HAL
open science

Estimating and monitoring laser induced damage size on glass windows with a deep-learning based pipeline

Isam Ben Soltane, Guillaume Hallo, Chloé Lacombe, Laurent Lamaignère, Nicolas Bonod, Jérome Néauport

To cite this version:

Isam Ben Soltane, Guillaume Hallo, Chloé Lacombe, Laurent Lamaignère, Nicolas Bonod, et al.. Estimating and monitoring laser induced damage size on glass windows with a deep-learning based pipeline. Journal of the Optical Society of America. A Optics, Image Science, and Vision, 2022, 39 (10), 10.1364/JOSAA.462367 . hal-03813714

HAL Id: hal-03813714

https://cnrs.hal.science/hal-03813714v1

Submitted on 13 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimating and monitoring laser induced damage size on glass windows with a deep-learning based pipeline

ISAM BEN SOLTANE[1,*], GUILLAUME HALLO[2], CHLOÉ LACOMBE[2], LAURENT LAMAIGNÈRE[2], NICOLAS BONOD[1], AND JÉROME NÉAUPORT[2]

[1] Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, 13013 Marseille, France
[2] CEA, CESTA, F-33116, Le Barp, France
[*] Corresponding author: isam.ben-soltane@fresnel.fr

Compiled October 13, 2022

Laser induced damage is a major issue in high power laser facilities such as Laser MegaJoule (LMJ) and National Ignition Facility (NIF) since they lower the efficiency of optical components and may even require their replacement. This problem occurs mainly in the final stages of the laser beamlines and in particular in the glass windows through which laser beams enter the central vacuum chamber. Monitoring such damage sites in high energy laser facilities is therefore of major importance. However, the automatic monitoring of damage sites is challenging due to the small size of damage sites and to the low resolution images provided by the onsite camera used to monitor their occurrence. A systematic approach based on a deep-learning computer vision pipeline is introduced to estimate the dimensions of damage sites of the glass windows of the LMJ facility. The ability of the pipeline to specialize in the estimation of damage sites of size less than the repair threshold is demonstrated by showing its higher efficiency than classical machine learning approaches in the specific case of damage site images. In addition, its performances on three datasets are evaluated to show both it robustness and accuracy. © 2022 Optica Publishing Group

http://dx.doi.org/10.1364/ao.XX.XXXXXX

1. INTRODUCTION

The Laser MegaJoule (LMJ) in France, the National Ignition Facility (NIF) in the United States and ShenGuang-IV (SG-IV) in China are high energy laser facilities designed to achieve fusion ignition experiments by inertial confinement [1–3]. The final optics, which have a side length of 40 cm, are illuminated by the laser beams and may thus suffer from laser-induced damage, defined as a permanent change of the optical components induced by laser beams [4–7]. Laser damage initiation is due to a combination of loading induced by the UV laser beam [8] on random defects [9, 10] or particulate contamination on the optics surface [11, 12] or even laser self-focusing [13]. Initiated damage may grow after each laser shot when the laser energy is greater than the growth threshold [14]. Damage growth on optics can thus limit the available laser energy for each experiments.

In order to mitigate damage growth, some techniques have been developed such as locally reducing the laser energy under the growth threshold at damage positions by shadowing [15] or using CO_2 laser optics mitigation [16, 17]. The effective and optimal use of these methods relies on the early detection of the growth of damage sites ranging from 50 µm to 750 µm in diameter. These damage diameters are close to the field of view of one pixel for observation systems of high energy laser facilities such as the Final Optics Damage Inspection (FODI) at NIF [18], the SG-III FODI [19] and the Chamber Center Diagnostic Module (MDCC) on the LMJ facility [15, 20]. In order to make damage sites visible, observed optics are illuminated by their edges, resulting in dark-field pictures on which damage sites appear as bright spots. For the LMJ, damage sites on the final optics are illuminated by two green LEDs. The MDCC imaging system acquires images of every 176 windows after each laser shot. These images have a definition of 4096 × 4096 pixels with a resolution of 100 µm/pixel and a pixel depth of 16 bits.

Image analysis methods have been developed to detect and track damage growth as early as possible. They are mostly based on two steps: damage detection using different algorithms and then precise reconstruction of damage shape. To identify the damage sites in images, an algorithm based on the analysis of the Local Area Signal-to-Noise Ratio (LASNR) for each pixel of each image was proposed [21]. Another method based on local area signal strength and 2D histogram was then proposed to detect damage sites [22]. To reconstruct the shape of damage sites precisely, a region growing algorithm is used [21]. In the case of the LMJ facility, this image processing was further improved by using image registration principles to extract, after motion, brightness and contrast corrections, 51 × 51 pixel sub-images of individual damage sites from the MDCC images (Fig. 1). This

approach offers sub-pixel resolution [23]. In this work, a deep-learning computer vision pipeline to detect "real" damage sites from potential false alarms in the aforementioned sub-image dataset and to measure damage site diameters is used.

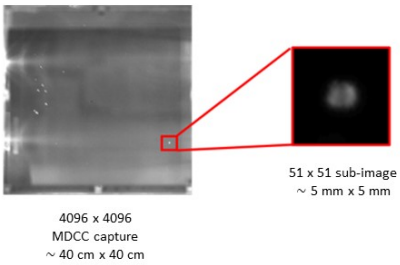


Fig. 1. Image of a glass window acquired by the MDCC camera. The white pixels correspond to potential damage sites detected by the image registration method to produce sub-images with 51 × 51 pixels such as the one displayed in the red box.

The interest for machine learning, and especially deep learning, over classical algorithms *i.e.* non-learning algorithms, arises from its ability to generalize well for unknown data by automatically generating relevant variables from a set of images to represent the data. Despite the growing amount of reviews highlighting this ability [24], the underlying mechanisms of such performance have not been unveiled [25]. Supervised machine learning algorithms for image processing usually work in two steps. The first one is a feature extraction step where variables of interest are manually or automatically inferred from the data (the images). The set of extracted features is usually referred to as the latent space representation. In the second step, these features are combined and used by a series of functions to produce new variables which are solutions to the task at hand. Machine learning has been increasingly used for studying laser induced damage [26]. Methods to classify damage sites have been developed using either manually extracted features directly linked to physical properties when enough data was available to discriminate between damage and artefacts [27], or automatically extracted ones when the data is composed of high resolution images only to separate the damage sites based on their morphology. Similarly, deep learning algorithms have been applied to damage detection problems [28, 29] in which binary segmentation maps of optical components are generated, where ones are pixels associated with laser-induced damage and zeros with the rest. However, research on damage site localization problems, which aim at detecting and estimating the size of damage sites using bounding boxes, is sparse and it is the main goal of the solution proposed in this article. In addition, in supervised learning the emphasis is usually put on the end result, *i.e.* the output of the algorithms, rather than on its relation with the extracted features or latent space. Exploiting this relation would diminish the black-box behaviour of the associated deep learning algorithms, which could lead to a better characterization of laser induced damages on sensitive optical components. Several very performant localization algorithms exist (such as Mask R-CNN [30]) or YOLO [31]), and while they rely on successive sub-components to obtain good results, they train them simultaneously and do not get rid of the aforementioned black-box behaviour. Inspired by these methods, we developed a pipeline or a sequence of algorithms, to isolate and optimize independently the relevant sub-tasks which we deemed essential to the localization and characterization of damage sites. We first describe this approach and then compare it to simple alternatives. This approach is based on damage site images provided by the on-site MDCC camera. Let us stress out that we also benchmark it on a third dataset generated in the MELBA facility, a centimeter aperture laser damage testing set-up equipped with high resolution diagnostics [8] giving access to images at a 6.5 µm resolution (damage size and fluence distribution).

2. MATERIAL AND METHODS

A. Datasets

Three datasets were used. The first to train the pipeline and evaluate its performances while the two others were exclusively used to test its precision and robustness. Each dataset was composed of labeled sub-images, *i.e.* input sub-images and target variables: a target class indicating if they contain real damage sites or artefacts (reflections, false alarms, or only noise), as well as the target bounding box around the damage site when relevant. Each bounding box was provided as a 4-dimensional vector (x_0, y_0, w, h), where (x_0, y_0) is the center of the bounding box on the sub-image, and w and h are respectively the width and height of the bounding box, which were used as estimations of the size of the damage sites. The first and main dataset (LMJ 1) was composed of 11620 sub-images of 51 × 51 pixels, corresponding to 166 damage site locations on one full glass window, captured at 70 different times. The captures were performed after one laser shot or after a sequence of laser shots. The intensities are displayed in gray levels ranging from 0 to 2^{16}, and the objects of interest are assumed to be at the center of the sub-images (Fig. 1). The targets were automatically generated using a skimage module [32], and their quality was manually assessed. Fig. 2 shows the distribution of the width and height of the target bounding boxes on damage site sub-images. The

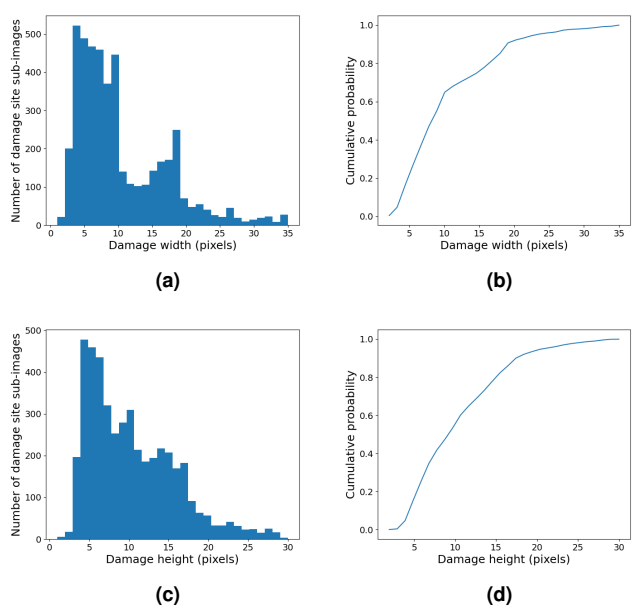


Fig. 2. Histogram (first column) and cumulative probability (second column) of the width (a, b) and height (c, d) of the target bounding boxes of damage images in the LMJ 1 dataset.

width and/or height of most damage sites were smaller than 7 pixels, which corresponds to the aforementioned limit on the size of reparable damages (750 µm). Note that a few damage sites (≈ 100) had a bounding box height or width larger than 20 pixels, or approximately 2 mm. These very large damage sites cannot be repaired. Damage sites should be detected as early as possible to prevent them from reaching such dimensions, especially when they have a rapid growth.

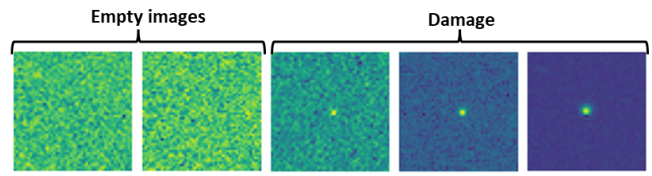
The main dataset (LMJ 1) was further split into 3 sets of damage site locations. For the learning phase of the networks, the first 100 site locations (7000 sub-images) were divided into 70 training locations (4900 damage site sub-images) and 30 validation locations (2100 sub-images). The remaining 66 site locations were kept for the testing phase. The validation data was used to keep track of effects such as overfitting during the training phase. The testing sites were used afterward to perform a statistical analysis of the performances of the trained network. The second dataset (LMJ 2) contained approximately 24000 sub-images acquired by the MDCC camera on different windows, after a sequence of laser beam shots or a single shot. The pixel intensities were similar to those from the main dataset. Since the damage sites were located on different windows, the date of acquisition by the MDCC camera varied from site to site, and there were irregularities such as the number of sub-images available for one particular site location. These irregularities of the data provided the means to test the pipeline with data more representative of the current operation of the MDCC on the LMJ facility. The associated targets were automatically generated. Despite the similarities of this dataset with the LMJ 1 dataset, its generated targets were not characterized as deeply and were thus used for performance assessment only. Finally, the third dataset was composed of more than 100000 sub-images obtained in the MELBA experimental set-up. It allowed the study of the growth of laser induced damage when irradiated with fluences matching those of the LMJ facility, with precise measurements providing high-resolution images (6.5 µm/pixel against 100 µm/pixel for the LMJ datasets). The purpose of this dataset was to test the robustness of the size estimation network and to exploit the high resolution of the sub-images to get an accurate validation database. In order to obtain accurate targets, the bounding boxes were automatically generated using the high-resolution sub-images. Additionally, a low-resolution version of this dataset was generated. The stability of the complete pipeline was assessed by comparing its output on this low-resolution version to the targets from the high resolution one. This dataset presented significant differences with the other two, which prevented us from training the network on it despite the high amount of data, as it might have led to a reduced performance for the LMJ data.

B. Description of the pipeline

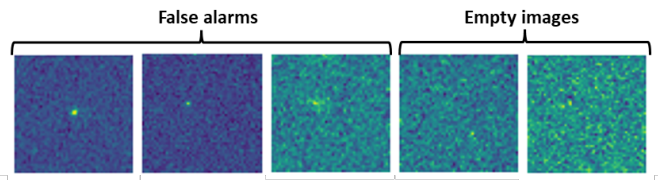
The pipeline is a sequence of algorithms using deep-learning to perform specific tasks which were deemed essential to the proper estimation of damage site size.

B.1. Sub-images pre-classification

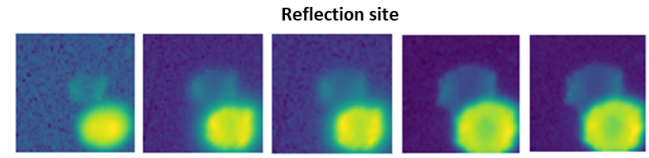
The first step was to classify the sub-images. Not all sub-images contained damage sites, and it was important to discriminate between real damage sites and artefacts. If the low resolution does not permit to obtain an accurate classification of the damage sites on the basis of their morphology, unlike the work proposed in ref. [33] where laser induced damage sites were classified into different morphologies using high-resolution sub-images, it is still possible to discriminate between sub-images of damage sites and artefacts. Three types of artefacts were considered: reflection sites, empty sub-images and false alarms. The first type is due to the MDCC LED lighting, while the two others are caused by the image registration method. Means to detect empty sub-images and reflection sites were developed and incorporated into the pipeline, but not for the false alarms which can be easily filtered out at the end by checking for locations with decreasing intensities or moving objects. In Fig. 3 are shown subsets of sub-images of three locations of potential damage sites in the main dataset LMJ 1, to illustrate the concept of empty sub-images false alarms and reflection sites.



(a) Damage site appearing after several shots. The first sub-images are empty



(b) False alarm detected as a damage site in the first sub-image. After some time, it disappears and the sub-images are empty.



(c) Reflection of a growing damage site at the center of the sub-images. The reflected damage is brighter, at the bottom.

Fig. 3. Evolution of three locations of potential damage sites with laser shots, from early acquisition by the MDCC camera on the left to more recent ones on the right.

A segmentation was then performed on the sub-images to identify the reflection-free ones containing damage sites. Bounding boxes around the damage sites were generated from the resulting segmented sub-images. They were fed as training targets to the the last damage size estimation network (or bounding box regression network).

Detection of empty sub-images A manual classification of the sub-images was performed to discriminate between empty and non-empty images. The set of sub-images resulting from this manual classification was used to train the neural network schematized in Fig. 4 to automatically detect empty sub-images. This simple convolutional neural network is composed of 4 convolutional layers, which produce $5\times5\times64$ images from one 51×51 sub-images, feeding into a sigmoid-activated artificial neuron outputting the probability of an input sub-image being empty. The input sub-images were augmented, *i.e.* rotated around the center and flipped. The use of data augmentation prevented the network from overfitting and improved its ability of generalization. The chosen loss function was the binary cross

entropy between the target and predicted values, optimized with the Adam optimizer with a default learning rate of 0.001. The performance of the network on the testing sites was estimated using the precision parameter, defined as the ratio between the number of correctly classified empty and non-empty sub-images over the total number of sub-images. A precision of 95.38 % was obtained.

Detection of reflection sites

The sub-images were then further classified to filter out those containing both a damage site and its reflection, resulting from the multi-reflection of light on the two sides of the vacuum window. Due to the slight tilting angle of the MDCC (up to 3°), the light beam illuminating the damage sites may be reflected on the front face of the vacuum window before going through the rear face again and reaching the MDCC camera. Both reflection and real damage sites appear on the sub-images. The intensity of the reflection site is significantly lower than that of the damage site. Thus only large damage sites with a higher scattered light intensity may lead to the generation of reflection sites. We call $< I_{damage} >$ the average light intensity of a damage site. The reflected site is obtained after a reflection on the front face of the glass window (reflection coefficient r) followed by a transmission through the rear face before reaching the MDCC camera. The expected ratio between the average reflection light intensity $< I_{reflection} >$ and the real damage light is therefore:

$$\frac{< I_{reflection} >}{< I_{damage} >} \approx r \quad \textbf{(1)}$$

In all three datasets, the objects of interest were always at the center of the 51×51 sub-images. This led to the following criterion for reflection:

$$\frac{I(25,25)}{max(I)} \leq 2.r, \quad \textbf{(2)}$$

where $I(25,25)$ denotes the intensity at pixel $(25,25)$, *i.e.* at the center of the sub-image. The factor 2 was arbitrarily added to take into account the high intensity variations which may be observed across the pixels of a single object. The sub-images were considered empty whenever the calculated ratio was below $2.r$, the sub-image is considered centered on a damage reflection. Despite its simplicity, this criterion provided an accuracy of 99.6% on the test dataset.

B.2. Sub-image segmentation

This task was performed with a UNET-like architecture [34], which had already proven its efficiency in damage detection (let us stress out that what is usually referred to as damage detection is actually damage site image segmentation) on large images [28]. The network architecture is shown in Fig. 5. The U shape of the network was kept. The left part of the U reduced the width and height of the images, added more channels and extracted useful features, while the right part combined different levels of features (three levels are visible in Fig. 5) to construct the target image from relevant information.

Reflection-free and non-empty sub-images only were used at this stage (around 2700 sub-images out of the 7000 present in the training subset of the LMJ 1 dataset), and the targets were automatically generated using the segmentation methods from the Python scikit-image module, a library of image processing tools [35]. Emphasis was put on the detection of small damage sites, with small grayscale values close to the noise level of the acquired sub-images. Although the network performed better than the classical machine learning approach, the price to pay for high performances was a detection of false damage sites in some sub-images. It was thus necessary to filter out irrelevant (*i.e.* non-centered) objects from the sub-images using the threshold-based tools provided by scikit-image (skimage). The results of the segmentation step for some damage sites with low pixel intensity values are presented in Fig. 6.

The first row displays 4 input sub-images of the algorithm, representing damage sites in different environments. a1 shows a damage site with an over-illuminated band on the right, which is considered as noise, b1 is an empty sub-image, c1 is a damage site of about 3 pixels ($\approx 300\,\mu$m), with background noise presenting a local maximum, and d1 corresponds to a larger damage site with locally high pixel intensities in the bottom-left corner (visually invisible due to the high pixel intensities of the damage site itself), which corresponds to a reflection. The second row shows the result of the segmentation process using the tools provided by the skimage module [32]. For these 4 specific examples, the threshold-based methods failed to extract useful information only. In a2, only the high intensity band shown in a1 was segmented. Despite the seemingly low noise levels in the sub-images b1 and c1, the result of the segmentation is inaccurate and even shows noise as damage sites in b2 and c2. As expected, the high intensity noise in d1 was difficult to isolate, and appeared as a damage in d2. The third row contains the result of the segmentation step using our pipeline. a3, b3,

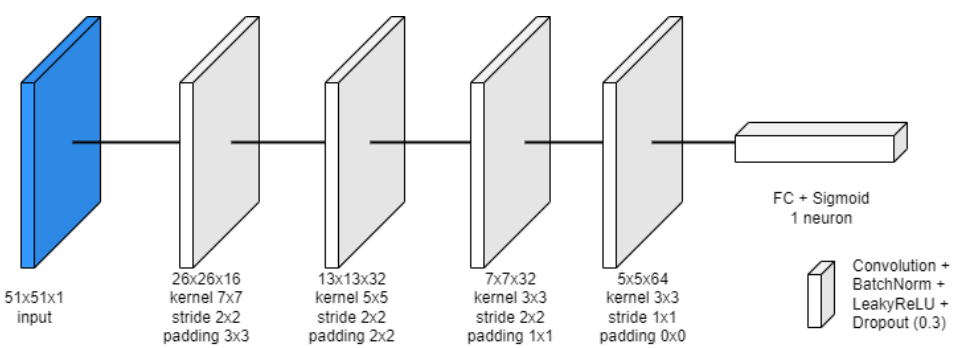


Fig. 4. Architecture of the convolutional network tasked with the detection of empty sub-images. The output dimensions, kernel size, stride and padding are indicated under the convolutional blocks, which are composed of a convolutional layer, a batch normalization, a droupout layer and a leaky ReLU activation function.

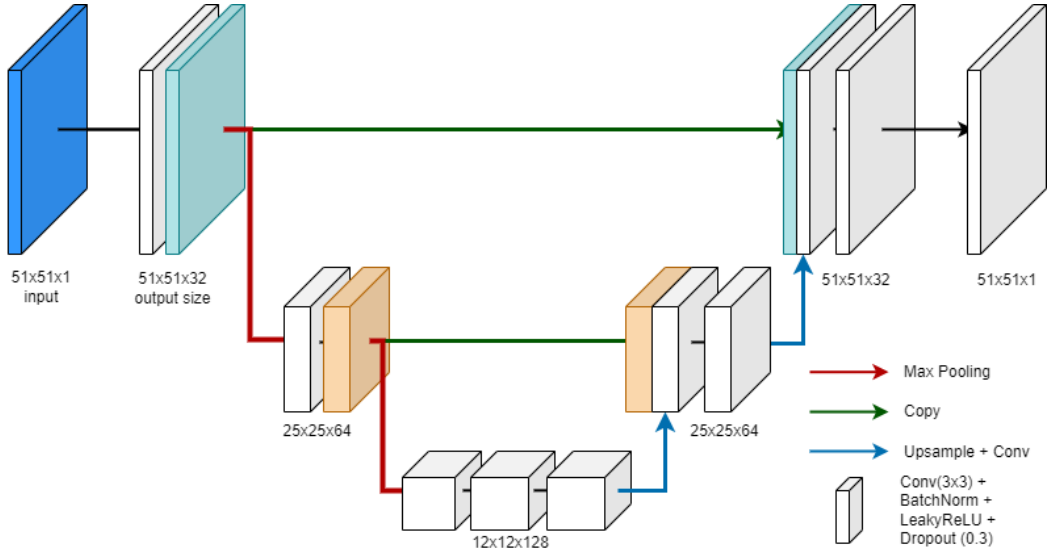


Fig. 5. Architecture of the UNET-like convolutional neural network used to generate the segmentations of the damage sub-images. The left part extracts features while reducing the size of the sub-images, and the right part combines these features to construct the expected output.

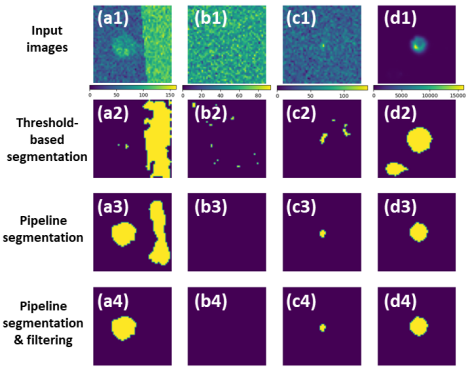


Fig. 6. Comparison of segmentation+filtering in the pipeline and using skimage threshold-based segmentation method. The first row displays 4 input sub-images. In a1 can be observed a high intensity vertical band due to LED light on the side of the glass windows. In b1 is displayed a sub-image without damage site. c1 is a small damage site with intensity levels close to the background. d1 is a large damage with locally high pixel intensities in the bottom-left corner. The second row shows the segmentation using skimage methods on the 4 input sub-images. The third row is the segmentation using the pipeline, and the fourth one combines this segmentation with an object filtering method.

c3 and d3 show an accurate segmentation of the central objects present in the input sub-images, despite the noisy background. In certain cases, other non-centered objects might appear, such as the large band in a3. The pipeline segmentation was therefore combined with a simple filtering method to keep only the centered objects. The results of this complete process for all 4 sub-images of the first row is presented in a4, b4, c4 and d4.

B.3. Size estimation of laser induced damage sites

At this stage, the target bounding boxes were used to train a convolutional neural network using the information obtained *via* the previous steps in order to enrich the loss function and maximize the performances. The variational neural network with the architecture shown in Fig. 7 was used. First, the 2-channel inputs, composed of the damage site sub-image as well as the corresponding segmented sub-image, propagate through convolutional layers followed by fully connected (FC) layers to produce latent vectors of dimension 128, *i.e.* intermediate representations of the inputs which constitute the latent space. In the specific case of object localization, the latent space is expected to embed relevant information about the damage sites in the 128 coordinates of the intermediate representations. To improve the efficiency of the information embedded into the latent vectors, the construction of the latent space was constrained by forcing each of the 128 coordinates to be normally distributed using two 128 coordinate vectors $\vec{\mu}$ and $\vec{\sigma}$, which are respectively the mean and standard deviations of the coordinates of the latent vectors. Finally, the latent vectors were used by another FC layer (the decoder made of 4 neurons) to output an approximation $(\tilde{x_0}, \tilde{y_0}, \tilde{w}, \tilde{h})$ of the target bounding box for each input.

The network was trained on augmented data, *i.e.* rotated and flipped sub-images. The loss function was designed to allow the network to perform better on damage sites smaller than 7 pixels, to discriminate between damage site sub-images and artefacts, and to force the normal distribution of the coordinates of the latent vectors. This was achieved with a combination of the Smooth L1 Loss with several other terms, each accounting for one of the specific wanted effects.

C. Performance metrics

Object detection aims at localizing objects (object localization), *i.e.* finding bounding boxes around them, and classifying them (classification). Although the means to evaluate the performances of algorithms specifically designed for the detection of objects have been exhaustively studied for problems different from that of laser-induced damage, little interest has been shown on the evaluation of the performance of object localization algorithms. Most object detection algorithms generate several candidate bounding boxes (regions in Mask R-CNN [30] or anchors in YOLO [31])

and filter out irrelevant ones using the Intersection Over Union score (IoU). We chose the IoU score as an object localization score which measures the overlap between the predicted and target bounding boxes instead of using it as a filtering tool. However, it did not provide sufficient information to properly characterize the object localization methods at play (the pipeline or its alternatives). Three metrics providing similar results but holding complementary information were thus added. An overview of the calculations of the four scores is displayed in Fig. 8. T is the set of pixels contained within a target bounding box, and P is the set of pixels contained within the associated predicted bounding box generated by an object localization algorithm.

C.1. Intersection over union (IoU)

The IoU is defined as:

$$IoU = \frac{|P \cap T|}{|P \cup T|}. \tag{3}$$

It measures how well two bounding boxes overlap, and yields consistent results for large objects, *i.e.* objects larger than 10 pixels, even with an error of 1 to 2 pixels on the dimensions or position of the predicted bounding boxes. For damage sites of width or height less than 7 pixels (750 µm), which are the damage sites of interest since they are those which can be repaired, a drop of the IoU score depending on the ability of the bounding box prediction solutions to find such objects should thus be expected. The IoU does not hold information on the distance between predicted and target bounding box centers. If the bounding boxes do not overlap, the score will remain 0, no matter how far their centers are. Therefore, studying failing cases may prove difficult, in particular in the case of small objects (with width or height less than 2 pixels). This is why the Generalized IoU (GIoU) [36] was also considered.

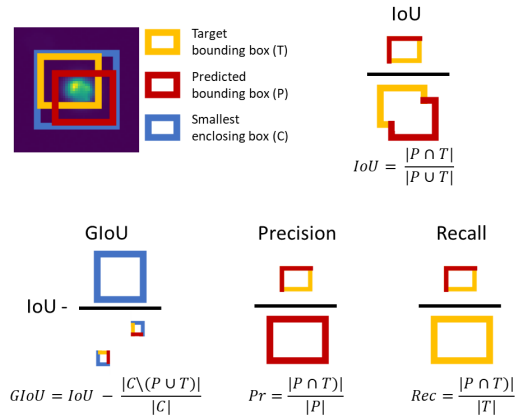


Fig. 8. Calculations of the Intersection over Union (IoU), Generalized IoU (GIoU), Precision (Pr) and Recall (Rec) using the Target bounding box (T) of a damage site, the Predicted bounding box (P) generated by an algorithm and the Smallest enclosing box (C) of the predicted and target bounding boxes.

C.2. Generalized intersection over union (GIoU)

The GIoU is defined as:

$$GIoU = IoU - \frac{|C \setminus (P \cup T)|}{|C|}. \tag{4}$$

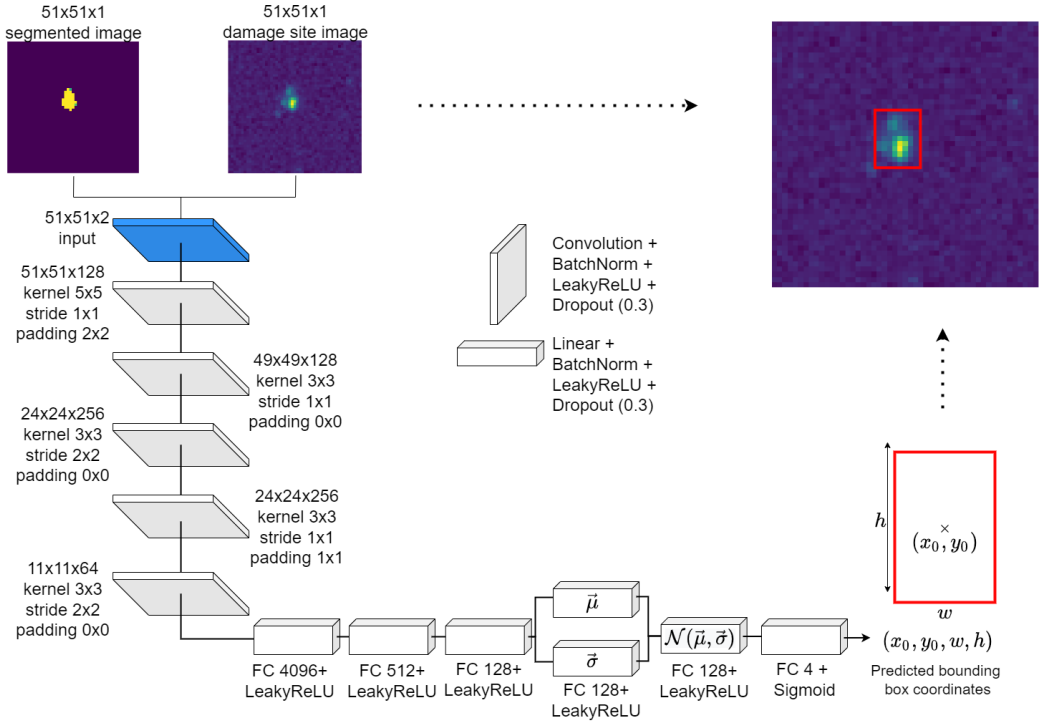


Fig. 7. Architecture of the variational bounding box regression neural network used to estimate the size of the damage sites. Each damage site sub-image and its segmentation are fed into convolutional layers to extract relevant features, before propagating through fully connected layers to generate the latent vector $\mathcal{L}$. The coordinates of L are all normally distributed, with means and standard deviations $\vec{\mu}$ and $\vec{\sigma}$. From the latent vector, the predicted bounding box target with coordinates (x_0, y_0, w, h) is produced (red rectangle).

where C is the set of pixels within the rectangle containing both P and T. The values of the GIoU range from -1 to 1, where the values in the [-1;0] range correspond to an IoU of 0. It shows a great agreement with the IoU for high values, while offering more information on damage sites with low IoU values. The GIoU could in theory replace the IoU, but it is generally less employed. Therefore, both metrics were chosen.

C.3. Precision and recall

Finally, the expression of precision is

$$pr = \frac{|P \cap T|}{|P|}. \tag{5}$$

and recall is defined as

$$rec = \frac{|P \cap T|}{|T|}. \tag{6}$$

Using these definitions, the precision indicates how well the algorithms are capable of finding the pixels which best represent the objects of interest, *i.e.* damage sites, while the recall provides information regarding the ability of the algorithms to encapsulate the entirety of a damage site. A high precision but low recall means that the tested algorithm may find some features to identify the damage sites, but is not capable of estimating their size. The IoU is positively correlated with the average of both metrics, and looking at each one of them separately is another means to have an in-depth view of sub-images with similar IoU values.

3. RESULTS AND DISCUSSION

A. Comparison of the different solutions

The performances of the pipeline were evaluated with the chosen metrics in the case of sub-images containing damage sites only. The test subset of the main LMJ dataset (LMJ 1) was thus reduced to only 1700 sub-images of real damage sites. However, training subset also contained sub-images associated with artefacts or even empty sub-images. In these cases, the coordinates of the bounding boxes were 0. All 4 metrics were computed for every sub-images, for the pipeline and the 6 alternatives presented in table 1: a simple convolutional neural network (CNN), the random forest algorithm (RF), and the ResNet50 [37] or Inception V3 [38] networks for the feature extraction, coupled with either a neural network (RN50+VNN, IV3+VNN) or the random forest algorithm (RF). Each alternative was optimized in order to provide the best possible results.

Despite the metrics supporting the results, they should be viewed as mostly qualitative. The comparison of machine learning algorithms is a complex task, which requires the consideration of many parameters. Several means to improve the results of the alternatives to the pipeline exist. However, the results should provide sufficient information to justify the design of a pipeline to tackle the damage site localization problem.

First, the mean GIoU for RF algorithm is well below 0, which evidences the fact that using the intensity of isolated pixels alone is not enough to approximate the dimensions of a damage site. Transfer-learning based methods, namely RN50 and IV3 + VNN or RF, provide better results. The choice of the RF algorithm was inspired by reference [33], which shows that feeding a set of relevant features extracted by a pre-trained backbone to decision trees (and thus the RF algorithm) should provide good results. However, low IoU values were obtained for the RN50+RF and IV3+RF solutions with our datasets. These results point out the influence of the specificity of the dataset. The resolution of the MDCC images is too low relative to the size of the damage sites in the 51 × 51 pixels sub-images, which is different from what is usually observed in the datasets with which the networks were trained. Still, the backbones provide a better representation of the damage sites than the sub-image pixel intensities (*i.e.* the RF algorithm only). Finally, combining RN50 and IV3 with the VNN instead provides better results, with an IoU exceeding 0.5, and precisions and recalls clearly demonstrating the ability of the overall architectures to find suitable bounding boxes for most damage sites. Finally, the CNN was trained specifically for datasets related to laser-induced damage. Therefore, the latent vectors were more relevant than in the case of pre-trained backbones RN50 and IV3, as evidenced by the increased values of the IoU and GIoU. However, the performances of the CNN drop for smaller damage sites, which are those of interest, as evidenced in Fig. 9, on which the GIoU scores for the damage site sub-images of the test subset as well as the diagonal length of the bounding boxes (referred to as damage size) are displayed. Therefore, the CNN does not seem to be a practical solution. In the end, the high values of the performance metrics observed for the pipeline justify its use over other solutions.

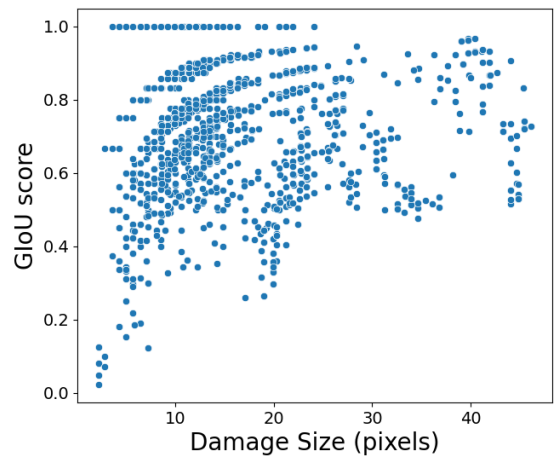


Fig. 9. Scatter plot of the GIoU score and damage size (bounding box diagonal length, $\sqrt{surface}$) for each of the damage sites from the test subset of the LMJ 1 dataset. The predictions were generated by the one-step CNN. Each point is a damage site.

B. Performances of the pipeline

Table 2 contains the mean values of the different metrics for the two LMJ datasets (LMJ 1, LMJ 2) and the MELBA dataset.

The IoU is greater than 0.7 for all three datasets, which indicates the ability of the network to properly predict the bounding boxes in all three cases. The small differences observed between the mean IoU and GIoU are evidence of a good consistency throughout all the damage sites. Finally, the recall values are greater than 0.9, which means that the predicted bounding boxes almost completely encapsulate the target bounding boxes. Thus, the discrepancies between the predicted and target bounding boxes are mostly due to an overestimation of the damage size by the predicted boxes compared to the target ones. This lowers the risk of wrongly assuming that a damage site is still repairable.

	Pipeline	CNN	RN50+VNN	IV3+VNN	RN50+RF	IV3+RF	RF
IoU	0.910	0.724	0.516	0.502	0.444	0.463	0.111
GIoU	0.908	0.715	0.489	0.463	0.364	0.417	-0.398
pr	0.940	0.883	0.830	0.719	0.618	0.612	0.189
rec	0.966	0.817	0.581	0.643	0.576	0.686	0.195

Table 1. Mean values of the Intersection over Union (IoU) and Generalized Intersection over Union (GIoU) calculated for the pipeline and the 6 alternatives: single-step convolutional neural network (CNN), ResNet50 with a variational neural network as decoder (RN50+VNN), Inception V3 with a VNN (IV3+VNN), ResNet50 with the random forest algorithm as decoder (RN50+RF), Inception V3 with the RF (IV3+RF), and the Random Forest only (RF). The scores were evaluated for the 1700 damage sites sub-images of the testing subset of the main LMJ dataset.

	LMJ 1	LMJ 2	MELBA
IoU	0.910	0.938	0.729
GIoU	0.908	0.937	0.729
precision	0.940	0.964	0.743
recall	0.966	0.971	0.976

Table 2. Mean values of the IoU, GIoU, precision and recall calculated on the 3 datasets: the test subset of the main dataset (LMJ 1), the second dataset (LMJ 2) and the MELBA dataset. The predicted bounding boxes were generated using the pipeline.

B.1. Efficiency based on damage size

In Fig. 10, the histogram of the IoU values for all the damage sites of the test subset (a) are plotted. The damage sites of interest, those with width or height less than 750 µm, were filtered out to generate (b).

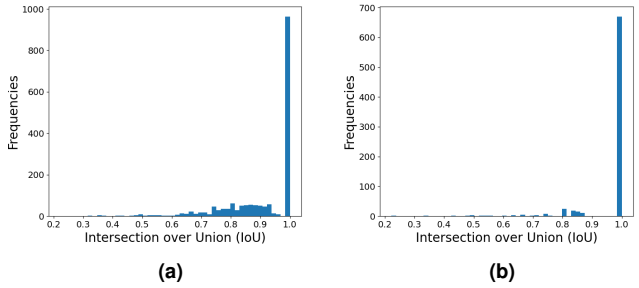


Fig. 10. Histogram of the computed IoU score for (a) the complete subset of the main dataset, and (b) only the damage sites with a predicted bounding box diagonal length less than 10 pixels.

Out of the 773 filtered damage sites present in Fig. 10 (b), 710 have an IoU value greater than 0.8, among which 674 have a perfect score. The distribution for all the damage sites showcase a few failing damage sites, 135 potential damage sites with an IoU less than 0.7 out of 1700 damage site sub-images. These 135 failing cases correspond to either empty sub-images (containing no damage site), false alarms which the pre-classification step failed to filter out, large damage sites (more than 10 pixels in height and width) or very elongated damage sites. Examples of these failures are provided in Fig. 11.

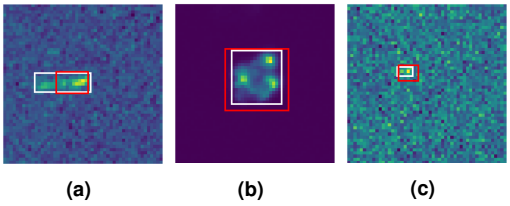


Fig. 11. Examples of damage sites for which the pipeline failed. In (a) an elongated damage site, in (b) a large damage site, and in (c) a false alarm wrongly labelled as damage site. The red rectangles are the predictions, the white ones are the targets.

In the case of large damage sites and elongated ones, the predicted bounding boxes widths and/or heights were greater than 7 pixels, thus above the size threshold of 750 µm, which is enough for us to monitor them carefully. Overall, the network performs better on smaller damage site sub-images. This is shown in Fig. 12.

In Fig. 12 (a), each point is a damage site sub-image of the test subset. They are positioned based on their predicted bounding box diagonal length (labelled as damage size) and their associated IoU score. Since the set of diagonal length values is finite and smaller than the number of damage sites, a 2D density plot is also displayed in Fig. 12 (b) to highlight the distribution of the damage sites in (a). The network generates more accurate predictions for damage sites with predicted diagonal lengths less than 10 pixels. As the size of the damage sites increases, good predictions are still obtained, but with lower IoU values than in the case of the smallest damage sites. These results agree with the way the bounding box regression network was trained: the loss function was designed in such a way that the pipeline would pay more attention to the smallest damage sites. This is highlighted in Fig. 13.

Fig. 13 (a) displays the norm of the latent vectors generated by the encoder of the bounding box regression network for all the sub-images of the damage site (approximately 4000 input sub-images), and the diagonal length of the predicted bounding boxes. The diagonal length of the predictions is highly correlated with the diagonal length of the target bounding boxes as shown in Fig. 13 (c). Fig. 13 (b) is a zoom in on the damage sites with a diagonal length less than 12 pixels. For real damage sites, the range of values taken by the norm of the latent vectors widens as the diagonal length increases. The network thus has a clearer representation of the smaller damage sites, and is capable of

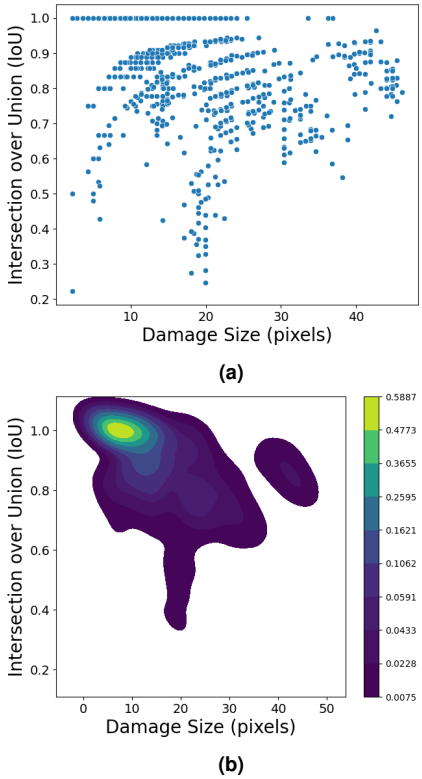


Fig. 12. Scatter plot (a) and 2D density plot (b) of the IoU score and damage size, defined as the diagonal length of the bounding box, for each of the damage site from the test subset of the LMJ 1 dataset. Each point corresponds to a damage site.

accurately associating bounding boxes to them. This is further highlighted by the high Pearson and Spearman correlations between the norm of the latent vectors of smaller damage sites (diagonal length less than 10 pixels) and the diagonal length in table 3.

	All sites	Sites < 10 pixels
Pearson corr.	0.557	-0.613
Spearman corr.	0.192	-0.750

Table 3. Pearson and Spearman correlation coefficients between the norm of the latent vectors of the damage sites in the main LMJ dataset, and the diagonal length of the bounding boxes.

In addition, the norm of the latent vectors of artefacts (reflections, unidentified objects or empty sub-images) is clearly separated from the norm of the real damage sites of smaller size. The artefacts, identified in the pre-classification step of the pipeline, were associated with out-of-distribution latent vectors, namely vectors with great norm values relative to the mean value and standard deviations of all latent vectors. The limit between real damage sites and artefacts becomes fuzzy as the size of damage sites increases. Even though the biggest damage sites are difficult to tell apart from artefacts from their latent vector norms only, Fig. 13 (d) shows that the artefacts with a vector norm close to the norm of bigger damage sites are reflections (green points), which can be identified with great precision in

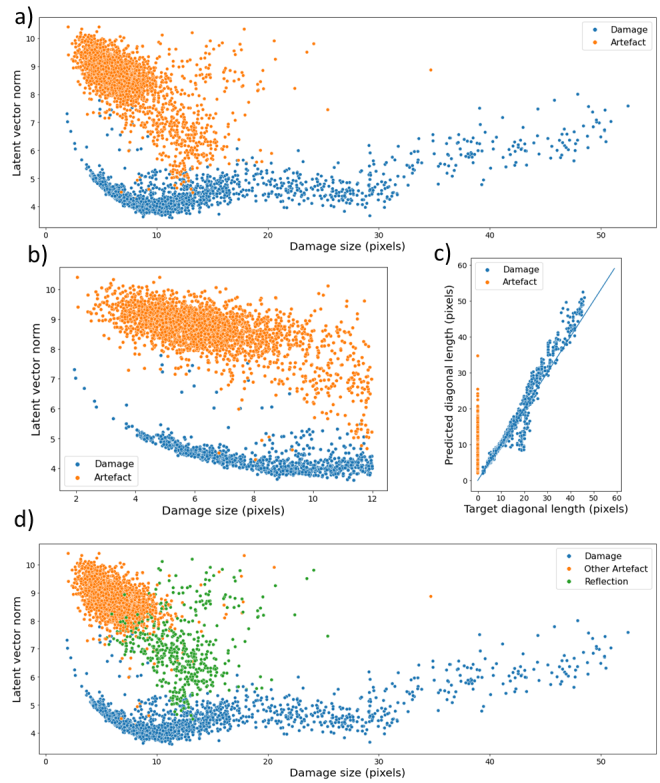


Fig. 13. (a) Scatter plot of the damage size, *i.e.* diagonal length of the predicted bounding box, and norm of the latent vectors generated by the size regression network on the sites of the test subset of the LMJ 1 dataset. Each point is a potential damage site. (b) Zoom of the scatter plot for site with damage size lower than 12 pixels ($\approx 1200\,\mu$m). (c) Scatter plot of the predicted damage size and target diagonal length for all the sites. (d) Same scatter plot as (a), where the artefacts were split into reflections (green points) and other artefacts (red points).

the pre-classification step.

Finally, Fig. 14 displays the results of the bounding box prediction of 9 randomly selected sub-images of the test subset. The red bounding boxes are predictions, the white ones correspond to targets.

B.2. Robustness of the pipeline

The performances of the size regression network were also studied with the LMJ 2 dataset. Unlike the LMJ 1 dataset, the sub-images of this dataset were extracted from the MDCC images of different glass-windows of the vacuum chamber of the LMJ facility, and not just one. The proportion of smaller damage sites was higher than it was in the LMJ 1 dataset. Therefore, all 4 metrics in the second column of table 2 indicate better results than in the first column.

The 2D density plot of the IoU and predicted diagonal length of all damage sites in the dataset is shown in Fig. 15 (a), and the histogram of the IoU score in (b). The results are very similar to those obtained with the test subset of the LMJ 1 dataset. However, there is a higher density of smaller damage sites in Fig. 15 (a), and only a few sub-images (379 out of 7041 damage sites) with an IoU smaller than 0.7. Most of these failures were artefacts which were wrongly labelled as damage sites, *i.e.* false alarms: sub-images containing moving objects classified as dam-

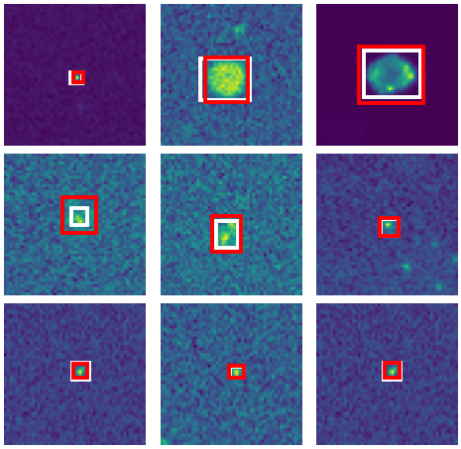


Fig. 14. Results of the bounding box regression obtained for 9 damage sites from the test subset of the main, LMJ 1 dataset. The white bounding boxes correspond to target bounding boxes, while the red bounding boxes are predictions generated by the pipeline.

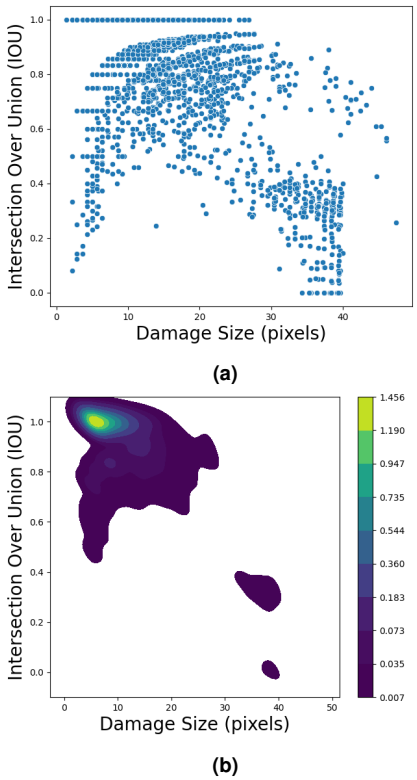


Fig. 15. Scatter plot (a) and 2D density plot (b) of the IoU score and damage size, which is the diagonal length of the bounding box, for each of the damage site from the LMJ 2 dataset. Each point corresponds to a damage site.

age. These artefacts may be removed by adding a filtering step based on the consistency of the position of the detected objects in the sub-images. Given these results, we expect an accurate prediction of the size of most damage sites appearing on the glass windows of the LMJ facility.

Similarly, the pipeline still performs well with the MELBA dataset despite the drop in efficiency observed in table 2. The high recall value shows that the target bounding boxes are at least contained within the predicted ones. As mentioned before, the images in the dataset have a higher default resolution and are bigger (771×771 pixels). In order to use them with the pipeline, they needed to be resized. This led to the disappearance of damage sites with width or height smaller than 50 µm and thus empty target bounding boxes, which is why the recall value remained high compared to the one obtained with LMJ datasets, while the IoU, GIoU and precision scores were significantly lower. However, the values of these metrics are still good enough to consider that the pipeline performs well, and robustly, with different datasets such as the MELBA one.

4. CONCLUSIONS

In this work, a deep-learning based pipeline designed to estimate the size of damage sites was presented. The efficiency of machine-learning approaches depends on the ability of the algorithms to generate a relevant intermediate representation, or latent space, of the input data, by incorporating metadata into it, *i.e.* additional information regarding the input data. As such, a lot of effort was put in the characterization of the data, and the resulting information was integrated in the form of sub-tasks, namely the detection of artefacts and empty images, and the segmentation of damage sites images. While common machine learning techniques usually rely solely on an input image to produce the output damage site size, it was demonstrated that, in comparison, the performances, which were measured through 4 complementary metrics measuring the overlap between predicted and target bounding boxes, were improved by considering these relevant sub-tasks. The emphasis was put on the ability of the pipeline to provide accurate estimations of damage sites with associated bounding box width or height smaller than 7 pixels, corresponding to the maximum size of $\approx 750\mu m$, below which the damage sites can still be repaired. We have shown that the pipeline is capable of generating robust predictions by evaluating it on two additional datasets. The LMJ 2 dataset, similar to the LMJ 1 dataset but more realistic, and the MELBA dataset, possessing many discrepancies with the LMJ ones. Accurate predictions were obtained for both datasets despite their difference with the training data.

Disclosures. The authors declare no conflicts of interest.

Data Availability Statement. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Acknowledgement. The authors thank François Hild for his careful reading of the manuscript and fruitful remarks.

REFERENCES

1. M. L. Andre, "Status of the lmj project," in Solid State Lasers for Application to Inertial Confinement Fusion: Second Annual International Conference, vol. 3047 (International Society for Optics and Photonics, 1997), pp. 38–42.
2. G. H. Miller, E. I. Moses, and C. R. Wuest, "The national ignition facility," Opt. Eng. **43**, 2841–2853 (2004).
3. X. He et al., "The updated advancements of inertial confinement fusion program in china," in Journal of Physics: Conference Series, vol. 688 (IOP Publishing, 2016), p. 012029.
4. S. G. Demos, M. Staggs, and M. R. Kozlowski, "Investigation of processes leading to damage growth in optical materials for large-aperture lasers," Appl. optics **41**, 3628–3633 (2002).
5. K. Manes, M. Spaeth, J. Adams, M. Bowers, J. Bude, C. Carr, A. Conder, D. Cross, S. Demos, J. D. Nicola et al., "Damage mechanisms

avoided or managed for nif large optics," Fusion Sci. Technol. **69**, 146–249 (2016).
6. P. Grua, L. Lamaignère, M. Chambonneau, R. Courchinoux, and J. Néauport, "Nanosecond laser damage initiation at 0.35μm in fused silica," Opt. Lett. **43**, 2692–2695 (2018).
7. T. Suratwala, R. Steele, M. Feit, L. Wong, P. Miller, J. Menapace, and P. Davis, "Effect of rogue particles on the sub-surface damage of fused silica during grinding/polishing," J. Non-Crystalline Solids **354**, 2023–2037 (2008).
8. M. Veinhard, O. Bonville, R. Courchinoux, R. Parreault, J.-Y. Natoli, and L. Lamaignère, "Quantification of laser-induced damage growth using fractal analysis," Opt. Lett. **42**, 5078–5081 (2017).
9. N. Bloembergen, "Role of cracks, pores, and absorbing inclusions on laser induced damage threshold at surfaces of transparent dielectrics," Appl. Opt. **12**, 661–664 (1973).
10. J. Neauport, L. Lamaignere, H. Bercegol, F. Pilon, and J.-C. Birolleau, "Polishing-induced contamination of fused silica optics and laser induced damage density at 351 nm," Opt. express **13**, 10163–10171 (2005).
11. D. M. Kane and D. R. Halfpenny, "Reduced threshold ultraviolet laser ablation of glass substrates with surface particle coverage: A mechanism for systematic surface laser damage," J. Appl. Phys. **87**, 4548–4552 (2000).
12. S. Palmier, S. Garcia, L. Lamaignère, M. Loiseau, T. Donval, J. L. Rullier, and et al., "Surface particulate contamination of the LIL optical components and their evolution under laser irradiation," in Laser-Induced Damage in Optical Materials: 2006, vol. 6403 G. J. Exarhos, A. H. Guenther, K. L. Lewis, D. Ristau, M. J. Soileau, and C. J. Stolz, eds., International Society for Optics and Photonics (SPIE, 2007), pp. 301 – 310.
13. M. J. Soileau, W. E. Williams, N. Mansour, and E. W. V. Stryland, "Laser-Induced Damage And The Role Of Self-Focusing," Opt. Eng. **28**, 1133 – 1144 (1989).
14. K. Manes, M. Spaeth, J. Adams, and M. Bowers, "Damage mechanisms avoided or managed for nif large optics," Fusion Sci. Technol. **69**, 146–249 (2016).
15. C. Lacombe, S. Vermersch, G. Hallo, M. Sozet, P. Fourtillan, R. Diaz, and et al., "Dealing with LMJ final optics damage: post-processing and models," in Laser-induced Damage in Optical Materials 2020, vol. 11514 C. W. Carr, V. E. Gruzdev, D. Ristau, and C. S. Menoni, eds., International Society for Optics and Photonics (SPIE, 2020).
16. P. Cormont, P. Combis, L. Gallais, C. Hecquet, L. Lamaignère, and J. L. Rullier, "Removal of scratches on fused silica optics by using a co2 laser," Opt. Express **21**, 28272–28289 (2013).
17. T. Doualle, L. Gallais, S. Monneret, S. Bouillet, A. Bourgeade, C. Ameil, and et al., "CO2 laser microprocessing for laser damage growth mitigation of fused silica optics," Opt. Eng. **56**, 1 – 9 (2016).
18. A. Conder, T. Alger, S. Azevedo, J. Chang, S. Glenn, L. Kegelmeyer, J. Liebman, M. Spaeth, and P. Whitman, "Final optics damage inspection (FODI) for the National Ignition Facility," in Laser-Induced Damage in Optical Materials: 2007, vol. 6720 (SPIE, 2007), pp. 346–360.
19. F. Wei, F. Chen, B. Liu, Z. Peng, J. Tang, Q. Zhu, and et al., "Automatic classification of true and false laser-induced damage in large aperture optics," Opt. Eng. **57**, 1 – 11 (2018).
20. G. Hallo, C. Lacombe, J. Néauport, and F. Hild, "Detection and tracking of laser damage sites on fused silica components by digital image correlation," Opt. Lasers Eng. **146**, 106674 (2021).
21. L. Kegelmeyer, P. Fong, S. Glenn, and J. Liebman, "Local area signal-to-noise ratio (lasnr) algorithm for image segmentation," Proc. SPIE - The Int. Soc. for Opt. Eng. **6696** (2007).
22. F. Bo, C. Fengdong, L. Bingguo, and L. Guodong, "Segmentation of small defects in final optics damage online inspection images," in 2012 International Conference on Image Analysis and Signal Processing, (IEEE, 2012), pp. 1–4.
23. G. Hallo, C. Lacombe, R. Parreault, N. Roquin, T. Donval, L. Lamaignère, J. Néauport, and F. Hild, "Sub-pixel detection of laser-induced damage and its growth on fused silica optics using registration residuals," Opt. Express **29**, 35820–35836 (2021).
24. N. O'Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan, and J. Walsh, "Deep learning vs. traditional computer vision," in Advances in Computer Vision, K. Arai and S. Kapoor, eds. (Springer International Publishing, Cham, 2020), pp. 128–144.
25. C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning (still) requires rethinking generalization," Commun. ACM **64**, 107–115 (2021).
26. L. Mascio-Kegelmeyer, "Machine learning for managing damage on NIF optics," in Laser-induced Damage in Optical Materials 2020, vol. 11514 C. W. Carr, V. E. Gruzdev, D. Ristau, and C. S. Menoni, eds., International Society for Optics and Photonics (SPIE, 2020).
27. F. Wei, F. Chen, B. Liu, Z. Peng, J. Tang, Q. Zhu, D. Hu, Y. Xiang, N. Liu, Z. Sun, and G. Liu, "Automatic classification of true and false laser-induced damage in large aperture optics," Opt. Eng. **57**, 1 (2018).
28. X. Chu, H. Zhang, Z. Tian, Q. Zhang, F. Wang, J. Chen, and Y. Geng, "Detection of laser-induced optical defects based on image segmentation," High Power Laser Sci. Eng. **7**, e66 (2019).
29. J. Kou, T. Zhan, L. Wang, Y. Xie, Y. Zhang, D. Zhou, and M. Gong, "An end-to-end laser-induced damage change detection approach for optical elements via siamese network and multi-layer perceptrons," Opt. Express **30**, 24084–24102 (2022).
30. K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in 2017 IEEE International Conference on Computer Vision (ICCV), (2017), pp. 2980–2988.
31. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2016), pp. 779–788.
32. S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, "scikit-image: image processing in python," PeerJ **2**, e453 (2014).
33. C. Amorin, L. M. Kegelmeyer, and W. P. Kegelmeyer, "A hybrid deep learning architecture for classification of microscopic damage on National Ignition Facility laser optics," Stat. Anal. Data Mining: The ASA Data Sci. J. **12**, 505–513 (2019). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sam.11437.
34. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds. (Springer International Publishing, Cham, 2015), Lecture Notes in Computer Science, pp. 234–241.
35. S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and the scikit-image contributors, "scikit-image: image processing in Python," PeerJ **2**, e453 (2014).
36. H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), (2019), pp. 658–666.
37. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2016).
38. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2015).