Response to the reviewer comments

Manuscript ID: Sensors-2436295
Manuscript Title: "Automated Categorization of Multiclass Welding Defects Using the X-ray Image Augmentation and Convolutional Neural Network "

Dear Editor,

We would like to thank you and the anonymous reviewers for the insightful and valuable comments concerning our manuscript," Automated Categorization of Multiclass Welding Defects Using the X-ray Image Augmentation and Convolutional Neural Network." (Manuscript ID: Sensors-2436295). We are also thankful to you for providing us with the opportunity to submit the revised version of paper while addressing the reviewers' comments.

We have revised the manuscript and addressed all the comments accordingly. For ease of reference, the responses to these comments are given below in an itemized fashion.

As per the authors best knowledge, we have made a considerable effort to take into account the interesting suggestions proposed by the reviewers. In any case, we are open to consideration of any further comments.

We hope that you will be satisfied with the quality of this revised version of manuscript, and you will recommend it for publication.

The authors,
(Dalila Say, Salah Zidi, Saeed Mian Qaisar, and Moez Krichen)

Responses to the Comments of Reviewer 2 :

A Brief Summary: The submitted article aims to better identify the different types of weld defects (as X-ray images) using an automated approach utilizing the techniques of data augmentation and deep learning algorithm (Convolution Neural Network). The accuracy/success of such automated techniques would revolutionize the detection of welding defects in various industries without the requirement of thorough inspection by experts. I believe the article has certainly tried to address this issue. However, there are many inaccuracies in the English language and formatting in addition to the deficiency in the clarification of some concepts, as mentioned in the comments attached below.

Response: The authors are thankful to the learned reviewer for a thorough review and valuable comments. It is the feedback of reviewer which has allowed us to effectively revise and improve the quality of our manuscript.

General Comments:

Comment 1 : In general, there are many inaccuracies in the English language and formatting throughout the paper. For example in "section 2: Related works" (page 4, line 127), the line says "The same researcher in[19]. Drew an 11-layer…". Similar formatting errors can be seen in lines 10, 12, 26, 74, 99, 464 etc. The other example is in "section 3.2: Data augmentation" (page 5, line 164), where it is said "One of the maximum common issues that training …". I think the more appropriate word would be "most" instead of "maximum". Another example is in the same section on page 5, line 187, where it says "it shows in equation (1)". I think it should rather be written as "as shown in equation (1)". Such language errors can be found in lines 20, 48, 153, 463 etc.}

Response for comment 1: Thanks for pointing it out. The manuscript is carefully revised while incorporating the recommended changes.

Comment 2 : There are many instances in the article, where the acronym is mentioned for the first time without giving the full form. For example in "section 1: Introduction" (page 1, line 26), NDT has been mentioned without the full form. I think it should be mentioned in line 22 after "Non-destructive testing" in brackets. Similar examples can be found in lines 41, 44, 106 etc.

Response for comment 2: Thanks for pointing it out. The manuscript is carefully revised while incorporating the recommended changes. Also a list of abbreviations ("Nomenclature") is presented on Page#1, in the revised version of manuscript.

Comment 3 : The conclusions section should be made more clearer and elaborated. Because the last few sentences (lines 465-467) speaks on some future investigations based on enhancing the precision. So, it should be made clear before on the limitation of the current approach and it's cause in the identification of the different types of weld defects..

Response for comment 3: Thanks for this comment. The manuscript is carefully revised while incorporating the recommended changes. The limitation of the devised approach and the future works are clearly presented in the conclusion section. The updates are highlighted in the revised version of manuscript in yellow colour and are also presented below for an ease of the review process.

In our research, we perform a multi-class classification method to detect weld flaws in X-ray images by employing a CNN model. The model's creation relied on GDX-ray image databases that are publicly available. This database contains only a minimum number of samples. Therefore, we have extracted the sub-images and created our dataset which are used for the training and test process. The technique of Data augmentation and fine-tuning strategies helped to achieve appropriate performance of the model. The multi-class classification has been combined with the optimizer algorithm. The model has achieved an accuracy score of 92% for the case of CNN classifier. This performance is comparable or superior compared to the counterparts. The applicability of proposed method is only tested for the intended dataset. In future, the performance of proposed approach will be studied for extended datasets in order to achieve an automated categorization of weld defects. Also its application to other industrial problems like Packaging Inspection, Defect Detection in Fabric and Textiles, and Printed Circuit Boards inspection will be investigated. An other future work is to investigate about the possibilities of enhancing the precision of our approach by incorporating features extraction, feature selection and ensemble learning techniques.

Specific Comments:

Comment 1: In Figure 3, the figures are not marked with (a), (b), … though it has been mentioned in the figure caption. That the topmost subfigure depicts the original image should also be mentioned. Also, the subfigure (d) for shear is not so clear as to what the "shearing" process does.

Response for comment 1: Thanks for pointing this out. The manuscript is revised to in order incorporate the recommended changes. We have taken your input into consideration and have implemented the necessary correction in our article. Specifically, we have addressed the issue by assigning designations (a, b, c, d, e) to each image presented in Figure 3. These designations can be found clearly indicated on page 7, within Section 3, and more specifically, in Subsection 3.2, where it is explicitly referred to as "data augmentation".

Comment 2: In "section 3.4: The evaluation metrics" (page 9, line 294), "recall" is mentioned. However, the "recall" does not appear in Table 1. So, it should be made clear in the table as to which parameter represents "Recall". Also from this table it is not so clear as to which quantity is used to calculate "Support".

Response for comment 2: Thanks for pointing this out. The manuscript is revised to in order incorporate the recommended changes. We recognize the interchangeability of "sensitivity" and "recall" within the context of our study. Additionally, we have diligently rectified any confusion by effecting the necessary corrections in Table 1. Moreover, we have taken the opportunity to provide a precise definition of "support" as the numerical count of images employed during the testing phase (cf. Section 3, immediately following the presentation of Table 1.)

Comment 3: In "section 4.1: The dataset preparation" (page 9, line 298), it is written as "BAM". However in "section 3.1: Dataset" (page 4, line 136), it is written as "BM". I guess both lines are talking about the same dataset.

Response for comment 3: Thanks for pointing this out. The manuscript is revised to in order incorporate the recommended changes. We acknowledge that the accurate abbreviation is true "BAM". We have promptly addressed this inconsistency and corrected the error in "section 3.1: Dataset".

Comment 4: The caption of Figure 5 says that it is an example of max-pooling. However, the figure seems to depict an “average-pooling”, if I am correct.

Response for comment 4: Thanks for pointing this out. The manuscript is revised to in order incorporate the recommended changes. The Figure 5 is updated for an accurate representation of "max-pooling". The updates are also presented below for an ease of review process.

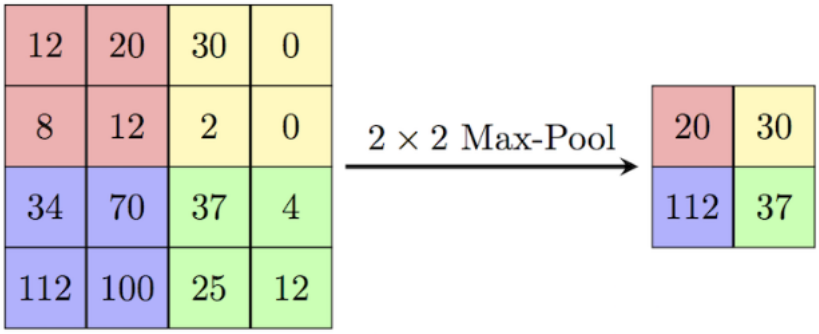

Figure 5. Example of the max-pooling with a 2 x 2 image.

Comment 5: The “ND”, “I”, “LOF”, etc. shown in Figure 6 are not explained in the text. I guess “ND” stands for “Normal Defect”, “I” for Inclusions etc. Also, for clarity, one could show these names also in Figure 7. The Figure 7 can also be improved by saying that the 80% of data goes to “Train” and the rest to “Test”, by putting the numbers above the boxes corresponding to “Train” and “Test”..

Response for comment 5: Thanks for pointing it out. The manuscript is revised in order to incorporate the recommended change. We have provided the corresponding explanations for each defect. In addition, we ensured consistency by incorporating the same notation in Figure 7 and including percentage values above the 'Training' and 'Test' boxes to illustrate the distribution of the data.

Comment 6: In “section 4.3: Discussion” (page 12, line 350), it is said that the images are resized to 128x128. However, in Figure 8, it is said that the input image has the dimensions 124*124.

Response for comment 6: Thanks for pointing it out. The manuscript is revised in order to incorporate the recommended change. The accurate and intended dimensions for the images are indeed 128x128. Consequently, we have promptly rectified this discrepancy and made the necessary correction in Figure 8 to reflect the accurate dimensions. For an ease of review process the updated version of Figure 8 is also parented below.

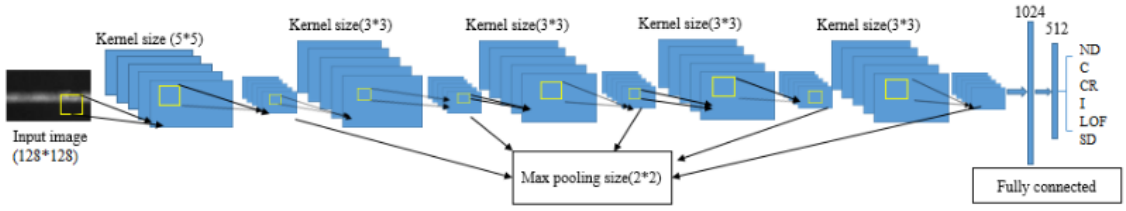

Figure 8. Designing a Convolutional Neural Network (CNN): proposed architecture.

Comment 7: In “section 3.1: Dataset”, the “normal defect” is shown in Figure 2, but is not mentioned in the text above (lines 147-149). I guess this is depicted as “ND” in Figure 6. However, in Figure 7, it is written as just “Normal”. And then, in Table 3, it is written as “N”. So, I think that these names must be named consistently so as not to confuse the reader.

Response for comment 7: Thanks for pointing it out. The manuscript is revised in order to incorporate the recommended change. We have taken ensured the consistency throughout the article. Firstly, we have revised Figure 6 to depict the "normal defect" as "ND" instead of "N". Additionally, in Figure 7 and Table 3, we have updated the reference to the "normal defect" to maintain consistency with the depiction in Figure 6.

Comment 8: The abstract says that the data augmentation performs the nearest neighbor interpolation, and several other operations on the X-ray images. However in “section 3.2: Data augmentation”, the nearest neighbor interpolation is not described. Also, line 8 in the abstract says that “augmentation with fine-tuning”. So, maybe in section 3.2, the fine-tuning procedure should also be explained to some extent.

Response for comment 8: Thanks for pointing it out. The manuscript is revised in order to incorporate the recommended change. We have discontinued the use of nearest neighbor interpolation and apologize for the oversight. Regarding the "augmentation with fine-tuning" mentioned in the abstract, we agree that it should be explained to a certain extent in section 3.2. We have made revisions to provide a more comprehensive explanation of this technique, as it is vital for optimizing our model and achieving improved results. You will find a detailed explanation of this technique in the "Results and Discussion" section.