resilient to camera and subject pitch and yaw.
- **Noise:** the incorporation of noise is instrumental in enhancing the resilience of our model to camera artifacts.

These data augmentation techniques will improve the robustness of the model, particularly the detection of small objects such as welding seams. The training data set will be expanded with several variations on the original images. In this way, the model will learn to generalize better and to be more resistant to changes in lighting, image quality, or orientation, among other things.

In this study, the size of the data set has been intentionally limited, since it is a proof of concept. However, at the production level, this data set must be increased in order to obtain a more robust model, better prepared for changes and to increase its precision and generalization.

2.4.5. Hyper-Parameters Selection (HS)

In the context of a YOLO system, a series of hyperparameters are configurable in accordance with the training to be performed. Table 1 lists the hyperparameters that have been used in each of the three experiments performed in this research.

Table 1. Hyperparameters used in each of the experiments performed.

Parameter	Experiment 1	Experiment 2	Experiment 3
<i>epochs</i>	105	300	300
<i>batch size</i>	16	16	16
<i>Learning rate</i>	0.01	0.01	0.01
<i>Optimizer</i>	SGD	SGD	SGD
<i>Input image size</i>	320 × 320	320 × 320	320 × 320
<i>Confidence Threshold</i>	0.75	0.75	0.75

In order to select the hyperparameters for each experiment, it has been taken into account that they are as homogeneous as possible so that a comparison and discussion of the results can be made as fair as possible later. Thus, only the number of epochs has been altered between one experiment and another, due to the complexity of the detection and the characteristics of the classes to be detected, which have led to a slower convergence of the model. In any case, the lowest possible number of epochs has been applied to avoid possible overfitting, taking into account that we have a limited dataset.

During the training process, a batch size of 16 was utilized, indicating that the data were updated following the concurrent processing of 16 images. The selection of this parameter was informed by the available resources. An increased batch size could enhance the efficiency of the training process; however, it may necessitate additional memory resources.

The learning rate was set to 0.01. This parameter controls the tuning of the model weights during each training iteration. Care must be taken in choosing it, as a higher learning rate might result in faster convergence, but at the same time it could lead to oscillations or exceeding the optimal weights.

The stochastic gradient descent (SGD) optimizer was employed. The SGD is a classic optimization algorithm that iteratively updates the model's weights based on the gradient of the loss function.

The input images were resized to 320×320 pixels, which is a lower resolution than the one commonly used for YOLO models. Due to the characteristics of the image to be detected, it was considered convenient not to invest in a larger image size in order to guarantee a consistent input size for the model and efficiency in terms of training time and memory allocated.

The confidence threshold is represented as the inverse of the significance threshold and is also often expressed as a percentage. Confidence determines how confident the model is that a prediction matches the true value of a class. The threshold determines the value to label a class as that class. Therefore, it can be posited that, in the event the confidence threshold is established at 0.6, the model will be required to attain a minimum confidence level of 60% in order to proceed with the classification of the object in question. In our case, we have required a confidence of 0.75, so we will have to ensure that the predictions are at least 75% certain. This value is interesting to use when we are sure that the model converges well with a high percentage of success.

Additionally, note that the MS COCO [44] weights have been used, which were passed as input parameters to the YOLO algorithm.

2.4.6. Deep Learning Model (DLM)

In the selection of a deep learning model, the primary criteria considered were the robustness, speed, accuracy, efficiency, and overall performance of the algorithm. Additionally, the datasets prepared for the various training sessions and the experiments to be conducted were taken into account. As delineated in Section 2.2, the architecture that optimally fulfills this function is Yolov8. In [34], the authors present the improvements that the different versions of YOLO have undergone over time, reaching version 8, which is the version employed in this research.

2.4.7. Performance Metrics (PM)

In order to demonstrate the rigor of this research and to be able to measure the performance of the YOLO algorithm used by making a fair and reliable comparison, the following metrics have been defined:

- Recall (R): It is also called sensitivity or TPR (true positive rate), representing the ability of the classifier to detect all cases that are positive, Equation (1).

$$Recall(R) = \frac{TP}{TP + FN} \quad (1)$$

TP (True Positive) represents the number of times a positive sample is classified as positive, i.e., correctly. On the other hand, FN (False Negative) tells us the number of times a negative sample is classified incorrectly.

- Precision (P): Controls how capable the classifier is to avoid incorrectly classifying positive samples. Its definition can be seen in Equation (2).

$$Precision(P) = \frac{TP}{TP + FP} \quad (2)$$

In this case, FP (false positive) tells us how many times negative samples are classified as positive.

- Intersection over union (IoU): is a critical metric in object detection as it provides a quantitative measure of the degree of overlap between a ground truth (gt) bounding box and a predicted (pd) bounding box generated by the object detector. This metric

is highly relevant for assessing the accuracy of object detection models and is used to define key terms such as true positive (TP), false positive (FP), and false negative (FN). It needs to be defined because it will be used to determine the mAP metric. Its definition can be seen in Equation (3).

$$\text{Intersection over union}(IoU) = \frac{\text{area}(gt \cap pd)}{\text{area}(gt \cup pd)} \quad (3)$$

- Mean Average Precision (mAP): in object detection is able to evaluate model performance by considering Precision and Recall across multiple object classes. Specifically, mAP50 focuses on an IoU threshold of 0.5, which measures how well a model identifies objects with reasonable overlap. Higher mAP50 scores indicate better overall performance.

For a more comprehensive evaluation, mAP50:95 extends the evaluation to a range of IoU thresholds from 0.5 to 0.95. This metric is appropriate for tasks that require precise localization and fine-grained object detection.

mAP50 and mAP50:95 are able to help evaluate model performance across multiple conditions and classes, thereby expressing information about object detection accuracy by considering the trade-off between Precision and Recall.

$$\text{Average Precision}(AP) = \int_{x=0}^{x=1} P(R) \cdot dR \quad (4)$$

Models with higher mAP50 and mAP50-95 scores are more reliable and suitable for demanding applications. These are appropriate metrics to ensure success in projects such as autonomous driving and safety monitoring.

Equation (4) shows the calculation of Average Precision, which is necessary to calculate the mean Average Precision (mAP), as demonstrated in Equation (5).

$$\text{Mean Average Precision}(mAP) = \frac{1}{N} \sum_{n \in N} AP(n) \quad (5)$$

- Box loss: this loss helps the model to learn the correct position and size of the bounding boxes around the detected objects. It focuses on minimizing the error between the predicted boxes and the ground truth.
- Class loss: the accuracy of object classification is of paramount importance. The system is designed to ensure that each detected object is correctly classified into one of the predefined categories.
- Object loss: object loss is responsible for choosing between objects that are very similar or difficult to differentiate, by better understanding their characteristics and spatial information.

3. Results and Discussion

In order to validate the suitability of the proposed methodology for recognizing welding seams will be validated according to three experiments presented in this research. The experimental environment and the obtained results are described and discussed in the following.

3.1. Experimental Environment

An experimental environment has been designed to execute the proposed methodology, with the methods and functions developed in Python. Additionally, a set of tools was utilized, including the algorithm YOLO v8 and Roboflow ([41] for labeling the images of the dataset created in this study). The labels in each experiment were awarded in accordance

with the visual assessment of two welding experts. These experts analyzed each of the welded plates and determined, based on their experience and criteria in visual inspection of welding beads, whether a weld was well manufactured. The welds were subsequently categorized as follows: GOOD, BAD, FCAW or GMAW, DEFECT (undercut (UNDER), lack of penetration (LOP), or rest of defects (OP)). Additionally, YOLO Ultralytics [45,46] has been used for performance metrics.

The dataset used in this study has been presented in Section 2.1. The suitably labeled images have been selected to be validated using the YOLO algorithm. This algorithm initially needs to learn the location of the object to be searched for within the set of images. To conclude the aforementioned points, 80% of the images have been provided for training and 20% for testing. From this set of 80%, 10% has been selected as the validation set.

Although more types of defects could have been detected in the images of the weld beads that make up the dataset, it has been decided to recognize the two most important ones, lack of penetration and underbite, leaving the rest for a class called “other problems (op)”. The rationale behind the exclusion of other defects is that the weld beads were made by a welding robot, ensuring their correct production. In the event of defective weld beads, typically attributable to improper parameter assignment, the defects produced are predominantly those that have been previously studied, with infrequent occurrences of other defects. The consideration of these defects individually would have resulted in a significantly reduced set of samples, potentially compromising the efficacy of the detection and classification of the weld seams.

The experiments presented in this study have been carried out on the facilities of the supercomputing center of the University of Cadiz, which has 48 nodes, each with two Intel Xeon E5 2670 2.6 GHz octa-core processors, equipped with 128 GB of RAM.

3.2. Experiment Results and Discussion

Throughout this subsection, each of the experiments that have been carried out to validate the effectiveness of using the YOLOv8 algorithm in detecting weld seams according to the proposed methodology will be explained. In each of the experiments, the obtained results are introduced both numerically and graphically.

Table 2 contains the numerical data for the metrics used to evaluate the performance of our study. For each of the experiments performed, Recall, Precision, and mAP are shown. The performance obtained for this last metric is shown both in the validation set and in the test set. Each of the experiments carried out is detailed below.

Table 2. Performance of the experiments performed in this study: Precision, Recall, and mAP. The latter taken on the validation set and on the test set.

Experiment	Class	Precision	Recall	mAP Val. Set	mAP Test Set
FCAW-GMAW weld seam	FCAW	0.951	0.979	0.99	0.99
	GMAW			0.99	0.97
GOOD-BAD weld seam	GOOD	0.982	0.985	0.99	0.93
	BAD			0.99	0.99
GOOD-LOP-UNDER-OP weld seam	GOOD	0.965	0.92	0.99	0.99
	LOP			0.77	0.94
	UNDER			0.99	0.92
	OP			0.99	0.99

3.2.1. Experiment 1: FCAW-GMAW Weld Seam

While it is possible to differentiate between FCAW and GMAW welding types at first glance, this task is difficult for those who are not trained in welding. For this reason, we have proposed the use of an automated system that can detect the type of welding used to create a particular welding bead.

For the present dataset [47], the weld seams were extracted from the plates illustrated in Figure 4, thereby ensuring that the model was trained on a labeled image predominantly containing the bead to be detected. The images in this dataset exhibit minimal additional space beyond the bead itself, and although the dimensions are standard (320×320), the appearance is irregular, as evidenced in Figure 9, which presents a set of these images. To address this challenge, two distinct classes of objects were identified for detection: FCAW and GMAW.

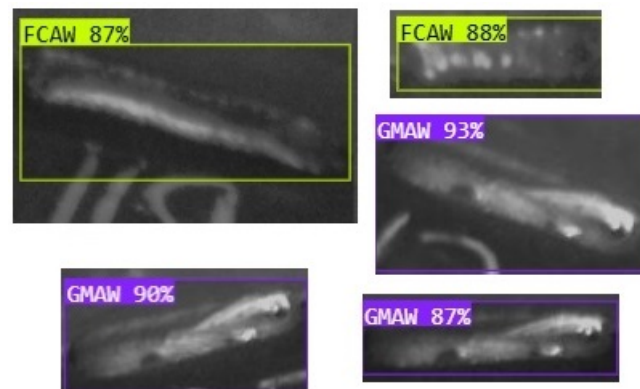


Figure 9. Set of images of the FCAW-GMAW dataset, where the predicted label and the percentage of that prediction can be observed. An irregular character of the image content can be observed, where the welding bead occupies practically all the space of the image.

As illustrated in the initial row of Table 2, the performance metrics demonstrate a consistent performance above 98% across all metrics. The test set of the mAP metric registered at 93%, which might have influenced the model's convergence time. However, it was observed that 100 epochs were sufficient for the model to reach convergence. The anticipated outcomes can be evaluated. Notwithstanding the presence of imbalance in the dataset, as demonstrated in Table 3, it is contended that the sample size is sufficient for the model to attain optimal performance.

Table 3. The number of labels classified by training, validation, and test sets is presented according to the number of classes in each experiment.

Proportion	Data	Experiment 1		Experiment 2		Experiment 3			
		Classes		Classes		Classes			
		FCAW	GMAW	BAD	GOOD	GOOD	LOP	UNDER	OP
80%	Train set 90%	1190	513	1182	507	564	313	539	322
	Val set 10%	132	57	131	56	62	35	60	36
20%	Test set	331	142	329	141	156	87	150	90
100%	Total data	1653	712	1642	704	782	435	749	448

Figure 10 shows the results after training of the first experiment, FCAW-GMAW detection. The graphs illustrate the training loss for the bounding box, segmentation,

classification, and confidence score predictions. The x-axis denotes the training epochs, with the y-axis representing the loss value (no units shown). At approximately 75 training epochs, a significant decrease is observed in all graphs, indicating that the model effectively learned to identify the weld bead. This is further confirmed by the decrease in loss related to the bounding box, segmentation, classification, and confidence score predictions. The model achieved good performance with high precision (around 0.98), recall (around 0.98), and mAP scores (reaching around 0.97 for mAP50). This suggests that the trained model was able to accurately detect the weld seam, whether it was created using the FCAW or GMAW technique. That is, the model was able to find characteristics specific to each welding method in order to differentiate them.

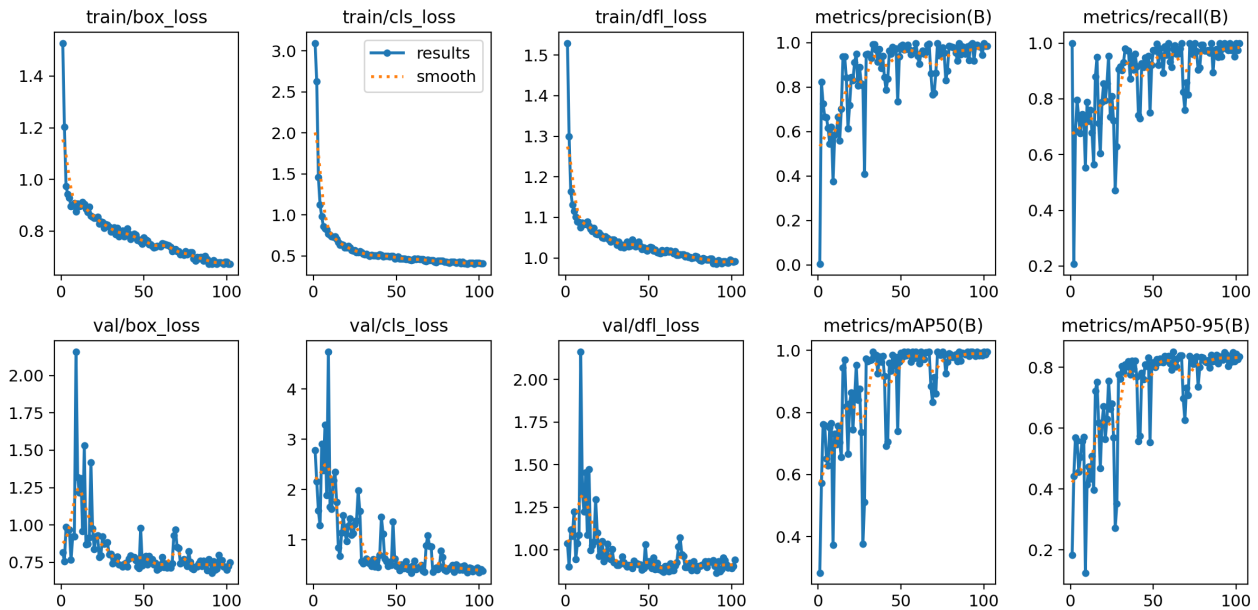


Figure 10. Training curves and performance metrics for the YOLOv8s object detection model trying to detect FCAW and GMAW weld seams. In all of them, we have the training epochs on the x-axis, while the y-axis represents the loss values, both without units. The curves show the learning of the model, observing a significant decrease in the loss while at the same time improving the precision, recall, and mAP50 scores, which leads us to think that the training has been effective.

3.2.2. Experiment 2: Good-Bad Weld Seam

In this second experiment, it was deemed expedient and intriguing to undertake a binary classification in which the model would discern a well-manufactured welding bead from a poorly manufactured one. This assessment was conducted without considering the welding process, a factor that was taken into account in Experiment 1.

In the dataset created for this experiment [48], labels (GOOD-BAD) have been established on each of the plates, which contain more than one welding bead. Correctly created beads and incorrectly created beads were included so that the model could obtain better learning, as can be seen in Figure 11.

As illustrated, in the second row of Table 2, the performance metrics obtained for this second experiment demonstrate consistent high scores, similar to those observed in the initial experiment. However, higher data in Precision and Recall, accompanied by lower data in the validation set of the “good” class, suggest that the model may encounter greater challenges in accurately classifying this category. This issue of detection, and the subsequent decline in performance, may be attributable to a reduced quantity of samples in relation to the ‘bad’ class, as illustrated in Table 3.

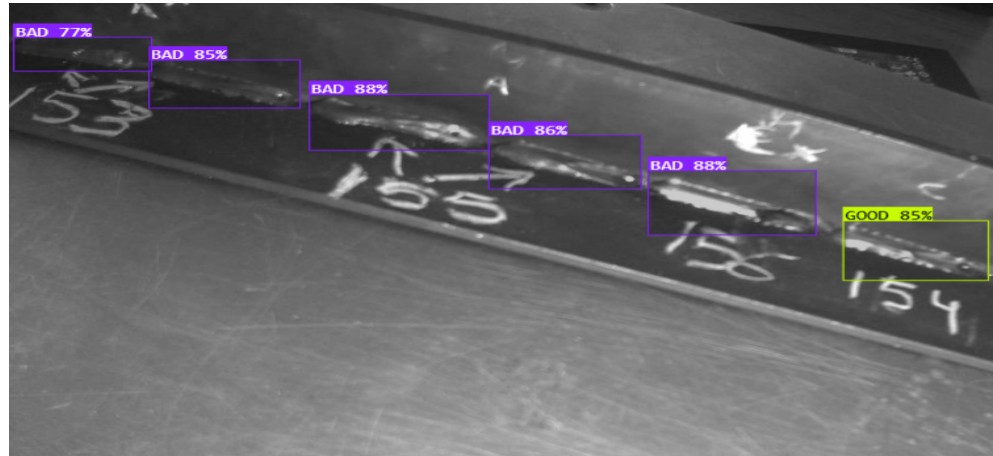


Figure 11. Plate of fillet weld beads where different beads can be seen, some labeled as GOOD and others as BAD, according to what the algorithm has learned once trained.

As illustrated in Figure 12 the training curves and performance metrics of the model trained in this second experiment demonstrate a convergence pattern that more regular and uniform way than in experiment 1. The curves in this case, as in Experiment 1, illustrate the training loss for the bounding box, segmentation, classification, and confidence score predictions. The x-axis represents the training epochs and the y-axis represents the loss value (unitless). However, it was necessary to execute 300 epochs in this experiment to achieve adequate model learning, with a significant decline being observed in all curves after 225 training epochs. This decline indicates the point at which the model commences optimal learning, thereby beginning to identify the weld bead within the environment of the full image. A decline in loss was once again noted for the bounding box, segmentation, classification, and confidence score predictions.

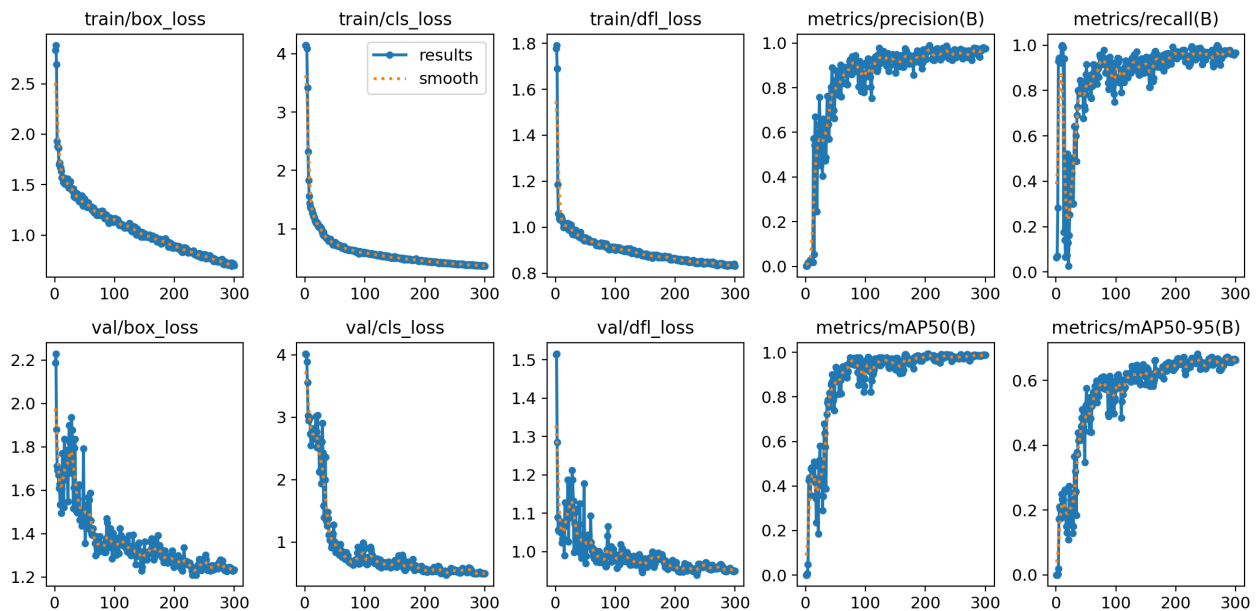


Figure 12. Training curves and performance metrics for the YOLOv8s object detection model trying to detect weld seams manufactured without defects (labeled as GOOD) and weld seams with some manufacturing defects (labeled as BAD). The x-axis shows the training epochs, while the y-axis shows the loss values, both without units. The curves show the learning of the model, observing a significant decrease in loss, while the precision, recall, and mAP50 scores improve, which leads us to think that the training has been effective.

3.2.3. Experiment 3: Good-Lop-Under Weld Seam

In order to carry out a more complex experiment, which could detect more than two classes and, at the same time, have an importance within the field we are investigating, we created this third experiment. The objective of this experiment was to conduct a preliminary study of the defects that may arise during the manufacturing process of weld beads, with a focus on identifying two of the most prevalent issues: undercuts and lack of penetration. This approach led to the establishment of an additional category for the detection of other welding problems, thereby distinguishing these welds from those that are well-manufactured. Consequently, a methodology was devised to detect a correctly manufactured weld, or the presence of defects, such as undercuts (UNDER), lack of penetration (LOP), and other issues (OP), in conjunction with the class of correct welds (GOOD).

It is acknowledged that there are other significant welding defects that warrant consideration, including but not limited to slag inclusions, porosity, incomplete fusion, spatter, overlap, and cracks. In this article, we sought to conduct a series of experiments (experiments 1 and 2) employing binary object detection, encompassing two distinct yet essential themes in welding operations. The third experiment was selected as a case study to demonstrate that, given the exposure of all experiments and methodologies, objects (weld beads) can be detected for more than two detection classes utilizing the same algorithm.

As illustrated in the second row of Table 2, the performance metrics for this second experiment demonstrate consistent high performance, similar to the results observed in the initial experiment. Notably, the precision and recall metrics exhibit higher values, accompanied by a decline in the validation set of the “good” class. This observation suggests that the model may encounter challenges in accurately classifying this category. The performance metrics for the remaining mAP metrics closely mirror those observed in prior experiments. An examination of the number of samples in each of the four classes, as presented in Table 3, reveals that the ‘LOP’ class contains the smallest number of samples. This could potentially result in the lowest performance in the mAP metric of the validation set; however, further analysis is necessary to ascertain the validity of this hypothesis.

In this third experiment, a dataset [49] was created for the purpose of detecting four distinct objects (GOOD-UNDER-LOP-OP), which correspond to well-manufactured welds (the first class) and various defects (the remaining three classes). Figure 13 illustrates a plate containing multiple fillet weld seams, each meticulously labeled with its respective model.

In this third experiment, a dataset was created for the purpose of detecting four distinct objects (GOOD-UNDER-LOP-OP), which correspond to well-manufactured welds (the first class) and various defects (the remaining three classes). Figure 2 illustrates a plate containing multiple fillet weld seams, each meticulously labeled with its respective model.

As illustrated in Figure 14, the training curves and performance metrics of the model trained in this third experiment demonstrate non-uniform trends, particularly in the initial epochs, where the model experiences difficulty adjusting the loss. This difficulty may be attributed to the model’s task of detecting four distinct classes, a challenge exacerbated by the limited number of samples available. The model achieves effective learning through a total of 300 epochs. As in the previous experiment, the learning stabilized after epoch 200 and reached maximum efficiency. As in the preceding experiment, an analysis of the loss revealed that the curves required more epochs to demonstrate effective training for bounding box, segmentation, classification, and confidence score predictions. However, a sufficient number of epochs were used in this experiment, as in the previous ones, to avoid overfitting.

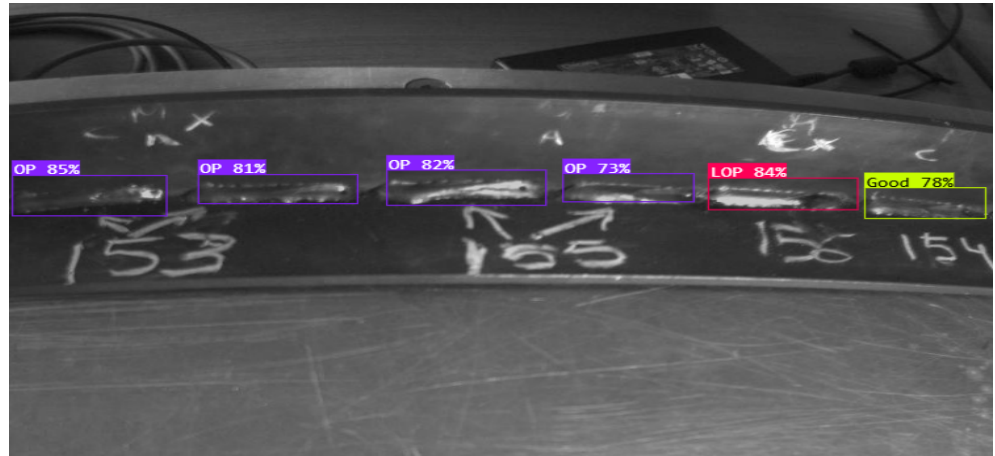


Figure 13. Plate of fillet weld beads analyzed with the model obtained in experiment 3. It shows three of the four types of weld beads (objects) for which the model of this experiment has been trained. In addition, the image shows other elements that the model is able to discard.

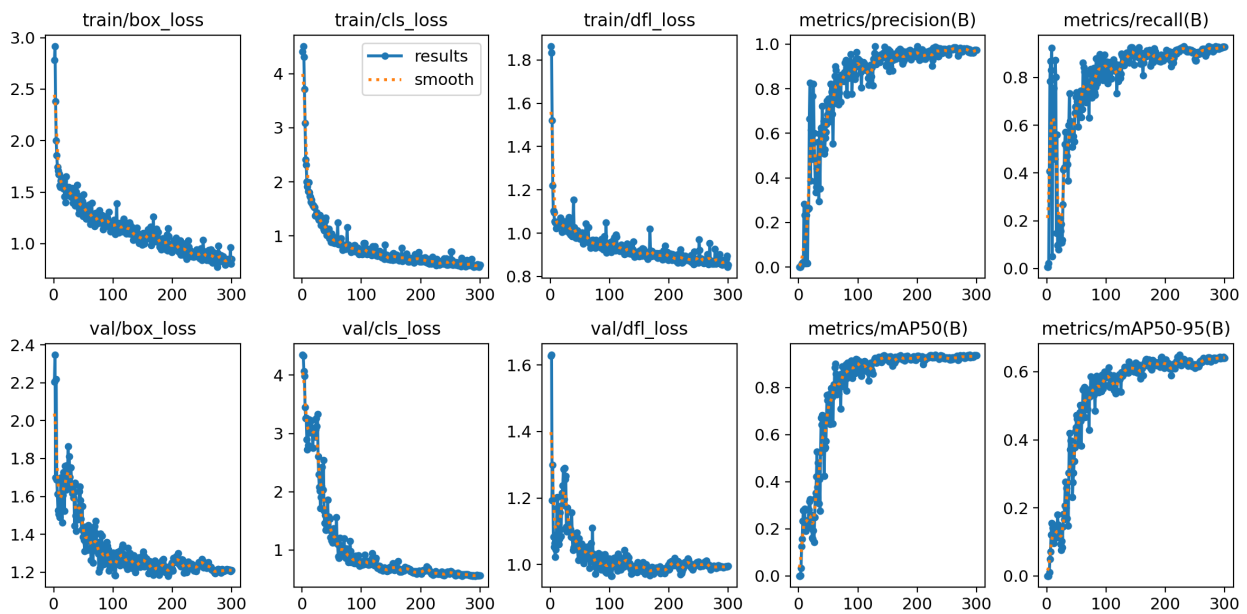


Figure 14. Training curves and performance metrics for the YOLOv8s object detection model, where we try to detect correctly made weld seams, without any defects (labeled as GOOD), and weld seams with some manufacturing defect, labeling and classifying several of these most common defects (labeled as UNDER for Undercuts, LOP for Lack Of Penetration, and OP for Other problems). The x-axis shows the training epochs, while the y-axis shows the loss values, both without units. The curves show the learning of the model; it is observed that the loss is significant, although somewhat milder than in the two previous experiments. Same case as in the precision, recall, and mAP50 scores that, although lower than before, we can deduce that the training has been effective.

4. Conclusions and Future Work

This paper presents the findings of an investigative study conducted to detect fillet weld beads in a series of 2D images, which were also captured during the course of this study. The images were subjected to a rigorous treatment process, resulting in the creation of multiple datasets. These datasets were then used to conduct a series of experiments aimed at detecting various types of welds, assessing the quality of the weld bead fabrication, and identifying defects. The object detection process has been focused using the YOLOv8

algorithm, with appropriate configuration of its hyperparameters and application of a specific methodology developed for this study.

Thus, after the three experiments carried out, it is demonstrated that the YOLOv8 algorithm is effective for the detection of fillet weld beads, since a prediction performance of over 97% is achieved in all the characteristics studied. However, as the complexity of detection has increased, with more challenging classes to detect or a greater number of class detections greater than two in the same experiment; performance has exhibited a notable decline. Nevertheless, there appears to be considerable potential for enhancing the performance of the model by developing a system capable of detecting a series of weld beads with multiple characteristics. Consequently, it can be deduced that this iteration of the algorithm is also efficacious in detecting smaller objects that differ from those typically identified by the algorithm.

This work signifies the genesis of the study of welding seam detection by implementing a geometric deep learning model with the YOLOv8 algorithm. The study encompassed two binary detection experiments and one multiclass detection experiment, which have been exhaustively delineated and deliberated. The outcomes substantiate the efficacy of the approach, signifying the potential for further research in this domain. It is evident that further refinement and expansion of this research are possible, as the scenarios can be rendered more complex to assess the efficacy of the algorithm in question. Additionally, conducting a wider range of experiments that encompass additional classes, such as welding defects, would further augment the dataset and facilitate more comprehensive investigation.

A potential limitation of the present study, akin to other related works on object detection, is the possibility that the images obtained from another system and subsequently inferred by the model trained here may exhibit defective detection or yield performance that is lower than that attained in this study. This is despite the implementation of the same methodology and precautions during image capture. Consequently, we propose the utilization of Transfer Learning techniques, grounded in the findings of this study, to identify additional types of welding beads produced by diverse processes. This approach would leverage the established checkpoints from this study to enhance the detection of welding beads in disparate experiments. Furthermore, in the context of the Data Augmentation phase, it would be advantageous to implement a more sophisticated technique, such as GAN networks, to assess its impact on the efficacy of defect detection in welding seams.

Author Contributions: Conceptualization, J.M.R.C., A.M.-E., I.D.-C. and P.M.-C.; data curation, B.S.-D. and I.D.-C.; formal analysis, M.A.-A. and J.M.R.C.; funding acquisition, A.M.-E. and M.A.-A.; investigation, P.M.-C., A.M.-E., I.D.-C. and B.S.-D.; methodology, B.S.-D. and M.A.-A.; project administration, A.M.-E. and I.D.-C.; resources, A.M.-E. and M.A.-A.; software, P.M.-C. and I.D.-C.; supervision, A.M.-E. and M.A.-A.; validation, J.M.R.C. and I.D.-C.; visualization, A.M.-E. and P.M.-C.; writing—original draft, B.S.-D. and I.D.-C.; writing—review and editing, I.D.-C., B.S.-D. and J.M.R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially support AUROVI Project (Vision-enabled robotic automation), developed in the framework, of the supported by the Ministry of Science, Innovation and Universities under grant EQC2018-005190-P.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used in the creation of this article can be found at the following addresses: [47–49].

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

YOLO	You Only Look Once
CNN	Convolutional Neural Network
RPN	Region Proposal Network
GMAW	Gas Metal Arc Welding
FCAW	Flux Cored Arc Welding
ANN	Artificial Neural Network
DA	Data Acquisition
FE	Feature extraction
DP	Data Preprocessing
DAU	Data Augmentation
DL	Deep Learning
NAS	Neural Architecture Search
PM	Performance Metrics
R	Recall
P	Precision
TP	True Positive
FP	False Positive
gt	Ground Truth
pd	Predicted Box
IoU	Intersection over Union
AP	Average Precision
mAP	Mean Average Precision
TEP940	Applied Robotics Research Group of the University of Cadiz
ROI	Region Of Interest
SSD	Single Shot Detector

References

- Oh, S.J.; Jung, M.J.; Lim, C.; Shin, S.C. Automatic detection of welding defects using faster R-CNN. *Appl. Sci.* **2020**, *10*, 8629. [[CrossRef](#)]
- Mohamat, S.A.; Ibrahim, I.A.; Amir, A.; Ghalib, A. The Effect of Flux Core Arc Welding (FCAW) Processes on Different Parameters. *Procedia Eng.* **2012**, *41*, 1497–1501. [[CrossRef](#)]
- Ibrahim, I.A.; Mohamat, S.A.; Amir, A.; Ghalib, A. The Effect of Gas Metal Arc Welding (GMAW) Processes on Different Welding Parameters. *Procedia Eng.* **2012**, *41*, 1502–1506. [[CrossRef](#)]
- Katherasan, D.; Elias, J.V.; Sathiya, P.; Haq, A.N. Simulation and parameter optimization of flux cored arc welding using artificial neural network and particle swarm optimization algorithm. *J. Intell. Manuf.* **2014**, *25*, 67–76. [[CrossRef](#)]
- Ho, M.P.; Ngai, W.K.; Chan, T.W.; Wai, H.w. An artificial neural network approach for parametric study on welding defect classification. *Int. J. Adv. Manuf. Technol.* **2021**, *1*, 3. [[CrossRef](#)]
- Kim, I.S.; Son, J.S.; Park, C.E.; Lee, C.W.; Prasad, Y.K. A study on prediction of bead height in robotic arc welding using a neural network. *J. Mater. Process. Technol.* **2002**, *130–131*, 229–234. [[CrossRef](#)]
- Bronstein, M.M.; Bruna, J.; Cohen, T.; Veličković, P. Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges. *arXiv* **2021**, arXiv:2104.13478.
- Liu, F.; Tao, C.; Dong, Z.; Jiang, K.; Zhou, S.; Zhang, Z.; Shen, C. Prediction of welding residual stress and deformation in electro-gas welding using artificial neural network. *Mater. Today Commun.* **2021**, *29*, 102786. [[CrossRef](#)]
- Zhang, Z.; Li, B.; Zhang, W.; Lu, R.; Wada, S.; Zhang, Y. Real-time penetration state monitoring using convolutional neural network for laser welding of tailor rolled blanks. *J. Manuf. Syst.* **2020**, *54*, 348–360. [[CrossRef](#)]
- Feng, T.; Huang, S.; Liu, J.; Wang, J.; Fang, X. Welding Surface Inspection of Armatures via CNN and Image Comparison. *IEEE Sens. J.* **2021**, *21*, 21696–21704. [[CrossRef](#)]
- Nele, L.; Mattera, G.; Vozza, M. Deep Neural Networks for Defects Detection in Gas Metal Arc Welding. *Appl. Sci.* **2022**, *12*, 3615. [[CrossRef](#)]
- Mery, D.; Riffo, V.; Zscherpel, U.; Mondragón, G.; Lillo, I.; Zuccar, I.; Lobel, H.; Carrasco, M. GDXray: The Database of X-ray Images for Nondestructive Testing. *J. Nondestruct. Eval.* **2015**, *34*, 42. [[CrossRef](#)]

13. Hartung, J.; Jahn, A.; Stambke, M.; Wehner, O.; Thieringer, R.; Heizmann, M. Camera-based spatter detection in laser welding with a deep learning approach. In *Forum Bildverarbeitung 2020*; Längle, T., Heizmann, M., Eds.; KIT Scientific Publishing: Karlsruhe, Germany, 2020; pp. 317–328. [[CrossRef](#)]
14. Nacereddine, N.; Goumeidane, A.B.; Ziou, D. Unsupervised weld defect classification in radiographic images using multivariate generalized Gaussian mixture model with exact computation of mean and shape parameters. *Comput. Ind.* **2019**, *108*, 132–149. [[CrossRef](#)]
15. Deng, H.; Cheng, Y.; Feng, Y.; Xiang, J. Industrial laser welding defect detection and image defect recognition based on deep learning model developed. *Symmetry* **2021**, *13*, 1731. [[CrossRef](#)]
16. Ajmi, C.; Zapata, J.; Martínez-Álvarez, J.J.; Doménech, G.; Ruiz, R. Using Deep Learning for Defect Classification on a Small Weld X-ray Image Dataset. *J. Nondestruct. Eval.* **2020**, *39*, 68. [[CrossRef](#)]
17. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
18. Wang, R.; Jiao, L.; Xie, C.; Chen, P.; Du, J.; Li, R. S-RPN: Sampling-balanced region proposal network for small crop pest detection. *Comput. Electron. Agric.* **2021**, *187*, 106290. [[CrossRef](#)]
19. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Advances in Neural Information Processing Systems*; Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., Garnett, R., Eds.; Curran Associates, Inc.: San Jose, CA, USA, 2015; Volume 28.
20. Wang, Y.; Shi, F.; Tong, X. A Welding Defect Identification Approach in X-ray Images Based on Deep Convolutional Neural Networks. In *Intelligent Computing Methodologies*; Huang, D.S., Huang, Z.K., Hussain, A., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 53–64.
21. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. In *Proceedings of the Computer Vision—ECCV 2016, Amsterdam, The Netherlands, 11–14 October 2016*; Leibe, B., Matas, J., Sebe, N., Welling, M., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 21–37.
22. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Los Alamitos, CA, USA, 27–30 June 2016*; pp. 779–788. [[CrossRef](#)]
23. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014*; pp. 580–587. [[CrossRef](#)]
24. Dai, W.; Li, D.; Tang, D.; Wang, H.; Peng, Y. Deep learning approach for defective spot welds classification using small and class-imbalanced datasets. *Neurocomputing* **2022**, *477*, 46–60. [[CrossRef](#)]
25. Girshick, R. Fast R-CNN. In *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 13–16 December 2015*; pp. 1440–1448. [[CrossRef](#)]
26. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)]
27. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 386–397. [[CrossRef](#)]
28. Yun, G.H.; Oh, S.J.; Shin, S.C. Image Preprocessing Method in Radiographic Inspection for Automatic Detection of Ship Welding Defects. *Appl. Sci.* **2021**, *12*, 123. [[CrossRef](#)]
29. Hobart. Choosing the Right Shielding Gases for Arc Welding | HobartWelders 2024. Available online: <https://www.hobartwelders.com/projects-and-advice/articles/choosing-the-right-shielding-gases-for-arc-welding> (accessed on 25 November 2024).
30. Lascentrum. Hyundai SC-420MC. 2024. Available online: <https://lascentrum.com/en/producten/hyundai-sm-70-eco/> (accessed on 6 December 2024).
31. Lascentrum. Hyundai SM-70 eco. 2024. Available online: <https://lascentrum.com/en/producten/hyundai-sc-420mc/> (accessed on 6 December 2024).
32. Murray_Steel. S275JR Steel Plate. 2024. Available online: <https://www.murraysteelproducts.com/products/s275jr> (accessed on 6 December 2024).
33. Kwon, J.E.; Park, J.H.; Kim, J.H.; Lee, Y.H.; Cho, S.I. Context and scale-aware YOLO for welding defect detection. *NDT E Int. Indep. Nondestruct. Test. Eval.* **2023**, *139*, 102919. [[CrossRef](#)]
34. Hussain, M. YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection. *Machines* **2023**, *11*, 677. [[CrossRef](#)]
35. Terven, J.; Córdova-Esparza, D.M.; Romero-González, J.A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 1680–1716. [[CrossRef](#)]
36. *ISO Standard 15792-3:2011; Welding Consumables—Test Methods*. ISO: Geneva, Switzerland, 2024.
37. American Welding Society. *Specification for Carbon Steel Electrodes for Shielded Metal Arc Welding*, 14th ed.; American Welding Society: Doral, FL, USA, 2004.
38. *ISO Standard 6947:2019; Welding and Allied Processes—Welding Positions*. ISO: Geneva, Switzerland, 2024.

39. IDS. Ensenso N Series. 2024. Available online: <https://www.ids-imaging.us/ensenso-3d-camera-n-series.html> (accessed on 29 October 2024).
40. Shinichi, S.; Muraoka, R.; Obinata, T.; Shigeru, E.; Horita, T.; Omata, K. *Steel Products for Shipbuilding*; Technical Report, JFE Technical Report; JFE Holdings: Tokyo, Japan, 2004.
41. Roboflow. Computer Vision Tools for Developers and Enterprises. 2024. Available online: <https://roboflow.com/> (accessed on 5 October 2024).
42. Puhan, S.; Mishra, S.K. Detecting Moving Objects in Dense Fog Environment using Fog-Aware-Detection Algorithm and YOLO. *NeuroQuantology* **2022**, *20*, 2864–2873.
43. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [[CrossRef](#)]
44. Lin, T.Y.; Maire, M.; Belongie, S.J.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common Objects in Context. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014.
45. Hermens, F. Automatic object detection for behavioural research using YOLOv8. *Behav. Res. Methods* **2024**, *56*, 7307–7330. [[CrossRef](#)] [[PubMed](#)]
46. YOLO-Ultralytics. Performance Metrics—Ultralytics YOLO Docs. Available online: <https://docs.ultralytics.com/es/guides/yolo-performance-metrics/> (accessed on 5 October 2024).
47. TEP940. Dataset Detection FCAW-GMAW Welding. 2024. Available online: https://universe.roboflow.com/weldingpic/weld_fcaw_gmaw (accessed on 25 October 2024).
48. TEP940. Dataset Detection WELD_GOOD_BAD Welding. 2024. Available online: https://universe.roboflow.com/weldingpic/weld_good_bad (accessed on 1 November 2024).
49. TEP940. GOOD-OP-LOP-UNDER Dataset. 2024. Available online: <https://universe.roboflow.com/weldingpic/good-op-lop-under> (accessed on 30 October 2024).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.