heritage

MDPI

Article

ImageOP: The Image Dataset with Religious Buildings in the World Heritage Town of Ouro Preto for Deep Learning Classification

André Luiz Carvalho Ottoni [1,*] and Lara Toledo Cordeiro Ottoni [2]

[1] Machine Learning and Robotics Research Group (MLBots), Department of Computing, Federal University of Ouro Preto (UFOP), Campus Morro do Cruzeiro, Ouro Preto 35400-000, MG, Brazil
[2] Department of Industrial Automation, Federal Institute of Minas Gerais (IFMG), Ouro Preto 35400-000, MG, Brazil; lara.ottoni@ifmg.edu.br
* Correspondence: andre.ottoni@ufop.edu.br

Abstract: Artificial intelligence has significant applications in computer vision studies for cultural heritage. In this research field, visual inspection of historical buildings and the digitization of heritage using machine learning models stand out. However, the literature still lacks datasets for the classification and identification of Brazilian religious buildings using deep learning, particularly with images from the historic town of Ouro Preto. It is noteworthy that Ouro Preto was the first Brazilian World Heritage Site recognized by UNESCO in 1980. In this context, this paper aims to address this gap by proposing a new image dataset, termed ImageOP: The Image Dataset with Religious Buildings in the World Heritage Town of Ouro Preto for Deep Learning Classification. This new dataset comprises 1613 images of facades from 32 religious monuments in the historic town of Ouro Preto, categorized into five classes: fronton (pediment), door, window, tower, and church. The experiments to validate the ImageOP dataset were conducted in two stages: simulations and computer vision using smartphones. Furthermore, two deep learning structures (MobileNet V2 and EfficientNet B0) were evaluated using Edge Impulse software. MobileNet V2 and EfficientNet B0 are architectures of convolutional neural networks designed for computer vision applications aiming at low computational cost, real-time classification on mobile devices. The results indicated that the models utilizing EfficientNet achieved the best outcomes in the simulations, with accuracy = 94.5%, precision = 96.0%, recall = 96.0%, and F-score = 96.0%. Additionally, superior accuracy values were obtained in detecting the five classes: fronton (96.4%), church (97.1%), window (89.2%), door (94.7%), and tower (95.4%). The results from the experiments with computer vision and smartphones reinforced the effectiveness of the proposed dataset, showing an average accuracy of 88.0% in detecting building elements across nine religious monuments tested for real-time mobile device application. The dataset is available in the Mendeley Data repository.

Keywords: deep learning; cultural heritage; dataset; building facade; image classification

check for updates

Citation: Ottoni, A.L.C.; Ottoni, L.T.C. ImageOP: The Image Dataset with Religious Buildings in the World Heritage Town of Ouro Preto for Deep Learning Classification. *Heritage* **2024**, *7*, 6499–6525. https://doi.org/10.3390/heritage7110302

Academic Editors: Peter J. Cobb and Dirk Spennemann

Received: 24 October 2024
Revised: 17 November 2024
Accepted: 18 November 2024
Published: 20 November 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The use of artificial intelligence (AI) represents a new phase and offers numerous opportunities for studies of cultural heritage [1–4]. An important aspect of this emerging research field focuses on the application of AI models in the visual inspection of historical buildings [3]. In this context, the study by Silva and Oliveira [4] highlights that AI techniques can be employed in conjunction with photography research for documentation, conservation, restoration, and education regarding cultural heritage.

In line with this, deep learning models are increasingly being adopted in research involving computer vision and cultural heritage [5,6]. In the literature, recent examples of applications include: monitoring of historic places using convolutional neural networks [7], semantic segmentation and photogrammetry for monitoring historic facades [8], remote

sensing of historical architecture [9], prediction and measurement of damage to architectural heritage facades [10], detection of disaster-affected cultural heritage sites from social media images [11], and statistics and location estimation of missing components in routine inspections of historic buildings [12]. Furthermore, Mishra and Lourenço [3] describe that applications of artificial intelligence in visual inspection for cultural heritage can be divided into several groups, such as detection of surface deterioration in cultural heritage, recognition of facade damages, crack detection, and degradation monitoring of stone cultural heritage. Another relevant research area described by Mishra and Lourenço [3] focuses on the detection of components and key architectural elements in cultural heritage using AI. Indeed, the recognition of building elements in historical monuments is essential in the processes of conservation and digital documentation of cultural heritage [3].

Recently, several studies in the literature have explored the classification of historical architectures and building elements using deep learning [13–15]. The paper by Llamas et al. [13] highlights that image classification is a relevant task for the digital documentation of cultural heritage, given that manual classification is a costly process that can involve many errors. Moreover, techniques to automate the digital documentation process can improve the management of cultural heritage archives, making searches more efficient and assisting in the study and interpretation of heritage [13]. Based on this, Llamas et al. [13] propose the Architectural Heritage Elements Dataset, which includes 10 types of architectural elements from heritage buildings, the majority of which are churches and religious temples. In another study, Siountri and Anagnostopoulos [15] propose the classification of cultural heritage buildings in Athens using deep learning. For this, the authors introduce a dataset with images collected from photographs, Google Street View, and online sources. Additionally, they conduct simulations in two phases to identify four building elements (doors, windows, balconies, and corbels) and the architecture style (neoclassical, neoclassical-eclectic, interwar-eclectic, interwar, apartment building).

However, the literature still lacks image databases for the classification of Brazilian cultural heritage, particularly regarding the historic town of Ouro Preto. Ouro Preto has been a UNESCO World Heritage Site since 1980 and features well-preserved historical buildings from the Brazilian colonial period, such as mansions, public buildings, and museums. In this context, this World Heritage Site is also notable for its large number of religious monuments, primarily reflecting Baroque and Rococo architectural styles. Therefore, the objective of this paper is to propose The Image Dataset With Religious Buildings in the World Heritage Town of Ouro Preto for Deep Learning Classification (ImageOP). This new dataset comprises 1613 images of 32 Catholic religious temples in the city of Ouro Preto, with the data divided into five classes: fronton, door, window, tower, and church. The simulations were conducted on the Edge Impulse platform using two deep learning models: MobileNet V2 [16] and EfficietNet B0 [17]. These artificial intelligence structures are convolutional neural network architectures widely discussed in the literature for computer vision applications using mobile devices. The experiments demonstrate that the dataset is suitable for deep learning experiments, as the F-score values were generally above 90% for the detection of all classes in simulated experiments. Additionally, accuracies exceeding 75% were achieved in real-time detection of building elements in experiments conducted in real-world environments using computer vision and mobile devices. Thus, this study aims to contribute to important areas of the literature, such as digital documentation of cultural heritage [3], classification of building components [13], and the creation of new datasets for deep learning applications [18,19]. Accordingly, the main contributions of this work are:

- A new image dataset featuring religious buildings in the World Heritage Town of Ouro Preto;
- Deep learning for the recognition of five classes pertaining to historical temples of Brazilian religious architecture: fronton, door, window, tower, and church;
- Simulated experiments and real-world applications using computer vision and mobile devices (smartphones) for the detection of components in historical religious buildings.

This paper is organized into five sections. Section 2 provides a description and comparison with related work. Section 3 presents a detailed description of the proposed new dataset. Section 4 discusses the experiments with deep learning and their results. Finally, Section 5 presents the conclusions and directions for future research.

2. Related Work

This section presents a comparison of the current proposal with related work in the literature. To this end, four papers from the field of machine learning for cultural heritage classification were selected: I—Llamas et al. [13], II—Janković [14], III—Siountri and Anagnostopoulos [15], and IV—Lamas et al. [20]. These articles were selected based on their high relevance to the research field and direct relationship with the theme of this paper. To achieve this, some necessary criteria to be met were defined. The first point was that related work should evaluate cultural heritage buildings. Another topic was that the analyzed papers should adopt deep learning methods for image classification. The third criterion for selecting studies for comparison was that the proposed or used datasets must be from the last ten years. Finally, relevance in the research field was assessed by the number of citations, with all selected studies being cited at least ten times (Google Scholar—accessed on 17 November 2024). Table 1 presents this comparison.

Table 1. Comparison of this proposal (Pr.) with four literature studies on machine learning for cultural heritage classification: I—Llamas et al. [13], II—Janković [14], III—Siountri and Anagnostopoulos [15], IV—Lamas et al. [20].

		Pr.	**I** [13]	**II** [14]	**III** [15]	**IV** [20]
Dataset	New image dataset	✓	✓	–	✓	✓
	Religious buildings	✓	✓	✓	–	✓
	Brazilian cultural heritage	✓	–	–	–	–
Classes	Fronton	✓	–	–	–	–
	Church	✓	–	–	–	–
	Window	✓	–	–	✓	–
	Door	✓	–	–	✓	–
	Tower	✓	✓	–	–	–
	Others	–	✓	✓	✓	✓
Experiments	Simulations with deep learning	✓	✓	✓	✓	✓
	Computer vision using smartphone	✓	–	–	–	✓

Initially, it is noteworthy that three selected studies [13,15,20] propose new image datasets for tasks related to cultural heritage classification. Llamas et al. [13] presented the Architectural Heritage Elements Dataset, which includes ten types of architectural elements from heritage buildings. The study by Siountri and Anagnostopoulos [15] introduces a new image database of cultural heritage buildings in Athens, and the work of Lamas et al. [20] presents MonuMAI (Monument with Mathematics and Artificial Intelligence) with the aim of identifying facade styles of monuments. However, only the ImageOP dataset is specifically intended for the classification of elements in religious buildings of Brazilian cultural heritage. This point highlights one of the main contributions of this paper: the development of a dataset specifically aimed at enhancing the application of artificial intelligence in the identification and digitization of Brazilian religious heritage, particularly from the 18th century, a period characterized by the development of Baroque and Rococo architectural styles.

The second criterion addressed pertains to the analysis of the image categories in the datasets. In this regard, it is important to note that the ImageOP dataset includes five classes (fronton, church, window, door, and tower). These labels are significant for identifying the characteristics of the facades of religious buildings in the historic town of Ouro Preto. The

other studies utilize combinations of different classes, which would be insufficient for the classification of the elements present in the architectural styles of the monuments evaluated in this paper. For example, Janković [14] applies machine learning models to classify five other elements of cultural heritage buildings: altars, gargoyles, domes, columns, and vaults.

Finally, it is important to highlight the characteristics of the experiments in this paper compared to the related work. Most studies conduct experiments solely in simulated environments using deep learning. In contrast, the present proposal addresses the application of artificial intelligence models in real-world situations, utilizing computer vision with smartphones for the classification of building components. Therefore, it is worth emphasizing the importance of proposing a robust dataset for practical applications of built heritage detection using mobile devices. This approach allows for the evaluation of the proposed dataset's performance in both simulated and practical real-time tests.

In summary, the following limitations of related studies that are improved in this paper are highlighted:

- Dataset: the literature lacks new datasets of religious buildings, especially with images of Brazilian Baroque and Rococo architecture;
- Classes: in general, previous work does not address the set of common building elements on the facades of Brazilian churches from the 18th century: fronton, window, door, and tower;
- Experiments: the literature also needs new deep learning approaches for practical experiments in real environments using smartphones.

3. ImageOP: Our Contribution

This section presents and describes the new dataset proposed in this paper: ImageOP—The Image Dataset With Religious Buildings in the World Heritage Town of Ouro Preto for Deep Learning Classification. Figure 1 summarizes the methodology used for the development of this new image database.

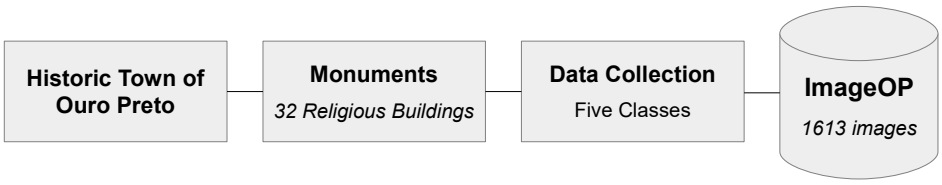


Figure 1. Methodology for the development of the ImageOP dataset.

Initially, the historic town of Ouro Preto was selected for the development of a new dataset featuring images of religious buildings because it was the first World Heritage Site in Brazil to be included on the UNESCO list in 1980. Subsequently, 32 religious monuments located in Ouro Preto were identified. The third stage involved data collection of building components, resulting in five classes for experiments with deep learning: pediment, tower, window, door, and church. Thus, the ImageOP dataset comprises 1613 varied images of religious buildings from the historic town of Ouro Preto. The following subsections provide a more detailed description of the development process and characteristics of the ImageOP dataset.

3.1. Scientific Motivation

The scientific motivation for the development of this study focuses on the improvement of artificial intelligence techniques for the processes of digital documentation [4] and visual inspection of cultural heritage [3]. Digital documentation using intelligent systems can contribute significantly to the automatic management of archives and the interpretation and study of heritage, and can avoid manual classification errors [13]. Following this line, the recognition of building elements in historical monuments is essential in the processes of visual inspection of cultural heritage. The correct detection of key architectural elements

in cultural heritage using AI can contribute to the process of quantifying these components and evaluating conservation [3].

Another point to highlight is that, for the improvement and precision of approaches with artificial intelligence, the availability of data is necessary to carry out effective training of deep learning models. More specifically, it is very important that the available dataset reflects the desired practical application. In this sense, the development of AI systems for digital documentation and visual inspection of cultural heritage with datasets that do not match the architectural reality of the usage environment can lead to serious interpretation errors. However, the literature still lacks datasets aimed at applying deep learning to classify components of Brazilian religious buildings. It is worth emphasizing that the state of Minas Gerais in Brazil has a significant number of religious temples built in the 18th and 19th centuries, with the historic town of Ouro Preto being one of these landmarks in the preservation and conservation of cultural heritage.

In this way, the objective of the proposed methodology is scientifically motivated by the importance of developing a new dataset of images for training deep learning models, and subsequent application in digital documentation and visual inspection of cultural heritage in Brazil, based on images of the historic town of Ouro Preto.

3.2. Historic Town of Ouro Preto

The historic town of Ouro Preto was the first city in Brazil added to the World Heritage List of UNESCO (United Nations Educational, Scientific and Cultural Organization) in 1980 [21,22]. Founded in the early 18th century, the city of Ouro Preto (meaning "Black Gold") was an important center for gold mining and served as the capital of the state of Minas Gerais until 1897. The development of the cities in Minas Gerais was greatly driven by the quest for gold during Brazil's period as a Portuguese colony (from 1500 to 1822) [21,23]. This phenomenon facilitated and financed the construction of significant monuments during this time, particularly in the Baroque architectural style, including squares, public buildings, residences, fountains, bridges, and churches [22]. Figure 2 presents photographs of the historic town of Ouro Preto.

Currently, the town of Ouro Preto preserves several examples of religious and civic buildings in terms of design and materials used in the 18th and 19th centuries [22]. As a result, the Brazilian government submitted the designation of the historic town of Ouro Preto as a World Heritage Site to UNESCO, which was approved on 5 September 1980. On the UNESCO World Heritage Convention page, several relevant criteria are highlighted for including Ouro Preto as a World Heritage Site, such as the aesthetic quality of the architecture represented by the religious monuments and administrative buildings in a remote and rugged landscape. Another criterion emphasizes the heritage constructed under Portuguese colonial rule during the mining period, leading to the construction of churches and chapels characterized by splendor, quality, and originality, blending European and Brazilian cultural traditions [22].

The historical and cultural significance of the Ouro Preto has stimulated scientific production in various fields of knowledge, such as architecture [24], biodiversity [25], microbiology [26], geotourism [27], and mining [23]. Another recent area of study involves the application of technologies in the analysis and documentation of religious monuments in the town of Ouro Preto [28,29]. However, the literature still lacks an image database intended for the classification of elements from the religious monuments of the historic town of Ouro Preto, which is the main motivation for this paper. This new dataset is described in the following subsections.



Figure 2. Historic Town of Ouro Preto. (**a**) Chapel of the Governors Palace and Museum of Inconfidence. (**b**) Church of Saint Francis of Assisi and Pico do Itacolomi. (**c**) Church of Saint Efigenia and historic houses. (**d**) Mountains and historic buildings.

3.3. Religious Buildings

In this paper, 32 Catholic religious buildings from the historic town of Ouro Preto were selected to comprise the proposed dataset. These buildings were chosen for their historical, artistic, and architectural significance. Furthermore, six regions (districts) of the city of Ouro Preto were included: Amarantina, Cachoeira do Campo, Glaura, Santo Antônio do Leite, São Bartolomeu, and the central district (referred to in this study as the Center of Ouro Preto). It is noteworthy that the historic center of Ouro Preto was home to the majority of the analyzed buildings (15 monuments), as this area features much construction in the Baroque and Rococo styles from the 18th century. Table 2 shows the number of photographed buildings by region of the city of Ouro Preto, and Figure 3 illustrates the locations of the visited regions for the development of the ImageOP dataset.

Table 2. Number of religious buildings photographed by region of the historic town of Ouro Preto. Distance: Distance in kilometers from the historic center of Ouro Preto to the other regions.

Regions	Distance (km)	Religious Buldings
Center of Ouro Preto	–	15
Cachoeira do Campo	18	5
St. Antônio do Leite	25	5
Glaura	22	3
Amarantina	25	2
São Bartolomeu	15	2
Total	–	**32**

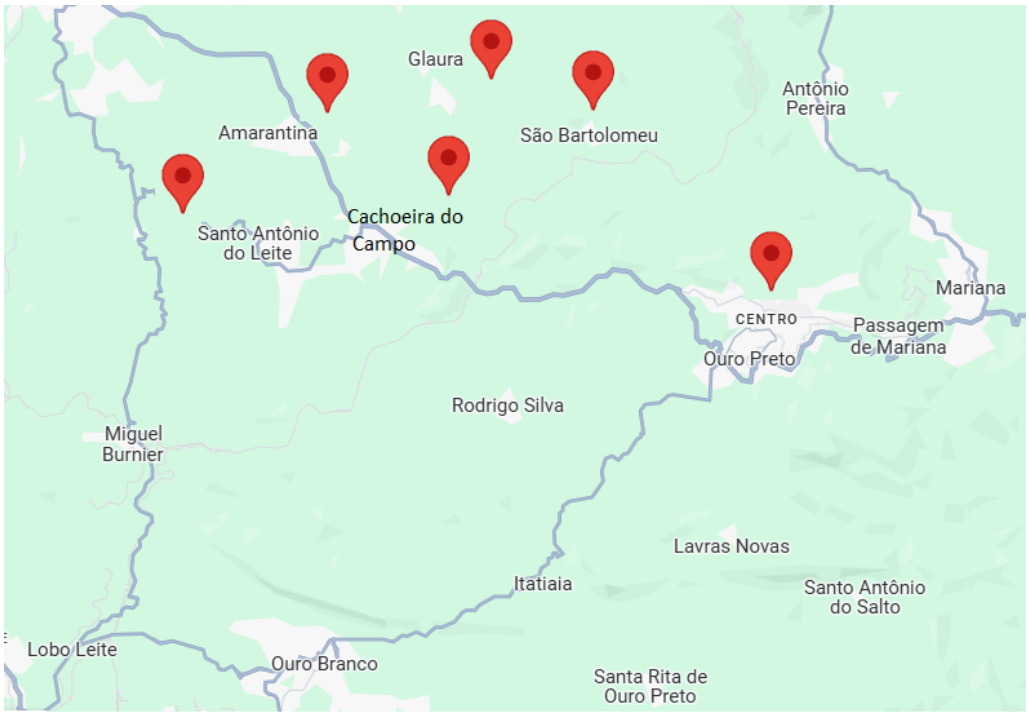


Figure 3. Regions of the historic town of Ouro Preto visited for the development of the ImageOP dataset. Source: modified from Google Maps.

Table 3 presents the complete list of the 32 photographed religious buildings (names in English and Portuguese). Figure 4 (small churches—chapels) and Figure 5 (churches) show examples of images of each of the photographed monuments. More information about these historical buildings can be obtained from the website of the Municipal Government of Ouro Preto[1,2].

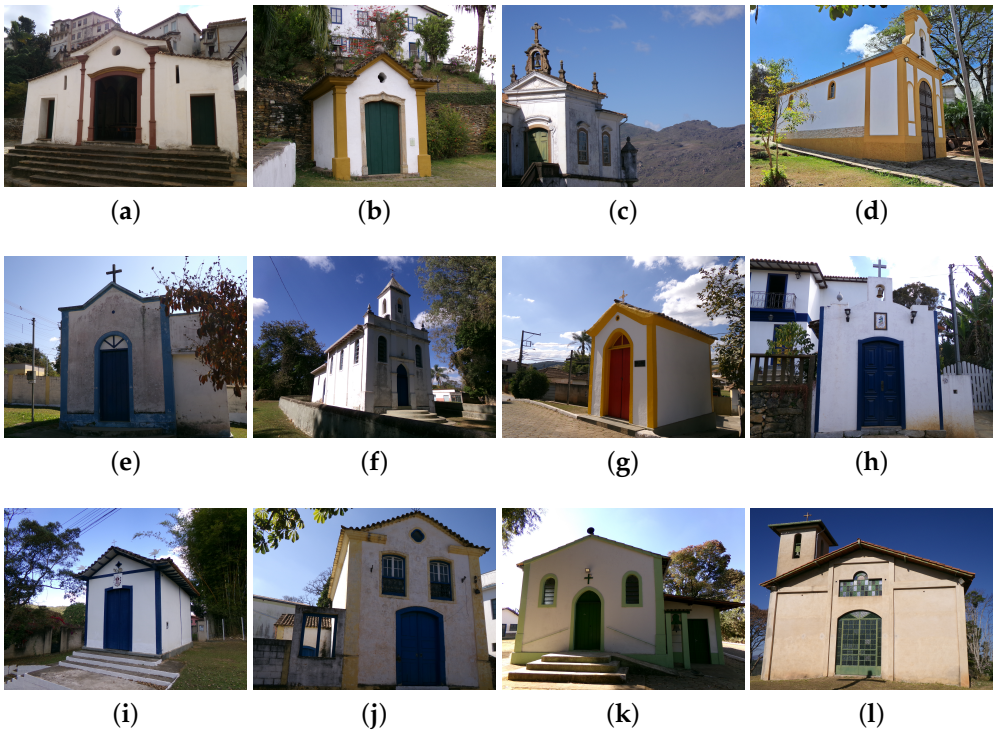


Figure 4. Religious monuments of Ouro Preto (Part I): (**a**) Chapel of Lord of Bonfim; (**b**) Chapel of the Dry Bridge Pass; (**c**) Chapel of the Governors Palace; (**d**) Chapel of Saint Anthony; (**e**) Chapel of the

Saint Kings; (**f**) Chapel of Saint Joseph; (**g**) Chapel of Our Lady of Piety; (**h**) Chapel of Our Lady of Conception; (**i**) Church of Our Lady of Mercy; (**j**) Chapel of Our Lady of Good Dispatch; (**k**) Chapel of Saint Luzia; (**l**) Church of Our Lady of Piety.



Figure 5. Religious monuments of Ouro Preto (Part II): (**a**) Church of Our Lady of Nazareth; (**b**) Church of Our Lady of Sorrows; (**c**) Basilica of Our Lady of Pilar; (**d**) Church of Saint Francis of Assisi; (**e**) Church of Our Lady of Mercy and Pardons; (**f**) Church of Saint Francis of Paula; (**g**) Church of Our Lady of Mount Carmel; (**h**) Sanctuary of Our Lady of Conception; (**i**) Church of Saint Anthony of Leite; (**j**) Church of Saint Anthony of Casa Branca; (**k**) Church of Good Jesus of Matosinhos and Saint Michael and Souls; (**l**) Church of Saint Efigenia; (**m**) Church of Our Lady of Mercy and Compassion; (**n**) Church of Saint Gonçalo; (**o**) Church of Our Lady of the Rosary; (**p**) Church of Saint Bartholomew; (**q**) Church of Our Lady of Mercy (Cachoeira do Campo); (**r**) Church of Our Lady of Sorrows of Mount Calvary; (**s**) Church of Saint Joseph; (**t**) Church of Our Lady of Mercy (São Bartolomeu).

Table 3. List of 32 religious buildings (names of monuments in English and Portuguese) photographed in the historic town of Ouro Preto for the development of the ImageOP dataset.

#	Religious Building	Region
1	Church of Our Lady of Nazareth	Cachoeira do Campo
2	Church of Our Lady of Sorrows	Cachoeira do Campo
3	Basilica of Our Lady of Pilar	Center of Ouro Preto
4	Church of Saint Francis of Assisi	Center of Ouro Preto
5	Church of Our Lady of Mercy and Pardons	Center of Ouro Preto
6	Church of Saint Francis of Paula	Center of Ouro Preto
7	Church of Our Lady of Mount Carmel	Center of Ouro Preto
8	Sanctuary of Our Lady of Conception	Center of Ouro Preto
9	Church of Saint Anthony of Leite	St. Antônio do Leite
10	Church of Saint Anthony of Casa Branca	Glaura
11	Church of G. Jesus of Matosinhos, St. Michael and Souls	Center of Ouro Preto
12	Church of Saint Efigenia	Center of Ouro Preto
13	Church of Our Lady of Mercy and Compassion	Center of Ouro Preto
14	Church of Saint Gonçalo	Amarantina
15	Church of Our Lady of the Rosary	Center of Ouro Preto
16	Church of Saint Bartholomew	São Bartolomeu
17	Church of Our Lady of Mercy	Cachoeira do Campo
18	Church of Our Lady of Sorrows of Mount Calvary	Center of Ouro Preto
19	Church of Saint Joseph	Center of Ouro Preto
20	Church of Our Lady of Mercy	São Bartolomeu
21	Chapel of Lord of Bonfim	Center of Ouro Preto
22	Chapel of the Dry Bridge Pass	Center of Ouro Preto
23	Chapel of the Governors Palace	Center of Ouro Preto
24	Chapel of Saint Anthony	Cachoeira do Campo
25	Chapel of the Saint Kings	St. Antônio do Leite
26	Chapel of Saint Joseph	St. Antônio do Leite
27	Chapel of Our Lady of Piety	Amarantina
28	Chapel of Our Lady of Conception	Glaura
29	Chapel of Our Lady of Mercy	Glaura
30	Chapel of Our Lady of Good Dispatch	Cachoeira do Campo
31	Chapel of Saint Luzia	St. Antônio do Leite
32	Chapel of Our Lady of Piety	St. Antônio do Leite

3.4. Data Collection of Building Components

In this paper, the data collection process for the development of the new ImageOP dataset was carried out in the following stages: (1) definition of the camera; (2) definition of the building components to be photographed; (3) photography conducted in the historic town of Ouro Preto; (4) data annotation. Figure 6 shows examples of the authors' engagement in the image collection process in the town of Ouro Preto.

Initially, the Kodak® PIXPRO AZ255 (Eastman Kodak Company, Rochester, New York, USA) digital camera was chosen for the image collection process. This equipment features a 25× optical zoom, which is suitable for capturing close-up photographs of the elements of historical buildings. Other characteristics of this digital camera model include a 16 MP CMOS sensor, optical image stabilization, 1080p full HD video, a 24 mm wide-angle lens, and a 3″ LCD (460K pixels). Figure 7 displays images of the Kodak® PIXPRO used.



(**a**) (**b**)

Figure 6. Image collection process in the historic town of Ouro Preto. (**a**) Data collection of small church. (**b**) Data collection of church.



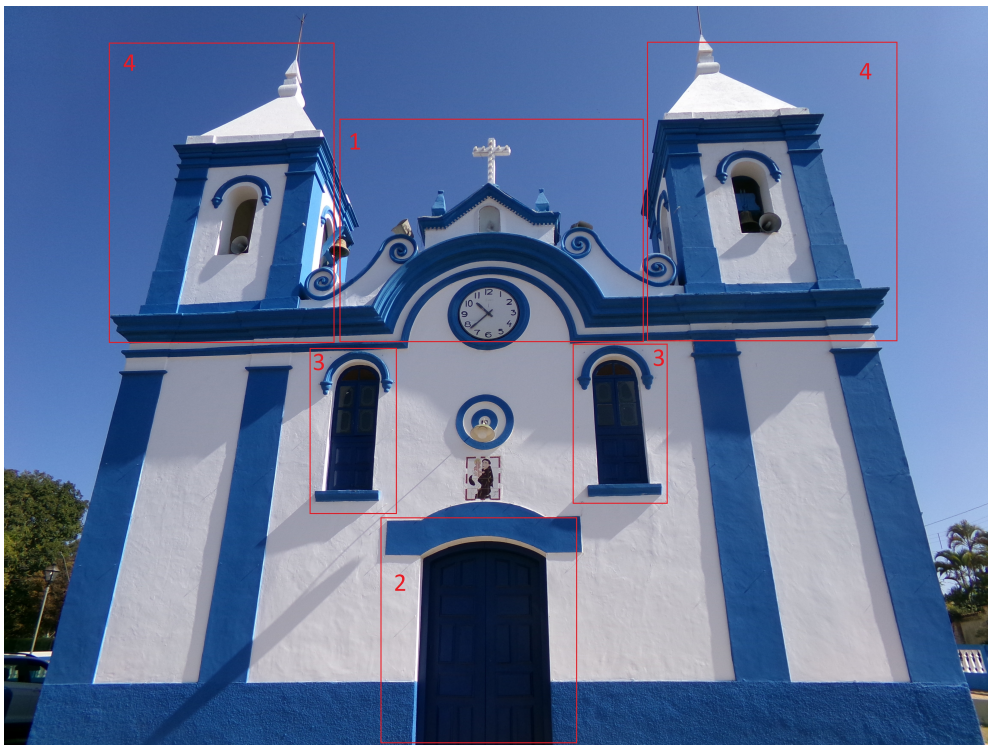
(**a**) (**b**)

Figure 7. Kodak® PIXPRO AZ255 digital camera. (**a**) Front view of the camera. (**b**) Camera display.

Next, the facade elements of the historical buildings were defined for data collection. In this regard, the characteristics of the religious constructions present in the historic town of Ouro Preto were observed. From the images in Figure 8, two common typologies of buildings can be identified: church (Figure 8a) and chapel (Figure 8b). Generally, the larger religious buildings in Ouro Preto, as exemplified in Figure 8a, are characterized by having a fronton with a cross (central upper part), doors (front and side), windows (front and side), and towers. In contrast, the smaller monuments, such as the one shown in Figure 8b, typically feature a central door and a fronton with a cross. Thus, the following building components were defined for photography: frontons, doors, windows, and towers. Additionally, images of complete religious monuments (containing more than one element in the photograph) were also collected, resulting in a total of five classes for the ImageOP dataset.

Subsequently, the image collection process was conducted in the historic town of Ouro Preto. To this end, the 5 regions (described in Table 2) and the 32 religious monuments (listed in Table 3) were visited. For each of the analyzed buildings, photographs of the facades were taken to gather data for each of the defined building components. In this context, the optical zoom feature of up to 25× was utilized to get close to the desired element of the building for photography. Finally, after recording images of the 32 religious buildings, the data annotation phase was conducted. For this purpose, each photograph

was analyzed and organized into digital folders according to the building component present in the image. The next subsection provides a summary of the resulting dataset.



(**a**)



(**b**)

Figure 8. (**a**) Building components in a church: (1) fronton; (2) door; (3) window; (4) tower. (**b**) Building components in a small church (chapel): (1) fronton; (2) door.

3.5. Dataset Result

In the data collection process, 1613 images of the analyzed religious buildings were recorded, as shown in Table 4. Furthermore, the final dataset obtained in this paper contains annotated images in five classes:

1. Fronton (Pediment): the upper part of religious buildings, usually accompanied by a cross;
2. Church: class containing photographs of religious monuments (normal and small size) featuring multiple building components in the same image;
3. Window: images of windows (front or side) of religious buildings;
4. Door: images of doors (front or side) of religious buildings;
5. Tower: photographs of towers of religious buildings, typically featuring a bell.

Table 4. Number of images per classes from the ImageOP dataset.

Class	Number of Images
Fronton	290
Church	353
Window	341
Door	292
Tower	337
Total	**1613**

Figures 9–13 show example images available in the ImageOP dataset for each of the five classes (fronton, church, window, door, and tower). The ImageOP dataset, containing all 1613 images, is publicly available on Mendeley Data repository[3].

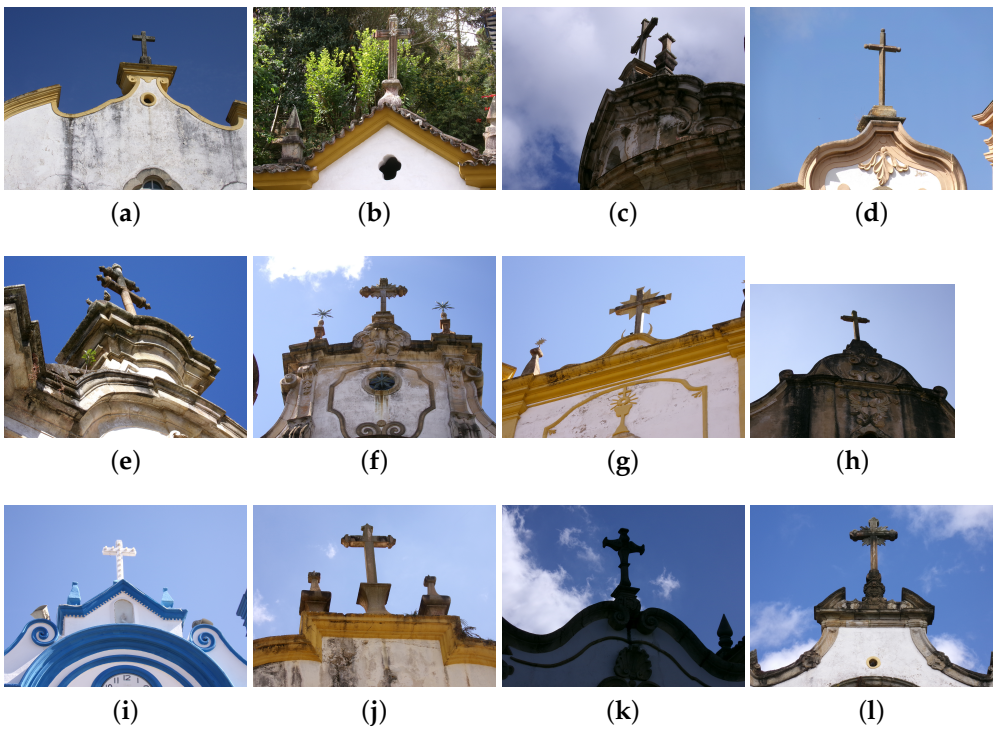


Figure 9. Examples of images from the fronton class. (**a**–**l**) Photographs of church pediments in the Historic Town of Ouro Preto.

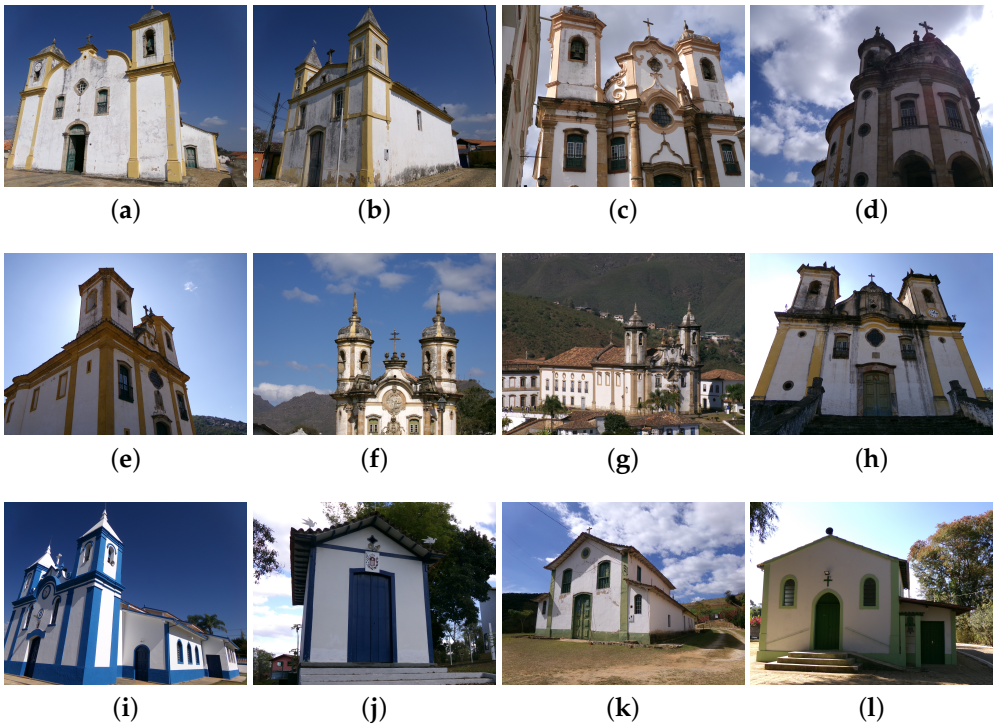


Figure 10. Examples of images from the church class. (**a**–**l**) Photographs of churches in the Historic Town of Ouro Preto.

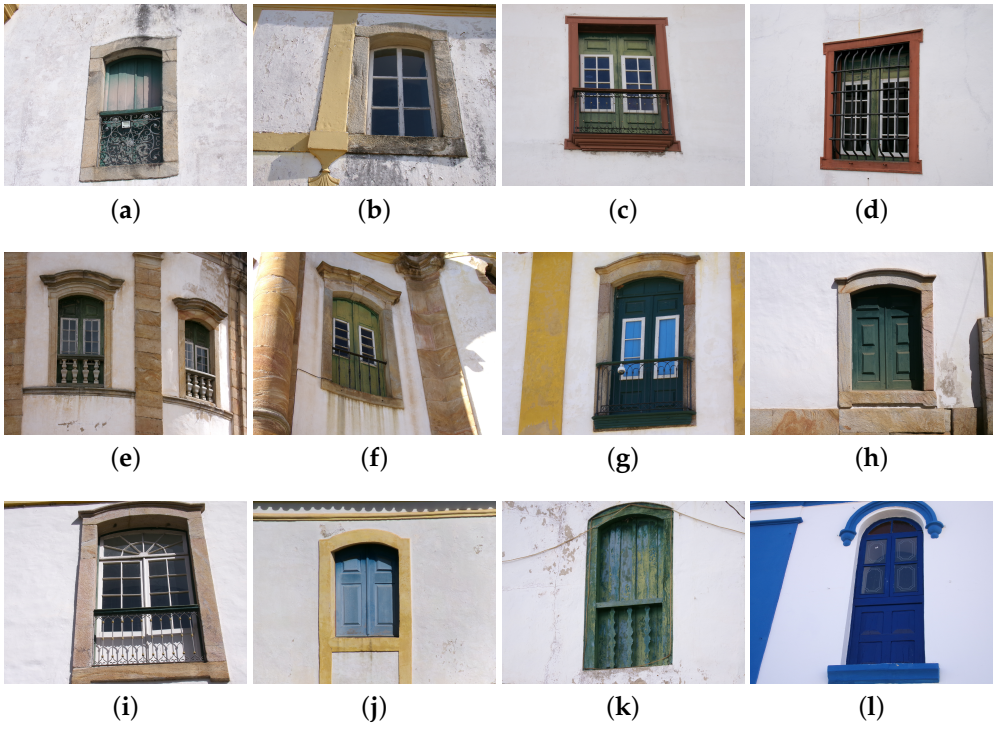


Figure 11. Examples of images from the window class. (**a**–**l**) Photographs of church windows in the Historic Town of Ouro Preto.

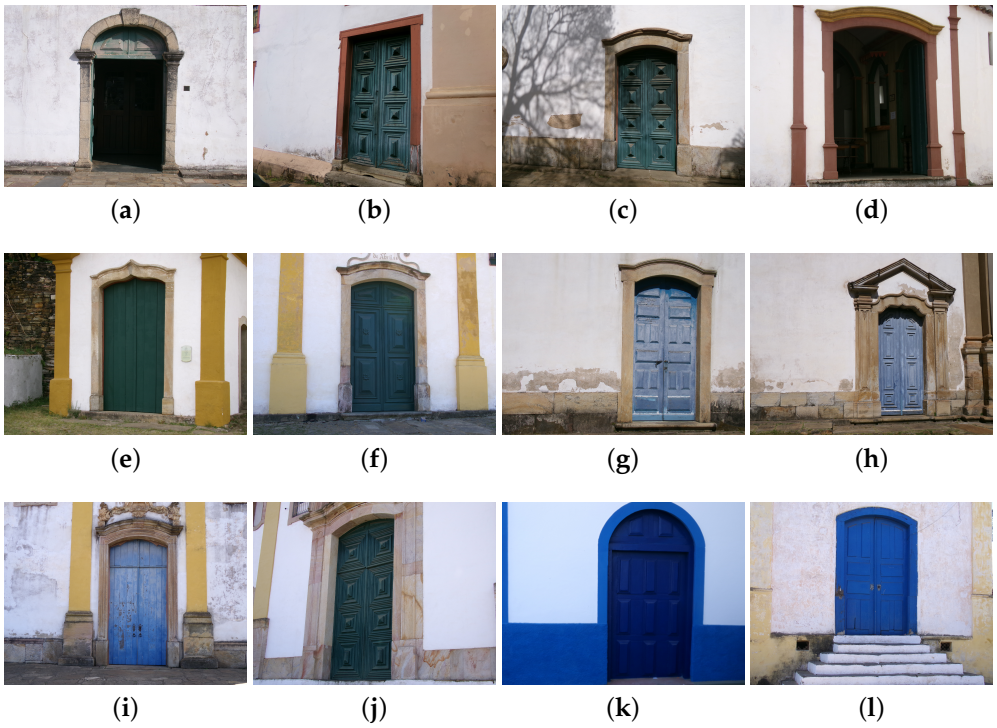


Figure 12. Examples of images from the door class. (**a**–**l**) Photographs of church doors in the Historic Town of Ouro Preto.

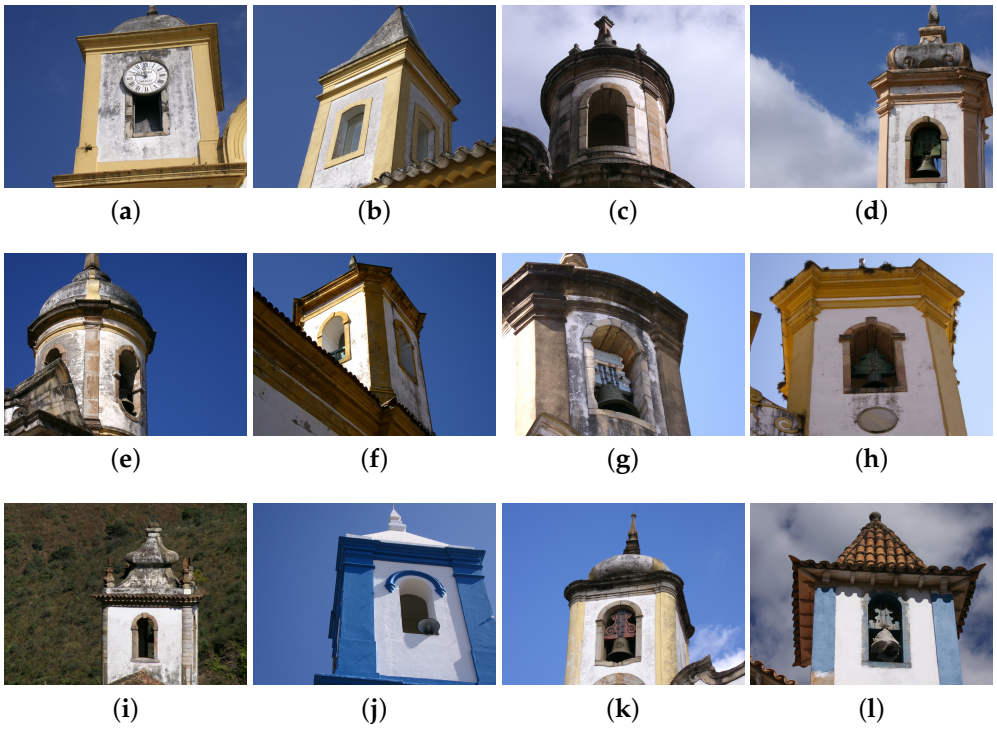


Figure 13. Examples of images from the tower class. (**a**–**l**) Photographs of church towers in the Historic Town of Ouro Preto.

4. Dataset Benchmarking for Deep Learning Classification

In this section, a benchmarking dataset for deep learning classification using ImageOP is proposed. Figure 14 illustrates the proposed method for the simulations and experiments conducted in this stage of the paper.

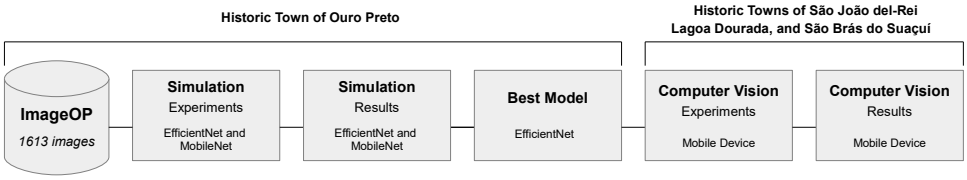


Figure 14. Method of dataset benchmarking for deep learning classification.

The benchmarking dataset was divided into two main phases: (i) simulations for training and testing deep learning models on the Edge Impulse platform using data from ImageOP (images of the historic town of Ouro Preto, as described in Section 3); (ii) application of the best model in a practical computer vision application with a mobile device for classifying building components of monuments from three other historic cities in the state of Minas Gerais, Brazil (São João del-Rei, Lagoa Dourada, and São Brás do Suaçuí). These phases are described in the following subsections.

4.1. Simulation Experiments

In this phase, simulated experiments were conducted for training, validation, and testing of deep learning classification using the ImageOP dataset. The Edge Impulse platform[4] was selected for this purpose. Edge Impulse is an online environment designed for training machine learning models, particularly for embedded systems or mobile devices. Additionally, the Edge Impulse platform features an interface for using deep learning models in computer vision applications on smartphones, which is one of the applications utilized in this paper. Subsequently, the ImageOP dataset was uploaded to the Edge Impulse software (Community plan version), and the 1613 images were divided into training (≈60%), validation (≈20%), and testing (≈20%) sets, as shown in Table 5.

Table 5. Division of dataset into training, validation, and testing for simulated experiments.

Class	Train (≈60%)	Validation (≈20%)	Test (≈20%)	Total (100%)
Fronton	176	59	55	290
Church	223	61	69	353
Window	232	44	65	341
Door	191	44	57	292
Tower	219	53	65	337
Total	1041	261	311	**1613**

Subsequently, two convolutional neural network architectures available on Edge Impulse were selected: MobileNet V2 [16] and EfficientNet B0 [17]. These architectures are extensively discussed in the literature [30–34], particularly for their performance in applications that require low computational cost, such as those using mobile devices. The hyperparameter settings were maintained according to the default configurations of the Edge Impulse software: number of training cycles = 20, learning rate = 0.0005, and batch size = 32.

The four performance metrics evaluated (accuracy, recall, precision, and F-score) are presented in Equations (1)–(4):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

$$Recall = \frac{TP}{TP + FN}, \quad (2)$$

$$Precision = \frac{TP}{TP + FP}, \quad (3)$$

$$F - score = \frac{2 \times Precision \times Recall}{Precision + Recall}, \quad (4)$$

where TP denotes true positives, TN denotes true negatives, FP denotes false positives, and FN denotes false negatives.

4.2. Simulation Results

This section presents the results of classifying historical monuments in Ouro Preto using convolutional neural networks (CNN). Two architectures were compared: MobileNet and EfficientNet. The evaluation was conducted using the dataset developed by this work, ImageOP, and the results were analyzed in terms of accuracy and F-score. Detailed comparisons between the two architectures are presented to identify which offers superior performance in the classification task.

Tables 6 and 7 present a summary of the results for the two analyzed neural architectures. It can be observed that the EfficientNet architecture achieved the best performance across all four evaluated metrics. Notably, the F-score attained by the EfficientNet model (96.0%) was 3% higher than the results from the MobileNet architecture (93.0%). Additionally, EfficientNet achieved the highest accuracy for the four evaluated classes: fronton (96.4%), church (97.1%), window (89.2%), and tower (95.4%).

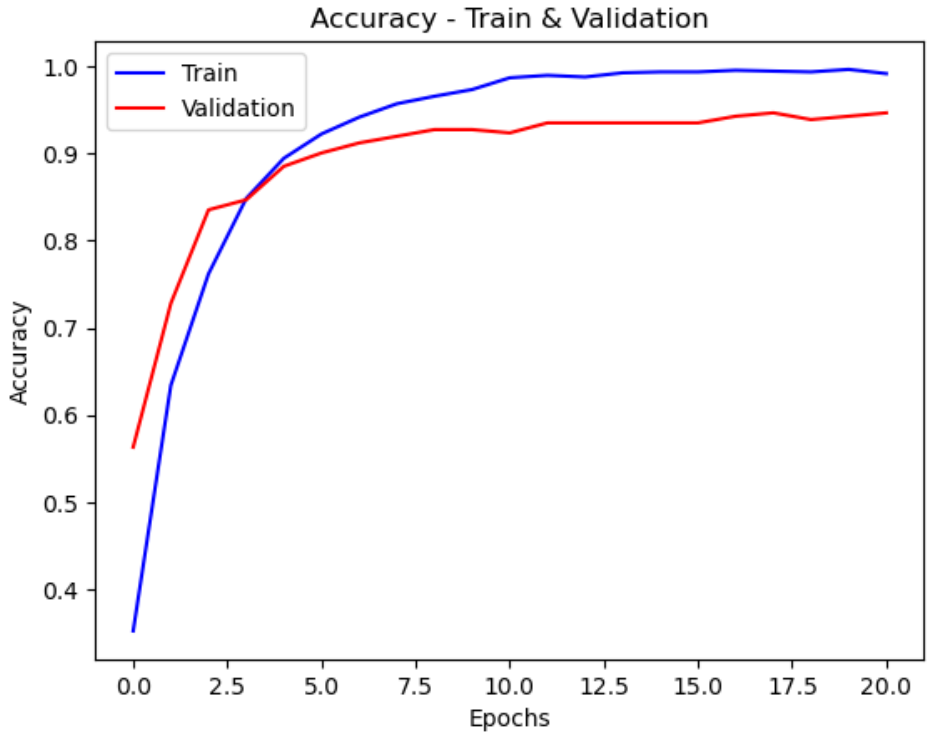
Table 6. Results for simulation experiments (%) using the MobileNet and EfficientNet architectures. Bold value indicates the best results.

Architecture	Accuracy	Precision	Recall	F-Score
MobileNet	92.6	94.0	94.0	93.0
EfficientNet	**94.5**	**96.0**	**96.0**	**96.0**

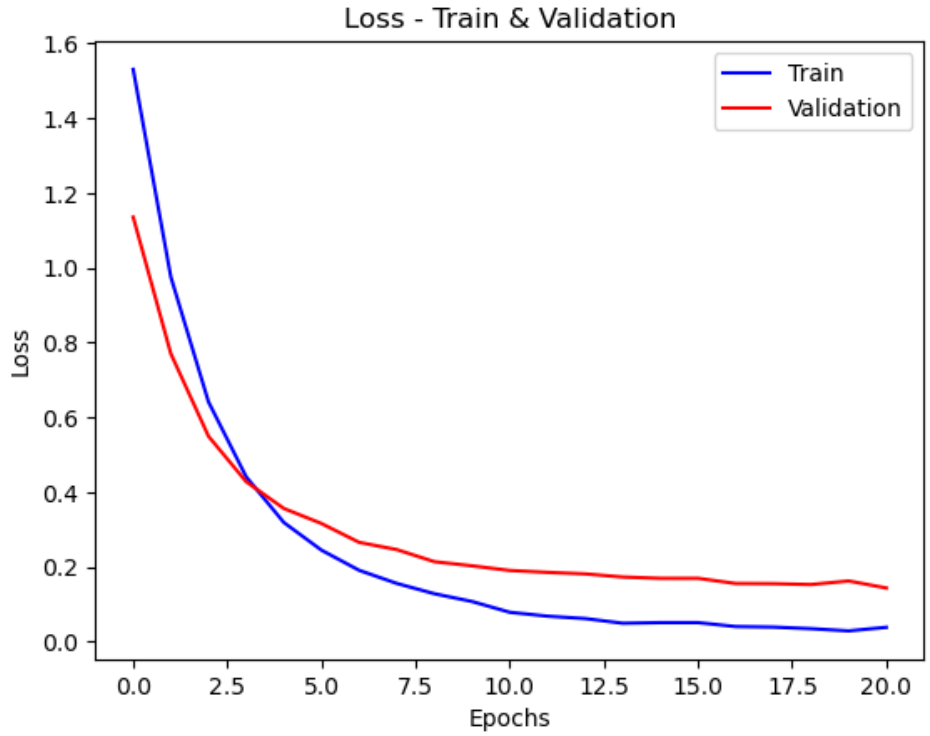
Table 7. Accuracy results (%) for the five classes in the simulation experiments. Bold value indicates the best results.

Architecture	Fronton	Church	Window	Door	Tower
MobileNet	92.7	**97.1**	81.5	**98.2**	93.8
EfficientNet	**96.4**	**97.1**	**89.2**	94.7	**95.4**

For the MobileNet architecture, Figure 15 displays the results from the training and testing history for (a) accuracy and (b) loss. Figure 16 presents the results in the confusion matrix for the experiments using the MobileNet architecture. As observed, the architecture achieved good performance, successfully generalizing the ImageOP dataset.



(**a**)



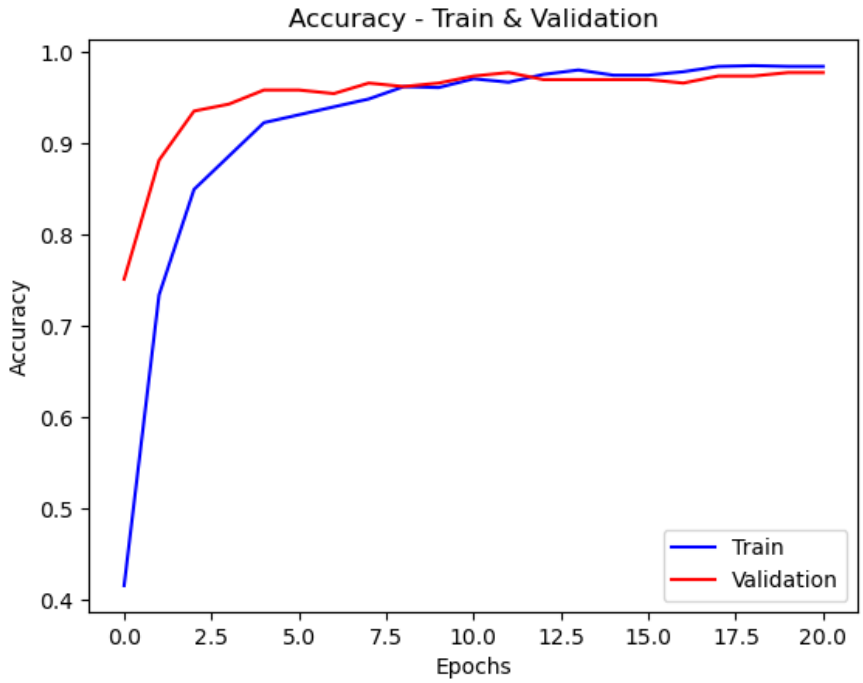
(**b**)

Figure 15. Graph of the train and validation history for the MobileNet architecture: (**a**) accuracy; (**b**) loss.

	Fronton	Church	Window	Door	Tower	Uncertain
Fronton	92.7%	3.6%	0%	0%	3.6%	0%
Church	1.4%	97.1%	0%	0%	0%	1.4%
Window	1.5%	1.5%	81.5%	7.7%	1.5%	6.2%
Door	0%	0%	0%	98.2%	0%	1.8%
Tower	3.1%	0%	1.5%	0%	93.8%	1.5%
F-score	93.0%	96.0%	89.0%	95.0%	95.0%	-

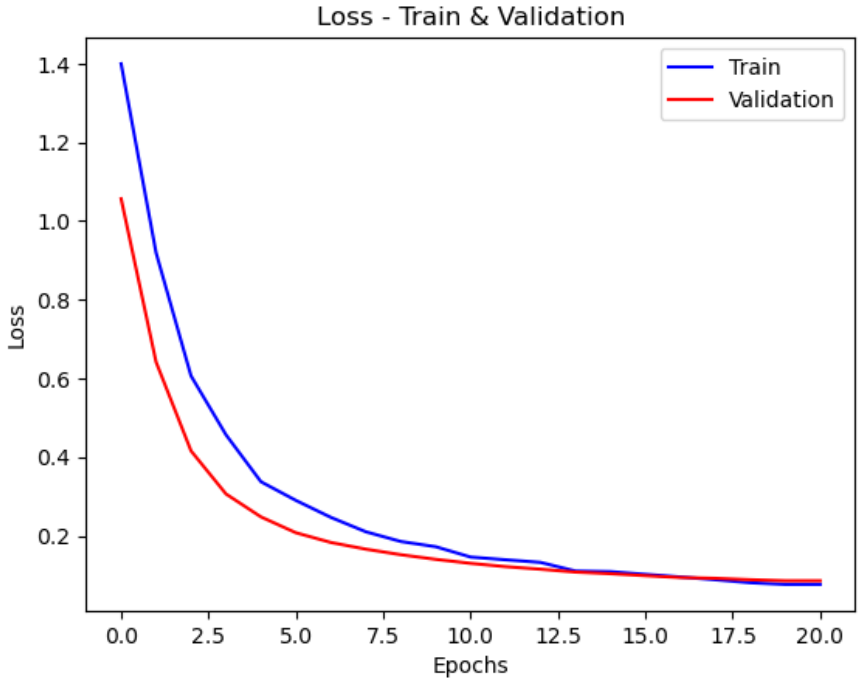
Figure 16. Confusion matrix for the MobileNet architecture.

Similarly, for the EfficientNet architecture, Figure 17 presents the training and testing history results for (a) accuracy and (b) loss. EfficientNet also performed admirably, showing consistent improvement in both metrics throughout the training process. Figure 18 shows the confusion matrix for the EfficientNet model. The EfficientNet architecture achieved an impressive F-score of about 92% across all classes, slightly outperforming MobileNet. This indicates EfficientNet's superior capability in accurately classifying the historical monuments in the ImageOP dataset.



(**a**)

Figure 17. *Cont.*



(**b**)

Figure 17. Graph of the train and validation history for the EfficientNet architecture: (**a**) accuracy; (**b**) loss.

	Fronton	Church	Window	Door	Tower	Uncertain
Fronton	96.4%	1.8%	0%	0%	0%	1.8%
Church	0%	97.1%	0%	1.4%	1.4%	0%
Window	0%	0%	89.2%	4.6%	1.5%	4.6%
Door	0%	0%	1.8%	94.7%	0%	3.5%
Tower	0%	0%	1.5%	0%	95.4%	3.1%
F-score	98.0%	98.0%	93.0%	94.0%	96.0%	-

Figure 18. Confusion matrix for the EfficientNet architecture.

Table 8 also presents examples of inference results for ten test images. It can be observed that the model using EfficientNet achieved a performance of 100% in correctly identifying the category of the building component in the photographs. Thus, the deep learning model trained with EfficientNet was selected for the practical experiments with a mobile device, as described in the next subsection.

Table 8. Examples of inference results (accuracy in %) for some images in the testing phase using the EfficientNet model. Bold value indicates the best results.

Class	Sample	Fronton	Church	Window	Door	Tower
Fronton		**100.0**	0.0	0.0	0.0	0.0
Fronton		**100.0**	0.0	0.0	0.0	0.0
Church		0.0	**95.0**	2.5	0.0	2.5
Church		0.0	**87.0**	0.0	4.0	9.0
Window		0.0	0.0	**100.0**	0.0	0.0
Window		0.0	0.0	**81.0**	3.0	16.0
Door		0.0	0.0	4.0	**96.0**	0.0
Door		0.0	0.0	0.0	**100.0**	0.0
Tower		2.0	0.0	2.0	0.0	**96.0**
Tower		1.0	2.0	0.0	0.0	**97.0**

4.3. Computer Vision Using Mobile Device

In this phase, real-world experiments were conducted to validate the proposed dataset for deep learning classification. The best model (with EfficientNet architecture) obtained during the training and testing phase was utilized. Additionally, a mobile device (Samsung A14 smartphone) was employed for real-time detection of building elements using computer vision. Access to the trained model on the Edge Impulse software was achieved through the smartphone. It is noteworthy that in this phase, nine religious buildings from three historical cities in the state of Minas Gerais, Brazil (São João del-Rei, Lagoa Dourada, and São Brás do Suaçuí), were selected. These cities feature historical monuments with architecture and construction periods (18th and 19th centuries) similar to the churches in Ouro Preto. The objective was to evaluate the performance of the proposed dataset and trained model in real-world environments not included in the ImageOP dataset images.

Figure 19 illustrates all the steps of the proposed process to apply computer vision using a smartphone to recognize the elements of religious buildings. As described in the previous sections, photographs were initially collected in the historic town of Ouro Preto for the development of the ImageOP dataset with 1613 images. Subsequently, the deep learning models were trained on the Edge Impulse platform. Then, the Edge Impulse software makes the application of the artificial intelligence system available on mobile devices by reading a QR code. On the smartphone, the application opens in a web browser. Thus, it is possible to direct the device's camera towards the church where you want to detect the building's components. Real-time recognition results are shown on the smartphone interface.

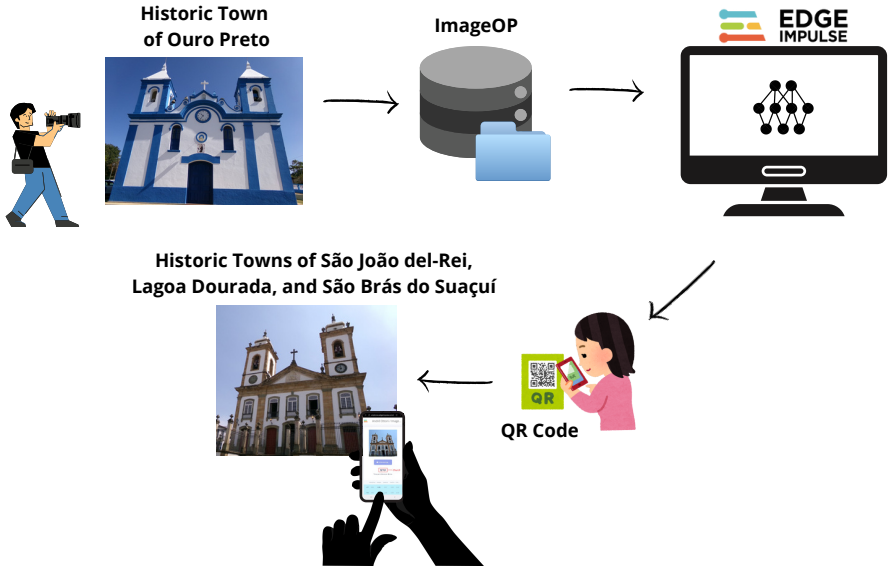


Figure 19. Proposed procedure for applying computer vision using smartphones to recognize elements of religious buildings.

Figures 20 and 21 presents examples of the computer vision process using a mobile device in the historic town of São João del-Rei. It can be observed (Figure 21) that when the smartphone camera is pointed at the element of the religious building, the inference is displayed on the Edge Impulse web interface. In this case, the classification was performed correctly, identifying the element as window (*janela* in Portuguese).



(**a**) (**b**)

Figure 20. Process of computer vision using mobile device in the historic town of São João del-Rei. (**a**) Church door detection. (**b**) Church Detection.



Figure 21. Example of deep learning classification using computer vision with a mobile device and Edge Impulse software. Detected class: window (*janela* in Portuguese).

Figure 22 displays examples of real-time classification through screenshots of the Edge Impulse graphical interface accessed on the mobile device. Five examples of correct inferences for the five classes discussed in this paper can be observed.

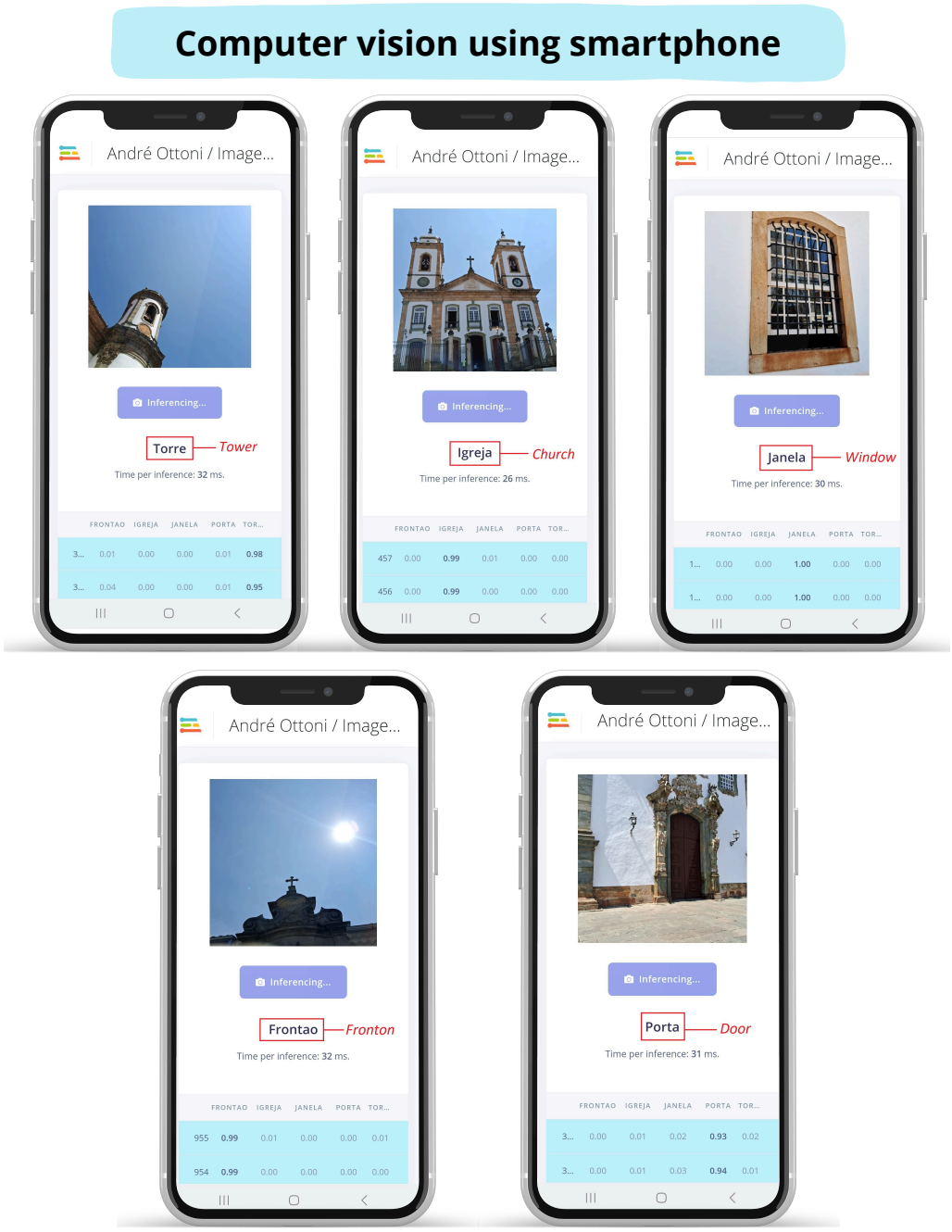


Figure 22. Examples of real-time classification from screenshots of the Edge Impulse graphical interface accessed on the mobile device.

Table 9 presents the real-time classification results for each of the nine religious monuments evaluated during the practical experiments using computer vision and a mobile device. It can be observed that an accuracy of up to 100% was achieved for the detection of building components in the case of the Church of Saint Francis of Assisi (São João del-Rei). Furthermore, the accuracy values in this phase were at least 75.0% for all the religious buildings assessed. Finally, it is noteworthy that the average accuracy reached 88.0%, indicating the effectiveness of the proposed dataset and models in classifying elements of religious buildings using computer vision.

Table 9. Accuracy results (Acc in %) for real-time classification of each religious monument evaluated during the practical experiments using computer vision and a mobile device.

Monument	Name	Historic Town	Acc
	Church of Saint Francis of Assisi	São João del-Rei	100.0
	Church of Our Lady of Mount Carmel	São João del-Rei	88.9
	Cathedral Basilica of Our Lady of Pilar	São João del-Rei	93.8
	Church of Our Lady of the Rosary	São João del-Rei	92.3
	Chapel of Divine Holy Spirit	São João del-Rei	75.0
	Chapel of Saint Anthony	São João del-Rei	80.0
	Church of Our Lady of Mercy	São João del-Rei	85.7
	Church of Saint Anthony	Lagoa Dourada	92.9
	Church of Saint Bras	São Brás do Suaçuí	83.3
		Avg. Accuracy	**88.0**

5. Conclusions

The main contribution of this work is the proposal of ImageOP: The Image Dataset With Religious Buildings in the World Heritage Town of Ouro Preto for Deep Learning Classification. This dataset comprises 1613 images of religious buildings from the historic town of Ouro Preto (State of Minas Gerais, Brazil). Ouro Preto was the first Brazilian World Heritage site recognized by UNESCO and houses numerous well-preserved historical monuments from the 18th and 19th centuries. Thus, this paper contributes to the field of heritage digitization through images for the application of artificial intelligence in recognizing five classes: fronton, door, window, tower, and church. Additionally, this study also contributes with a new methodology for visual inspection using deep learning for recognition of architectural elements in religious buildings. For this purpose, experiments are proposed in simulation and practical inspection stages using computer vision and mobile devices for the detection of components in cultural heritage.

The proposed experiments evaluated two traditional convolutional neural network architectures from the literature: MobiletNet V2 and EfficientNet B0. The results from the simulations showed that the model trained with the EfficientNet architecture achieved the best performance, with accuracy = 94.5%, precision = 96.0%, recall = 96.0%, and F-score = 96.0%. Furthermore, this model achieved an accuracy greater than 89% in detecting the five classes: fronton (96.4%), church (97.1%), window (89.2%), door (94.7%), and tower (95.4%). In general, the simulated results showed that the MobileNet architecture performed worse than EfficientNet, however, with metrics also superior to 90%: accuracy = 92.6%, precision = 94.0%, recall = 94.0%, F-score = 93.0%, fronton (92.7%), church (97.1%), window (81.5%), door (98.2%), and tower (93.8%). Thus, simulated results show the performance capacity of the proposed method to provide detection of components in religious buildings using artificial intelligence.

The results from real-world experiments with the EfficientNet architecture using computer vision and smartphones also confirmed the effectiveness of ImageOP and the deep learning models in classifying building components. In this phase, accuracies exceeding 75.0% were obtained for detection of elements in nine religious monuments. It is noteworthy that the buildings evaluated in this final phase are from three other historic cities in the state of Minas Gerais, Brazil, meaning they do not have data stored in ImageOP, thereby reinforcing the dataset's potential for generalization and inference of building components.

In future work, it is anticipated that the proposed dataset will be utilized in various applications. The adoption of ImageOP is suggested for the development of an automatic categorization and digitization system for Brazil's historic religious heritage. Additionally, the development of a computer vision system for detecting pathologies in the elements (fronton, window, door, and tower) of historical buildings is expected. Other technological devices using the ImageOP dataset will also be proposed, such as a mobile robot prototype that recognizes and describes religious buildings for tourists with visual impairments. Furthermore, it is expected to improve deep learning architectures for recognizing historical religious monuments. To this end, we intend to apply methods for hyperparameter tuning to increase the accuracy of convolutional neural networks, and, thus, enhance the performance of structures aimed at computer vision tasks on mobile devices such as MobileNet and EfficientNet. Furthermore, the data collection process can be improved in the development of the next datasets with the application of unmanned aerial vehicles to take photographs and visual inspection of the religious building roofs.

Author Contributions: Methodology, A.L.C.O. and L.T.C.O.; Software, A.L.C.O.; Writing—original draft, A.L.C.O.; Writing—review and editing, A.L.C.O. and L.T.C.O. All authors have read and agreed to the published version of the manuscript.

Funding: The author André Luiz Carvalho Ottoni is grateful for the funding received under Edital PROPPI 18/2024—UFOP.

Data Availability Statement: The dataset proposed during the current study is available in the Mendelay Data repository, https://doi.org/10.17632/hdz96zts8j.1 (accessed on 19 September 2024).

Acknowledgments: The authors are grateful to UFOP (Edital PROPPI 18/2024) and IFMG.

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

Notes

1 https://www.ouropreto.mg.gov.br/turismo/atrativo/113 (accessed on 24 October 2024).
2 https://www.ouropreto.mg.gov.br/turismo/atrativo/139 (accessed on 24 October 2024).
3 https://doi.org/10.17632/hdz96zts8j.1.
4 https://edgeimpulse.com/.

References

1. Piccialli, F.; Casolla, G.; Cuomo, S.; Giampaolo, F.; Prezioso, E.; Di Cola, V.S. Unsupervised learning on multimedia data: A Cultural Heritage case study. *Multimed. Tools Appl.* **2020**, *79*, 34429–34442. [CrossRef]

2. Spennemann, D.H. Will Artificial Intelligence Affect How Cultural Heritage Will Be Managed in the Future? Responses Generated by Four genAI Models. *Heritage* **2024**, *7*, 1453–1471. [CrossRef]
3. Mishra, M.; Lourenço, P.B. Artificial intelligence-assisted visual inspection for cultural heritage: State-of-the-art review. *J. Cult. Herit.* **2024**, *66*, 536–550. [CrossRef]
4. Silva, C.; Oliveira, L. Artificial Intelligence at the Interface between Cultural Heritage and Photography: A Systematic Literature Review. *Heritage* **2024**, *7*, 3799–3820. [CrossRef]
5. Karimi, N.; Mishra, M.; Lourenço, P.B. Deep learning-based automated tile defect detection system for Portuguese cultural heritage buildings. *J. Cult. Herit.* **2024**, *68*, 86–98. [CrossRef]
6. Li, Y.; Zhao, M.; Mao, J.; Chen, Y.; Zheng, L.; Yan, L. Detection and recognition of Chinese porcelain inlay images of traditional Lingnan architectural decoration based on YOLOv4 technology. *Herit. Sci.* **2024**, *12*, 137. [CrossRef]
7. Maria, M.D.; Fiumi, L.; Mazzei, M.; V, B.O. A system for monitoring the environment of historic places using convolutional neural network methodologies. *Heritage* **2021**, *4*, 1429–1446. [CrossRef]
8. Liu, Z.; Brigham, R.; Long, E.R.; Wilson, L.; Frost, A.; Orr, S.A.; Grau-Bové, J. Semantic segmentation and photogrammetry of crowdsourced images to monitor historic facades. *Herit. Sci.* **2022**, *10*, 1–17. [CrossRef]
9. Yazdi, H.; Sad Berenji, S.; Ludwig, F.; Moazen, S. Deep learning in historical architecture remote sensing: Automated historical courtyard house recognition in Yazd, Iran. *Heritage* **2022**, *5*, 3066–3080. [CrossRef]
10. Samhouri, M.; Al-Arabiat, L.; Al-Atrash, F. Prediction and measurement of damage to architectural heritages facades using convolutional neural networks. *Neural Comput. Appl.* **2022**, *34*, 18125–18141. [CrossRef]
11. Kumar, P.; Ofli, F.; Imran, M.; Castillo, C. Detection of disaster-affected cultural heritage sites from social media images using deep learning techniques. *J. Comput. Cult. Herit. (JOCCH)* **2020**, *13*, 1–31. [CrossRef]
12. Zou, Z.; Zhao, X.; Zhao, P.; Qi, F.; Wang, N. CNN-based statistics and location estimation of missing components in routine inspection of historic buildings. *J. Cult. Herit.* **2019**, *38*, 221–230. [CrossRef]
13. Llamas, J.; Lerones, P.M.; Medina, R.; Zalama, E.; Gómez-García-Bermejo, J. Classification of architectural heritage images using deep learning techniques. *Appl. Sci.* **2017**, *7*, 992. [CrossRef]
14. Janković, R. Machine learning models for cultural heritage image classification: Comparison based on attribute selection. *Information* **2020**, *11*, 12. [CrossRef]
15. Siountri, K.; Anagnostopoulos, C.N. The classification of cultural heritage buildings in athens using deep learning techniques. *Heritage* **2023**, *6*, 3673–3705. [CrossRef]
16. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
17. Tan, M. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv* **2019**, arXiv:1905.11946.
18. Abayomi-Alli, O.O.; Damaševičius, R.; Misra, S.; Abayomi-Alli, A. FruitQ: A new dataset of multiple fruit images for freshness evaluation. *Multimed. Tools Appl.* **2024**, *83*, 11433–11460. [CrossRef]
19. Ali, A.; Sarkar, R.; Das, D.K. IRUVD: A new still-image based dataset for automatic vehicle detection. *Multimed. Tools Appl.* **2024**, *83*, 6755–6781. [CrossRef]
20. Lamas, A.; Tabik, S.; Cruz, P.; Montes, R.; Martínez-Sevilla, Á.; Cruz, T.; Herrera, F. MonuMAI: Dataset, deep learning pipeline and citizen science based app for monumental heritage taxonomy and classification. *Neurocomputing* **2021**, *420*, 266–280. [CrossRef]
21. de Alvarenga Pereira Costa, S.; Barci Castriota, L.; Salgado, M. The World Heritage site of Ouro Preto. *Facilities* **2011**, *29*, 339–351. [CrossRef]
22. UNESCO. *United Nations Educational, Scientific and Cultural Organization—World Heritage List: Historic Town of Ouro Preto*; United Nations Educational, Scientific and Cultural Organization: 2024. Available online: https://whc.unesco.org/en/list/124 (accessed on 19 September 2024).)
23. Barbosa, V.d.S.B.; de Lima, H.M.; Fonseca, B.M. Assessing risks of abandoned urban mines in the UNESCO World Heritage City of Ouro Preto, Brazil. *Appl. Geogr.* **2022**, *139*, 102648. [CrossRef]
24. Valerio, M.A. Architects of their own humanity: Race, devotion, and artistic agency in Afro-Brazilian confraternal churches in eighteenth-century Salvador and Ouro Preto. *Colon. Lat. Am. Rev.* **2021**, *30*, 238–271. [CrossRef]
25. Bhakti, T.; Rossi, F.; de Oliveira Mafia, P.; de Almeida, E.F.; Fujaco, M.A.G.; de Azevedo, C.S. Preservation of historical heritage increases bird biodiversity in urban centers. *Environ. Dev. Sustain.* **2021**, *23*, 8758–8773. [CrossRef]
26. Boniek, D.; de Abreu, C.S.; Dos Santos, A.F.B.; de Resende Stoianoff, M.A. Evaluation of microbiological air parameters and the fungal community involved in the potential risks of biodeterioration in a cultural heritage of humanity, Ouro Preto, Brazil. *Folia Microbiol.* **2021**, *66*, 797–807. [CrossRef]
27. Santos, B.H.d.; Fonseca Filho, R.E.; Castro, P.d.T.A. Potential Urban Geotourist profile of Ouro Preto (MG, Brazil). *Anu. Inst. Geociencias* **2021**, *44*, 1–14.
28. Froner, Y.A.; de Barros Gonçalves, W.; Souza, L.A.C.; Costa, A.G.; Rosado, A.; Mizrahy Cuperschmid, A.R.; de Oliveira Walter, G.; Neves Miranda, A.C.; Michelin, G.; Montalvão, A.C.; et al. Data Collection for Cultural Heritage Risk Management: The Damage Map through Heritage Building Information Modeling (HBIM) Project Applied to the Façade of St Francis of Assisi, Ouro Preto, Brazil. *Stud. Conserv.* **2024**, *69*, 98–107. [CrossRef]

29. Cuperschmid, A.R.M.; Neves de Oliveira, G.; Froner, Y.A. Exploring the Use of LiDAR in Smartphones: Documenting the Frontispiece of Saint Francis of Assisi Church in Ouro Preto, Brazil. *Int. J. Archit. Herit.* **2024**, 1–18. [CrossRef]
30. Ottoni, A.L.C.; Souza, A.M.; Novo, M.S. Automated hyperparameter tuning for crack image classification with deep learning. *Soft Comput.* **2023**, *27*, 18383–18402. [CrossRef]
31. Almujally, N.A.; Alzahrani, A.; Hakeem, A.M.; Attiah, A.; Umer, M.; Alsubai, S.; Polsinelli, M.; Ashraf, I. Selective feature-based ovarian cancer prediction using MobileNet and explainable AI to manage women healthcare. *Multimed. Tools Appl.* **2024**, 1–22. [CrossRef]
32. Liu, L.; Wang, X.; Bao, Q.; Li, X. Behavior detection and evaluation based on multi-frame MobileNet. *Multimed. Tools Appl.* **2024**, *83*, 15733–15750. [CrossRef]
33. Marques, G.; Ferreras, A.; de la Torre-Diez, I. An ensemble-based approach for automated medical diagnosis of malaria using EfficientNet. *Multimed. Tools Appl.* **2022**, *81*, 28061–28078. [CrossRef]
34. Nair, A.; Vadher, H.; Patel, P.; Vyas, T.; Bhatt, C.; Bruno, A. Lung sound disease detection using attention over pre-trained efficientnet architecture. *Multimed. Tools Appl.* **2024**, 1–35. [CrossRef]

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.