Artificial Intelligence at the Interface between Cultural Heritage and Photography: A Systematic Literature Review

Carmen Silva 1,2, * and Lídia Oliveira 1

1 DigiMedia—Digital Media and Interaction Research Center, University of Aveiro, 3810-193 Aveiro, Portugal; lidia@ua.pt
2 Institute of Art Sciences, Museology, and Master Programs in Cultural Heritage Sciences, Federal University of Pará, Belém 66075-110, Brazil
* Correspondence: carmenluciasilva@ua.pt

Abstract: Artificial intelligence has inspired a significant number of studies on the interface between cultural heritage and photography. The aims of these studies are, among others, to streamline damage monitoring or diagnoses for heritage preservation, enhance the production of high-fidelity 3D models of cultural assets, or improve the analysis of heritage images using computer vision. This article presents the results of a systematic literature review to highlight the recent state of these studies, published in the last five years and available in the Scopus, Web of Science, and JSTOR databases. The aim is to identify the potential and challenges of artificial intelligence through the connection between cultural heritage and photography, the latter of which represents a relevant methodological aspect in these investigations. In addition to the advances exemplified, the vast majority of studies indicate that there are also many obstacles to overcome. In particular, there is a need to improve artificial intelligence methods that still have significant flaws. These include inaccuracy in the automatic classification of images and limitations in the applications of the results. This article also aims to reflect on the meaning of these innovations when considering the direction of the relationship between cultural heritage and photography.

Keywords: artificial intelligence; cultural heritage; photography; machine learning; deep learning; neural network

1. Introduction

Over time, practices and studies on the preservation, documentation, and enjoyment of cultural heritage have incorporated emerging technologies. This can be observed in the advances in image reproduction and processing devices, which have been integrated into these activities since the advent of photography. In other words, just a few years after Joseph Nicéphore Niepce recorded the first recognized photograph in 1826, or the announcement of the daguerreotype in 1839 was made. Over the past two centuries, humanity has observed the evolution of these technologies, which have been employed in various ways to document, diagnose, and promote cultural heritage. These developments have contributed to the knowledge and recognition of these cultural assets, which are now facing the challenge of artificial intelligence (AI).

This was the case with the “heliographic journeys” carried out between 1850 and 1855 by pioneering photographers who were also researchers into photographic capture processes. On these excursions, they recorded images of important monuments in France, compiling a large collection that is still largely in archives in Paris [1] (p. 99). Additionally, at the beginning of the second half of the nineteenth century, traveling photographers, some of whom were members of scientific expeditions, visited North Africa, the Middle East, and the Americas, among other destinations. Their objective was to portray archaeological sites such as the Mayan and Aztec ruins, whose images were disseminated globally.
With photography, it was thus possible to provide greater realism to the imagery of cultural assets. This resource expanded inventories and broadened the ways in which knowledge was socialized, which had previously been mostly textual. For example, photographic albums became commonplace at the end of the 19th century, many of which were commissioned by public managers to promote their governments. These publications were illustrated with photographs of what would later, in the following century, be considered heritage sites [2] (p. 384).

The valorization of cultural, historical, and architectural heritage was a growing phenomenon throughout the 20th century, with the involvement of political institutions worldwide. In this context, photographs served as an invaluable apparatus in the dossiers demanding recognition of these assets. However, there are numerous applications of photography in the field of heritage, based on the premise that “photography is, at the same time, a form of expression and a means of information and communication based on reality”. As such, “it can be considered a document of historical life” [3] (p. 131). Therefore, photography is relevant in the development of heritage actions, as well as musealization, which was consolidated in the 1970s. This concept encompasses the visual representation of heritage in museum and heritage environments, which is commonly referred to as “visualization” [4] (p. 22). Consequently, photography is also employed in the fields of heritage preservation (selection, acquisition, management, conservation), research (cataloguing), and communication (exhibitions, publications, etc.).

The advent of a new revolution in studies involving photography and cultural heritage at the end of the 20th century led to significant changes in these processes. Digital technologies and the commercial Internet facilitated, for example, the documentation of cultural assets in a more instantaneous manner. The circulation of these images also became wider and faster, including through social networks. Furthermore, there was a broadening of perspectives to favor other immersive heritage experiences through photography, such as the 360-degree tour. Over the past two decades, AI has emerged as a further paradigm shift in the activities and research emerging from this interface.

However, this reality is the result of an accumulation of studies that began in the 1940s. These studies were initiated by pioneers such as Warren McCulloch and Walter Pitts, who published the first recognized work on AI in 1943 [5] (p. 16). Another notable figure was Alan Turing, whose research became famous after the “Turing test”, which challenged anyone to “distinguish between a machine and a person based on the answers given during a blind conversation” [6] (p. 209). In the 1950s, the term was first introduced by John McCarthy, who defined it as a “science and engineering of making intelligent machines, especially intelligent computer programs” [7] (p. 2).

Nevertheless, the advancement of research in this field has prompted the debate and formulation of numerous definitions of AI. Russell and Norvig [5] (pp. 1–5) enumerate eight historical definitions of AI, categorized into four approaches, proposed during the 1980s and 1990s. The approaches categorized by the authors as “Thinking Humanly” and “Thinking Rationally” are “concerned with thought processes and reasoning”. The other two, “Acting Humanly” and “Acting Rationally”, are associated with “behavior”. Furthermore, the approaches “Thinking Humanly” and “Acting Humanly” discuss the measure of “success in terms of fidelity to human performance”. The approaches “Thinking Rationally” and “Acting Rationally” refer to “an ideal performance measure, called rationality”. In other words, rationality can be described as doing “the ‘right thing’, given what it knows”. Russell and Norvig highlight that the four approaches have been pursued by disparate groups, who have provided assistance or critiqued one another. Consequently, the routes to the definition of AI are multifaceted and complex, comprising a multitude of interrelated factors. However, Russell and Norvig proffer a suggestion, predicated on “the idea of an intelligent agent”: AI is “the study of agents that receive percepts from the environment and perform actions” [5] (p. viii).

As contemporary developments in AI research indicate, according to Cozman and Neri [8] (p. 26), it is possible to organize this area around three general axes. The first of
these axes is that of “knowledge representation”, which is associated with the “domain of epistemology”. The second area of focus is “decision-making”, which is related to fields such as psychology, economics, engineering, and law. The third area of focus is “machine learning”, which is being investigated by several fields, including pedagogy and statistical techniques for data processing. All three axes have an interdisciplinary vocation; however, it is the third that is currently most engaged in dialogue with research into cultural heritage, an equally interdisciplinary field.

Machine learning is “a sub-field of AI focused on the creation of algorithms that use experience with respect to a class of tasks and feedback in the form of a performance measure to improve their performance on that task” [9] (p. 11). In more precise terms, the field of heritage has a strong interest in studies in deep learning, which is a subset or subfield of machine learning. Cozman and Neri [8] (p. 25) posit that a “deep neural network” represents “a function composed of layers of artificial neurons” and that “this type of learning can extract patterns of surprising complexity from data, making tasks that are difficult to automate possible”. In this context, numerous heritage studies that employ images have gravitated towards a specific type of deep learning, namely, convolutional neural networks (CNNs). This is due to the fact that CNNs are frequently employed in the field of computer vision, where they are utilized for the purpose of pattern recognition and image classification. CNNs have demonstrated efficacy in object recognition tasks, including cultural assets, which they are able to detect, classify, and segment.

Processes and experiments are being reinvigorated through innovations in this field that can redirect practice and thinking at the confluence of the areas of focus in this work. Therefore, this systematic literature review examines how AI is employed in cultural heritage studies that utilize photography or its derivatives as a foundational element of their methodological approach. Derivatives of photography include other image-creation and recording technologies such as cinematography and photogrammetry. Consequently, this review also encompasses research that focuses on these variations.

The research covered in this review was classified in the databases searched as being in the areas of arts, humanities, social science, imaging science photographic technology, and museum studies. This refinement was necessary due to the objective proposed in this research, which was to identify the potential and challenges of AI, as a way of thinking about the meaning of these innovations when considering the direction of studies that articulate cultural heritage, especially architectural heritage, and photography, with an emphasis on the specified areas. In consequence, the aforementioned fields of interest represent the convergence of interdisciplinary studies, with socio-technical effects on the relationship between humanity and technology.

Therefore, the review was conducted with the following problem in mind: How is AI investigated in the link between cultural heritage, especially architectural heritage, and photography, and its derivatives, with a view to the areas of knowledge classified in databases as arts, humanities, social science, imaging science photographic technology, and museum studies? In other words, how is research currently being conducted in these areas, which investigate and work with cultural heritage and photography, and its derivatives, in terms of AI?

Based on these questions, this article presents, as a general result, a recent overview of research on the proposed theme and reflects on this scenario in an attempt to contribute to the advancement of knowledge in this interdisciplinary area. It is justified insofar as it highlights the investigative issues underway in contemporary times and discusses the orientations in this field of study.

2. Materials and Methods

This literature review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, in order to answer the research problem and achieve the proposed objective. Firstly, a search of the databases was conducted between the 13 November and the 12 December 2023. The databases selected
were Scopus, Web of Science, and JSTOR, as they are multidisciplinary or focused on the arts and social sciences. This approach enabled the identification of publications classified within the areas of research interest. Furthermore, the selected publications engage in interdisciplinary dialogues with other fields of knowledge, particularly Computer Science.

The following composition of the query was outlined in all the databases searched: “Artificial Intelligence” OR “Neural Network” OR “Machine Learning” OR “Deep Learning” AND Photo* AND “Cultural Heritage”. The terms employed in the query were selected by the proposed problem, and potential variations in the concepts utilized in the researched texts were also considered. In addition to the term “Artificial Intelligence” (AI), which is a more general concept, the variation “Neural Network” (NN) was included in the searches using the Boolean operator OR, as it is an approach within this field that has made significant progress in recent years [10] (p. 1646). The search also included variations “Machine Learning” and “Deep Learning”, which were introduced using the Boolean operator “OR”. Machine learning (ML) is considered to be the dominant field in AI research, while deep learning (DL) has sparked a revolution in AI in more recent studies [6] (p. 320).

In order to obtain variants and thus collect texts that refer to photography, its processes, and derivatives, the Boolean operator AND was employed to include the term “Photo*”, with an asterisk at the end of the word. This option facilitated a more expansive collection of texts encompassing terminology related to photography, including terms such as “photograph” as a verb and “photographer” as a noun. Furthermore, it permitted the aggregation of texts on techniques derived from photography, such as photogrammetry. The Boolean operator AND was employed in order to identify works that could be considered to lie at the intersection between AI, photography, and cultural heritage.

Ultimately, the term “Cultural Heritage” was incorporated into the search using the Boolean operator “AND”, with the objective of assembling a corpus of texts pertaining to cultural heritage in its general sense, irrespective of the typology. This option also permitted the collection of a larger number of texts, which could be reviewed in subsequent stages to identify those that pertained to architectural heritage, the focus of this review. As is detailed in the following subsections, these texts were evaluated with a view to the final selection of works to be studied in this research review, in accordance with the main criteria of relevance to the proposed theme and research problem, confirmation of the use of photography in the methodological process or as an object of study, and insertion in the areas of interest to the research.

2.1. Identification and Screening

For the identification of the texts, each database was initially consulted, then, the texts classified in the areas relevant to the research were screened. It is important to emphasize that the databases do not use the same thematic areas. However, in each of them, we endeavored to identify those of interest to the research.

In the Scopus database, the search returned 175 documents, published between 2002 and 2023. Of these, 54 were classified as belonging to the field of social science. In the “Arts and Humanities” area, 25 documents were identified. Ten publications were identified in both areas. In total, therefore, 69 texts were identified in Scopus in these two areas.

The search on the Web of Science revealed 74 documents published between 2008 and 2023. Of these, 19 were classified as “Art” or “Arts and Humanities other topics”. In the field of “Imaging Science Photographic Technology”, 17 documents were identified. Three texts appeared in both lists, resulting in a total of 33 publications in the areas of research interest.

The search in JSTOR identified 461 documents, published between 1955 and 2023. Of these, a total of 39 documents were identified in the area of “Art & Art History”. In the “Museum Studies” area, 19 documents were classified. As 2 papers appeared in both areas, 56 publications were totaled in this database in these segments.
Subsequently, 140 titles from a range of areas of interest, retrieved from the three databases, were listed after duplicates had been removed. Figure 1 presents the flow of the research process.

From this total of 140 documents, only conference papers or scientific journal articles were selected, resulting in the exclusion of 34 texts. To gain access to the most recent research, texts published in the last 10 years were also selected. For the criteria, this period was defined as extending from November 2013 to October 2023, the month preceding the start of the research. This selection is justified by the constant updates on this subject. Therefore, 28 papers published between 1966 and October 2013 were excluded. A list of 78 papers was then compiled.

The 78 texts were subjected to a preliminary analysis to ascertain their potential relevance to the research theme. This entailed a sift through the titles, abstracts, and keywords of the texts to identify any indications that they might be suitable for inclusion in the study. In addition, the texts were examined to determine if they were primarily concerned with photography or its derivatives, rather than any type of image. Similarly, it was verified whether the items indicated that AI and cultural heritage were the primary focus of the studies conducted, not merely mentioned as possibilities for future research. Following this screening, 42 texts were excluded, and 36 were selected for eligibility assessment, which are discussed below.

2.2. Eligibility and Inclusion

In the eligibility stage, the 36 texts selected for qualitative evaluation were read in full. This stage aimed to assess and ratify whether the papers reported research that had as its object of study AI, applied to the cultural heritage domain. Furthermore, it was necessary to ascertain whether the investigations reported also connected these two strands to photography, which should therefore be present as an object of study or as a relevant component in the methodologies. Upon completion of the reading, it became evident that photography was incorporated into the methodologies of all the publications under review.
The studies were also evaluated for their potential to provide insights into the proposed problem and to contribute to the achievement of the outlined objective. The papers were also assessed for clarity and consistency in the presentation of the theme and problem, methodology, results, and discussion. Based on these criteria, four texts were excluded at this stage.

A total of 32 texts published between November 2018 and October 2023 were preliminarily eligible for inclusion in the review. These publications were then subjected to analysis to ascertain the types of heritage they dealt with. The objective was to identify which studies focused on immovable cultural heritage, with a particular emphasis on architectural heritage [11,12], which was considered to be an additional aspect of the research's delimitation. Consequently, a subset of 22 publications with this focus, published between August 2019 and October 2023, were included in the review.

It is crucial to underscore that this review is constrained by an investigative delimitation, delineated by the proposed theme, problem, and methodology, taking into account the areas of interest. Accordingly, this review is limited to the presentation of the works returned from the database searches and selected according to the above-mentioned route, in accordance with the PRISMA protocol. Consequently, this review and the corresponding presentation of the results are not intended to be definitive on the subject; rather, they are intended to provide an overview, with representative study perspectives, to reflect on AI at the interface between cultural heritage and photography.

In accordance with the aforementioned methodology, the following sections present and discuss the 22 selected studies, in consideration of the proposed theme and problem for this review. Potential biases observed in these investigations are also highlighted in order to contribute to the interdisciplinary studies mobilized in this approach. In light of this, it is important to understand that the basis of this aspect of the analysis is the “epistemic notion of bias”. This notion “understands biased science not as science deviating from some ideal outcome, but as science that we have good reasons to suspect could have been (done) systematically better” [13]. In consequence, the organization of the presentation of the results, as well as the discussion and conclusion, is based on the principal activities and procedures developed in the heritage field. Accordingly, the biases to be highlighted are intended to identify potential avenues for enhancement in the systematic conduct of research, with a particular focus on the foundational principles and practices of the heritage field.

3. Results

To facilitate the visualization of the profiles of the qualitatively analyzed texts, Table 1 was constructed, comprising six categories: authors; publication year; publication type, which can take the form of an article (A) or a conference paper (CP); the location of the architectural heritage studied; the application of AI in the convergence between heritage and photography; and the main purpose for heritage studies.

Table 1 was organized in descending chronological order, with the most recent research identified first. As can be seen, the majority of the articles analyzed, fourteen [14–27], were published in scientific journals, while the remaining eight were presented at conferences [28–35]. In terms of location, most of the texts deal with heritage assets, whether they still exist or not, in countries on the European continent. However, there are also articles dealing with buildings located in other parts of the world, particularly in Asia. This category also includes, where possible, a list of the heritage sites studied and the cities in which they are located. The succinct presentation of the discussions proposed in the texts analyzed were undertaken to highlight trends in recent research in this area.
Table 1. The qualitative profiles of the analyzed texts are presented in a table that summarizes the key aspects of the studies.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Publication Year</th>
<th>Publication Type</th>
<th>Location of the Architectural Heritage Studied</th>
<th>Application of AI in the Convergence between Heritage and Photography</th>
<th>Main Purpose for Heritage Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artopoulos et al. [14]</td>
<td>2023</td>
<td>A</td>
<td>Cyprus</td>
<td>It is based on deep neural networks and Support Vector Machine to identify stylistic influences in architectural heritage from 3D images</td>
<td>Communication–Education</td>
</tr>
<tr>
<td>Murtiyoso and Grussenmeyer 28</td>
<td>2023</td>
<td>CP</td>
<td>Switzerland (Reformers’ Wall, Geneva, and Bern Minster, Bern)</td>
<td>A preliminary test was conducted to assess the feasibility of utilizing the Neural Radiance Field (NRF) method to recreate two instances of cultural heritage objects</td>
<td>Documentation</td>
</tr>
<tr>
<td>Notarangelo et al. [15]</td>
<td>2023</td>
<td>A</td>
<td>Italy (Sassi di Matera)</td>
<td>The development of an immersive virtual tour of cultural heritage with a Hand Gesture Recognition (HGR) system based on deep learning (DL)</td>
<td>Communication–Education</td>
</tr>
<tr>
<td>Panagiotopoulou et al. [16]</td>
<td>2023</td>
<td>A</td>
<td>Greece (Omorofokklisia, Galatsi)</td>
<td>The utilization of super-resolution techniques based on DL in the context of 3D reconstruction of cultural heritage</td>
<td>Documentation</td>
</tr>
<tr>
<td>Azizifard et al. [17]</td>
<td>2022</td>
<td>A</td>
<td>Various</td>
<td>A convolutional neural network (CNN) was utilized to perform scene recognition and to assign labels to the photographs according to each country participating in the Wiki Loves Monuments contest</td>
<td>Documentation</td>
</tr>
<tr>
<td>Gujski et al. 29</td>
<td>2022</td>
<td>CP</td>
<td>Italy (Temple of Neptune, Rome)</td>
<td>Proposition of an outlier detection method that employs two clustering machine learning algorithms: Self Organizing Map (SOM) and K-means. It utilizes data obtained through aerial photogrammetry.</td>
<td>Conservation–Restoration</td>
</tr>
<tr>
<td>Pellis et al. [30]</td>
<td>2022</td>
<td>CP</td>
<td>Italy (Spedale del Ceppo, Pistoia; Ospedale Sant’Antonio, Lastra a Signa; Basilica della Santissima Annunziata, Certosa del Galluzzo and Cappella Buontalenti, Florence)</td>
<td>State-of-the-art neural networks on heritage scenarios are employed to enhance automation in the generation of 3D heritage models</td>
<td>Conservation–Restoration</td>
</tr>
<tr>
<td>Liu et al. [18]</td>
<td>2022</td>
<td>A</td>
<td>Scotland (Bothwell Castle, South Lanarkshire)</td>
<td>The use of DeeplabV3+, a deep learning-based semantic segmentation algorithm, was employed to extract data on the damaging growth of plants on heritage buildings from photographs</td>
<td>Conservation–Restoration</td>
</tr>
<tr>
<td>Authors</td>
<td>Publication Year</td>
<td>Publication Type</td>
<td>Location of the Architectural Heritage Studied</td>
<td>Application of AI in the Convergence between Heritage and Photography</td>
<td>Main Purpose for Heritage Studies</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Maiwald et al. [19]</td>
<td>2021</td>
<td>A</td>
<td>Germany (Frauenkirche, Hofkirche, Moritzburg, Semperoper, Sophienkirche, Stallhof, Crowngate, in the vicinity of Dresden)</td>
<td>The presentation of a workflow for the retrieval and estimation of poses in historical terrestrial images, utilizing DL methods</td>
<td>Communication–Education</td>
</tr>
<tr>
<td>Kimura et al. [20]</td>
<td>2021</td>
<td>A</td>
<td>Cambodia (Bayon Temple, Angkor)</td>
<td>A proposal is put forth for the implementation of a preventive preservation system for World Heritage sites, which would employ 3D photogrammetry and AI-based image discrimination functions</td>
<td>Conservation–Restoration</td>
</tr>
<tr>
<td>Croce et al. [31]</td>
<td>2021</td>
<td>CP</td>
<td>Italy (three cloisters in the Province of Pisa)</td>
<td>Implementation of semantic segmentation via machine learning to increase automation in the recognition and classification of element classes in both 2D- and 3D-based heritage survey data</td>
<td>Communication–Education</td>
</tr>
<tr>
<td>Kawato et al. [32]</td>
<td>2021</td>
<td>CP</td>
<td>Indonesia (Borobudur Temple, Central Java)</td>
<td>The utilization of deep learning techniques to derive 3D data from two-dimensional monocular photographs of concealed relief panels</td>
<td>Communication–Education</td>
</tr>
<tr>
<td>Garozzo et al. [21]</td>
<td>2021</td>
<td>A</td>
<td>Not specified</td>
<td>A method based on the Generative Adversarial Network (GAN) is proposed to automatically synthesize unrealistic photographs. These images can be used to train systems for the classification and retrieval of images of cultural heritage sites.</td>
<td>Documentation</td>
</tr>
<tr>
<td>Croce et al. [22]</td>
<td>2021</td>
<td>A</td>
<td>Italy (Grand-Ducal Cloister of the Pisa Charterhouse)</td>
<td>Semantic segmentation of heritage via the application of supervised Machine Learning (ML)</td>
<td>Communication–Education</td>
</tr>
<tr>
<td>Felicetti et al. [23]</td>
<td>2020</td>
<td>A</td>
<td>Jordan (St. Stephen’s Church in Umm ar-Rasas)</td>
<td>The utilization of deep learning and image segmentation techniques for the purpose of obtaining a digital (vector) representation of a mosaic is hereby proposed</td>
<td>Documentation</td>
</tr>
<tr>
<td>Hatir et al. [24]</td>
<td>2020</td>
<td>A</td>
<td>Turkey (11 historical buildings in Konya)</td>
<td>The development of models based on DL and an Artificial Neural Network (ANN) for the recognition of weathering in images of heritage stone buildings</td>
<td>Conservation–Restoration</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Publication Year</th>
<th>Publication Type</th>
<th>Location of the Architectural Heritage Studied</th>
<th>Application of AI in the Convergence between Heritage and Photography</th>
<th>Main Purpose for Heritage Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrone et al. [25]</td>
<td>2020</td>
<td>A</td>
<td>Italy (Trompone Church, Palace of Pilato of the Sacred Mount of Varallo, portico of the Sacred Mount of Ghiffa, Piedmont)</td>
<td>A comparison of Machine Learning (ML) and DL methods for the classification of 3D cultural heritage.</td>
<td>Communication–Education</td>
</tr>
<tr>
<td>Kumar et al. [26]</td>
<td>2020</td>
<td>A</td>
<td>Nepal</td>
<td>Training deep neural networks to identify photographs of damaged cultural heritage among images sourced from the internet</td>
<td>Conservation–Restoration</td>
</tr>
<tr>
<td>Condorelli et al. [33]</td>
<td>2020</td>
<td>CP</td>
<td>France (Tour Saint Jacques, Les Halles, Paris)</td>
<td>The presentation of an open-source match-moving method that exploits DL and Structure from Motion (SfM) to document lost monuments based on identification in film frames</td>
<td>Documentation</td>
</tr>
<tr>
<td>Condorelli et al. [27]</td>
<td>2020</td>
<td>A</td>
<td>France (Tour Saint Jacques, Les Halles, Paris)</td>
<td>A proposed workflow for the automatic detection of lost assets in films utilizing DL is presented</td>
<td>Documentation</td>
</tr>
<tr>
<td>Condorelli and Rinaudo [34]</td>
<td>2019</td>
<td>CP</td>
<td>France (Tour Saint Jacques, Paris)</td>
<td>The application of ML techniques to the identification of architectural heritage within historical footage</td>
<td>Documentation</td>
</tr>
<tr>
<td>Condorelli, et al. [35]</td>
<td>2019</td>
<td>CP</td>
<td>France (Tour Saint Jacques, Paris)</td>
<td>A neural network is trained to identify film frames in which architectural heritage is present</td>
<td>Documentation</td>
</tr>
</tbody>
</table>

As indicated in Table 1, the 22 works are on research that can be aimed at studies in documentation, conservation–restoration, or communication–education of cultural heritage. This classification was also made based on the activities involved in the processes of musealization and patrimonialization [36] (p. 48). Furthermore, some of these studies posit collaboration as part of the methodological processes or aim to improve user accessibility in the technological environments developed. For better organization of the results, the presentation was segmented according to these aspects, also taking into account the areas of interest emphasized in the problem and objective. However, this organization should not be seen as fragmented but as integrated and connected in the composition of studies on AI, cultural heritage, and photography.

3.1. Documentation

The documentation of cultural assets is a fundamental aspect of the process of building collections or registering heritage. This involves the identification, categorization, and classification of these objects based on information obtained through research. In addition, the objective is to produce knowledge that can be used as a source of research. It is “the whole of the object and all the apparatus that documents it, that brings knowledge about it and the world it comes from, that constitutes the heritage object, or what we commonly call heritage” [37] (p. 53).

The cycle of knowledge generation, catalyzed by documentation, is also a process of symbolic valuation, as it contributes to heritage recognition, which is affected by power relations [38] (pp. 9–10). It can therefore be argued that “document and heritage are values and must therefore be understood as virtual constructions” [39] (p. 45). The
The process of “virtual construction” of heritage documents has been significantly affected by technological developments, particularly over the past two decades of the 20th century. This has involved the automation of information systems in documentation centers, the expansion of management, and access to databases through digital technologies and the Internet.

This context has prompted political action to protect and democratize the digital heritage, which is defined in the Charter on the Preservation of the Digital Heritage, published by the United Nations Educational, Scientific and Cultural Organization (UNESCO) in 2003, as “unique resources of knowledge and human expression [. . .] created digitally or converted into digital format from existing analogue data” [40] (p. 1). This charter reaffirms the commitments made by UNESCO through the Memory of the World program, created in 1992 with the objective of “guaranteeing the preservation and universal accessibility of the world’s documentary heritage”. During the first decades of the 21st century, there was a continued expansion of political, technical, and research initiatives aimed at using digital media to safeguard and enhance heritage.

The development of AI has given rise to a new set of circumstances. This research examined three distinct approaches in the texts analyzed, all of which involved in the process of documenting images of a cultural asset. The first approach aimed to collaborate with heritage documentation by contributing to the process of three-dimensional (3D) reconstruction of cultural assets. The second study explored the application of deep learning and image segmentation techniques for the documentation of mosaics. The third perspective examined the use of generative AI to synthesize unrealistic images for training in the classification and retrieval processes of photographic data from cultural heritage sites using DL [21], which is discussed in greater detail in the Collaboration and Accessibility subsection. This aspect is also covered in the study in question, as well as one of the articles from the first approach [17].

In the first perspective, the texts under discussion relate to the use of AI in association with or as an alternative to photogrammetry techniques. These techniques are already widely used for heritage documentation purposes because they allow a more comprehensive perception of the object in its three-dimensionality. This provides more information for databases, as well as favoring the dissemination, education, and preservation of cultural assets [41]. In association with that, these techniques are used to recognize and recover images, or for improving their visual and metric properties, in a process of recreating heritage in 3D. and as an alternative, to compare AI and photogrammetry methods that can be used in the 3D documentation of a heritage asset.

The associated study, a set of four texts, published between 2019 and 2020, reports on research into the application of AI in the detection of architectural heritage in historical film frames for subsequent photogrammetric application. In the first of these [35], a case study was presented on the training of neural networks to search for images of the Tour Saint Jacques, located in the French capital, in the film Études sur Paris [42]. The ML technique was proposed as the initial phase in a photogrammetric process based on the Structure from Motion (SfM) pipeline due to its considerable reduction in search time, especially when compared to manual collection. Furthermore, the images found were of sufficient quality for the successful 3D reconstruction of the heritage site in question.

The aforementioned building serves as the subject of a case study in the second paper in this series [34], which also addresses the automatic search for architectural heritage images in historical footage to identify frames suitable for photogrammetric processing, employing SfM. Furthermore, this contribution proposes a methodology for researching and evaluating the accuracy of 3D photogrammetric reconstruction, developed from heritage images identified using ML.

In addition to the Tour Saint Jacques, the other two texts from the same group [27,33] contribute to the case studies by including the Parisian pavilions of Les Halles, which were demolished in 1971, as narrated in the film chosen for the experiment, La Destruction des Halles de Paris [43]. Both texts compare the application of the methodologies presented.
above to existing and non-existent heritage. The research concluded that the proposed automatic workflow can be effective even in the face of limitations in the study material, such as the quality of the film, the absence of information about the camera used in the filming, and precise metric references in the case of the missing heritage.

In the context of photogrammetric reconstruction of heritage sites, Panagiotopoulou et al. [16] highlight that the availability of high-quality images is not always guaranteed, which is a prerequisite for the creation of accurate 3D models in terms of detail and integrity. This is due to several reasons, including the fact that these are historical images with limited spatial resolution, due to the quality and resources of the cameras used, or even, in the case of capture using unmanned aerial vehicles (UAVs), due to the height of flight and the resolution selected in 3D space.

To contribute to this process, the authors conducted a study to assess the suitability of different DL-based image super-resolution (SR) techniques, including generative AI, namely, RankSRGAN, Densely Residual Laplacian Super-Resolution, and Hybrid Attention Transformer Super-Resolution. The images used in the study were captured by a UAV and depict the 13th-century Greek Byzantine church Omorfokklisia. The study concluded that the SR techniques were effective in terms of spatial resolution, visual quality, and image detail, as well as in improving the reconstruction of 3D heritage scenes made using Multi-View Stereo (MVS) photogrammetry, carried out using the Agisoft Metashape tool. However, the work does not address the legitimacy of the images produced for heritage documentation studies. For instance, numerous “flaws” in the original images are “corrected” with synthetic data, which could compromise their status as a “document”. This introduces a potential bias in the research, particularly when considering their application in the heritage field.

Murtiyoso and Grussenmeyer [28] opted to compare the MVS photogrammetric solution with the Neural Radiance Field (NeRF) method, which employs neural networks to recreate radiance fields, in the process of 3D reconstruction of cultural heritage. In this study, the Reformers’ Wall in Geneva and the main entrance to Bern Minster in Bern, Switzerland, were selected as case studies. A set of photographs of the objects in question was used to recreate the heritage scenes using the Nerfstudio platform’s Nerfacto architecture. The results were then compared with image processing carried out using Agisoft Metashape. The NeRF method was found to be faster at 3D heritage reconstruction than MVS, although it was not superior in terms of accuracy and density. This is a qualitative issue that directly affects the documentation of cultural assets. Therefore, this research suggests that the MVS would remain a superior choice for documentary heritage practices, due to the quality of the model generated. However, the absence of discussion on this topic may be perceived as a bias, as it is not evident from the work whether the principles of heritage documentary practice were observed.

In the second perspective, another study [23] examined the benefits of utilizing Mo.Se. (Mosaic Segmentation), an algorithm that leverages deep learning and image segmentation techniques. The methodology employed a combination of the U-Net 3 Network and the Watershed algorithm. The objective was to define a workflow that delineates the steps for a segmentation, ultimately leading to the generation of a digital (vector) representation of the mosaic of St. Stephen’s Church in Umm ar-Rasas (Jordan).

As is noted in the majority of these publications, studies into the documentation of cultural heritage using AI are also relevant because they provide information that is necessary for investigations and conservation and restoration activities. This is discussed in the following subsection.

3.2. Conservation and Restoration

In addition to documentation, conservation is a fundamental process that aims to ensure the longevity of heritage. It can be defined as “an activity that consists of adopting measures to ensure that a particular asset experiences the least number of alterations for as long as possible” [44] (p. 24). About restoration, it can be stated that “it consists of
returning something to its original or authentic state”, although it should be noted that on many occasions it is not possible to differentiate between the two activities [44] (p. 24). In both instances, photography helps identify alterations and diagnose damage, as it provides the visual data necessary to examine and monitor the state of the object.

This practice has been a fundamental aspect of conservation and restoration work since the advent of photography. For instance, the French Historical Monuments Commission organized “heliographic missions” in the 19th century to tour the country and photograph the “most important national buildings” to develop a restoration program [45] (p. 29). Imagery research has been refined throughout the 20th century and continues into the 21st century, with the incorporation of increasingly accessible digital technologies and the utilization of computer simulation tools with greater frequency [46] (pp. 103–104).

AI contributes to this process, as evidenced by the publications analyzed. In this respect, there are two main research directions. As discussed in the previous subsection, one group of researchers proposes the combined use of DL with photogrammetry, in a process that results in the possibility of assessing the condition and state of deterioration in heritage sites through observation of the 3D model. This subsection focuses on two of the selected texts [29,30], with the other two [18,20] discussed later, in the Collaboration and Accessibility subsection, as they place greater emphasis on this aspect of the research process. The same is true of one of the articles [26], which, like the other article [24] highlighted shortly afterwards, is in the field of studying the use of AI to recognize damage, degradation, and other alterations to the structures of heritage buildings using images.

In light of the aforementioned considerations, Hatir et al. [24] conducted a comparative analysis between CNNs and Artificial Neural Networks (ANNs) in the context of weathering recognition in 11 historic buildings constructed from Sille stone in the Konya region of Turkey (Catholic Church, Seyfettin Kara Sungur Mausoleum, Hoca Ahmet Fakih Mosque, Teacher Training College, Ateş Baz-ı Veli Mausoleum, Sakahane Mosque, Şeyh Osman Rumi Mausoleum, Halka Begüş Mosque, Ak Mosque, Mevlana Mausoleum, and Ottoman Bank).

That research aimed to reduce the subjectivity associated with the recognition of deterioration, which is susceptible to “experts’ errors in restoration work”, by developing an AI-based methodology that can accelerate and improve the accuracy of this process. In a field image study, fresh stone and eight types of weathering common in the buildings studied (flaking, contour scaling, cracking, differential erosion, black crust, efflorescence, higher plants, and graffiti) were identified and photographed, based on the characterization drawn up by the International Scientific Committee for Stone (ISCS) of the International Council on Monuments and Sites (ICOMOS). The 8598 images generated were classified using the CNN and ANN methods. The study found that the CNN method was more successful than the ANN in the classification process. This was due to the speed and reliability of the CNN method.

In a different direction of research, a case study [29] was carried out of the Temple of Neptune (Rome, Italy), surveyed by aerial photogrammetry with the aid of an UAV. To ascertain the accuracy of the final model, an outlier detection method was devised which considered each and every calculated parameter. This method employed two clustering machine learning algorithms—Self Organizing Map (SOM) and K-means—to arrive at a compromise model between the available point data and the noise reduction associated with 3D definition.

Conversely, Pellis et al. [30] presented a workflow for the automatic semantic segmentation of 3D point clouds of architectural heritage. The research considered a series of image datasets sourced from several Italian locations: Spedale del Ceppo in Pistoia, Ospedale Sant’Antonio in Lastra a Signa, and Basilica della Santissima Annunziata, Certosa del Galluzzo, and Cappella Buontalenti, in Florence. However, not all of these datasets were utilized in all of the tests conducted.

The process consisted of two main stages. In the initial stage, the images were segmented utilizing the DeepLabv3+ CNN architecture, implemented with MATLAB software,
employing the basic classification architecture, namely, the CNN ResNet18 pre-trained on the ImageNet visual database set. Subsequently, in the second stage, the classified pixels were back-projected onto the 3D point cloud utilizing a masking method, photogrammetry, and the principles of dense image matching. Although the results were deemed promising, further tests were required to evaluate the entire workflow and identify potential improvements. These included optimizing the automatic image segmentation and back-projection procedures, particularly to address the identified issues of segmentation prediction precision and label overlapping.

In that study, there are indications of potential biases, particularly with regard to greater transparency in the semantic segmentation process carried out by AI. Without such transparency, it is challenging to assess the quality of the results and their applicability to conservation and restoration practices.

In the next subsection, entitled Communication and Education, two additional articles on semantic segmentation are discussed [14,25]. Although they do not explicitly state this as their purpose, they have been classified as such because they signal the need to promote cultural heritage and socialize knowledge. The other two texts [19,32], which are also discussed below, are more clearly aimed at communication and education. A fifth article [15] in this category is discussed in the subsection on Collaboration and Accessibility, as it places greater emphasis on these aspects.

3.3. Communication and Education

In the field of cultural heritage, the areas of communication and education are mutually reinforcing, not only in terms of sharing information and socializing studies but also in the process of constructing meaning from heritage. In this regard, heritage can be understood as a “cultural production” in which its objects are “significant social beings” [47], whose images are therefore symbolic representations that mobilize multiple space–time relationships.

This understanding extends beyond the instrumental dimension of communication and education applied to heritage activities aimed at the public. It primarily considers the relevance of (re)signification, mediation, interaction, and the experience of heritage in environments where knowledge is exchanged in society and, consequently, continuously transformed. The social appropriation of cultural heritage through heritage education seeks a socio-historical understanding of cultural references and aims to recognize, value, and preserve these assets [48] (p. 19). Furthermore, the communicational perspective on heritage necessitates an understanding of communication as “a socio-symbolic activity” This implies that “the variety of modes of existence of heritage assets must be taken into account”, or “the place they occupy in the cultural life of our societies” [47].

This diversity encompasses the expansion of heritage to other environments, as enabled by the development of technology, and the cultural representation of these objects in the contemporary era. AI introduces further elements that prompt reflection on these meanings. One such element is semantic segmentation, which has been described as “one of the most important research methods for computer vision” and which “has the task of classifying each pixel or point in the scene into classes that have specific characteristics” [25] (p. 1).

Matrone et al. [25] recognized the significant advances that AI has made possible for research into the semantic segmentation of 3D point clouds in cultural heritage. They compared and combined ML and DL methods in the development of an architecture called DGCNN-Mod+3Dfeat for this purpose. In that research, tests were conducted on three heritage scenes: the Trompone Church, the Palace of Pilato of the Sacred Mount of Varallo, and the portico of the Sacred Mount of Ghiffa, all located in the Italian region of Piedmont. The study concluded that neither ML nor DL demonstrated superior classification accuracy. Consequently, both approaches had the potential to classify data collected by various techniques, such as LiDAR or photogrammetry.
In three additional studies, semantic segmentation was employed at the initial stage of the methodological process. In the first study [14], semantic segmentation was considered at a preliminary stage, to assist the process of classifying the styles of Cypriot architectural heritage. This classification was carried out with a CNN’s 3D and knowledge transfer, based on data obtained through photogrammetry or Terrestrial Laser Scanning (TLS). Additionally, a Support Vector Machine (SVM) was employed to identify stylistic influences in the images. In the case of Cypriot heritage, these influences were diverse and even combined (e.g., Gothic and Ottoman, or Ottoman and colonial).

The initial phase of the second research [22] involved the semantic segmentation of the 3D point cloud on the case study of the Grand-Ducal Cloister of the Pisa Charterhouse in Italy. The methodology commenced with the raw 3D point cloud, which was utilized as input data in the application of supervised ML. The process resulted in obtaining the semantic point cloud, in which the classes of architectural elements had been identified and labeled.

A third study [31] was developed by a similar group of researchers. This study commenced with the implementation of semantic segmentation via machine learning. The objective was to increase automation in the recognition and classification of element classes in both 2D- and 3D-based heritage survey data. In the second phase of the study, a parametric reconstruction of the classes of elements was conducted, employing the use of visual programming languages. The case study encompassed three cloisters in the Province of Pisa (Italy). As a result, the predictive ML model enabled the semantic organization of the information present in the raw data, subsequently facilitating the derivation of a conceptual representation.

The objective of these three research works was to facilitate the segmentation and annotation process for Heritage or Historic Building Information Modelling (HBIM), a modelling system for cultural heritage, based on the Building Information Modelling (BIM) procedure developed by the Architecture, Engineering and Construction (AEC) industry. This procedure creates intelligent 3D models of buildings that can be added to libraries for reuse. However, this task has not been straightforward in HBIM studies [49]. In light of this consideration, a potential bias may be discerned in the aforementioned studies, as they do not discuss in sufficient depth the viability of these AI methods with respect to HBIM.

On the one hand, AI is present in the processes of classifying and identifying the characteristics of heritage, which contributes to the weaving of its meanings in contemporary societies. On the other hand, it can also collaborate with studies that emphasize the interaction and enjoyment of cultural heritage in technologically created environments. This could involve travelling “through space and time, and thus, making cultural heritage tangible” on platforms such as hand-held Virtual Reality (VR) or geographic information systems (GIS) [19] (p. 2).

This is the possibility identified by Maiwald et al. [19] when they present a workflow using DL for retrieving and estimating the precise positioning and orientation of historical images, especially those that are sources of visual information about destroyed or modified buildings. The study defines heritage buildings in the vicinity of Dresden, Germany, as objects of interest, namely Frauenkirche, Hofkirche, Moritzburg, Semperoper, Sophienkirche, Stallhof, and Crowngate. It tests a CNN in the recovery of relevant historical images, whose position and orientation of the camera is defined by photogrammetry. Heritage scene reconstruction is carried out using the SFM COLMAP software.

For a similar purpose, Kawato et al. [32] developed a VR system for the Borobudur temple in Indonesia. To do this, they used photogrammetry to obtain 3D point clouds of the entire temple, with data obtained by a UAV-coupled camera, and of the selected parts of the building. They also used DL with a depth estimation neural network to recover 3D data from Buddhist relief panels. As these panels are situated behind stone walls, the recovery was based on photographs taken by Kassian Cephas in 1890, which are scarce records of the works that are no longer visible in situ.
From these texts, it can be seen that the technological strategies emerging in classification, interaction, and immersion studies have AI as a relevant resource for updating heritage communication and educational processes. This is because AI enables new experiences, methods, and results, especially with the use of images. Furthermore, the studies indicate the potential of AI methods to contribute to collaborative and accessibility procedures, which can be developed in all the activities of the heritage process. This is examined in more detail below.

3.4. Collaboration and Accessibility

This final subsection of the presentation of results lists publications that refer to collaboration or accessibility resources as strategies that, through AI, can provide answers to contemporary problems in research using images about cultural heritage. These problems involve Big Data, or massive data, a term coined in 2005 to describe the wide range of data available on the Internet [6]. Furthermore, they refer to the production of cyborg networks, traced through the interaction between people and machines. In addition, they affect the reconfiguration of “common memory, the inextricable mass of relationships between data” [50] (p. 84), including images and heritage.

In this context, the six articles included in this subsection intertwine not only because they mobilize these issues, but also because they represent a meeting of different possibilities for applying these studies in the various patrimonialization actions. These actions should also be understood in an integrated way. Based on these considerations, three perspectives emerge from this research. They are inferred more from the issues highlighted in this subsection than from the main purposes in patrimonialization activities, as discussed above.

The first perspective is based on collaboration through the generation of massive data shared on the Internet, especially photographs. Two articles are presented in this vein, each with its purpose. Garozzo et al. [21] used 1853 images of classical order columns collected through web scraping on various sites, including Flickr and Google Images, to train a Generative Adversarial Network (GAN) structure. The training prepared the GAN structure to generate synthetic images on this theme, which could also be used in other classifier training and image retrieval systems on cultural heritage, thereby increasing the volume of data needed in these processes. Despite the potential demonstrated, the study recognized that progress needed to be made in generating better-quality images, which requires more training and tests with other network architectures. Notwithstanding this acknowledgment, a potential bias may be discerned in that research. That article did not address the document status of these synthetic images. This is a pertinent issue, given that they underscore the absence of a connection with reality, a link that is highly valued in heritage documentation.

Another study [26] sought to detect images of cultural heritage damaged by disasters on social networks, with a particular focus on the X network, previously known as Twitter. Kumar et al. [26] presented an automated method for classifying images, to reduce the professional’s effort in the task. In particular, they conducted a case study with images from the Nepal earthquake in 2015, which were classified using a CNN. Despite the potential of using images from social networks, which provide instantaneous records of disasters, and of the method, which significantly reduces and qualifies the quantity of data to be examined by technicians a posteriori, the authors noted that the model created was not yet generalizable for use in other types of disasters and may not work with other social networks. Furthermore, a potential bias that can be observed in that research and others involving social networks and other Internet spaces is the failure to make explicit the validity of the use of these images from these sources. This includes both the permission to use the images and the authenticity and integrity of the images themselves. These are fundamental considerations in the field of heritage activities and studies.

On the other hand, that article also pointed in the direction of the second perspective mentioned, namely, cyborg collaboration between human and machine, as it indicated...
that people shared images on a network that could be used in an AI model which, in turn, was collaborative with human expertise. In a similar vein, but with more specific sources, three other articles deal with the AI processing of photographic data of damage to heritage structures, provided by visitors and tourists when they explore these sites. Such data would be of use in preservation actions carried out by specialists in the field.

The first piece of research [17] utilizes images of wiki sites, a Hawaiian word meaning “quick”. Consequently, wiki represents a rapid and easy way for non-technical individuals to create and edit websites [51], enabling the collaborative management of content on the Internet, including image banks. That study examined the photographic collection generated through crowdsourcing in 11 editions of the Wiki Loves Monuments (WLM) international photography competition. The images were utilized to train the visual descriptor by country, with the objective of identifying the characteristics of a digital representation of each country’s national heritage in order to facilitate comparison with the other nations participating in the competition. A CNN was employed to perform scene recognition and to label the photographs according to each country. However, the study acknowledged the existence of limitations, including “sample biases that are intrinsic to the nature of the contest and the platform used for their collection”. These biases included the valorization of European and colonial heritage images, as well as imprecision about the definition of each nation’s heritage characteristics.

The second study [18] employs crowdsourced images shared by visitors via email or social networks following invitations from the Monument Monitor project, run by Historic Environment Scotland (HES) and the Institute of Sustainable Heritage at University College London. The research proposes the combined use of computer vision, through DeeplabV3+, a convolutional semantic segmentation model, with photogrammetry in the processing of crowdsourcing images of Bothwell Castle, located in Scotland, the target of the case study. Despite the limitations of the study, it points to achievements, such as the feasibility of being applied to other damage monitoring actions in this type of building.

The third study in this group [20] evaluates the possibility of tourists participating in a preservation system using AI, which, in the case study, is aimed at the Bayon Temple in Cambodia. Also through crowdsourcing, participation would be through the recording of images via an app, which would send them to storage for the project. The article’s focus is on evaluating the participation of volunteers, with approximately half of those interviewed indicating a willingness to collaborate with the system. While it does not detail how AI would be used, it indicates that it would be combined with SfM to create the 3D model, with the heritage expert assessing the state of deterioration and damage to the structure.

While previous studies have examined collaborations in the generation and analysis of image data from two perspectives, the latest article takes a different approach, focusing on expanded accessibility in immersive heritage memory, woven by images and neural, biological or artificial networks. In this third perspective, Notarangelo et al. [15] discussed the potential of a virtual tour application for cultural heritage. The application was developed through the combined use of DL, in a Hand Gesture Recognition (HGR) system, and photogrammetry, with RealityCapture software to generate 3D representations from photographs taken of an alleyway within the Sassi di Matera heritage site in Italy. In a test carried out as part of the study, users accepted the proposed prototype, including the use of hand gestures, which was considered intuitive and conducive to immersion in the interactive platform.

In addition to the educational purposes that Notarangelo et al. [15] highlight, this experience refers to the impact of AI on the reconfiguration of heritage memory through the expansion of immersive environments. This is not only due to advances in technology, including accessibility, but mainly because it mobilizes other heritage experiences that increasingly contribute to reshaping “common memory”, which is strongly imagistic. This perspective is added to the other two, which lead us to reflect on the emerging involvement of cyborg collaborative practices in cultural heritage preservation processes and on the (massive) generation of image data in a network of points of view, with diverse perspectives
on the heritage issue. The discussion of this technological conjuncture, which is constantly changing, continues in the next section to understand the meaning of these approaches in AI studies, in the convergence between cultural heritage and photography.

4. Discussion

The objective of the 22 publications analyzed in this review was to propose AI methods that would result in solutions to the challenges encountered in cultural heritage preservation work routines. These publications emphasize the main purposes of heritage studies, namely, documentation, conservation and restoration, communication, and education. However, many studies indicate the application of solutions to more than one procedure [18, 21, 25, 28, 30, 32, 34, 35], and sometimes even to all three angles of action [14, 20]. This illustrates the motivational force behind the study of AI to establish integrated practices in heritage preservation, which can enhance the efficiency of the process, both in terms of its parts and as a whole, which can be woven together in its complexity [52].

This issue was briefly highlighted in the concluding subsection of the presentation of results, Collaboration and Accessibility. It is reinforced again to signal that this integration for heritage preservation can go beyond expert circles and also involve society. It can therefore be noted that there is an interest in investigating AI in the use of the mass production and sharing of images, facilitated by digital technologies, in heritage preservation procedures previously reserved for specialized expertise.

On the one hand, these initiatives have the potential to lead to progress, not only by expanding study materials and developing AI methods to process them, which would save time and professional effort, as has been pointed out in various publications analyzed [19, 21, 26, 27, 30, 34, 35]. However, they also have the potential to make society aware of the spontaneous preservation of cultural assets, a subject not explored in the publications studied.

Conversely, other issues that are not dealt with, or only briefly touched upon, in the texts indicate potential biases or problems that could be discussed in greater depth. One such issue is the very idea of the image as a representation or sign, which underpins the attribute of a heritage image document. This is further affected by AI technologies, especially generative AI. Another is ethics in image processing and creation. Furthermore, there is the question of permission to use, authenticity, and integrity of images included in investigations, particularly those taken directly from the Internet.

Another issue is the consequences of using AI in heritage preservation and enhancement procedures, not only for the job market in the sector but also in terms of the transformations that may occur in the (re)organization of the area itself. In addition, the recognition of new heritages or the rethinking of established ones based on discoveries made possible by the use of AI is also a possibility. It should be noted that these issues may not have been the primary focus of the studies analyzed, as the majority of them were conducted by researchers in the fields of Computer Science and architecture. Nevertheless, they may potentially give rise to further research in the fields that are of interest for this review as previously mentioned.

About the methodological highlights, it is notable that some of the publications analyzed report the development of solutions using CNNs [14, 17, 19, 26, 30], which have proven to be a productive approach in the recovery of historical images, the automatic classification of images, and the recognition of alterations or damage to heritage structures. However, there has also been research into other methods, such as ANNs [24] or NeRF [28], which have not generally shown greater efficiency compared to the use of CNNs and photogrammetry, respectively.

Another trend in the research analyzed was the combined study of AI and photogrammetry in processes aimed at reconstructing 3D heritage scenes [14–16, 18, 22, 27, 28, 30, 32–35]. These processes yielded promising results in the three heritage study purposes highlighted here. However, the use of generative AI has demonstrated the need for greater research
efforts, at least in terms of improving the quality of the synthetic images generated by GAN trained on the imagery dataset collected from the Internet [21].

As well as images retrieved from the internet, successful studies with DL have also used other sources. These include historical archives [34,35], which are often the only place to obtain photographs, especially if the heritage site no longer exists [27,33] or is inaccessible [32].

In addition to the acquisition of images by technical staff [24] or through collaboration [18,20], another method of obtaining photographs has been through the use of cameras attached to UAVs [16,29,32]. These aerial images can be satisfactorily adjusted using DL techniques, including generative AI [16], which was used with greater success in this case. The combined use of photographic material, including analogue images recorded over a hundred years ago, and digital images obtained with UAVs has been employed in techniques that combine DL with photogrammetry [32].

These strategies demonstrate that photography continues to be an important methodological component of heritage studies. However, the changes in techniques show transformations in the relationship between photography and cultural heritage, the nuances of which are visualized in contemporary times. Although the publications analyzed in this review do not address this issue directly, they demonstrate that AI is driving the process of reconstituting photography, through updates, as a reference point for reconstructing the meaning of heritage.

Photographic representations of heritage are no longer fixed solely in the two-dimensionality of the frame; rather, they are modelled on the three-dimensionality of the environment, where humanity is more accustomed to perceiving objects in the world. Furthermore, the discourse surrounding heritage imagery has been extended to encompass the four-dimensionality of temporal displacement, with the link between space and time. As a result of this, photography with AI is becoming an increasingly effective means of learning about the space–time machine of heritage, where the journey leaves no one untouched. The knowledge that emerges from this trip, made dynamic by AI, is therefore doubly transformative. On the one hand, it updates photography through extensions of image production that preserve its essence. On the other hand, it re-signifies cultural heritage through discoveries that are substantially possible as a result of changes in this relationship.

5. Conclusions

This review provides a contemporary overview of AI research at the interface between cultural heritage and photography, to contribute to future research, particularly, but not exclusively, in the areas of interest highlighted here. The studies analyzed respond to the proposed problem of observing how this theme is currently approached in the fields of interest to the research. Furthermore, the publications analyzed, with their methods and results, indicate potential avenues for the continued exercise of interdisciplinarity or transdisciplinarity through this interface in the context of these innovations, with implications for the relationship between humanity, and its technologies, cultural assets, and their images.

Concerning the proposed objective of the study, in summary, the publications analyzed generally show two facets under discussion. On the one hand, in terms of potential, they seek to present AI techniques that automate, speed up, and increase the efficiency of heritage preservation processes carried out by professionals, mostly in manual or semi-manual tasks. The aim is not only to improve these practices, including reducing human error but also to collaborate with academic research. They also carry out studies using materials and methods that are unconventional or unusual for the heritage field, for instance, the utilization of images sourced from the Internet or generated by generative AI, and the integration of cyborg collaborative practices, involving non-specialists such as tourists and visitors to heritage sites, in processes facilitated by AI. These methodologies could prompt intense debate in the technical domain and the field of cultural heritage studies.
On the other hand, three key aspects emerge from the challenges identified in the publications analyzed. Firstly, the quality and quantity of the material considered in the research is a significant factor. For instance, there are cases where there are few images of poor quality, which makes it challenging for the neural network to be trained effectively. In other cases, the large volume of images makes it difficult for the AI to detect the data of interest due to their dispersion and heterogeneity. Secondly, it is recognizing the necessity to improve AI methods, which still have considerable shortcomings. For example, the significant percentage of inaccuracy in automatic image classification. This requires retraining of models and renewed testing to re-evaluate workflows in search of improvements and solutions to the problems identified. Finally, several of the case studies cannot be generalized to apply the methods to other investigations.

In the context of challenges and potential, it is imperative to consider the ethical responsibility in the use of these innovations. Spinoza posits that ethics concerns the “path that leads to freedom” [53], where knowledge of the affections is fundamental. For Spinoza, “those who attempt to regulate their affections and appetites exclusively for the sake of freedom will endeavor to gain knowledge of the virtues and their causes, and to fill their minds with the joy that arises from a true knowledge of them, rather than considering the defects of men, humiliating them, or rejoicing in a false appearance of freedom” [53]. Thus, ethics represents an exercise in the pursuit of knowledge that commences with the individual, as a matter of priority. This process may subsequently result in the formulation (if only provisionally) of a collective understanding of ethics and the ethical freedom to act. Given the potential of AI technologies, the debate on ethics in this area of research is extensive and still in its nascent stages. According to Bartneck et al. [9], “there is a lack of practical, agreed guidelines and rules regarding systems that are much more autonomous in their calculations, actions and reactions than what we have been used to in the past”. These authors observe that, at present, “most jurisdictions around the world have only just started to investigate regulatory aspects of AI” [9] (p. 103).

With regard to the weaving of consensus and norms, the UNESCO Recommendation on the Ethics of Artificial Intelligence, adopted on November 23, 2021, indicates among the areas for political action, in the field of culture, that the Member States should “incorporate AI systems, where appropriate, in the preservation, enrichment, understanding, promotion, management and accessibility of tangible, documentary and intangible cultural heritage” [54] (p. 32). However, as per the recommendation, such actions should be undertaken with caution and accompanied by studies. Among the various recommendations, the document points out that Member States should “promote and support AI research, notably AI ethics research” and “promote general awareness programmes about AI developments, including on data and the opportunities and challenges brought about by AI technologies” [54] (pp. 34–35).

This represents a pressing need for research and the socialization of knowledge that affects cultural heritage studies, including in the interface with photography and its derivatives. Nevertheless, the inappropriate use of AI can, among other things, compromise the veracity of the information obtained or even the credibility of the results generated. For example, it can generate false images about heritage sites or distort heritage image data, which can lead to threats to human rights and disrespect for cultural diversity, among other harms. This can result in a lack of trust and a potential compromise to the reliability of the “document” and “representation” of cultural assets, which have been increasingly questioned since the advent of digital technologies. Despite this, these aspects remain fundamental to heritage studies. As indicated in UNESCO’s recommendation for addressing this issue, ensuring the transparency and explainability of AI systems is crucial for “ensuring the respect, protection and promotion of human rights, fundamental freedoms and ethical principles” [54] (p. 22). Transparency and explainability, in equilibrium with other principles such as privacy, safety, and security, are essential for promoting an understanding of these innovations and their impact on cultures and their heritage. Furthermore, these
actions must be aligned with the objective of combating disinformation and promoting inclusion.

Therefore, the ethical issue, which has not been addressed by the research reviewed here, represents a significant opportunity for further investigation in the interface between AI and cultural heritage, particularly in the context of the use of photography and its derivatives. The image, through the use of AI, provides a process of resignification of the link with reality. It reveals remembrances that were never recorded imagistically, such as through the creation of images from textual descriptions. It generates new imagistic memories, whether or not they are based on images from the past. However, with ethical care, it can be preserved as a “document of historical life” [3], albeit in transformation. Memory is a fleeting permanence, made up of changeable frames.

An examination of the significance of these innovations leads to the perception of photography and heritage in the “duration that flows”, given that “a self that doesn’t change doesn’t last” [55] (p. 18). Consequently, it is essential to “deepen the nature of time” to “better understand that duration means invention, the creation of forms, and the continuous elaboration of the absolutely new” [55] (p. 25). This process actively affects the relationship between cultural heritage and photography, which is woven into memory as a “prolongation of the past in the present, in other words, active and irreversible duration” [55] (p. 31).

Heritage imagery memory forms networks, made up of layers, through which data increasingly accumulated in archives flow, whether analogue or digital, personal or communal, private or public. These networks are updated with AI, which enables new connections, processing, and experimentation of data. As it expands in time and space, it is reconfigured to remain in duration, which is always in flux and undergoing constant change.

Author Contributions: Conceptualization, C.S.; methodology, C.S. and L.O.; validation, C.S. and L.O.; formal analysis, C.S. and L.O.; investigation, C.S.; resources, C.S. and L.O.; data curation, C.S.; writing—original draft preparation, C.S. and L.O.; writing—review and editing, C.S. and L.O.; visualization, C.S. and L.O.; supervision, L.O.; project administration, C.S. and L.O.; funding acquisition, L.O. All authors have read and agreed to the published version of the manuscript.

Funding: This work is financially supported by national funds through FCT—Foundation for Science and Technology, I.P., under the project UIDB/05460/2020, Portugal; and by Federal University of Paraíba, Brazil.

Data Availability Statement: Not applicable.

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

References
4. Desvallées, A.; Mairese, F. Conceitos-Chave de Museologia; Comitê Brasileiro do Conselho Internacional de Museus; Pinacoteca do Estado de São Paulo; Secretaria de Estado da Cultura: São Paulo, Brazil, 2013.

de%20Crac%C3%B3via%202000.pdf (accessed on 1 April 2024).


15. Notarangelo, N.M.; Manfredi, G.; Gilio, G. A collaborative virtual walkthrough of Matera’s Sassi using photogrammetric reconstruction and hand gesture navigation. J. Imaging 2023, 9, 88. [CrossRef]


27. Condorelli, F.; Rinaudo, F.; Salvador, F.; Tagliaventi, S. A neural networks approach to detecting lost heritage in historical video. ISPRS Int. J. Geo-Info. 2020, 9, 297. [CrossRef]


49. Lovell, L.J.; Davies, R.J.; Hunt, V.L. The Application of Historic Building Information Modelling (HBIM) to Cultural Heritage: A Review. Heritage 2023, 6, 6691–6717. [CrossRef]
52. Morin, E. Introdução ao Pensamento Complexo; Sulina: Porto Alegre, Brazil, 2005.
53. de Spinoza, B. Ética; Autêntica Editora: Belo Horizonte, Brazil, 2009.

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