ISSN 1870-4069

Gas Metal Arc Welding Dataset for Computer Vision Quality Assessment

José A. López-Islas[1,2], Oscar Camacho-Nieto[2], Yenny Villuendas-Rey[2]

[1] Instituto Politécnico Nacional,
Escuela Superior de Ingeniería Mecánica y Eléctrica, Unidad Azcapotzalco,
Mexico

[2] Instituto Politécnico Nacional,
Centro de Innovación y Desarrollo Tecnológico en Cómputo,
Mexico

{jlopezi;ocamacho;yvilluendasr}@ipn.mx

Abstract. Robotic Gas Metal Arc Welding is commonly used for several industrial purposes and requires expedited quality assurance feedback due to the nature of the robotic welding. Usually, in small and medium enterprises, this feedback comes from a human expert. There is an interest in digitalizing the experts' knowledge regarding welding quality to train computer vision systems for automatic quality assessment. Having the correct data is crucial for this task. In this paper, we introduce a novel dataset of robotic Gas Metal Arc Welding images belonging to four categories: good welding, welding that can be reworked by the robot itself, welding that can be reworked by a human expert, and welding with unsalvable fails (usually named as scrap). The proposed dataset includes the experts' knowledge and annotates the individual welding seam quality. It is publicly available and can be used to train computer vision systems for welding quality assurance.

Keywords: Robotic gas metal arc welding, quality assessment, computer vision dataset.

1 Introduction

Robotic welding is widely used in the metalworking industry and is most often supervised by qualified and trained personnel for weld seam verification. With increasing technological development worldwide, artificial intelligence is being integrated into industrial processes. It seeks to digitize the knowledge of specialized supervisors, both to verify process quality and for its optimization.

Among the various types of welding, robotic gas metal arc welding (GMAW) stands out for its wide-ranging applications in diverse industries such as automotive and manufacturing, among others. GMAW consists of a wire-shaped welding electrode, which, through the application of a controlled electrical charge by means of the welding

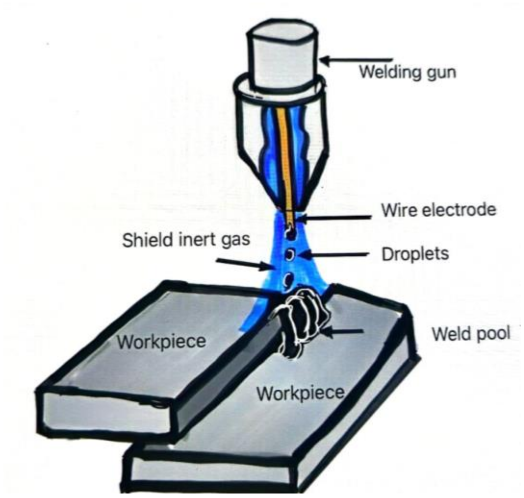


Fig. 1. Main elements of GMAW.

gun [1], causing it to reach its melting point and drip continuously towards the material (weld pool). The process is carried out progressively, so it requires an automatic wire feeder. When this welding process is carried out, an active or inert gas (shield of inert gas) is used to protects it from atmospheric contamination [2], as shown in Figure 1.

The application of robotic GMAW welding has undergone significant development in terms of automation and supervision, seeking to create intelligent welding systems. This approach to applying artificial intelligence models is known as "Intelligent Welding Manufacturing (IWM)" [3]. These systems can include sensing technology, followed by an image processing method based on prior knowledge, and acquire weld characteristics by measuring pixels obtained from the images in real time.

Some research has also focused on collecting electrical and mechanical data from the welding process in addition to capturing images for the design of an intelligent system [4].

However, most research is conducted with process data from different environments, which are not made public for research. Therefore, it is difficult to find images of robotic GMAW weld seams to train computer vision systems.

To address this limitation, this article proposes a robotic GMAW image dataset. This dataset stands out for its low acquisition cost and for its ability to include human expert labeling of weld seam characteristics.

The remainder of the paper is as follows: Section 2 covers some of the existing works in image-based automatic quality assurance for robotic GMAW. Section 3 provides an explanation of the robotic process and the characteristics of the welding parts. Section 4 describes the proposed dataset, and Section 5 covers the conclusions of the paper.

Fig. 2. Metallic parts to be welded.

Table 1. Chemical properties of the metallic parts.

Chemical element	Composition
C	0.100
Mn	0.600
P	0.030
S	0.035
Cu	0.200
Ni	0.200
Cr	0.150
Mo	0.060
V	0.008
Cb	0.008
Ti	0.008

2 Related Works

Several researchers have addressed the topic of GMAW quality assurance by computer vision. Li et al. used Convolutional Neural Networks (CNNs) for defect prediction [5]. They recorded the welding process with a welding camera. Then, they used the video frames as a molten pool of images to construct the dataset. Unfortunately, in [5] is stated that “The datasets generated and analyzed during the current study are not publicly available due the confidentiality of the data”.

Kim et al. also created a dataset for classifying penetration conditions in GMAW processes by CNNs. They also stated that “The data are not publicly available due to

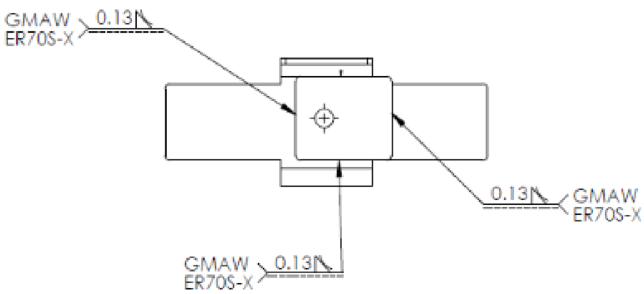


Fig. 3. Welding specifications.

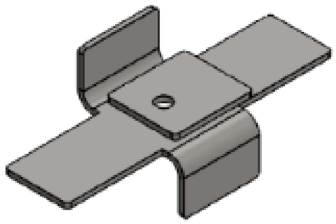


Fig. 4. 3D isometric view of the parts.

privacy", although it can be provided by the corresponding author upon reasonable request [6].

Diaz-Cano et al. [7] created a set of images of weld seams considered acceptable, as well as seams with defects such as lack of penetration or undercuts. Their images correspond to FCAW and GMAW welding processes and were captured with a high-resolution camera (Ensenso model N35) positioned on the robotic arm. The camera was moved in seven poses for each weld seam, each with different luminosities. The images captured are publicly available at the following links:

- https://universe.roboflow.com/weldingpic/good-op-lop-under/dataset/1
- https://universe.roboflow.com/weldingpic/weld_fcaw_gmaw/dataset/1
- https://universe.roboflow.com/weldingpic/weld_fcaw_gmaw/dataset/2

However, in their work, the images correspond to straight lines of welding beads, with the purpose of measuring the thickness and height (in the best of cases) in a controlled, smooth relief without any distortion in its trajectory, therefore, the deposition of the liquid metal is only dispersion.

Research has taken into account different cameras for the welding process and certain additional elements such as the distance, width, area, and angle of the welding electrode. When using low-cost cameras, such as a webcam, it is necessary to apply

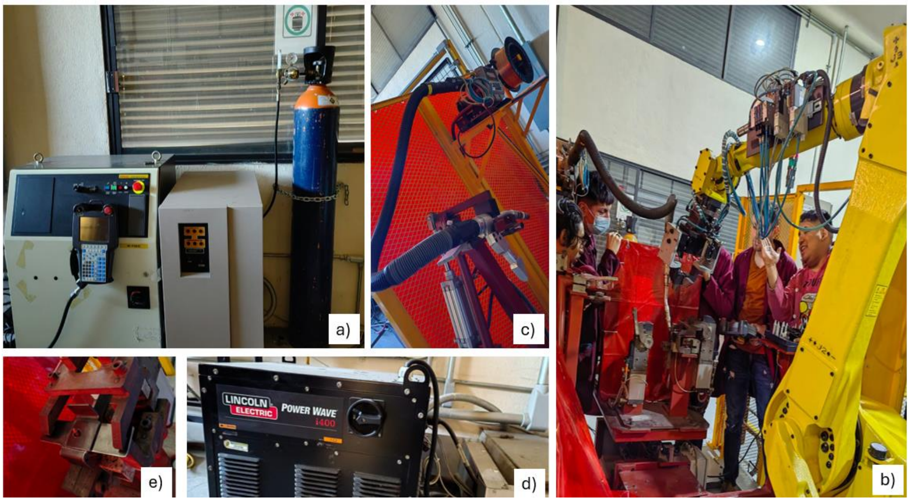


Fig. 5. GMAW Welding Station. (a) Fanuc R30ia Controller with Ipendant, Vogar Voltage Regulator and Gas Tank, (b) Fanuc M70iC/50 Robot, (c) AutoDrive 4R90 Automatic Microwire Feeder and Lincoln Electric Welding Torch, (d) Lincoln Electric Power Wave i400 Welding Controller, and (e) Automatic Automotive Workholding System.

more processing methods such as YOLOv5 (You Only Look Once), PAN (Path Aggregation Network), CNN (Convolutional Neural Network), and FPN (Feature Pyramid Networks) [6].

In Spain, one of the investigations used two different robotic welding processes: Flux-Cored Arc Welding (FCAW) and Gas Metal Arc Welding (GMAW). The inspection technique used in this case was to train a neural network with a series of 2D images of welding beads, taken by special equipment to offer different shades of light, and a high-resolution industrial camera, placed on the end effector on another robotic arm, the technique, follows a deep learning geometric model in CNN and YOLOv8, the results were achieved in binary and multiclass classification [8]. In India, the focus of research is on a machine vision-based algorithm for robotic weld path detection, gap measurement, and weld length calculation in the GMAW process. The camera used is a low-cost webcam located on the robotic arm, and the captured images are converted to grayscale, and YOLOv5, PAN, CNN, and Feature Pyramid Network (FPN) are used for variance scaling in object detection [9]. In this research, we focused on the use of low-cost cameras.

In addition, our proposal corresponds to a higher level of complexity. It is derived from the fact that it is not only the deposition of the electrode material, but it also entails penetration into both materials for the purpose of a permanent union. During the welding path, there will always be an imaginary line between both joining materials that will not be homogeneous. In addition, the technique in cutting the material will make the edges rough, not smooth. This causes the electric arc that is generated to not be the same throughout its path, and the movement of the materials due to their thermal expansion generates different welding patterns along their path.

The next section presents the details regarding the materials and methods used in our research.

3 Materials and Methods

This section details the elements necessary to obtain the welding seams, the capture medium, the number of images, and the parameters used for the classification of the process quality.

3.1 Metallic Parts

The parts to be welded are two and are made of 0.25-inch thick black sheet, whose classification is hot-rolled steel sheet, designated as Commercial Steel Type A (CS Type A), as illustrated in Figure 2. These pieces were machined with a water jet. The mechanical properties of the parts are "Yield Strength" (Ksi): 30 to 50, and Elongation in 2 in. [50 mm] % ≥25. The chemical properties of the parts are provided in Table 1.

The welding implemented was by Pulse Spray Transfer, whose characteristics are: Wire feed: 145, Trim: 1.05, Voltage: 29V, Travel speed: 36, Pulse: 0.045, Feedback Current: 199.4, and Air mix: Ar + CO2.

The welding specifications are detailed in Figure 3, and how both parts must be superimposed is shown with a 3D isometric view in Figure 4.

3.2 Welding Equipment

We used the welding equipment of the Automotive Cell laboratory at the School of Mechanical and Electrical Engineering, Azcapotzalco Unit, of the Instituto Politécnico Nacional. It includes a robotic welding station (Figure 5). It contains an Industrial robotic system. Figure 5a shows a model r30iA controller commanded by an IPendant, as well as its Vogar voltage regulator; Figure 5b shows a Fanuc robot model M710iC/50 with a Schunk brand gripper end effector to hold the welding torch.

In addition, the station has an automatic welding system, including a GMAW welding torch with an AutoDrive 4R90 automatic feeder (Figure 5c), a Power Wave i400 controller from Lincoln Electric (Figure 5d), and the active gas tank (MAG) with a combination of Ar + CO2 (Figure 5a). Finally, the station also includes an automatic clamping system, as shown in Figure 5e), is used to hold the automotive part while it is being welded.

3.3 Image Acquisition

For welding, both parts must be aligned using a locating pin in the center, as shown in Figure 6a. Both ends of the lower piece, which is cross-shaped, are held by an automatic clamping device designed for the piece. This prevents movement of the piece at the time of welding, which may occur due to thermal expansion caused by the process.

(a) (b)

Fig. 6. Metallic parts to be welded. (a) In clamps and (b) With welding seams.

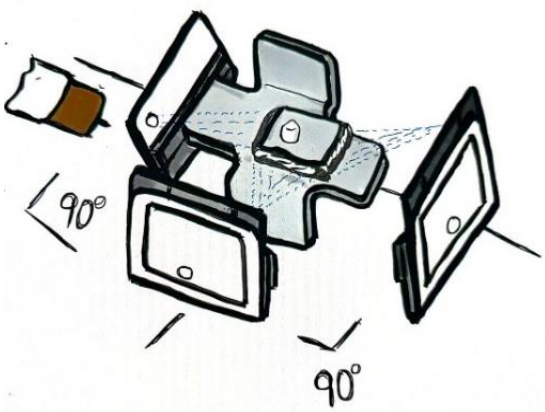
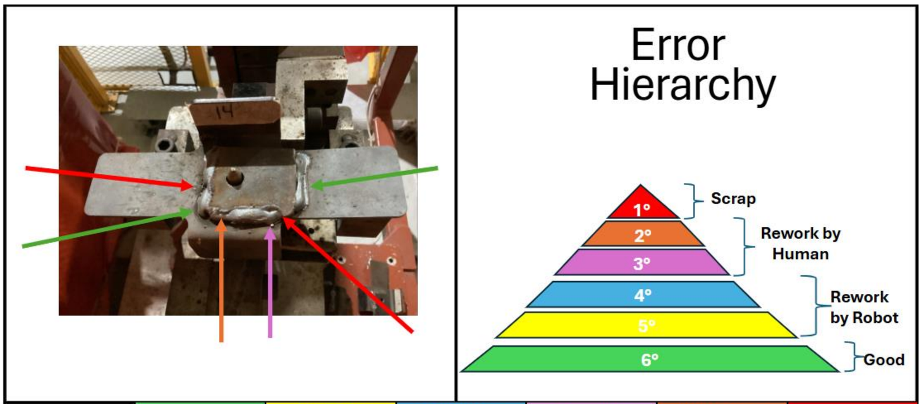


Fig. 7. Image acquisition of the welded parts. Note how the capturing device is arranged in three different positions to acquire a good view of the weld seams.

The first weld bead must be made from the front, in order to be able to remove the part's holding element and continue with the side beads, ensuring that the part will not move again, leaving only the central locating pin responsible for the correct position of the part when it is released, as shown in Figure 6b.

The images were captured after applying the three welding paths, ensuring a parallelism with the weld bead to avoid shadows that could be reflected between the piece to be welded and the weld bead (Figure 6). The device used to capture the images was a 9th Generation iPad, with an 8MP wide-angle lens with an f/2.4 aperture and 5x digital zoom, and HDR for photos.



ITEM	Good	Rework by robot	Rework by robot	Rework by human	Rework by human	Scrap
14F	Homogeneus Geometry	Incorrect Trajectory	Lack of Welding	Remove splatter welding	Remove welding seam	Check if it's a hole

Fig. 8. Example of welded parts (image 14F) with the annotated characteristics (left) and the proposed error hierarchy for the quality assessment of the weld seams (right) in a colorimetric scale.

4 Results

A crucial aspect of this proposal is the human expert's identification of potential flaws in the weld seam. To this end, a hierarchy of errors is proposed (Figure 7). These errors range from the most serious (top of the pyramid) to the good weld (base of the pyramid), including the less serious errors. Thus, six elements are considered that can be visually recognized by the weld seam geometry. As can be seen, if there is at least one first-degree error in any of the weld seams of the parts to be welded, the entire part must be scrapped, as it would no longer be suitable for rework.

Similarly, if there is at least one error that requires the part to be worked by a human operator, it would be labeled as such. Something similar happens, for example, with a part that has two good weld seams and one that requires robotic rework. In this case, the entire part is labeled as "robotic rework," not as good. Even if most weld seams were correct, the presence of an error leads to labeling the part according to the highest-ranking error.

These elements represent a significant contribution to the state of the art, as the proposed dataset not only offers image labeling but also includes the knowledge of the human expert, identifying, for each weld seam, its geometric characteristics and corresponding errors on a colorimetric scale.

Depending on the characteristics of the weld seams, the images were labeled into four categories, as in Figure 7: Good, Rework by Robot (Robot), Rework by Human (Human), and Scrap.

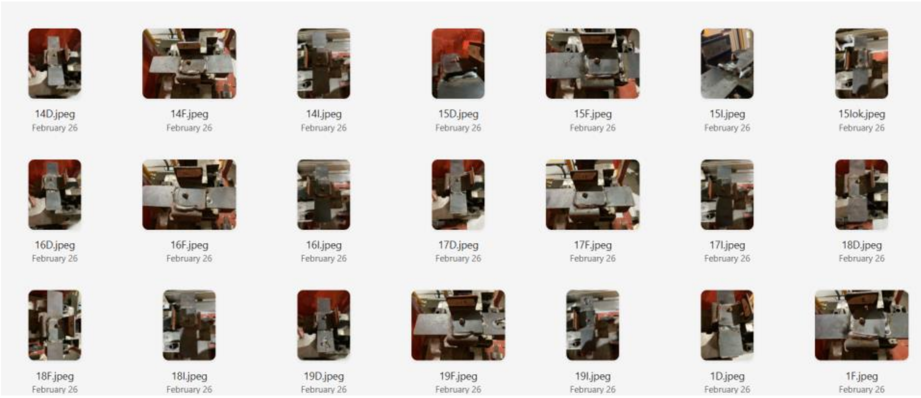


Fig. 9. Example of images in the proposed dataset.

The developed dataset included the analysis of weld seams from 30 automotive parts. Three weld seams were created on each part, for a total of 90 images. Figure 8 shows some of the images obtained. In addition to the images labeled into the four aforementioned categories, the database includes, for each image, labeling of the weld seam characteristics, according to the proposed error hierarchy.

The dataset will be donated to the Machine Learning Repository of the University of California at Irvine, to be publicly available worldwide. In the meantime, interested readers can find it in the following institutional link: https://correoipn-my.sharepoint.com/:f:/g/personal/yvilluendasr_ipn_mx/EmD7pChS4zpLkIZRsc6ldE0BgIzsgX9F4sErXxgCFxswlQ?e=NLZDo9

In the following, we summarize the criteria considered in the labeling process. First, A Good weld is homogeneous, without any porosity or black shadow that could indicate a perforation or spatter when welding (Figure 10a).

The weld that can be corrected (reworked) by the robot (Robot label) appears to be homogeneous but contains spatter in the form of small spheres around the weld bead (Figure 10b). These are generally visible as part of the unwelded side of the part, so the robot can re-pass through them, correcting the path.

As shown in Figure 10c, the weld bead is not homogeneous; it is thin in some places, bulged in the middle, and has a lot of spatter. However, the color is light, meaning it was not overheated, so the weld can be removed using a chisel and hammer. Thus, a human operator can rework the weld bead and not discard the part. Such seams are labelled as Human.

Finally, as shown in Figure 10d, round-shaped shadows appear on the weld bead, so there is a risk of overheating of the part and possible perforation of the part, so it would not withstand a second welding application. Therefore, they are labeled as Scrap.

Fig. 10. Examples of weld seams belonging to the four categories. (a) Good, (b) Robot, (c) Human, and (d) Scrap.

As shown in Figure 10c, the weld bead is not homogeneous; it is thin in some places, bulged in the middle, and has a lot of spatter. However, the color is light, meaning it was not overheated, so the weld can be removed using a chisel and hammer. Thus, a human operator can rework the weld bead and not discard the part. Such seams are labelled as Human. Finally, as shown in Figure 10d, round-shaped shadows appear on the weld bead, so there is a risk of overheating of the part and possible perforation of the part, so it would not withstand a second welding application. Therefore, they are labeled as Scrap.

Of the 90 images analyzed, we have a classification of 36 Good, 19 manual rework (Human), 30 robotic rework (Robot), and five bad seams for discard (Scrap). The images were divided into train and test sets by a human expert. It allows the replicability of the experiments researchers can make in the future and guarantees the representativity of the weld seams.

5 Conclusions

This paper focused on quality assessment of GMAW in automotive parts. This process entails high complexity derived from the fact that it is not only the deposition of the electrode material but also the penetration into both materials for the purpose of a permanent union. In addition, the technique in cutting the material will make the edges rough, not smooth. This causes the electric arc that is generated to not be the same throughout its path, and the movement of the materials due to their thermal expansion generates different welding patterns along their path.

We proposed a novel image dataset from real-world welding images, that includes the analysis of weld seams from 30 automotive parts. In addition, the dataset not only offers labelled images but also includes the knowledge of the human expert, identifying, for each weld seam, its geometric characteristics and corresponding errors on a hierarchical colorimetric scale. The images were labeled into four categories: Good, Rework by Robot (Robot), Rework by Human (Human), and Scrap. In addition to the images labeled into the four aforementioned categories, the database includes, for each image, labeling of the weld seam characteristics according to the proposed error hierarchy.

In future work, we want to increase the number of images. We also want to add a camera to the robotic arm to have more degrees of liberty for image capturing.

Acknowledgments. The authors would like to thank the Instituto Politécnico Nacional (Secretaría Académica, Secretaría de Investigación y Posgrado, ESIME Azcapotzalco, and Centro de Inno-vación y Desarrollo Tecnológico en Cómputo), the Secretaría de Ciencia, Humanidades, Tecnolo-gía e Innovación, and Sistema Nacional de Investigadores for their economic support to develop this work.

References

1. url: https://www.keyence.com.mx/ss/products/measure/welding/
2. url: https://www.millerwelds.com/resources/article-library/the-history-of-welding
3. Hou, Z., Xu, Y., Xiao, R.: A Teaching-Free Welding Method Based On Laser Visual Sensing System in Robotic GMAW. The International Journal of Advanced Manufacturing Technology, 109, pp. 1755–1774 (2020).
4. Wang, X., Chen, Q., Sun, H.: GMAW Welding Procedure Expert System Based On Machine Learning. Intelligence & Robotics 3, pp. 56–75 (2023).
5. Li, H., Ma, Y., Duan, M.: Defects Detection of GMAW Process Based On Convolutional Neural Network Algorithm. Scientific Reports 13, 212–219 (2023).
6. Kim, D.-Y., Lee, H.W., Yu, J.: Application of Convolutional Neural Networks for Classifying Penetration Conditions in GMAW Processes Using STFT of Welding Data. Applied Sciences 14, 4883 (2024).
7. Diaz-Cano, I., Morgado-Estevez, A., Rodríguez Corral, J.M.: Automated Fillet Weld Inspection Based on Deep Learning From 2D Images. Applied Sciences 15, 899 (2025).
8. Pico, L.E.A., Marroquín, O.J.A., Lozano, P.A.D.: Application of Deep Learning for the Identification of Surface Defects Used in Manufacturing Quality Control and Industrial Production: A Literature Review. Ingeniería, 28, pp. 1–20 (2023).
9. Reddy, J., Dutta, A., Mukherjee, A.: A Low-Cost Vision-Based Weld-Line Detection and Measurement Technique for Robotic Welding. International Journal of Computer Integrated Manufacturing, pp. 1–21 (2024).