



**HAL**  
open science

# Creating a Dataset for the Detection and Segmentation of Degradation Phenomena in Notre-Dame de Paris

Laura Willot, Kévin Réby, Adeline Manuel, Valerie Gouet-Brunet, Dan Vodislav, Livio de Luca

► **To cite this version:**

Laura Willot, Kévin Réby, Adeline Manuel, Valerie Gouet-Brunet, Dan Vodislav, et al.. Creating a Dataset for the Detection and Segmentation of Degradation Phenomena in Notre-Dame de Paris. 6th Workshop on AnalySis, Understanding and ProMotion of HeritAge Contents (SUMAC '24), ACM Multimedia, Oct 2024, Melbourne, Australia. pp.5-12, 10.1145/3689094.3689473 . hal-04727818

**HAL Id: hal-04727818**

**<https://hal.science/hal-04727818v1>**

Submitted on 2 Dec 2024

**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.

# Creating a Dataset for the Detection and Segmentation of Degradation Phenomena in Notre-Dame de Paris

Laura Willot\*  
laura.willot@cyu.fr  
ETIS, LASTIG, MAP  
France

Kévin Réby\*  
kevin.reby@map.cnrs.fr  
MAP, UPR 2002 CNRS  
Marseille, France

Adeline Manuel  
MAP, UPR 2002 CNRS  
Marseille, France

Valerie Gouet-Brunet  
LASTIG, Université Gustave Eiffel,  
ENSG, IGN  
Saint-Mandé, France

Dan Vodislav  
ETIS, UMR 8051, ENSEA, CNRS, CY  
Cergy Paris Université  
Cergy-Pontoise, France

Livio De Luca  
MAP, UPR 2002, CNRS  
Marseille, France

## Abstract

After the fire that destroyed most of the Notre-Dame de Paris cathedral's roof and vaults, scientists gathered in an effort to help the restoration process of the cathedral. Several digital methods and heterogeneous data acquisitions were introduced in the process, including many images and annotations. Part of this data focuses on stone degradation phenomena, a crucial element when evaluating the damages caused by the fire and the state of the cathedral before the restoration started. In this paper, we present the first implementation of a dataset creation pipeline with the aim of training AI models to automatically detect and segment stone alteration patterns in images taken in the context of the restoration of Cultural Heritage buildings. Our resulting dataset will be improved in a near future with more data, while conforming with the ambition to provide our experts and researchers with reliable, structured data.

## CCS Concepts

• **Computing methodologies** → **Image segmentation; Object detection.**

## Keywords

Dataset; Deep Learning; Cultural Heritage; Notre-Dame de Paris

## ACM Reference Format:

Laura Willot, Kévin Réby, Adeline Manuel, Valerie Gouet-Brunet, Dan Vodislav, and Livio De Luca. 2024. Creating a Dataset for the Detection and Segmentation of Degradation Phenomena in Notre-Dame de Paris. In *Proceedings of the 6th workshop on the analySis, Understanding and proMotion of heritAge Contents (SUMAC '24)*, October 28-November 1 2024, Melbourne, VIC, Australia. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3689094.3689473>

\*Both authors contributed equally to this research.

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only. Request permissions from owner/author(s).  
SUMAC '24, October 28-November 1 2024, Melbourne, VIC, Australia  
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 979-8-4007-1205-0/24/10  
<https://doi.org/10.1145/3689094.3689473>

## 1 Introduction

In April 2019, a devastating fire swept through the Notre-Dame de Paris cathedral, causing extensive damage to the roof, vaults, and nave, and the collapse of the 93-meter-tall spire. In response to this tragic event, a consortium of scientists mobilised to study the cathedral and contribute to its restoration. This effort was organised into nine working groups, each focusing on a specific aspect of the cathedral, from stained glass windows and metal artefacts to building structure and acoustics. Among these, the Digital Data Working Group (WG) [9] was tasked with coordinating the scientific project and managing digital data. To this end, this WG combines computer science and artificial intelligence to build a collaborative knowledge system for digital humanities experts [10], also called the n-dame ecosystem<sup>1</sup>. Our work within this group has been made possible thanks to the European Research Council (ERC), as part of the 'nDame Heritage' project<sup>2</sup>.

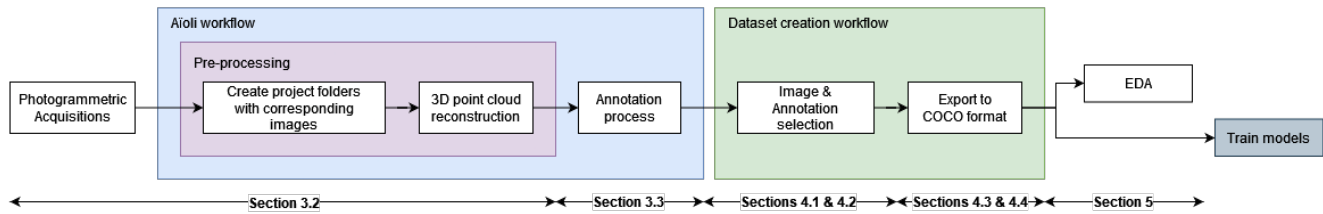
Degradation phenomena refer to the processes or mechanisms by which materials or substances break down or lose their functionality or quality over time due to various internal or external factors. These factors can include environmental conditions (such as temperature, humidity, light, etc.), chemical reactions (such as oxidation, hydrolysis, or corrosion), physical stresses (such as erosion), or biological activity (such as microbial degradation) [5]. Detecting and characterising degradation phenomena is important for predicting and mitigating the effects of degradation. One of the key concerns raised by the architects in their post-fire condition reports was the significant deterioration of the cathedral's stones [14, 25]. The annotation and location of these observations in post-fire images were critical for documenting and analysing the cathedral's condition. Fast and precise annotation methods accelerate the documentation and therefore the analysis of Cultural Heritage sites.

Although computer vision and deep learning have demonstrated significant promise across various applications, they rely on large amounts of labelled data for supervised learning. While numerous datasets exist for object detection and segmentation in common scenes, such as MS COCO [19] and Cityscapes [8], as well as for damage detection, like dacl10k [13], there is a notable absence of datasets specifically focused on Cultural Heritage buildings.

In this paper, we present the creation process of a new dataset of annotated images for the detection and segmentation of stone

<sup>1</sup><https://www.notre-dame.science/>

<sup>2</sup><http://www.ndameheritage.map.cnrs.fr/>



**Figure 1: Global schema of the data export process – The numbers indicated below refer to this article’s sections**

degradation phenomena. Figure 1 offers a general overview of the different processes and steps involved in the creation of our dataset:

- (1) Collection of a large quantity of heterogeneous data related to the Notre-Dame de Paris cathedral.
- (2) Integration of the data into a digital ecosystem, including the Aioli platform for 2D/3D annotation.
- (3) Selection of a subset of images showing clear signs of stone degradation.
- (4) Expert annotation of the selected images using the Aioli platform.
- (5) Access to Aioli data through a custom Python script.
- (6) Transformation of Aioli data to comply with the COCO standard.
- (7) Writing the COCO files.

This dataset, derived from images of the restoration of the Notre-Dame de Paris cathedral, aims to automate and facilitate the annotation process typically performed by experts. The purpose of our dataset is twofold: get enough data to train a model to automatically recognise stone alteration degradation in the context of built heritage sites and structure data from the n-dame project with the aim to leverage this data in this broader context. Through this work, we want to contribute to the preservation and restoration of Cultural Heritage buildings by providing a valuable resource for training deep learning models that will help experts identify degradation phenomena in those buildings.

Section 2 is a brief review of the state-of-the-art methods and existing datasets used for the automatic detection of degradation phenomena in the context of Cultural Heritage building restoration. Section 3 will introduce the collection of data that has been gathered so far and the methods that led to its acquisition. Then, Section 4 will introduce our main contribution, the dataset creation process. Some general and more specific statistics are presented in Section 5 as an analysis of the dataset. Finally, in Section 6 we introduce a few improving features that we aim to implement in a near future, before concluding in the last section.

## 2 Related Work

In the field of computer vision, several datasets and methods have been developed for various tasks, including image retrieval, object detection [28], semantic segmentation [16], and instance segmentation [17]. These tasks have been extensively studied for common objects. The Alegoria dataset<sup>3</sup>, is an example of a database that can

be used for image retrieval for Cultural Heritage. This dataset contains over 20,000 high-resolution images of artworks from various museums and cultural institutions in France [15]. The ArCH dataset [21] focuses on 3D point cloud segmentation of architectural elements<sup>4</sup>. As for the detection of stone degradation phenomena, some works have addressed this task in various contexts. For instance, the DACL workshop [13] focused on damage assessment and location in images<sup>5</sup>. CrackDet [6] is another example that focuses on crack segmentation in surfaces such as roads and pavements. However, to the best of our knowledge, only a few datasets and methods focus on Cultural Heritage applications, particularly for damage detection and stone degradation phenomena.

Deep learning has been increasingly used for the restoration and preservation of Cultural Heritage [12]. For example, DGCNN (Dynamic Graph Convolutional Neural Network) was used for point cloud segmentation on the ArCH dataset [22], and DGCNN-Mod+3Dfeat was used for 3D cultural semantic segmentation in point clouds [20]. However, these methods focus on 3D point clouds and do not address stone degradation phenomena. Foundation models like CLIP [24] and the Segment Anything Model (SAM) [18] have shown promising results in computer vision tasks due to their zero-shot learning and generalisation capabilities. Zero-shot learning enables models to recognize and classify objects or concepts they have never encountered during training by leveraging semantic information or descriptions of unseen classes, by generalizing knowledge without additional labeled data [23]. Foundation models, which are large-scale models pre-trained on extensive and diverse datasets, demonstrate strong generalization capabilities, allowing them to understand and adapt to a wide range of tasks and domains, even those they were not explicitly trained for, without fine tuning [3]. In [26], the authors demonstrated the usefulness of foundation models for zero-shot semantic segmentation of Cultural Heritage data. However, these foundation models are limited to standard objects and do not specifically address stone degradation phenomena. When tested on degradation phenomenon, SAM is missing some parts, even if it works when bounding boxes are created by hand. In Figure 2b is an example of automatic segmentation using the SAM model, from the original image shown in Figure 2a. Although SAM is capable of detecting and segmenting some architectural elements and parts of the painting, the model is missing many pieces of stone degradation on the wall. Therefore, we have decided to create a specific dataset to test and fine-tune models for this particular task.

<sup>3</sup>[https://www.alegoria-project.fr/en/Alegoria\\_dataset](https://www.alegoria-project.fr/en/Alegoria_dataset)

<sup>4</sup><https://archdataset.polito.it/>

<sup>5</sup><https://dacl.ai/workshop.html>

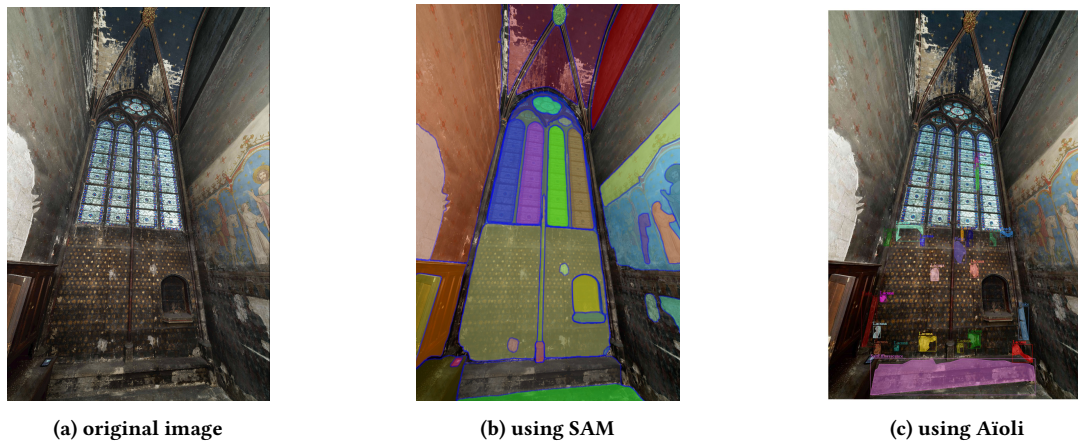


Figure 2: Example of segmented areas

### 3 Dataset collection

In this section, we will briefly introduce the data acquisition context and tools set up for the study of the Notre-Dame de Paris cathedral, before giving more details on the image acquisition process. The last subsection presents the current human-driven annotation process.

#### 3.1 Data production

After the Notre-Dame fire, the Digital Data WG gathered diverse data for restoration, including bibliographic materials, technical drawings, photographs, and 3D models, in collaboration with institutions and companies. These data were incorporated into a digital ecosystem, featuring dedicated web services for data management, content indexing, and integration with thesauri and the CIDOC-CRM ontology [1, 11], including Aioli<sup>6</sup>, a 2D/3D annotation platform [2]. Aioli enables annotating observations on 2D images and propagating them to a 3D point cloud, connecting images and 3D reconstructions. Figure 3 illustrates this process, with coloured regions in the main viewer representing the projections of 2D annotations onto the 3D point cloud, which were created directly on the images. To study stone degradation, experts selected images showing clear deterioration, annotated them using Aioli, and propagated the annotations to the 3D point cloud for comprehensive analysis.

#### 3.2 Image acquisition & pre-processing

To ensure correct data processing on the Aioli platform, all images were acquired using a photogrammetric process [7]. The acquired images revealed unconventional, overlapping views of the cathedral. This kind of pictures has rarely been used to train the models and thus, justifies the necessity of creating a specific dataset for object detection and segmentation based on these images.

The large building size required batch processing of the dataset [25]. The cathedral was segmented into smaller spatial volumes, like side chapels, each forming an individual project. Each project includes a subset of images, a 3D point cloud, and annotations. For example, our dataset contains a folder with all annotated images

captured after the fire and before restoration in chapel 19, one of the cathedral’s side chapels (see Figure 4). In the end, each project consists of a space-related subset of the image dataset, a 3D point cloud, and a series of annotations.

#### 3.3 Annotation process

After creating projects in Aioli, a member of our research team generates annotations based on the monument’s condition reports, digitising expert knowledge using the platform [25]. The annotation system offers a hierarchical structure to organise annotations around topics like statuary, furniture, electrical systems, and post-fire damage, emphasising stone degradation (see the upper-left corner of Figure 3). Annotation labels, defined according to condition reports, are consistently used across all cathedral-related projects in Aioli. The annotation process consists of the following steps:

- (1) The expert’s observation is identified in the condition report, and corresponding labels are created in Aioli.
- (2) The annotation is drawn on one of the images using the appropriate pre-defined label.
- (3) An automatic process is initiated in Aioli to propagate the annotation from its original ‘support image’ to the reconstructed point cloud and back to all project images. This results in the creation of additional annotations of the same elements on images with different views.

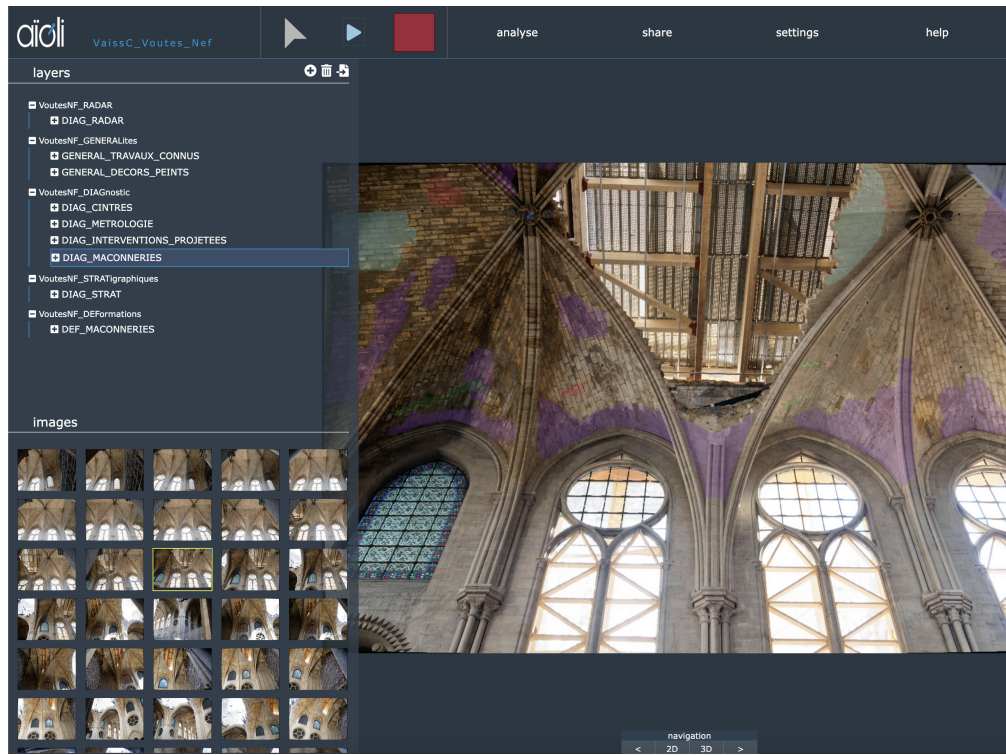
This semi-automatic process offers two significant advantages. First, it enables the use of consistent, reliable expert observations in a digital format within our ecosystem. Second, the automatic generation of multiple annotations from a single manual annotation increases the quantity of data that can be used in our dataset. However, it is important to note that this process does not incorporate a data cleaning step.

### 4 Stone degradation dataset creation

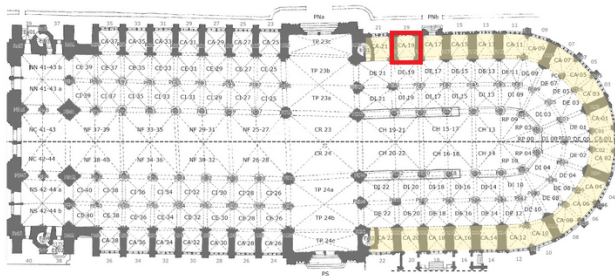
In this section, we present our dataset creation process, starting with the selection of images and annotation categories, namely, the stone alteration patterns. Then, we introduce the COCO standard

<sup>6</sup>[www.aioli.cloud/](http://www.aioli.cloud/)





**Figure 3: Example of annotations created for the images of the vaults with Aioli – The main viewer, on the right, presents the 3D point cloud generated from the images displayed on the bottom left corner. An excerpt of the hierarchical naming system used to annotate the images is shown on the upper left corner – Source: Roxane Roussel**



**Figure 4: Map of the cathedral – The red rectangle refers to chapel 19, one of the nave chapels (in yellow)**

and demonstrate how we adapted it to our needs. The last subsection presents several features we implemented to export the data introduced in section 3.

#### 4.1 Data selection

In the context of the restoration of a Cultural Heritage monument such as the Notre-Dame de Paris cathedral, experts need efficient, accurate methods for automatic segmentation and object detection.

To create a dataset for detecting and segmenting stone degradation phenomena, a representative subset of post-fire Notre-Dame images of various parts of the cathedral (interior and exterior) with

stone alteration patterns annotations was selected. We took care to select the undistorted images used during the annotation process, instead of the raw images. Indeed, the small size and precision required for damage evaluation necessitates accurate region contour associations with images. To maintain data cleanliness, only the first expert-drawn annotation was selected, excluding propagated regions computed by Aioli's procedure, as no cleaning step was implemented in Aioli, so far.

#### 4.2 Stone alteration patterns

Interoperability is one of the main criteria for good data management. In order to comply with it, we manually aligned the annotations' labels created in Aioli with well-established taxonomies and glossaries [4, 5, 27]. Generally speaking, an alteration is defined as a 'modification of the material that does not necessarily imply a worsening of its characteristics from the point of view of conservation' [5]. The 9 degradation phenomena we used in our dataset are summarised in Table 1

#### 4.3 Custom COCO files

To create our dataset, we chose to follow the COCO standard [19] which versatility enables researchers to use any COCO-like dataset for multiple computer vision tasks. The widespread use of this standard within the Machine and Deep Learning communities also provides existing datasets in this format. This facilitates comparison

Category	Definition
Bleaching	Gain in value due to chemical weathering of minerals or extraction of colouring matter, or loss of polish
Salt Efflorescence	Formation of salts forming efflorescence on the surface of sandstone masonry
Lead Deposit	Accumulation of exogenic lead material of variable thickness
Erosion	Loss of original surface, leading to smoothed shapes
Lacuna	Empty space, obviously located in the place of some formerly existing stone part
Scaling	Detachment of stone as a scale or a stack of scales, not following any stone structure
Flake	Scaling in thin flat or curved scales of submillimetric to millimetric thickness, organised as fish scales
Crack	Individual fissure, clearly visible by the naked eye, resulting from separation of one part from another
Humidity	The amount of moisture present in the air or in a porous material

**Table 1: Definitions of the categories of alteration patterns, based on [5] and [27]**

and integration with different models, promoting collaboration and accelerating research progress.

Written in a JSON-based textual format, this standard allows the introduction of custom fields to capture specific features or metadata of the dataset. This extensibility ensures that the characteristics of our dataset can be effectively described and used in an easy integrated approach with other tools and systems.

For our dataset, we based our schema on the 'object detection' and 'segmentation' tasks that include the following **standard** fields, to which we added our *custom* fields:

- Info – General information of each project
  - **url**: link to the public Aioli viewer; directly loads the project's 3D scene and annotations
- Images – List of images, included in the project, with following attributes:
  - **id**: unique integer identifier
  - **width and height** (in pixels): size of the image
  - **file\_name**: file name of the image
  - **aioli\_url**: link to the image used in Aioli
- Categories – List of object categories (degradation phenomena) with following attributes:
  - **id**: unique integer identifier
  - **name**: name of the category, e.g., "Bleaching", "Erosion"
- Annotations – List of annotations with following attributes:
  - **id**: unique integer identifier
  - **image\_id**: id of the image the annotation refers to
  - **category\_id**: id of the category the annotation belongs to
  - **bbox**: list of four floating-point numbers representing the bounding box of the object
  - **segmentation**: optional list of pixel-wise segmentation masks for the object

## 4.4 Data export

The export into the COCO format is done directly from the data saved in Aioli, with a dedicated Python script. There were a few steps worth mentioning that we implemented to correctly process the data. This process was repeated for each of the 6 projects we previously selected.

**4.4.1 Folder hierarchy.** Accessing the Aioli annotations is a project-led process. Thus, even though the models are to be trained on the overall dataset, we kept this project-led organisation for the creation of our dataset. This decision is also guided by two other reasons.

First, due to the lack of an application programming interface (API) to interact with the Aioli platform, the data access is limited to a project-led access. Furthermore, in our much broader context of data production, it was necessary to keep the dataset consistent with the different tools used in the n-dame project.

**4.4.2 Image ID.** The images are stored on the Aioli servers with a filename given at the acquisition of the image. Combined with the unique project ID that the images are associated with, this *project ID-image name* fusion creates an identifier comparable to a unique ID. We implemented two simple functions to convert this character string into –and back from– an integer, used as the unique image ID.

**4.4.3 Coordinate convention systems.** Each annotation is defined in Aioli by a list of coordinates ranging from 0 to 1. Given the size of the images and knowing the coordinate convention systems used in Aioli and in the COCO standard, we were able to convert the annotations' contours' coordinates into the appropriate COCO standard. The bounding box is computed 'on the fly' by looking for the maximum and minimum values of each *x* and *y* components of the annotation. This data is then transformed into the expected list of [*x*, *y*, *width*, *height*].

## 5 Dataset Analysis

This dataset, derived from 6 projects, contains 45 images with 180 annotations labelled under 9 categories (see Table 2). The current limited size is due to two restrictions: the dataset is based on a few Aioli projects, and only expert-drawn annotated regions were used, excluding those propagated by Aioli's internal mechanism. This decision aimed to minimise the data cleaning step and test the initial version of the dataset creation pipeline.

Figure 5 shows some examples of the annotated images that were obtained through our pipeline. The resulting images were obtained thanks to an adapted implementation of a COCO visualiser<sup>7</sup>. The example shown in Figure 2c is created from the original image in Figure 2a.

### 5.1 Projects and annotations overview

Table 2 also highlights a significant heterogeneity between the projects that can be explained by the nature of the projects: they

<sup>7</sup>original notebook available at: <http://www.immersivelimit.com/tutorials/create-coco-annotations-from-scratch>



Figure 5: Examples of annotated images, obtained through our pipeline

Project title	#ann	#img
IE_NF_CB_ExtradosEst	39	17
IE_NF_CA30	14	4
IE_CH_CA19	78	10
E_NF_TA_TAC_CouverturesNord	3	2
E_NF_T31_ContrefortF31	30	4
IE_NF_CA25	16	8
<b>TOTAL</b>	<b>180</b>	<b>45</b>

Table 2: Number of images (#img) and annotations (#ann), project-wise and for the entire dataset (TOTAL)

were created to mirror the cathedral’s interior and exterior spaces of the cathedral. As the fire did not spread the same way everywhere in the cathedral, it affected differently the spaces of the cathedral. Thus, it is likely that some stone degradation phenomena will be found more often in some spaces than others, leading to projects with a higher or lower number of annotations.

The distribution of the size of the annotations per category is shown in Figure 6. Our images are both portrait- and landscape-oriented, and range between (3200, 4800) and (6336, 9504) pixels while the annotations have an average size of about 270417 pixels. This means that most degradation phenomena are small compared to the size of the images. The significant gap between the size of images and annotations is also the result of the photogrammetric

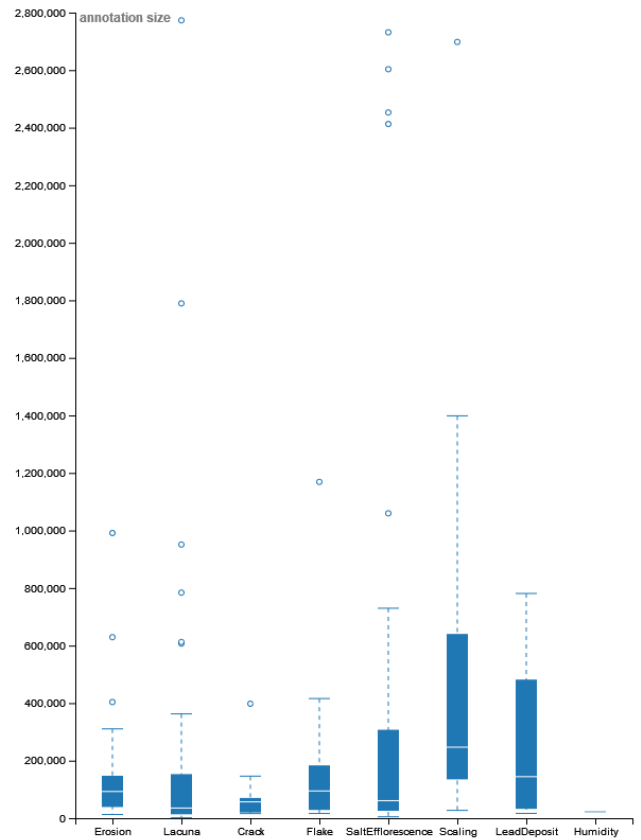


Figure 6: Distribution of annotation’s sizes per category (in pixels)

acquisition process: the annotations are based on large views of the walls, ceilings, and roofs rather than on detailed portions of them. Images with large views are also a good support on which to get both generic categories (walls, statues, furniture, etc.) and those of the alteration patterns. Finally, these images are also useful for large annotations. Some degradation phenomena could be larger than the ones we found in the cathedral and would thus be annotated on a single image rather than several of them.

## 5.2 Annotations and categories analysis

Most images contain few annotations, with an average of 4 annotations per image, as shown in Table 3 (note that even though we defined the ‘Bleaching’ category, no annotation of our dataset is associated with this label). We explain this number of annotated areas on each image by two factors: the nature of the images and the annotation process. Indeed, as indicated before (see Section 3), the photogrammetric acquisitions led to images with a very high image-to-image coverage. This was considered an opportunity for the expert user to choose the *best view* (that is, the best image) to create each annotation. Thus, many different images were used to draw all annotations (see also Figure 7 for a distribution of the number of images with at least one annotation for each category, in blue).

category name	minimum	maximum	average
Bleaching	0	0	0
Salt Efflorescence	1	8	2.43
Lead Deposit	1	2	1.2
Erosion	1	15	5.33
Lacuna	1	14	6.14
Scaling	1	3	1.92
Flake	1	7	5
Crack	1	2	1.25
Humidity	1	1	1
<b>OVERALL</b>	<b>1</b>	<b>23</b>	<b>4</b>

Table 3: Number of annotations per image

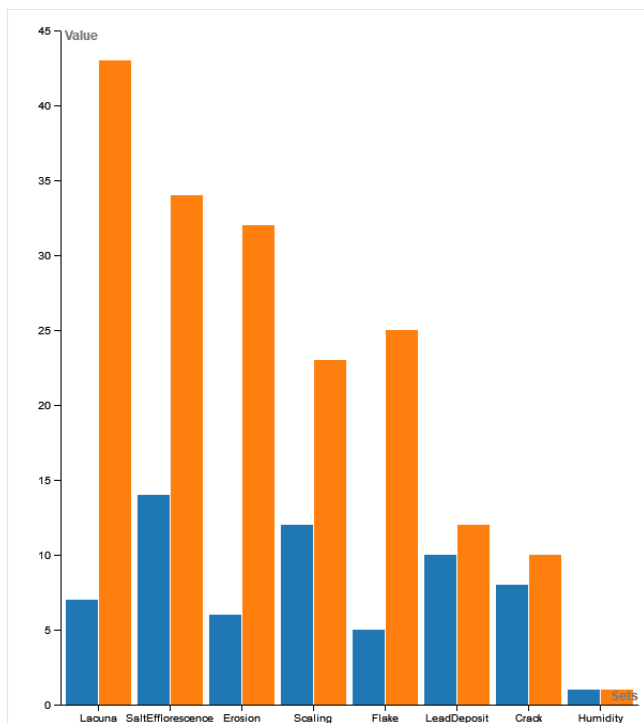


Figure 7: Number of annotations per category – blue: number of images with at least one annotation of the category; orange: number of annotations of the category

Figure 7 also presents the distribution of annotations per category. This significant imbalance between categories is explained by the dataset content: each annotation is created from a real stone alteration phenomenon that was noticed in the cathedral and described by the architects in the condition reports. Furthermore, the number of annotations per category also depends on the nature of the alteration patterns: for instance, 'lacuna' and 'erosion' phenomena are represented as many small areas over a larger surface. The variability of the categories is thus nothing more than the variability of the actual alterations to the cathedral's stones.

## 6 Perspectives

This pipeline is likely to evolve in the future as the amount of data to be handled will drastically increase. We hope to process over tens of thousands of images and as many annotations, distributed over dozens of Aioli projects covering the entire cathedral. The future dataset expansion will also integrate Aioli's re-propagated regions. The number of classes will also be reevaluated to include more degradation phenomena.

To maintain reliability, a data cleaning step will be added to the pipeline to remove incorrectly re-projected annotations.

In the future, we aim to leverage this dataset to fine-tune existing deep learning models and conduct a comparative analysis of their performance, including but not limited to CNN-based models, U-Net, and Vision Transformer (ViT) models. It will then be possible to use those trained models to detect stone alteration patterns in other parts of the cathedral thus decreasing the amount of time spent by the experts to describe those degradation phenomena.

We also plan to improve the integration of our dataset and any future annotations obtained from the automatic image segmentation into the n-dame ecosystem by working on the interoperability. Creating a dedicated thesaurus for stone alteration patterns and linking its terms to annotation categories will enhance dataset quality and provide experts with well-structured data. To this end, a direct link between the annotation categories and a thesaurus (as those managed with Opentheso<sup>8</sup>) will be established and a dedicated field will be added in the relevant section of the COCO files.

Furthermore, as we can notice from the comparison between the human-annotated image of stone degradation phenomena and the SAM-annotated image (Figure 2b), these semantic image segmentation methods complement each other: in the context of built Cultural Heritage restoration, it is important to notice not only the degradation phenomena but also *what structures* are damaged. The consequences of a crack on a wall or on the ceiling might not be the same.

Thus, by combining this data from semantic image segmentation methods together with other known data of the n-dame ecosystem, such as the size and location of the annotations, deduced from the photogrammetric acquisition process will enable higher-level queries on researchers' data, e.g., "what's the coverage of a certain degradation phenomenon on the walls of a given chapel?"

## 7 Conclusion

In this paper, we presented the first implementation of a pipeline used to create a COCO standard-compliant dataset designed for the detection and segmentation of stone alteration patterns in Cultural Heritage buildings, from image acquisition to file creation. This dataset offers a comprehensive array of images that capture various degradation phenomena, enabling the development of robust and accurate models. Improvements to our dataset creation pipeline is currently underway to produce a more complete, more structured dataset that will be used to fine-tune existing deep learning models. The results provided by those models will also be leveraged to increment the data produced in the n-dame project and answer higher-level queries. The availability of the dataset, its evolution and trained models is still under discussion between the several

<sup>8</sup><https://opentheso.huma-num.fr/opentheso/>



institutional partners gathered around the reconstruction of the cathedral, with the objective of making them available as open data, at least for the community of digital heritage architects. Meanwhile, this research will offer valuable insights into new approaches for detecting and analysing degradation phenomena in Cultural Heritage buildings.

## Acknowledgments

All pictures of the Notre-Dame de Paris cathedral presented in this work are credited under the © Chantier Scientifique Notre-Dame de Paris / Ministère de la Culture / CNRS.

This work benefits from a State aid managed by the French National Research Agency (Agence Nationale de la Recherche) under the future investment programme integrated into France 2030, with reference ANR-17-EURE-0021 – Paris Seine Graduate School of Humanities Creation Heritage (École Universitaire de Recherche Paris Seine Humanités, Création, Patrimoine) – Foundation for Cultural Heritage Science (Fondation des Sciences du Patrimoine).

This work was funded since 2019 by the CNRS and the French Ministry of Culture within the framework of the national scientific action Notre-Dame de Paris, and since 2022 by the European Research Council (ERC Advanced Grant nDame\_Heritage: n-Dimensional analysis and memorization ecosystem for building cathedrals of knowledge in Heritage Science).

## References

- [1] 2021. Definition of the CIDOC Conceptual Reference Model (Version 7.1.1 Produced by the CIDOC CRM Special Interest Group). <https://cidoc-crm.org/get-last-official-release>
- [2] Violette Abergel, Adeline Manuel, Anthony Pamart, Isabelle Cao, and Livio De Luca. 2023. Aioli: A reality-based 3D annotation cloud platform for the collaborative documentation of cultural heritage artefacts. *Digital Applications in Archaeology and Cultural Heritage* 30 (Sept. 2023), e00285. <https://doi.org/10.1016/j.daach.2023.e00285>
- [3] Rishi Bommasani, Drew A Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, et al. 2021. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258* (2021).
- [4] Philippe Bromblet. 2010. Guide “Altérations de la pierre”. [https://www.pierres-info.fr/biblio-taille\\_de\\_pierre/alteration-pierre-naturelle.pdf](https://www.pierres-info.fr/biblio-taille_de_pierre/alteration-pierre-naturelle.pdf)
- [5] Anson Cartwright, Elsa Bourguignon, Philippe Bromblet, Jo Ann Cassar, A. Elena Charola, Eddy De Witte, José Delgado Rodrigues, Vasco Fassina, Bernd Fitzner, Laurent Fortier, Christoph Franzen, José-Maria Garcia de Miguel, Ewan Hyslop, Marie Klingspor-Rotstein, Daniel Kwiatkowski, Wolfgang E. Krumbain, Roger Lefevre, Ingvál Maxwell, Andrew McMillan, and David Young. 2008. *ICOMOS-ISCS: Illustrated glossary on stone deterioration patterns*. ICOMOS.
- [6] Zhuangzhuang Chen, Jin Zhang, Zhuonan Lai, Guanming Zhu, Zun Liu, Jie Chen, and Jianqiang Li. 2023. The devil is in the crack orientation: A new perspective for crack detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 6653–6663.
- [7] F. Comte, A. Pamart, K. Réby, and L. De Luca. 2024. Strategies and Experiments for Massive 3D Digitalization of the Remains after the Notre Dame de Paris’ Fire. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLVIII-2-W4-2024 (Feb. 2024), 127–134. <https://doi.org/10.5194/isprs-archives-XLVIII-2-W4-2024-127-2024>
- [8] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Scharwächter, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. 2015. The cityscapes dataset. In *CVPR Workshop on the Future of Datasets in Vision*, Vol. 2. 1.
- [9] Livio De Luca. 2024. A digital ecosystem for the multidisciplinary study of Notre-Dame de Paris. *Journal of Cultural Heritage* 65 (Jan. 2024), 206–209. <https://doi.org/10.1016/j.culher.2023.09.011>
- [10] Livio De Luca, Violette Abergel, Anaïs Guillem, Olivier Malavergne, Adeline Manuel, Ariane Néroulidis, Roxane Roussel, Miled Rousset, and Sarah Tournon. 2022. L’écosystème numérique n-dame pour l’analyse et la mémorisation multidimensionnelle du chantier scientifique Notre-Dame-de-Paris. Lyon, France, 1. <https://hal.science/hal-03826931>
- [11] Martin Doerr, Christian-Emil Ore, and Stephen Stead. 2007. The CIDOC Conceptual Reference Model - A New Standard for Knowledge Sharing. In *Conferences in Research and Practice in Information Technology (CRPIT)*, Vol. 83. Auckland, New Zealand, 51–56. <https://doi.org/10.13140/2.1.1420.6400>
- [12] Marco Fiorucci, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Travaglia, Alessio Del Bue, and Stuart James. 2020. Machine learning for cultural heritage: A survey. *Pattern Recognition Letters* 133 (2020), 102–108.
- [13] Johannes Flotzinger, Philipp J Rösch, Christian Benz, Muneer Ahmad, Murat Cankaya, Helmut Mayer, Volker Rodehorst, Norbert Oswald, and Thomas Braml. 2024. dacl-challenge: Semantic Segmentation during Visual Bridge Inspections. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 716–725.
- [14] Yves Gallet. 2024. Of stone and mortar: Studying Notre-Dame Cathedral’s core materials. *Journal of Cultural Heritage* 65 (Jan. 2024), 46–48. <https://doi.org/10.1016/j.culher.2023.09.014>
- [15] Dimitri Gominski, Valérie Gouet-Brunet, and Liming Chen. 2021. Connecting images through sources: Exploring low-data, heterogeneous instance retrieval. *Remote Sensing* 13, 16 (2021), 3080.
- [16] Yanming Guo, Yu Liu, Theodoros Georgiou, and Michael S. Lew. 2018. A review of semantic segmentation using deep neural networks. *International Journal of Multimedia Information Retrieval* 7, 2 (June 2018), 87–93. <https://doi.org/10.1007/s13735-017-0141-z>
- [17] Abdul Mueed Hafiz and Ghulam Mohiuddin Bhat. 2020. A Survey on Instance Segmentation: State of the art. *International Journal of Multimedia Information Retrieval* 9, 3 (Sept. 2020), 171–189. <https://doi.org/10.1007/s13735-020-00195-x> arXiv:2007.00047 [cs, eess].
- [18] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. 2023. Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 4015–4026.
- [19] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. 2015. Microsoft COCO: Common Objects in Context. arXiv:1405.0312 [cs]
- [20] Francesca Matrone, Eleonora Grilli, Massimo Martini, Marina Paolanti, Roberto Pierdicca, and Fabio Remondino. 2020. Comparing machine and deep learning methods for large 3D heritage semantic segmentation. *ISPRS International Journal of Geo-Information* 9, 9 (2020), 535.
- [21] E Pellis, A Masiero, G Tucci, M Betti, and P Grussenmeyer. 2021. Assembling an image and point cloud dataset for heritage building semantic segmentation. *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 46 (2021), 539–546.
- [22] Roberto Pierdicca, Marina Paolanti, Francesca Matrone, Massimo Martini, Christian Morbidoni, Eva Savina Malinverni, Emanuele Frontoni, and Andrea Maria Lingua. 2020. Point cloud semantic segmentation using a deep learning framework for cultural heritage. *Remote Sensing* 12, 6 (2020), 1005.
- [23] Farhad Pourpanah, Moloud Abdar, Yuxuan Luo, Xinlei Zhou, Ran Wang, Chee Peng Lim, Xi-Zhao Wang, and QM Jonathan Wu. 2022. A review of generalized zero-shot learning methods. *IEEE transactions on pattern analysis and machine intelligence* 45, 4 (2022), 4051–4070.
- [24] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. 2021. Learning transferable visual models from natural language supervision. In *International conference on machine learning*. PMLR, 8748–8763.
- [25] R. Roussel and L. De Luca. 2023. An Approach to Build a Complete Digital Report of the Notre Dame Cathedral after the Fire, using the Aioli Platform. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLVIII-M-2–2023 (June 2023), 1359–1365. <https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-1359-2023>
- [26] Kévin Réby, Anaïs Guillem, and Livio De Luca. 2023. Semantic Segmentation using Foundation Models for Cultural Heritage: an Experimental Study on Notre-Dame de Paris. In *4th ICCV Workshop on Electronic Cultural Heritage*, Guillaume Caron (Ed.), Computer Vision Foundation, Paris, France. <https://hal.science/hal-04275484> Backup Publisher: Computer Vision Foundation.
- [27] Angela Weyer, Pilar Roig Picazo, Daniel Pop, JoAnn Cassar, Aysun Özköse, Jean-Marc Vallet, and Ivan Srša. 2015. EwaGlos - European illustrated glossary of conservation terms for wall paintings and architectural surfaces. <https://openarchive.icomos.org/id/eprint/1706/>
- [28] Zhengxia Zou, Keyan Chen, Zhenwei Shi, Yuhong Guo, and Jieping Ye. 2023. Object Detection in 20 Years: A Survey. (Jan. 2023). <https://doi.org/10.48550/arXiv.1905.05055>