buildings

MDPI

Article

Public Database of Cracks Images in Mortar Coating with Different Types of Surface Finishes

Renner de Assis Garcia Sobrinho [1,*], Franklin Piauhy Neto [2] and Henrique Fernandes [2]

[1] Faculty of Mechanical Engineering, Federal University of Uberlândia, Uberlândia 38408-100, Brazil
[2] Faculty of Computing, Federal University of Uberlândia, Uberlândia 38408-100, Brazil
* Correspondence: renner_tt@hotmail.com

Abstract: The use of technology, such as artificial intelligence (AI), in production processes has been optimizing several industrial realities. In civil construction, AI can be used in different applications, one of which is building inspection. One of the difficulties in developing this type of study is the low number of public image databases that represent more general aspects of building wear. In view of this, the main objective of this research was to set up a public database of images of cracks in mortar coating, considering different types of surface finish—smooth type, scrapped type, and rough type. A database was created with 33,088 images that went through a systematic labeling process based on classes defined in the study. Network training was carried out through transfer learning using the VGG16 in different groupings of finishes. It was found that the training accuracy varies according to surface finish and data balancing. The finish of the scrapped type was the one that presented the lowest accuracy. The database presented several types of noise and was unbalanced in all categories defined in the labeling. In this way, it was possible to create a database that represented possible situations to be found in real inspections.

Keywords: artificial intelligence; mortar coating; public database; artificial neural networks; transfer learning; convolutional neural network

check for updates

Citation: Garcia Sobrinho, R.d.A.; Piauhy Neto, F.; Fernandes, H. Public Database of Cracks Images in Mortar Coating with Different Types of Surface Finishes. *Buildings* **2023**, *13*, 1872. https://doi.org/10.3390/buildings13071872

Academic Editor: Ahmed Senouci

Received: 24 May 2023
Revised: 7 July 2023
Accepted: 17 July 2023
Published: 24 July 2023

Copyright: © 2023 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

Technology use in production processes, such as advanced robotic incorporation and systems integration, is modernizing several industrial realities. In view of this scenario, it is considered that the fourth industrial revolution is taking place, a concept called industry 4.0 [1]. Among the innovations used in the development of products and services in industry 4.0, there is artificial intelligence (AI). This technology aims to develop the autonomy of machines and computational systems so that they simulate human reasoning, especially in decision-making. In this way, it is possible to optimize long-scale repetitive processes and activities, which would require human analysis [2,3].

In civil construction, there are already studies with the objective of applying AI to different approaches, ranging from project production to building operation [4–7].

One of the realities in which the application of AI in civil construction is feasible is the building inspection process. The issue is that buildings suffer wear and loss of performance throughout their useful life, thus requiring periodic assessments that are carried out through services such as inspections, surveys, and expertise. These, on the other hand, tend to generate a large number of images that must be cataloged and classified [8,9]. Garcia Sobrinho et al. [10] carried out a study based on building inspection in which they generated 4481 images that had to be classified manually, an activity that could have been undertaken by a network trained to classify images into categories.

Visualizing the optimization of this type of service, it is possible to train a neural network to catalog the problems that could be found in buildings. One of the types of networks most used in image classification is the convolutional neural network (CNN).

One of the important points in applying a CNN to a classification problem is to teach it to correctly identify the image classes. For this, one must have a representative database with a lot of images and with a balance between the number of images of each class. One of the most used databases for image recognition applications is ImageNet. It contains over 14 million images divided into over 20,000 categories, all in the public domain [11]. However, for some specific problems that are objects of scientific investigation, the current public databases do not usually have information of sufficient scale or representativeness. Given this situation, some authors, such as Silva et al. in [12], Katija et al. in [13], and Leenhardt in [14], conducted studies with a focus on elaborating public image databases for research and AI applications. Researchers in some areas of engineering often find it difficult not having public sources of data for classification studies. Furthermore, even when a specific database for a given application exists, they tend to be small when compared to those used for training the current CNNs. To work around this situation, it is possible to use learning transfer.

Researchers are making an effort to train classifier networks for different civil construction problems. However, most of these studies work with homogeneous materials such as concrete and paving. Not all materials used in civil construction have surfaces with a homogeneous appearance, as is the case with facade coatings [15–18].

Facades are elements of buildings that are exposed to environmental weather conditions such as insolation, drying, and wetting cycles. Therefore, periodic inspections and maintenance are necessary in these regions of the building. Facades can be made of different materials; one of the most used for finishing is mortar coating. This type of finish can be applied in several ways, by varying the thickness of its components and type of texture, for example. This can result in different characteristics of the surface finish of coatings [19–21].

The most used surface finish is smooth. In this type of coating, leveling and straightening are carried out on the surface, obtaining a smooth appearance of the coating. Another possible texture is a scrapped finish. For its execution, ground rock grains, derived from a selective grinding process by color and granulometry, are added to the mortar. This results in a grooved appearance of the finish. The rough cast finish is another type of finalization of the mortar coating. It can be executed by spraying mortar or with a texture roller. Its visual finish gives a rough appearance where it is applied.

Regarding the categorization of defects in mortar coating using CNN, Lee et al. [22] used Faster R-CNN to detect multi-class defects in facades with mortar cladding. The average accuracy achieved by the authors was 62.7%. Guo et al. [23] performed the classification of defects in mortar coating from an unbalanced data set, using a VGG16. To deal with the unbalanced dataset, the authors used meta-learning. They obtained an accuracy of 82.86%. Guo et al. [24] proposed a semi-supervised learning based on ResNet 101 and an uncertainty filter for the classification of defects in facades, achieving an accuracy of 84.36%.

However, these studies only addressed the smooth coating with a paint finish. They did not consider different types of surface finish of coating, which reduces the capacity of networks to generalize.

Cracks are common problems in mortar coverings in buildings. From the visual analysis of this type of discontinuity, it is already possible to arbitrate important information about the problem, such as, for example, what caused the crack or its severity. Regarding this type of defect, studies related to the detection of cracks in concrete already show good results. Chow et al. [25] elaborated a pipeline with artificial intelligence for image-based inspection of concrete structures. The authors obtained an average accuracy of 95.6% in the test set. Ali et al. [26] evaluated the performance of CNN-based deep crack location and detection techniques for concrete structures. In the detection method proposed by the authors, they achieved 96.7% accuracy. Islam et al. [27] developed a CNN approach based on transfer learning models using data augmentation and transformation for crack detection in concrete. The best research results were achieved using VGG16 and AlexNet,

both with 99.9% accuracy. Chaiysarn et al. [28] carried out integrated crack detection, from CNN-FCN, at pixel level via photogrammetric mapping of the 3D texture of concrete structures. The authors achieved 99.8% accuracy.

The surface of mortar coatings is more heterogeneous when compared to that of reinforced concrete, which is not so explored in the literature. These are intended for multi-class classification of defects and do not consider different types of surface finish. Another point is that, in the case of concrete, there are already public databases on problems in this type of structure, which may justify the incidence of more research on this topic and better results in assertiveness metrics when compared with studies of this type on mortar coating. Thus, it is possible to verify a research gap in relation to the application of CNN for the classification of mortar coatings with different types of surface finishes. The absence of public databases of images of this type of material may be a reason for the low incidence of studies on this subject.

Therefore, the main objective of the present study was to assemble a public database of cracks in mortar coating considering some different types of surface finish: smooth coat finish, rough cast finish, and scrapped finish. This database is publicly available and, as far as we know, it is the first of its kind. An application of the database is also presented. Images in the database were classified with a pre-trained VGG16 network. Different subsets of the database were used to investigate how representative the individuals in the database are, so most of the peculiarities of the application domain are covered.

Convolutional Neural Network

CNN is a neural network architecture that is widely applied in image recognition. It is basically divided into convolutional layers, pooling layers, and densely connected layers. In convolutional layers, a group of pixels is assigned to a neuron, and this is carried out for all pixels in the image. Based on this configuration, filters are applied to these clusters in order to extract the most relevant features from the image. In this way, in the convolution process, feature maps are generated in every image, through which it is possible to identify important information for the classification process, such as edges and geometries, considering the dependence and positioning of the pixels. In the pooling process, the dimensionality of the image is reduced. After the process of following the feature maps in the convolution layer, this information can be simplified, which reduces the amount of weights to be learned and can avoid overfitting. The last layers are the densely connected ones, in which the image classification process takes place based on information on the variant of characteristics defined in the convolutional and pooling layers [29,30].

One relevant factor for network performance is the database that will be used in the training process. The database is usually divided into three groups: training, validation, and testing. In the training group, one keeps the images that will be used to teach the network which class should be assigned to each image. The validation group is used to verify if the network is really learning during the training phase or it is simply memorizing the training group. Finally, in the test group are the images that will be used to verify the ability of the network to succeed, i.e., to check whether it is capable of correctly classifying an individual never seen before. The whole database is important for the success of training. The database must be large enough for the network to be able to generalize and cover as much as possible the aspect of the desired classification. Some networks like AlexNet and GoogleNet were trained with millions of images, for example. Another important point to be considered is the balance between classes. If some class has many more images than others, the network can present a tendency to adjust the weights to favor the dominant class [29,31].

To work around this situation, it is possible to use learning transfer. In this process, the convolution layers of already-trained networks are frozen and the fully connected layer is changed. In this case, it is understood that the features extracted in the convolution process do not usually vary much between different objects. Thus, a network trained to extract features from millions of images has great generalization capacity. So, when using

pre-trained networks, one enjoys the feature extraction capably and just has to train the network to classify the individuals in the database in question. This type of approach has been applied in classification studies and obtained good results [32,33].

2. Materials and Methods

This research was carried out in five main stages: the delimitation of the environment and the collection procedure; the pre-processing of the images; data labeling; the analysis of the number of images and the database balancing; and, finally, the training of neural networks from the transfer of learning using the VGG16. The flowchart of the steps can be seen in Figure 1.



Figure 1. Flowchart of the steps taken.

The process of delimitation of the environment, collection of images, and pre-processing will be explained in the methodology. The other processes will be described in this section; however, more detail will be given in the sections on the results and the discussion. This setting was made because the process and decision-making in the assembly of the database were important parts of the research objectives.

2.1. Delimitation of the Environment and Collection Procedure

The data collection environment defined for this study was one of the units of the Federal University of Uberlândia (UFU), called Campus Santa Mônica. This campus is located in the city of Uberlândia—MG in Brazil. Currently, it has 73 blocks named by letters and numbers. The blocks are administrative classifications of buildings, or a set of them, defined by the university. They serve various purposes, such as classrooms, laboratories, community centers, university restaurants, and test sheds, among other features.

The buildings that make up the collection environment are diverse. They vary in relation to their formats, sizes, and constituent materials. In Figure 2, it is possible to observe four buildings with different compositions of materials on their facades.

On an initial visit to UFU, 43 units were identified with their facades finished with mortar. Other structures with mortar coating were also observed, such as retaining walls, ramp walls, stair walls, and low walls.

From the cataloging of constructions, inspections were carried out to survey the database. Some criteria were established to standardize, as far as possible, the acquisition of images. The first point was the inclusion and exclusion criteria for the façades. Only the facades of buildings that had some direct contact with the environmental weather were considered. In this way, only images of the elements that were directly exposed to the sun or protected with short sections of awnings or roof eaves were obtained. Walls that were in the center of the buildings were not considered, even if positioned on the outside, or with protection from long spans of coverage or upper floors. In Figure 3, it is possible to observe examples of closures that were or were not part of the database.



Figure 2. Facades with different types of constituent materials. (**a**) Facade with mortar coating; (**b**) Exposed brick façade; (**c**) Facade with metallic and glass frames; (**d**) Facade with ceramic coating.



Figure 3. Example of faces of buildings considered or not in the data collection. (**a**,**b**) Walls with short coverage protections, considered in database protection; (**c**,**d**) External walls protected by large areas of roofs or upper floors, not considered in the assembly of the database.

Walls, low walls, and any other type of construction that had the type of coating studied and that was exposed to the environment were included in the database. In the preliminary treatment of the data, the images of the buildings were associated with the

nomenclature of their respective blocks. Buildings considered as annexes, when they were included in the same block as the main building, were separated into distinct groups. Images of retaining walls, ramp walls, stair walls, low walls, and other elements were grouped to the nearest block dataset. An access ramp was classified as a distinct region from the others, having its own group of data. In the end, 44 groups of images were obtained, according to the defined regions. Most of them were named after the block they were obtained from.

The photos were taken using the camera of a cell phone. The device can obtain 48 Mp or 5 Mp images, has a $\frac{1}{2}''$ sensor, aperture size of F 1.8 + F 2.2, and autofocus. Images were obtained at a resolution of 8000 × 6000 pixels. The distance between the camera and the photographed objects was approximately 1 m (±0.2 m). This distance was limited in order to have greater standardization of the images obtained and for them to be closer to the structural performance criteria of facades of the Brazilian standard [34]. The distance between the operator and the mortar coverings was controlled by means of a laser tape measurement.

The camera was kept parallel to the inspected object. For this, horizontal angulations between the camera and the coating were avoided and little vertical angulation was allowed. Instruments were not used to measure the angles between the camera plane and the coating. The control in question was carried out by the sensitivity of the operator regarding the positioning of the camera. In Figure 4, it is possible to verify what was considered as horizontal and vertical angulation in this study.

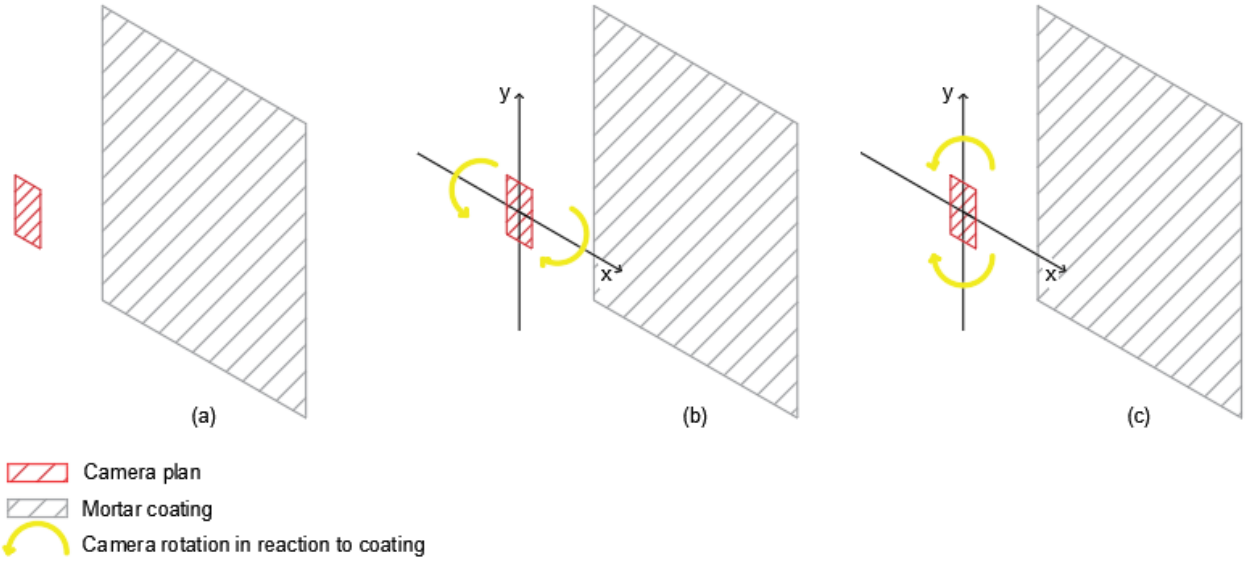


Figure 4. Angular variations between the camera and the mortar coating. (**a**) Camera plane parallel to the casing plane; (**b**) Rotation that generates horizontal angular variation between the camera and the coating occurring around the y-axis; (**c**) Rotation that generates the vertical angular variation between the camera and the casing occurring around the x-axis.

Finally, there were 4092 images in the 44 specified clusters. In Table 1, it is possible to verify the number of images for each grouping. These images underwent pre-processing and were classified.

Table 1. Number of images per grouping.

Grouping	N. of Images	Grouping	N. of Images	Grouping	N. of Images	Grouping	N. of Images
1C	4	1V	3	3Z	9	5OA	55
1D	15	1X	46	5B	96	5OB	183
1E	94	3C	47	5D	26	5P	10

Table 1. *Cont.*

Grouping	N. of Images	Grouping	N. of Images	Grouping	N. of Images	Grouping	N. of Images
1EA	162	3E	256	5DA	24	5RA	101
1I	49	3J	84	5F	167	5RB	176
1K	38	3K	60	5G	96	5T	91
1M	160	3KA	90	5H	54	5U	46
1N	112	3L	43	5I	88	5V	29
1P	20	3O	92	5J	77	R	186
1Q	99	3P	388	5K	128		
1S	142	3Q	228	5L	92		
Total		4.091					

2.2. Pre-Processing of Images

In pre-processing, the photos were resized and cropped into image patches. Resizing was carried out with the aim of optimizing training since the original size of the images was too large. The usual size of the images that are generally used in the training of current networks is 224 × 224 pixels, based on the dimensions present in the ImageNet database, which is much smaller than the original size of the images obtained. The division does not affect the semantics of what is presented in the image; on the contrary, it is even advantageous as it generates a greater number of images in the database.

The resizing process was carried out using the Image Resizer tool for Windows. The photos were reduced from the size of 8000 × 6000 pixels to 2000 × 1500 pixels. This dimension was chosen so that patches smaller than 500 × 500 pixels could be cut. Thus, each photo generated a total of 12 image patches.

Dividing images into patches is a method commonly used in classification studies of problems in buildings, as can be seen in the studies of Lee et al. [22], Guo et al. [23], Guo et al. [24], and Chow et al. [25]. The size of 500 × 500 pixels was initially chosen based on the study of Guo et al. [23]. Image division was performed using the PhotoScape v3.7 program. In Figure 5, it is possible to visualize the pre-processing carried out.

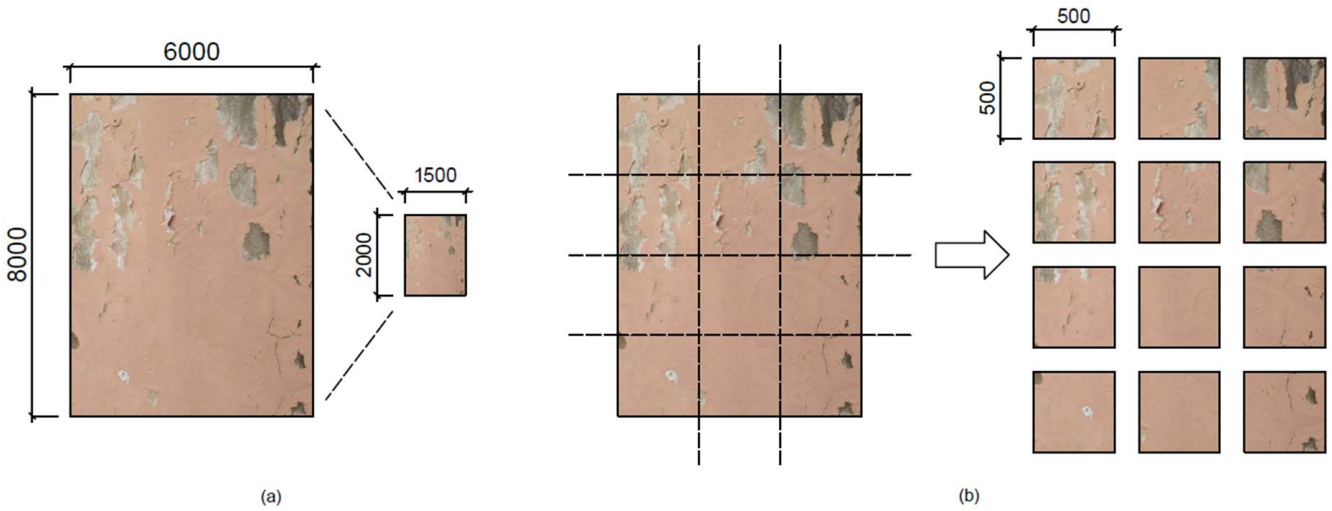


Figure 5. Image pre-processing scheme. (**a**) Resizing; (**b**) Crop into patches of 500 × 500 pixels images.

After pre-processing, the database was left with a total of 49,092 patches of 500 × 500 pixels images.

2.3. Data Labeling

Data labeling followed six steps: noise elimination; division of images between different types of surface finish; division of the images into those with or without noise; division

of the images into those with or without cracks; division of images with cracks according to the level of visibility of the problem; and naming the image according to all divisions performed. In Figure 6, it is possible to verify the flowchart of the data rotation process carried out.

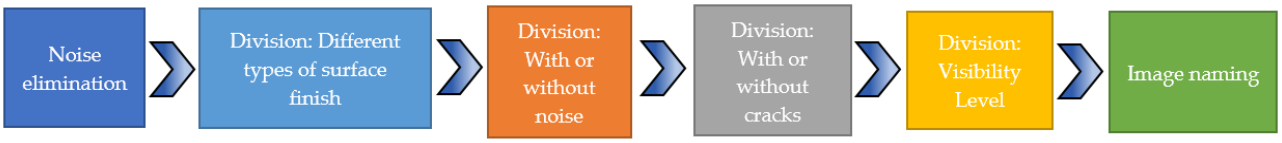


Figure 6. Data labeling procedure.

In the noise elimination process, the patches for which the main information was not the mortar coating or that did not have the finish were removed from the bank. After this step, the bank had a total of 33,088 image patches.

In the division in relation to the type of surface finish, the data were grouped into three types of mortar coating texture: smooth coat finish, rough cast finish, and scrapped finish. In Figure 7 it is possible to verify examples of the three main types defined.

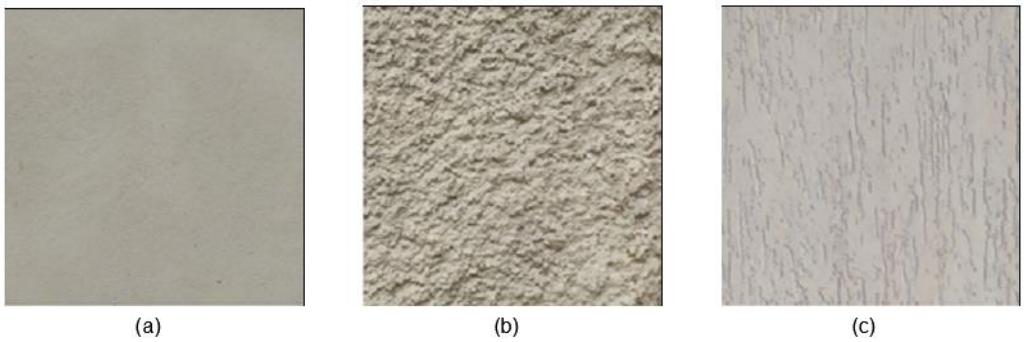


Figure 7. Different types of surface finish. (**a**) Smooth coat finish; (**b**) Rough cast finish; (**c**) Scrapped finish.

In the third step, image patches were divided into noiseless and noisy images. In general, every image that had some information that was not related to the coating was considered as having noise. Then, the data were classified between the coatings that contained cracks and those that did not.

The last grouping process was about the visibility of the cracks. It was observed that, in some images, the discontinuities were more visible while in others they were difficult to identify. Thus, the patches were identified according to three levels of identification in relation to the cracks: high, medium, and low. Finally, the images were named using a code developed for the database in question. The detailed decision-making process during data labeling will be explained in Section 3.1. In Figure 8, it is possible to see a flowchart of the data classification procedure.

2.4. Analysis of the Number of Classes and the Balance of the Database

With the database divided and classified, the quantity of each category was tabulated and analyzed using spreadsheets.

To verify the balance of the database, a ratio between the number of images in each of the classes that were compared was used. The equation of the respective relationship can be seen below.

$$\text{Relationship between the amount of image of each class } (R_c) = \frac{M-}{M+}. \quad (1)$$

In this case, $M-$ is the number of images of class with the smallest amount in comparison and $M+$ is the number of images in the class with the largest amount of data. In

this way, there is always a ratio greater than 0 and less than or equal to 1. This means that, when the ratio between the classes in a comparison is closer to or equal to 1, we have a database that is well or completely balanced. On the other hand, the smaller the ratio, the more unbalanced the analyzed situation.

Based on this parameter and data tabulation, the following relationships were analyzed:

- Between the images containing cracks or not of the total dataset and of the separate surface finishes;
- Among the images containing cracks or not considering the situations of having noise or not;
- Among the images containing cracks or not considering the types of finishes and the presence of noise;
- The number of images with cracks considered with high, medium or low visibility according to each type of surface finish.

In this way, points that can influence the assertiveness-ability metrics of training the database were created.

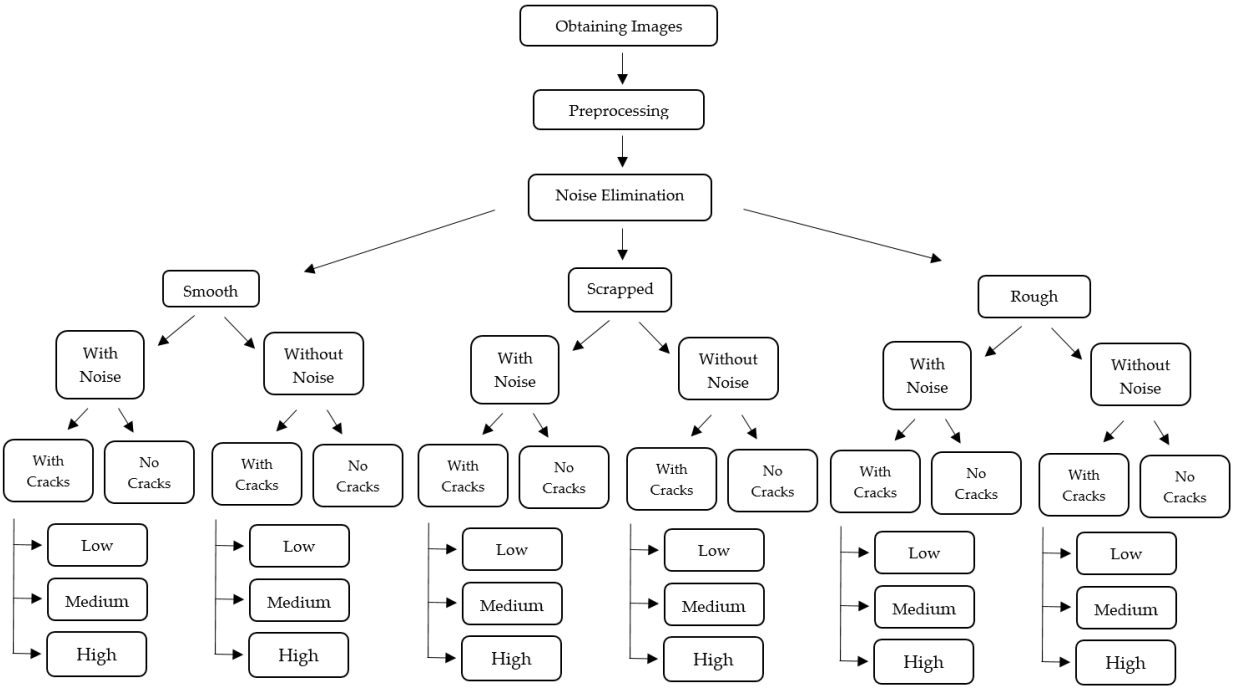


Figure 8. Classification process in data labeling.

2.5. Training of CNNs from Transfer of Learning

The last stage of the investigation aimed to compare some assertiveness metrics of networks trained with different image clusters. The focus of this part was to test whether the type of finish has any influence on assertiveness during training on the network and, from this, raise points that can be the basis of research that uses the database elaborated in this study. For this, the VGG16 was used, based on transfer of learning.

The transfer learning process was carried out by removing the last layer of the network, which had 1000 classes, and replacing it with one with two output classes. A new layer was created to replace the hidden layer of the network so that the pre-trained model of VGG16 could classify the presence or absence of cracks in the mortar coating. The network was trained in 14 epochs and with a learning rate of 0.001. The entire process was carried out in the Google Colab Pro+ environment using the Python language and the Pytorch library. In the execution environment, an A100 GPU and high RAM were used.

In the pre-processing of data at the input of the network, the patches were resized to a size of 224 × 224 pixels, which corresponds to the dimensions of the images used for

training the VGG16. The inputs were transformed to the given Pytorch type using the transforms.ToTensor function and normalized using the transforms.Normalize function, which uses the standard deviation and the mean of the RGB channels. The values of the standard deviations and averages used in the normalization correspond to the ImageNet database, being, in each RGB channel, (0.485; 0.456; 0.406) for the averages and (0.229; 0.224; 0.225) for the standard deviations.

The images were divided into datasets for training and validation. The metric used for comparison was accuracy. The accuracy curves and the loss function in relation to the number of epochs were also verified. The accuracy is the metric most used to evaluate the performance of AI training models; it indicates the percentage of correct answers in the network obtained from Equation (2). Accuracy indicates the percentage of images that the network classified correctly out of the total number of images.

$$\mathrm{A} = \frac{V = TP + TN}{TP + TN + FP + FN}. \quad (2)$$

In this case, the letter A corresponds to accuracy. *TP* is the number of images in the positive class that the network managed to get right and *FP* is the number of positive images that the network missed. *TN* is the number of images in the negative class that were classified correctly and *FN* is the number of this class that were classified wrongly.

Initially, tests were carried out with a small group of 446 images of the smooth surface finish, without noise, and in the proportions of 50% of images containing cracks and 50% without discontinuities. Two tests were performed, one with the manual selection of images from the training and the validation groups and the other with this division being randomly generated from the splitfolders library. In the tests, the image patches were organized in the proportions of 80% for the training set and 20% for the validation set.

From these tests, the use of the splitfolders library was defined to split the training and validation sets of the analysis groups. The proportions were 70% for training and 30% for validation. For comparisons, the database was divided into two main groups. One of them contained all the images, generated in unbalanced data sets; this group was identified with the letter A. In the other grouping, the data were balanced from the minority class. That is, the classes that had more images were reduced for an amount of data from the class with the lowest value. This process was carried out through the splitfolder library. The balanced group was identified by the letter B.

Within groups A and B, four different approaches were performed for each one of them. The four different types of data organization were:

- A dataset containing all images, identified by the number 1;
- A dataset containing only images with a smooth surface finish, identified by the number 2;
- A data set containing only the images with surface finish of the type scrapped, identified by the number 3;
- A data set containing only the images with a rough surface finish, identified by the number 4;

In this way, eight training sessions were carried out. In Table 2, it is possible to verify the groupings that were performed for the study.

Table 2. Data clusters analyzed.

	Grouping According to Data Balancing	
Set of Images According to the Type of Finish	**Unbalanced Database (Group A)**	**Balanced Database (Group B)**
All the types—1	A1	B1
Smooth type—2	A2	B2
Scrapped type—3	A3	B3
Rough type—4	A4	B4

Each of the groups had two classes: the images containing cracks and those without. The objective of the training was to verify the ability of the networks to classify the presence or absence of discontinuity in the groupings performed. From this configuration, the datasets were processed and the assertiveness metrics of the networks analyzed.

3. Results and Discussions

The results are divided into three parts—data labeling, database analysis, and training.

3.1. Data Labeling

Labeling was carried out manually by two civil engineers who are experts in the field. The division was carried out based on the professionals' perception.

The first process carried out was the elimination of noise. The initial photographs had regions that did not contain the studied coating. Therefore, some of the patches had information that was not interesting for the database. These patches were therefore excluded.

Two main types of images were identified that were considered to be, exclusively, noise. In the first type, there were the patches that had no relation to the mortar coating, such as: background images, floors, frames, vegetation, and other objects. Some examples of this type of noise can be seen in Figure 9.

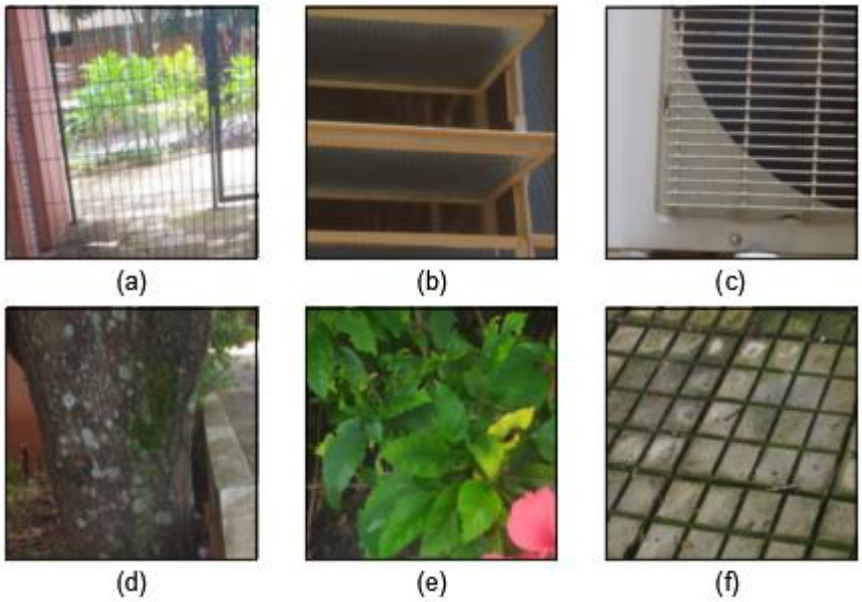


Figure 9. Example of patches with information unrelated to grout coating. (**a**) Background image; (**b**) Glass window part; (**c**) Air conditioning part; (**d**) Tree trunk; (**e**) Vegetation; (**f**) Floor.

The other type of noise identified was the patches whose main image information was not the coating. In Figure 10, it is possible to verify some examples of this type of noise. Patches of images that were repeated were also removed.

After eliminating the noise, the images remained divided according to the three types of surface finish defined, namely: smooth finish, rough finish, and scrapped finish. Within each of the finishing categories, it was possible to observe some variability, for example, in the surface texture. The issues observed will be maintained below.

In coating execution, the finishing process can be applied using different techniques. This also varies according to the purpose of the building. For example, some buildings need to have a greater aesthetic performance than others. In this way, the finishing process tends to be more careful. These variations can be seen in Figure 11. In images (a) and (b) of the figure, it is possible to verify finishes with a really smooth appearance. In images (c) and (d), however, local problems can be observed in the execution of the coating. In the case of patches (e) and (f), there are scratches resulting from the passage of the trowel in

the leveling process, probably because the finalization was not carried out using a damp sponge. The possible absence of the last-mentioned executive procedure can also be seen in (g) and (h). However, the result in this case was a rough look. Finally, the texture of image (i) may have been executed using a specific technique that was not intended to make the surface completely smooth.



Figure 10. Example of patches where grout coating is not the main information. (**a**) Piping; (**b**) Concrete pillar; (**c**) Grade; (**d**) Floor and piping; (**e**) Air conditioning wiring; (**f**) Window corner (**g**) Floor and door; (**h**) Vegetation; (**i**) Image Background.

Another variability found in the smooth surface finish was the paint colors. In Figure 12, it is possible to observe some of these differences. One of the visually verified points is that painted surfaces usually have a more homogeneous appearance than unpainted ones. This can be justified because the paint seals the porosity of the mortar, which reduces the rough appearance of the coating.

Regarding the scrapped type finish, it was also possible to verify different aspects of the surface texture. Image (a) in Figure 13 shows a more uniform distribution between orientations of the grooves. In (b), it is possible to observe a smooth region between the scrapped finish. In the image patch (d), the finish was left with a smooth appearance in a large part of the image, probably due to some incorrect execution procedure. In image (d), the grooves orientations are thinner and with a higher density. This contrasts with image (e), where the grooves are thicker. In the case of (f), the finished appearance was closer to holes than grooves, which may be the result of problems in dragging the stone grains in the execution process.



Figure 11. Different textures in the smooth surface finish; (**a**,**b**) Smooth finish; (**c**,**e**) Local flaws in textures; (**d**,**f**) Texture with scratches due to the passage of the trowel and possible absence of a finishing process; (**g**,**h**) Texture with roughness due to the possible absence of a finishing process; (**i**) Finishing not intended to be completely smooth.

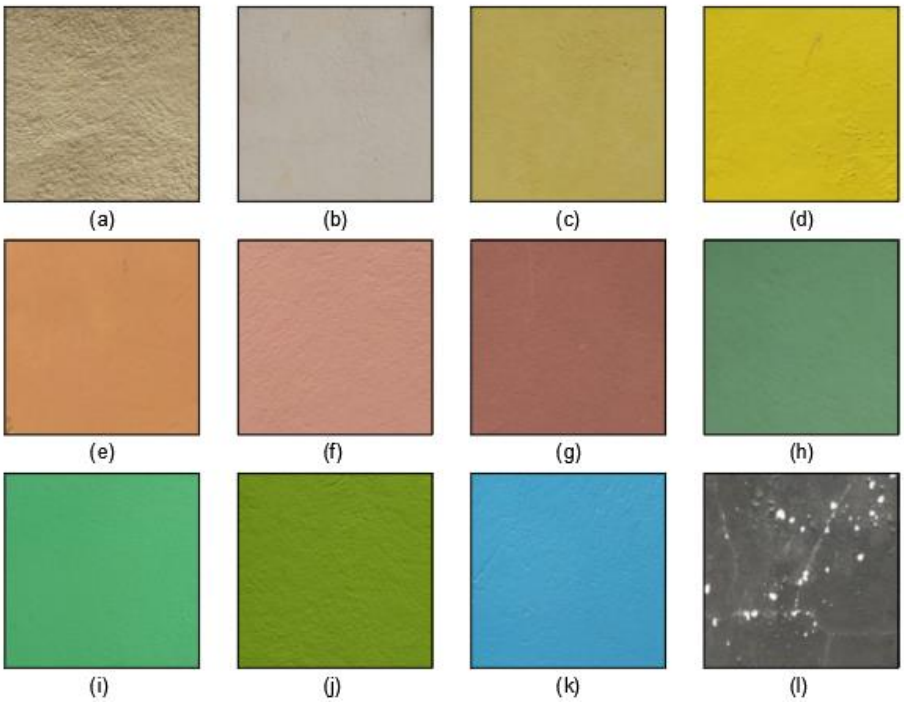


Figure 12. Different paint colors on smooth surface finish; (**a**) Unpainted; (**b**) White; (**c**) Beige; (**d**) Yellow; (**e**) Orange; (**f**) Rose; (**g**) Red; (**h**–**j**) shades of green; (**k**) Blue; (**l**) Black.

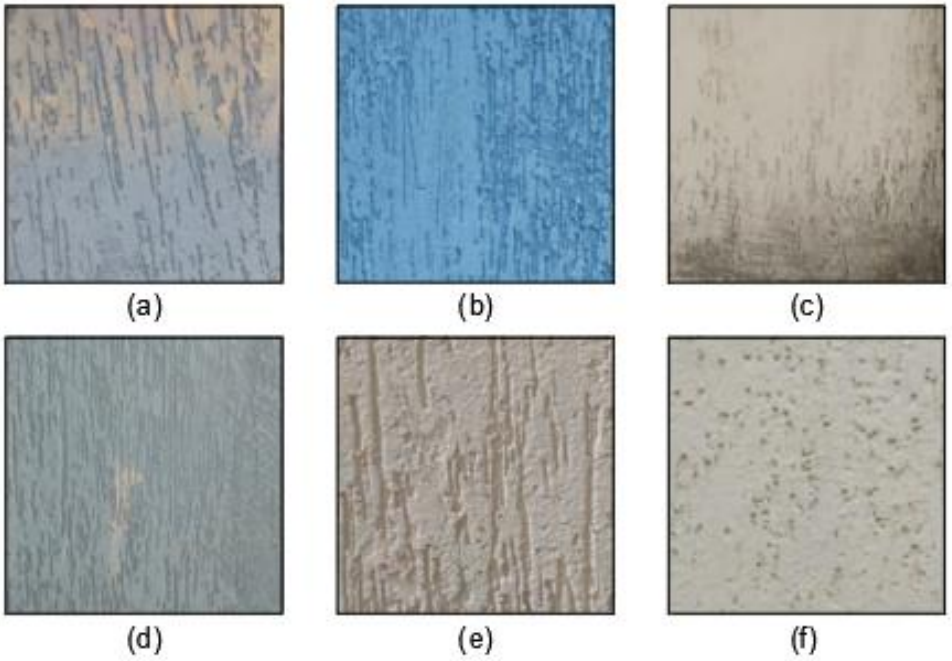


Figure 13. Scrapped surface finish textures; (**a**) Uniform distribution; (**b**) With region without grooves; (**c**) With part of the smooth area; (**d**) With thin grooves and with greater density; (**e**) With thick grooves; (**f**) Without adequate dragging of the stone grains, giving an appearance more of holes in the coating.

Regarding the colors of this type of finish, they were between white, gray, blue, and yellowish tones, not varying much in relation to the colors developed in Figure 13.

In the case of the rough finish, in Figure 14, it is possible to observe differences in roughness between the images. In the cases of (c) and (d), the roughness is more prominent, with larger ripples than in images (a) and (b). As for patches (e) and (f), the roughness is lighter, the result of an application using a textured roller. In images (g) and (h), a finer roughness texture is observed. Finally, image (i) has a flaw in the application of the rough surface finish.

Like the scrapped finish, the rough cast finish did not have much variation in terms of paint colors, which were found in white and in shades of gray and yellowish, similar to what is shown in Figure 14.

After classification in relation to the types of surface finish, the patches were divided into groups with or without noise. Unlike the images that were noise, those that were defined as having noise were those that had the mortar coating as the main information of the image; however, they also had some other information that was not directly related to the finish.

An example is the images in which the background appears at the edges, as is the case with the corners of buildings. This can be seen in Figure 15.

Another situation considered as noise was the encounter between walls, wall and slab, walls and floor, and surfaces that were misaligned. The main issue of this delimitation is that these configurations generate lines in the images that, theoretically, can be confused with geometric fissures in the training of a network, mainly in the vertical and horizontal directions. In Figure 16 it is possible to verify examples of this situation.

Some images contained corners and parts of frames such as windows and doors. This type of noise appears in Figure 17.

The presence of devices and elements of hydraulic, electrical, logic, and air conditioning installations was another type of noise found next to the coverings. Examples of this type can be seen in Figure 18.



Figure 14. Rough cast finish textures; (**a**,**b**) With slight roughness; (**c**,**d**) With large roughness; (**e**,**f**) Applied with a texture roller; (**g**,**h**) With fine appearance; (**i**) With flaws in the finish.

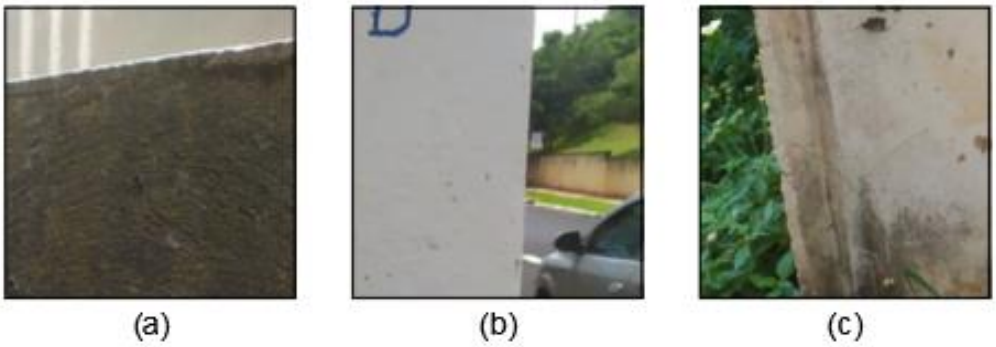


Figure 15. Coating images with background (**a**–**c**).

A noise that appeared frequently was the incidence of vegetation. This can be seen in the patches in Figure 19.

Another type of noise considered was shadows that outlined images on the coating surfaces and intense light variations, which impaired the visualization of the finish. Examples of this type can be seen in Figure 20. Focus oscillations in the images were not considered with noise.

Drawings, symbols, or letters resulting from painting or graffiti on the coating were also classified as noise. Figure 21 shows examples of these.

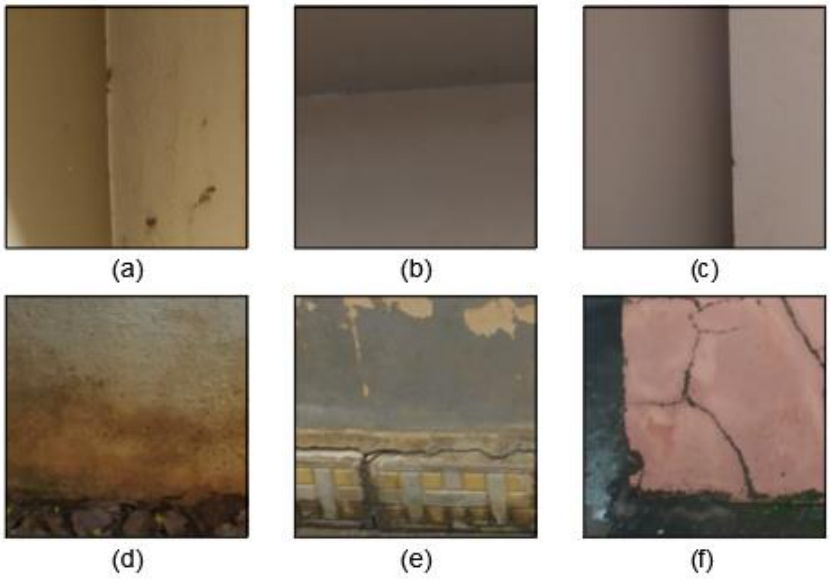


Figure 16. Encounter pictures; (**a**) Encounter between two walls; (**b**) Encounter between wall and slab; (**f**) Misaligned surfaces; (**c**–**e**) Meeting of wall and floor.

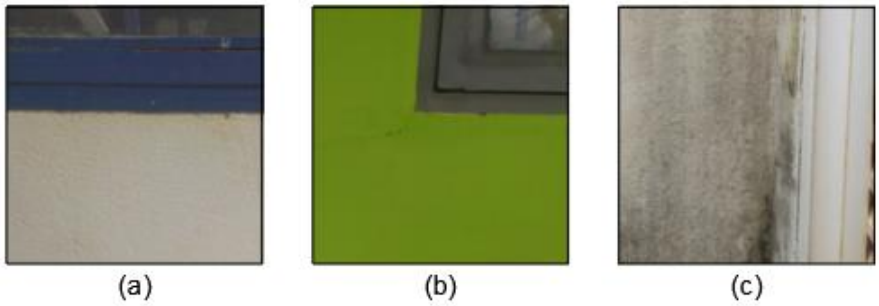


Figure 17. Images of frame parts; (**a**) Window at the top; (**b**) Window corner; (**c**) Part of the door in the right corner.

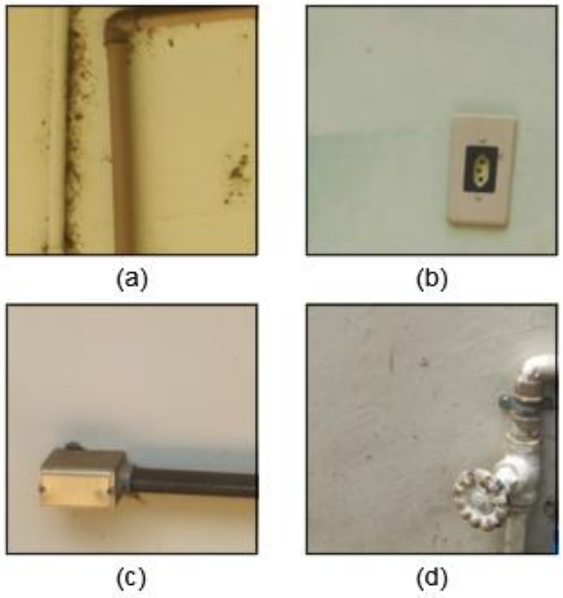


Figure 18. Images with parts of building installations; (**a**) Piping; (**b**) Power socket; (**c**) Energy point; (**d**) Gate valve.

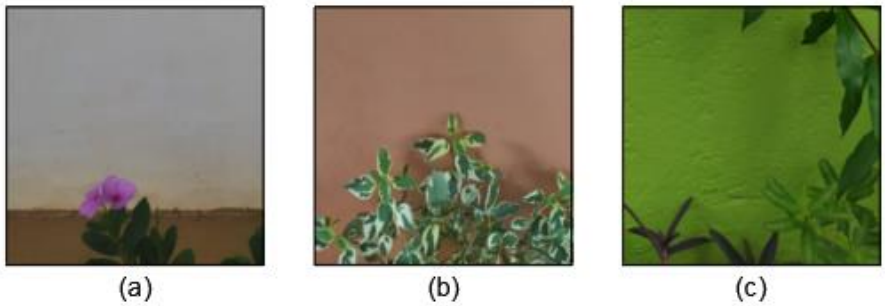


Figure 19. Images with vegetation (**a**–**c**).

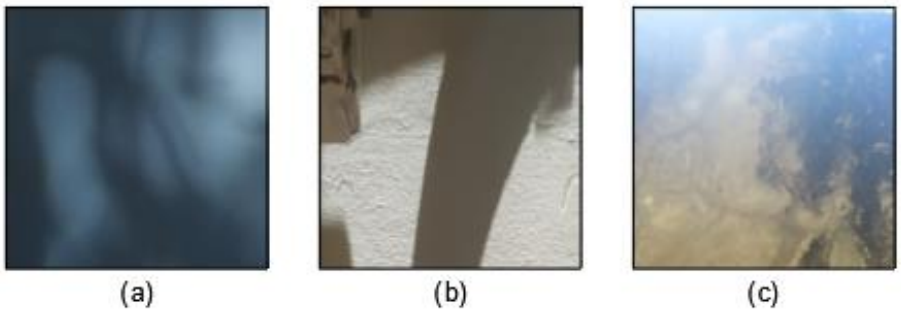


Figure 20. Images of shadow and lighting; (**a**,**b**) Shadows that outlined images on the coating; (**c**) Intense brightness at the top.

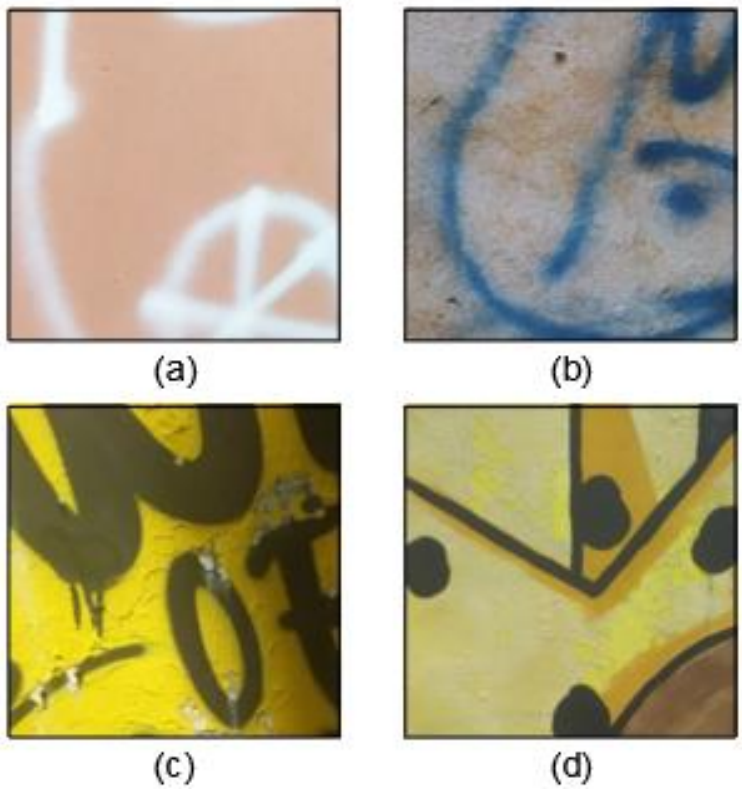


Figure 21. Drawings, symbols or letters made on the mortar coating; (**a**,**b**) Graffiti; (**c**) Letter painting; (**d**) Drawing painting.

Finally, any patches with objects other than the cladding were placed in the class having noise. Figure 22 shows some types of images with these settings.

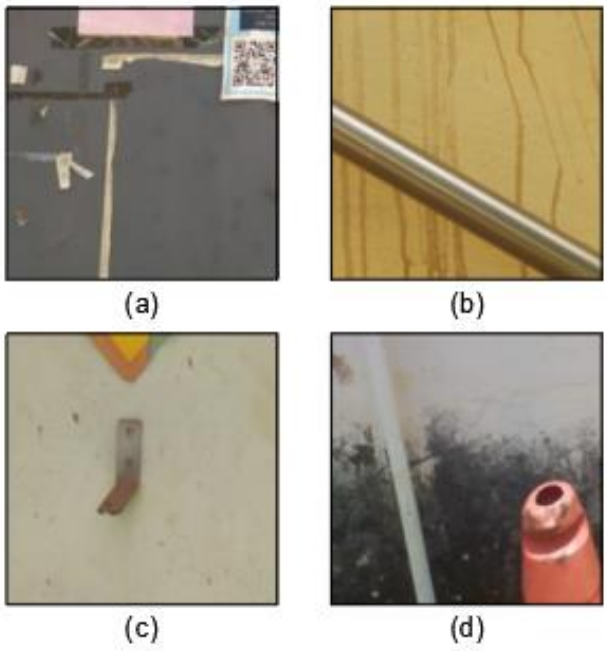


Figure 22. Images with objects; (**a**) Papers and adhesive tape nailed to the coating; (**b**) Handrail part of a ramp; (**c**) Bench support; (**d**) Miscellaneous objects.

Pathological manifestations or defects arising from natural wear and tear or the use and operation of buildings were not considered noise. It was considered that the network should be able to classify cracks correctly even with this type of information. Some pathological manifestations found can be seen in Figure 23.

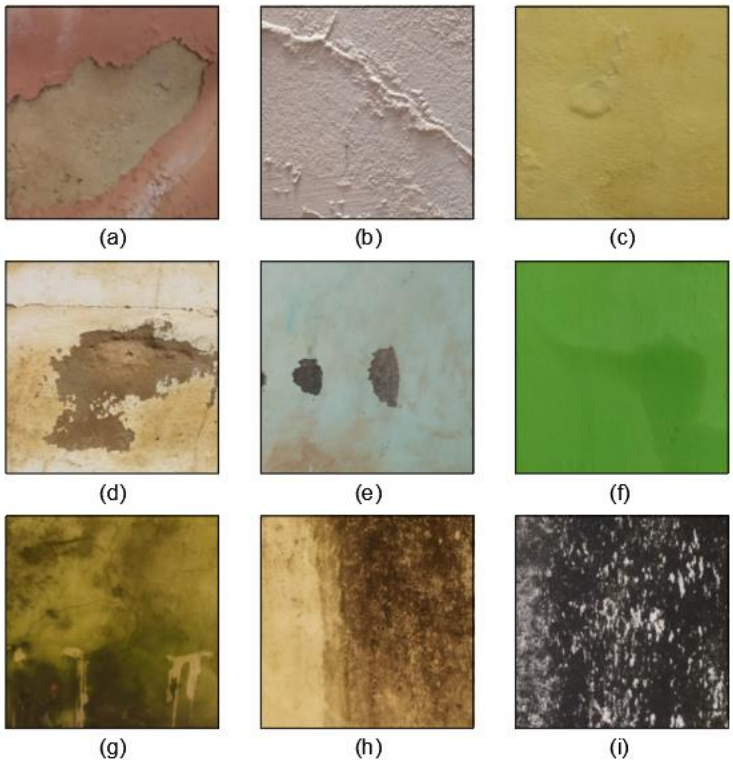


Figure 23. Pathological manifestations; (**a**) Paint displacement; (**b**) Paint blistering; (**c**) Bubbles in the paint; (**d**) Disaggregation of the coating; (**e**) Impact damage; (**f**) Humidity; (**g**) Moss proliferation; (**h**) Dirt from rainwater runoff; (**i**) Intense biological proliferation.

It was possible to observe several types of pathological manifestations in the database. Studies like Lee et al. [22], Guo et al. [23], and Guo et al. [24] aimed to classify these types of defects in facades from a CNN multi-class approach. In this way, it is observed that the elaborated database can undergo a labeling process considering the defects as classes and be used in classification problems of this type.

Following this labeling step, the images were separated into those with or without cracks. It was observed that, in some images, the discontinuities were difficult to identify or opened up room for doubt. Thus, it was established that any image that suggested the presence of a crack would be included in the crack class. Examples of cracking by each type of surface finish can be seen in Figure 24.

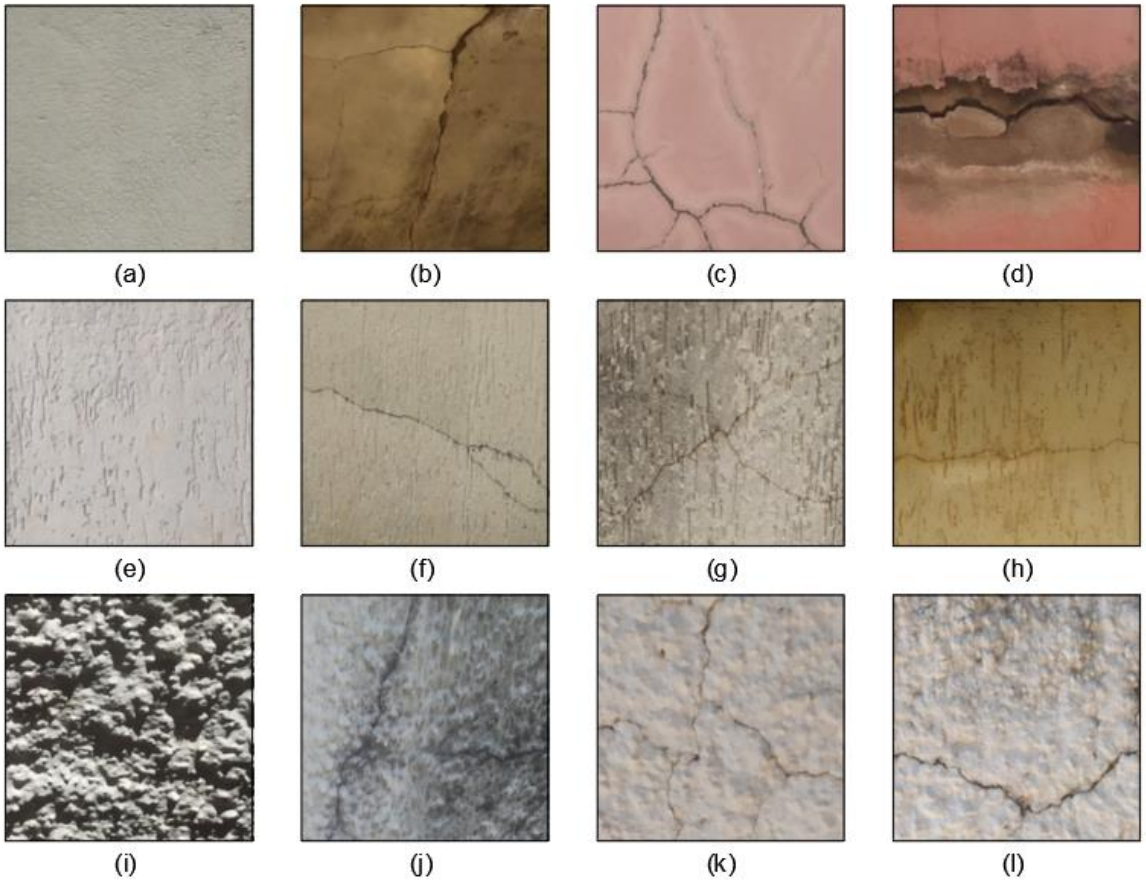


Figure 24. Cracks by type of finish; (**a**) Smooth finish without cracks; (**b**–**d**) Smooth finish with cracks; (**e**) Scrapped finish without cracks; (**f**–**h**) Scrapped finish with cracks; (**i**) Rough cast finish without cracks; (**j**–**l**) Rough cast finish with cracks.

It was possible to observe that the cracks presented different types of configuration, as can be seen in Figure 24. This feature is known to be important for the visual inspection of this type of discontinuity. Figure 25 presents examples of different configurations of cracks found.

From the configuration of the cracks, it is possible to arbitrate the possible causes of their occurrence. For example, horizontal cracks tend to occur in connections between structure and masonry, cracks inclined at 45° can indicate structural settlement or the absence of lintel and slat in windows. In the case of the mapped cracks, these tend to be associated with problems of performance or coating fatigue. In this way, the database can also be labeled in relation to crack configurations and applied for this purpose.

Due to the criterion of including all images that raised doubts in the group of patches with cracks, this class was divided according to visibility levels: High, medium, and low. Thus, it is possible to group the database according to the visibility of the cracks, for

example, remove images with low visibility to test a network training. This part of the labeling can be seen in Figure 26.

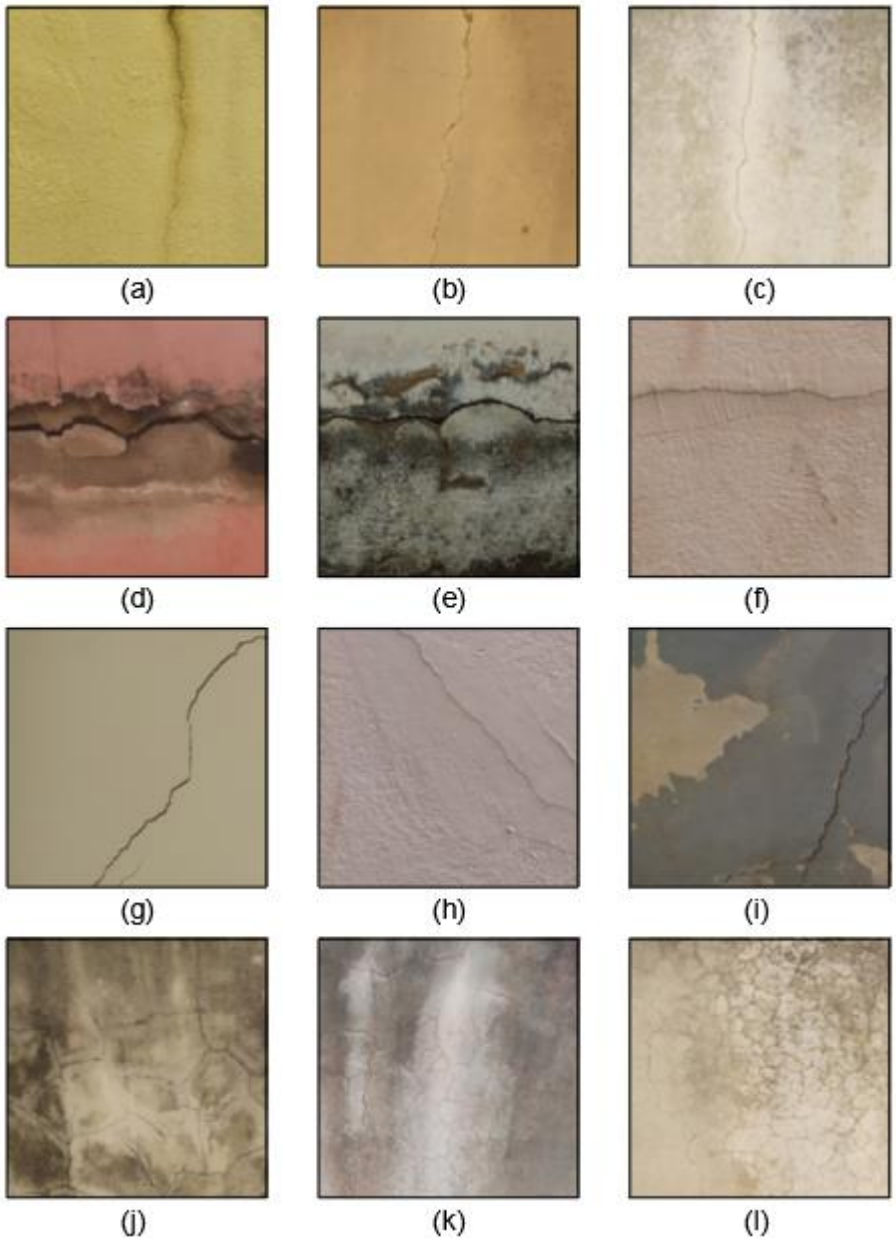


Figure 25. Different configurations of cracks; (**a**–**c**) Geometric cracks in the vertical direction; (**d**–**f**) Geometric cracks in the horizontal direction; (**g**–**i**) Inclined cracks; (**j**–**l**) Cracks mapped.

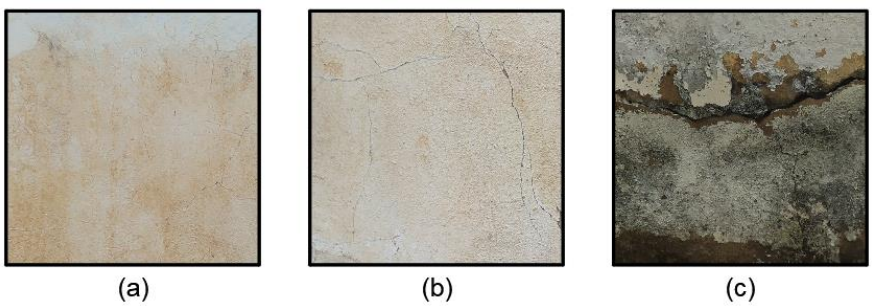


Figure 26. Different levels of visibility of cracks; (**a**) Low; (**b**) Medium; (**c**) High.

The last procedure performed in the rotation was naming the images. For this, a codification was created using letters that were:

Regarding the presence or absence of cracks:

- F: Images with fissures;
- N: Images without cracking.

Regarding the presence of noise:

- R: Images with noise;
- S: Images without noise;

As for the visibility of the crack, when it was in the image:

- A: High visibility;
- M: Medium visibility;
- B: Low visibility.

Regarding the type of surface finish:

- L: Images with a smooth surface finish;
- G: Images with a scrapped surface finish;
- C: Images with rough cast surface finish;

The name of the group used to obtain the images, according to Table 1, was followed by a number to differentiate images with the same classifications.

Thus, the name of the image corresponded to the sequence of specific codes. In Figure 27, it is possible to observe two examples of image encoding.

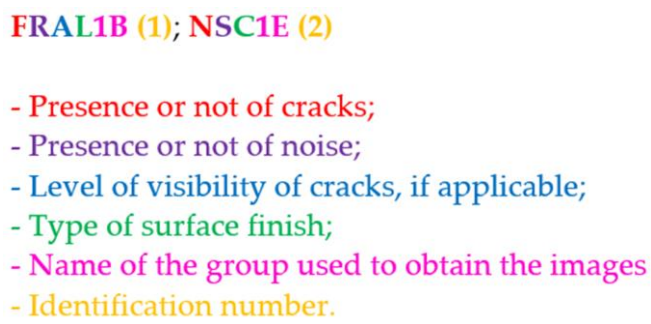


Figure 27. Examples of patch coding.

It was possible to verify variability within the defined categories during the rotation process. In the surface finish, the three types presented different textures in their images, generated in different topographical aspects. In situations where this feature is important, this setting can have an influence. For example, image (c) of Figure 13, in which the coating with a scrapped finish has a large smooth part, can be confused by a network as belonging to a classification of smooth coating.

The database presented various types of noises. Some of them generate lines in the images and shapes that may resemble cracks in the process of feature extraction, such as the encounter between walls and the incidence of shadows. Another point is that some noises appear frequently related to cracks. This was possible to observe during the labeling process when there was part of the floor, since many of the images with this type of noise were concomitant with cracks. This situation was also verified when there were pathological manifestations. Some types of problems were associated with the incidence of fissures.

The cracks showed different sizes and configurations. In some images, the discontinuities were difficult to detect, raising doubts even among professionals responsible for manual labeling.

The database elaborated in this research represents a real situation of problems within a delimited universe. Therefore, finding solutions to work with data configuration and obtaining good assertiveness results can result in networks with greater generalization capacity, in relation to the investigated problem, and that are more suitable for real situations.

Thus, the data labeling process was completed.

3.2. Analysis of the Number of Data in the Classes and the Balance of the Database

In this part of the research, the data were tabulated and, based on this, comparisons were made between the number of images of the analyzed classes and groups. For this, the R_m relation was used, which was defined in Section 2.1 of the methodology.

Table 3 shows the number of images, divided into those with or without cracks, for each type of finish and for the entire database.

Table 3. Number of images per surface finish.

Surface Finishing	With Cracks	No Cracks	$R_m = \frac{M-}{M+}$
Smooth	11,648	11,147	0.957
Scrapped	985	4739	0.208
Rough	1140	3429	0.332
Total	13,773	19,315	0.713

It is possible to verify in Table 3 that the smooth finish dataset is almost completely balanced, having a ratio of 0.957. The others have an unbalanced configuration. The scrapped type has the greatest difference between the number of images in the compared classes, with a ratio of 0.208. The bank considering all the finishes together was also unbalanced; however, with a ratio value of 0.713, it was higher than the groups of scrapped and rough type coatings.

Table 4 shows the number of images with or without noise, in relation to the presence of cracks, considering the entire database.

Table 4. Number of images in relation to noise.

	With Cracks	No Cracks	$R_m = \frac{M-}{M+}$
With noise	4867	3779	0.776
Without noise	8906	15,536	0.573

It is observed that the images are also not balanced regarding the presence of noise. With this situation in mind, the data were organized by surface finish, noise, and the presence or absence of cracks. This can be seen in Table 5.

Table 5. Number of images in relation to surface finish and noise.

	No Cracks			With Cracks		
Surface Finishing	**Without Noise**	**With Noise**	$R_m = \frac{M-}{M+}$	**Without Noise**	**With Noise**	$R_m = \frac{M-}{M+}$
Smooth	8973	2174	0.242	7462	4186	0.561
Scrapped	3887	852	0.219	673	312	0.464
Rough	2676	753	0.281	771	369	0.479

Thus, it is possible to perceive that the amount of noise present in the images also varies within each surface finish grouping. This was verified because in all the configurations in Table 5 the relations have values between 0.6 and 0.2. It is assumed that, if noise is an important factor for the assertiveness of the network, even when balancing the bank in relation to the presence or absence of cracks, noise may affect the training performance.

In Table 6, it is possible to verify the number of images containing cracks according to the visibility classes. Next to each value, the percentage of images in relation to the total number of images with cracks for each surface finish is listed.

Table 6. Number of images in relation to the level of visibility of the cracks.

	Number of Images (Percentage *)			
	High	**Average**	**Low**	**Total**
Smooth	1839 (16%)	5668 (49%)	4141 (36%)	11,648 (100%)
Scrapped	123 (12%)	309 (31%)	553 (56%)	985 (100%)
Rough	367 (32%)	518 (45%)	255 (22%)	1140 (100%)

* Percentage between the number of images in relation to the level of visibility and the total number of images with cracks in each surface finish.

It appears that the scrapped type finish has the most images in which the cracks are difficult to detect, representing 56% of the cracks in this grouping. The smooth finish comes next, with 36% of images with low visibility. The rough type was the finish with fewer types of these images.

The database showed an imbalance in all categories delimited in the labeling, that is, in relation to surface finish, presence of noise, presence of cracks, and visibility of discontinuities. In general, there was a greater number of images without the incidence of cracks than those with the problem. The rough database was the one with the lowest number of images. On the other hand, the scrapped type had the greatest imbalance in the category associated with the presence or absence of cracks. The smooth coating was the finish that had the highest volume of data and the most balanced relationship between the incidence of discontinuities in their images that contained cracks or not.

With the dataset developed in this study, it was possible to verify the statements of Lee et al. [22], Guo et al. [23], and Guo et al. [24] about mortar coating generating databases with a lot of variability and that are unbalanced, mainly in comparison with concrete and pavement. The situations raised in the labeling process correspond to scenarios that can be found in real inspections. Therefore, the elaborated database has the potential to be worked on according to different approaches of computer vision, as it is able to optimize the generalization capacity of networks.

3.3. Training of VGG16 from Transfer of Learning

Initially, training was carried out with a small group that was manually selected and balanced between images with or without cracks. The division of patches for training and validation was performed in two ways, by manual selection and by randomly using the splitfolders library. The training graphs can be seen in Figure 28.

It is observed that training using random selection was more regular than manual selection. This can be seen in image (c) with the validation error decaying along with training, which is the ideal situation in the process. Manual selection can generate a tendency in the choice of images, which would harm the network performance and the representation between the division of categories. Thus, it was defined that the other groupings and divisions would be random and carried out using the splitfolders library.

From this, the majority classes of group B had their quantities of images reduced to the value of the minority class. In sequence, group A and B were divided into training and validation sets with proportions of 70% and 30%, respectively. The data measures used in training can be seen in Table 7.

In the reduction process, the library ended up reducing groups B1 and B2 below the minority class; however, it allowed for the balancing of the data. The experiments continued with the groups in Table 4.

In Figure 29, it is possible to verify the training of A1 and B1. The charts for each group were plotted with different cores to facilitate identification. This procedure was repeated in the other graphs.

Table 7. Number of images after balancing and split between training and validation.

Groupings	With Cracks	No Cracks	$R_m = \frac{M-}{M+}$	Total	Training (70%)	Validation (30%)
Group A1	13,773	19,315	0.713	33,088	23,144	9944
Group A2	11,648	11,147	0.957	22,795	15,938	6857
Group A3	985	4739	0.208	5724	4002	1722
Group A4	1140	3429	0.332	4569	3194	1375
Group B1	11,761	11,756	1.000	23,517	16,460	7057
Group B2	9636	9621	0.998	19,257	13,471	5786
Group B3	985	990	0.995	1975	1381	594
Group B4	1140	1145	0.996	2285	1598	687

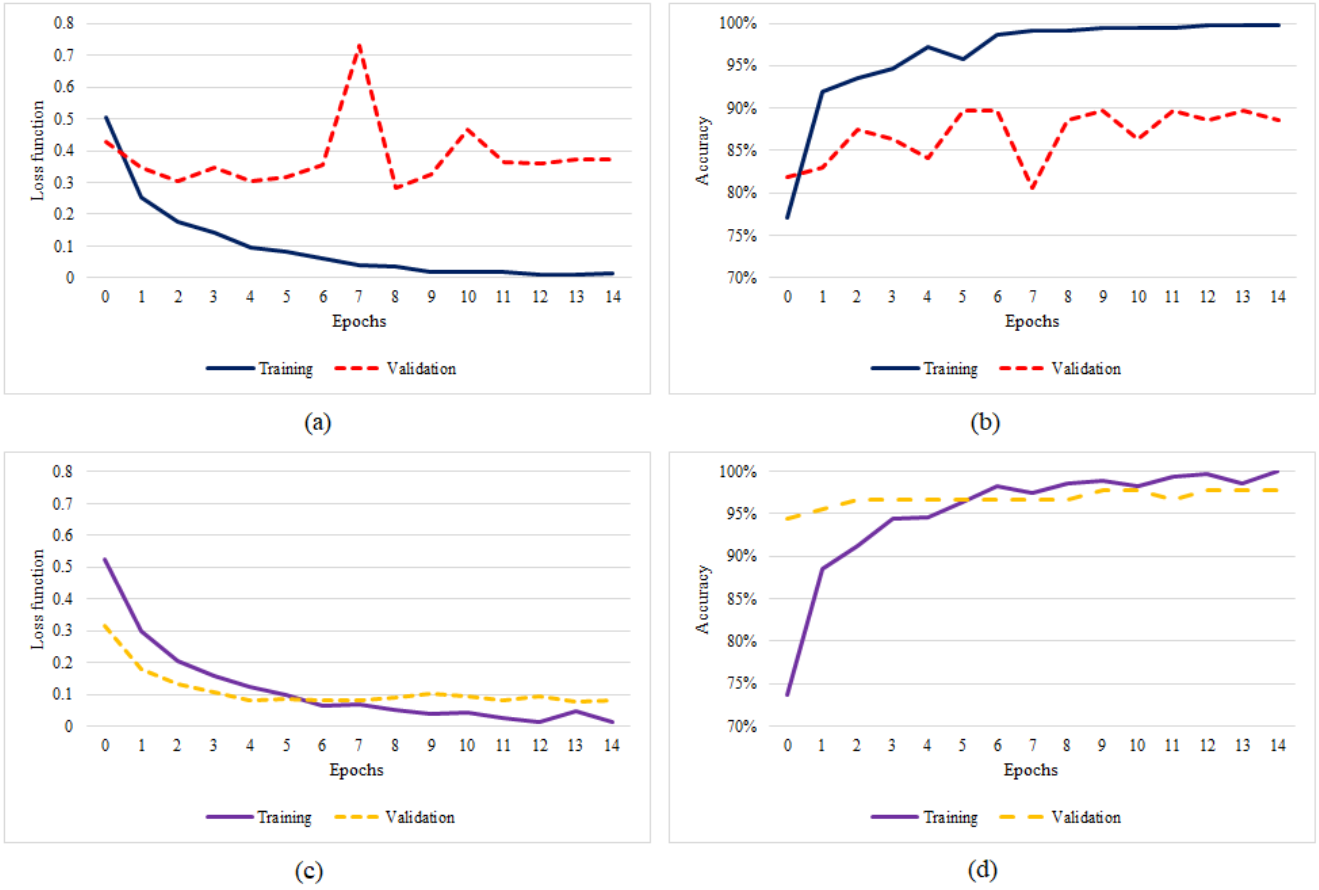


Figure 28. Graphs of loss function and accuracy in test training; (**a**) Group epoch loss function with manual selection; (**b**) Group epoch accuracy with manual selection; (**c**) Epoch loss function of the group with random selection; (**d**) Accuracy by group epochs with random selection.

It is possible to verify that the graphs related to the training of groups A1 and B1 almost did not vary. The accuracy value of group B2 was in a slightly lower range than that of group A1, being close to and below 85%. The accuracy of group A1 was, for almost all training, above and close to 85%. Regarding the loss function, the validation error in both groups decreases to a certain point and then starts to increase. This occurs because the network stopped learning and from that point on it starts to adjust its own training error, which can lead to overfitting. One way to avoid this would be to use early stopping to stop training as soon as the network is no longer learning.

Figure 30 shows the graphs referring to the training of the groups containing only the smooth type coating.

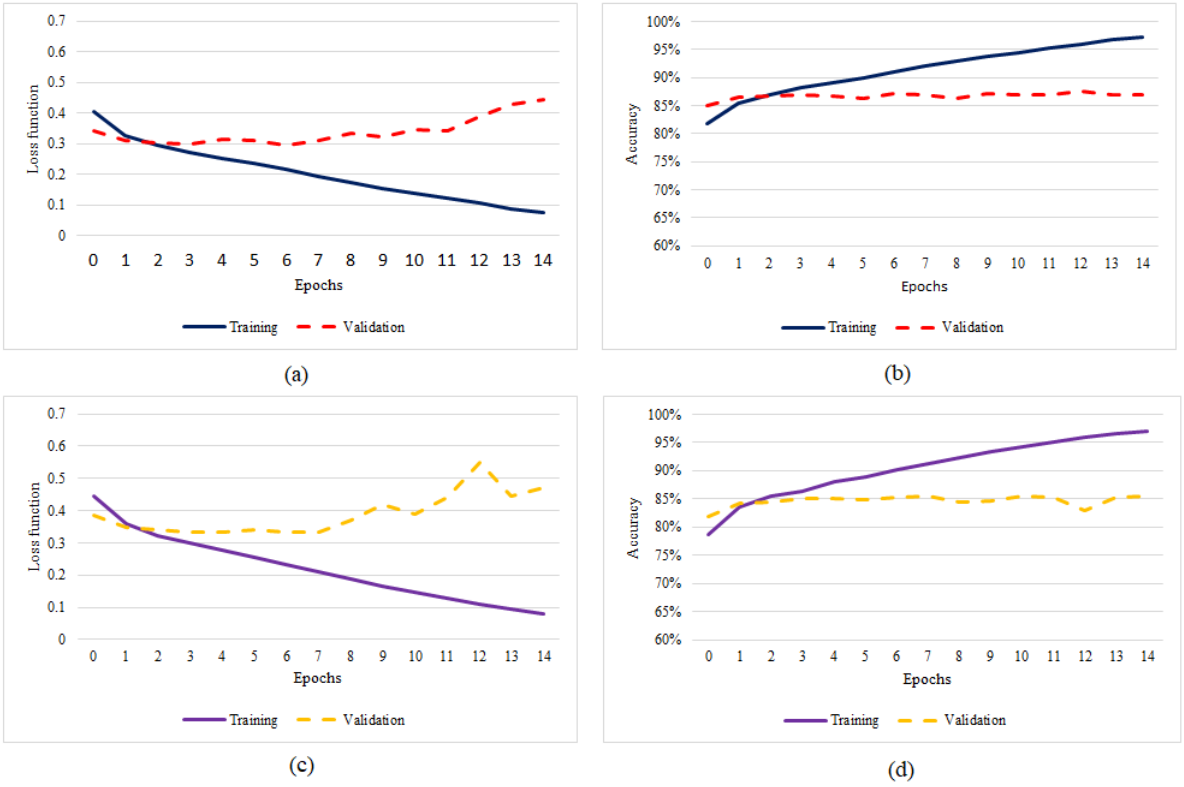


Figure 29. Graphs of loss function and accuracy groups with the whole database; (**a**) Epoch loss function of group A1 (unbalanced); (**b**) Accuracy by epochs of group A1 (unbalanced); (**c**) Epoch loss function of group B1 (balanced); (**d**) Accuracy by epochs of group B1 (balanced).

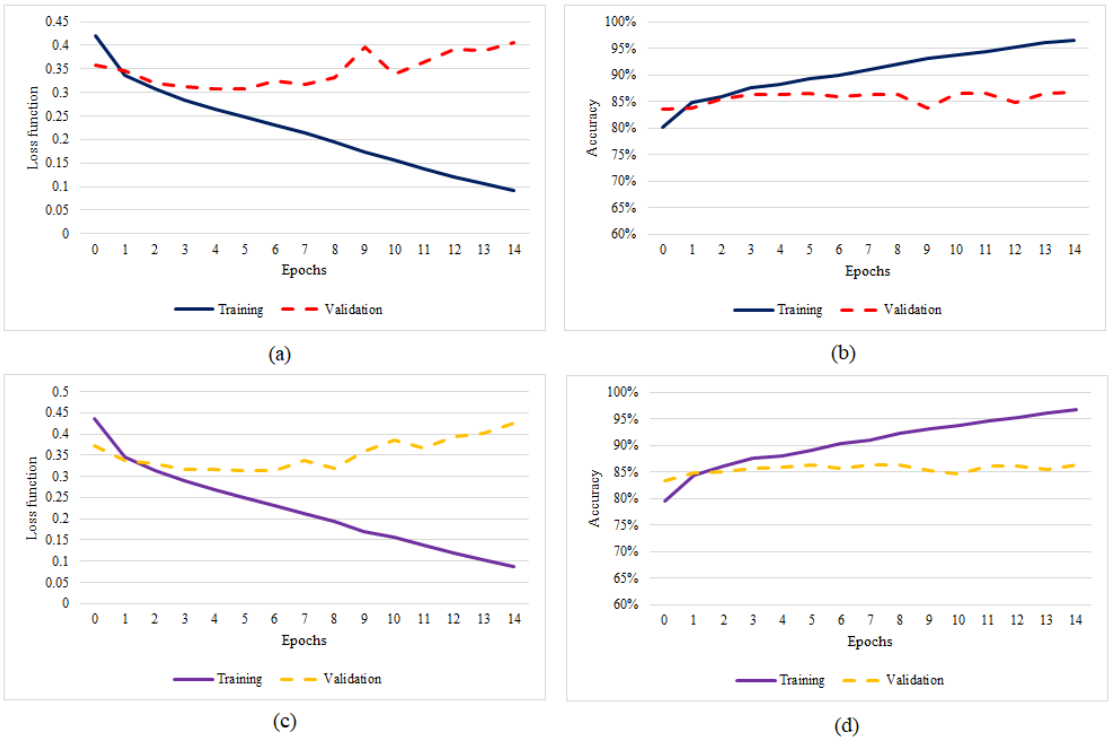


Figure 30. Graphs of the loss function and accuracy of groups containing only the smooth coating; (**a**) Epoch loss function of group A2 (unbalanced); (**b**) Accuracy by epochs of group A2 (unbalanced); (**c**) Epoch loss function of group B2 (balanced); (**d**) Accuracy by epochs of group B2 (balanced).

In the case of the smooth type coating, there was almost no variation between the A2 and B2 groups. The graphs of the loss and accuracy functions show the same behavior in both situations with accuracy being slightly above 85% during training. It turns out that the A2 group was already well balanced. Its R_m ratio was 0.957, very close to 1. In this way, the balancing process by the minority class caused almost no changes.

Another observed point is that the results of groups A2 and B2 are also similar to those of groups A1 and B1. This can be explained by the fact that the smooth type coating is the surface finish with the highest number in the database. It represents a total of 22,795 images in a universe of 33,088. In this way, it is possible that the smooth finish generates a tendency of the configurations in the networks trained with the whole database.

Figure 31 shows the data referring to the groupings of the scrapped type finish.

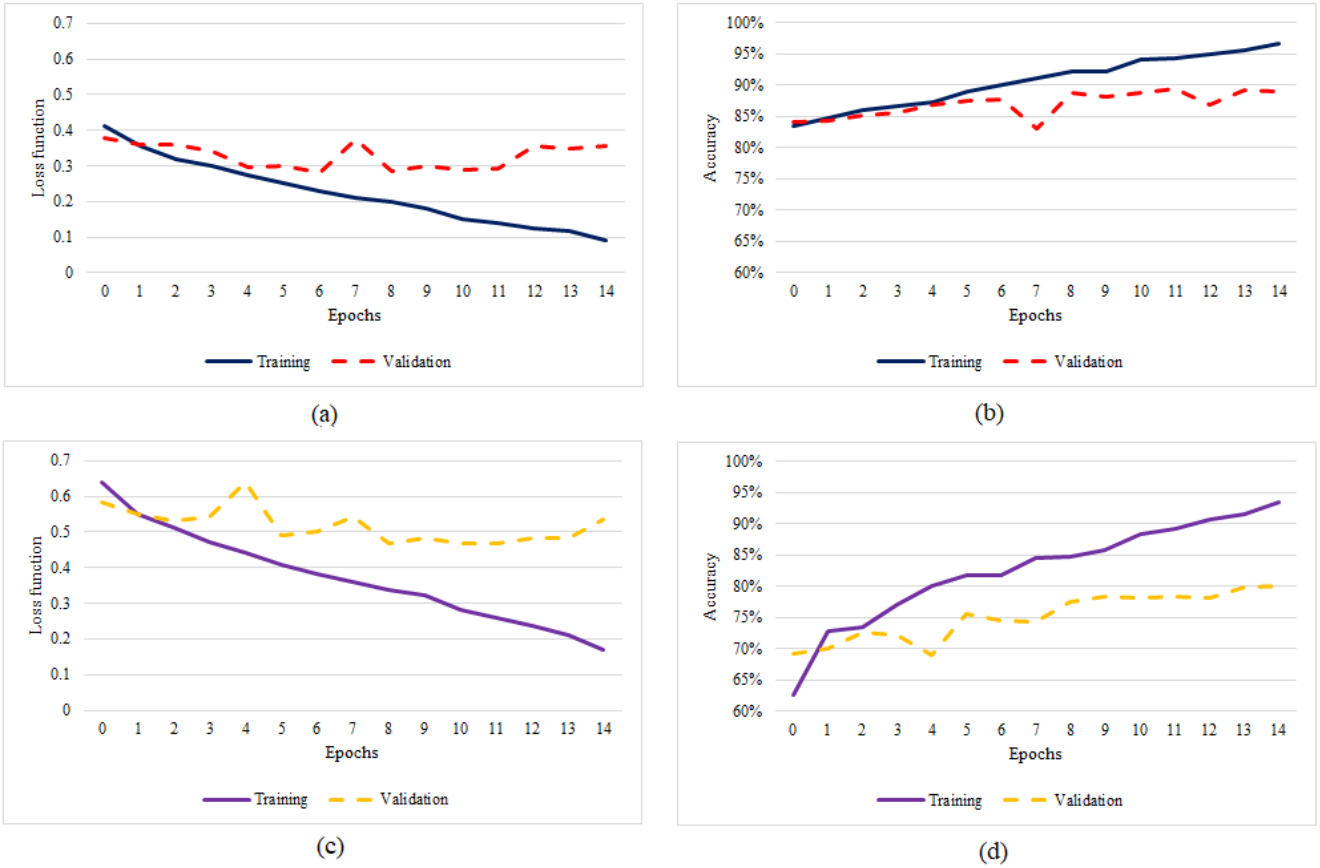


Figure 31. Graphs of the loss function and accuracy of groups containing only the scrapped coating; (**a**) Epoch loss function of group A3 (unbalanced); (**b**) Accuracy by epochs of group A3 (unbalanced); (**c**) Epoch loss function of group B3 (balanced); (**d**) Accuracy by epochs of group B3 (balanced).

In the case of scrapped finish, there was a difference between groups A3 and B3. The accuracy of group A3 was higher than that of group B3. While the A3 reached results close to 90%, the B3 group did not reach 80%. This difference can be explained by the imbalance of data. In the case of A3, its database has 4739 images without cracks against 985 images with cracks. In this way, the network can specialize in correcting images that do not have a crack. Thus, it will end up having better accuracy, as the volume of images that do not have cracks is much greater than those that do. In this way, the network will have more hits, even if it fails to classify the images with cracks. This situation is problematic, because the network ends up not having the ability to generalize. In this way, we have a weak classification model. A suitable process for evaluating this would be to also have a test group with images that were not used during training.

Another point is that the scrapped type finish has the highest number of images with low visibility, corresponding to 56% of the total number of images with clefts in the group.

The data referring to the training of the rough type finish can be seen in Figure 32.

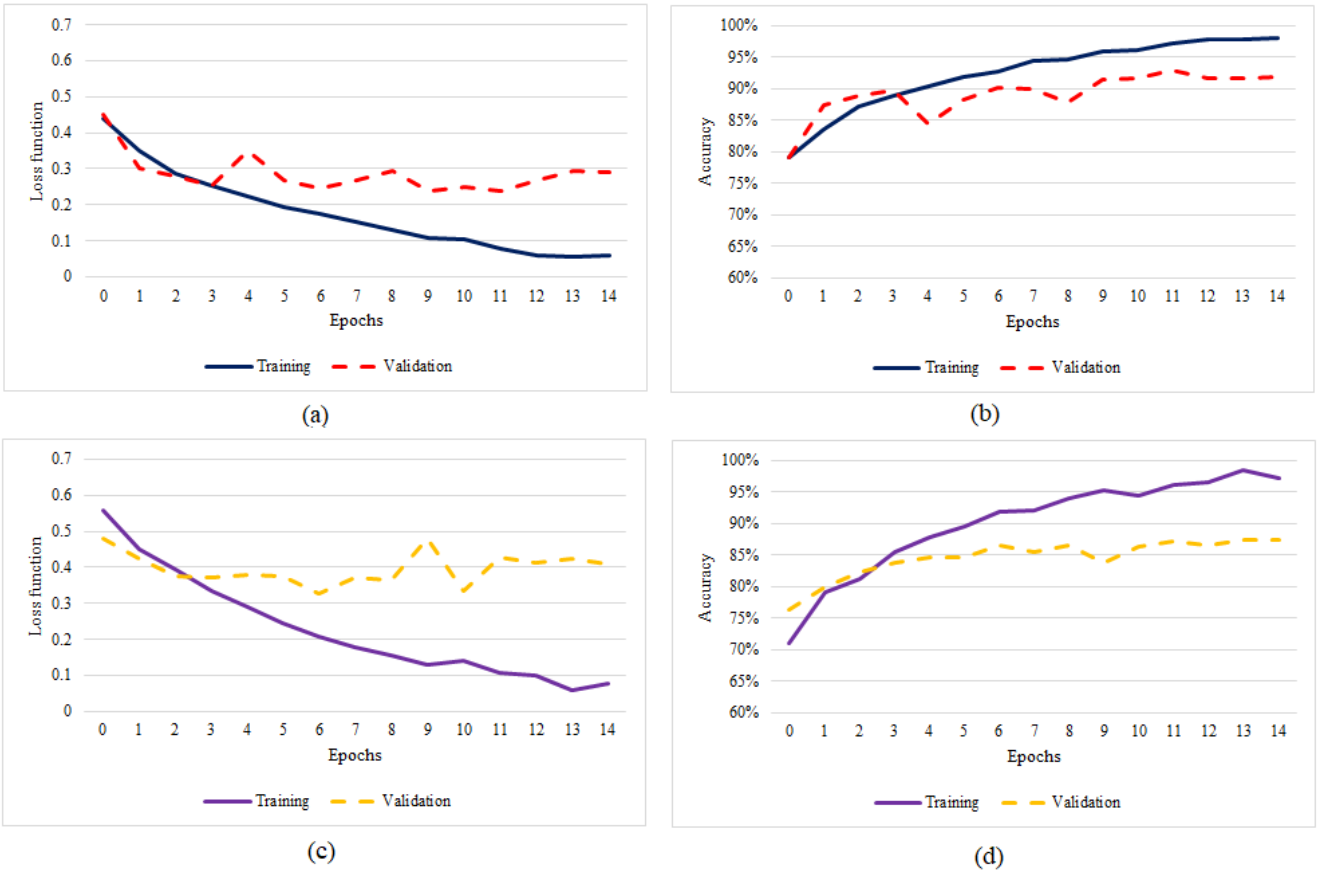


Figure 32. Graphs of the loss function and accuracy of groups containing only the rough type coating; (**a**) Epoch loss function of group A4 (unbalanced); (**b**) Accuracy by epochs of group A4 (unbalanced); (**c**) Epoch loss function of group B4 (balanced); (**d**) Accuracy by epochs of the B4 group (balanced).

The rough cast finish group also had a difference in accuracy between the A4 and B4 groups. In this case, the accuracy of the A4 group was above 90% while that of the B4 was below this value. The same condition of data imbalance can be observed in the scrapped group, although with a smaller variation in accuracy between A4 and B4. In this case, the A4 had a total of 1140 images with cracks and 3429 without discontinuities.

Regarding training, the loss function graph showed the same behavior for all groups. The error decreases to a certain point and then starts to increase. This indicates that overfitting is taking place. Thus, the data presented cannot be used as a criterion with which to evaluate the performance of the VGG16 for the elaborated database; however, they serve the purpose of the research to compare the different groupings and raise points that can be investigated. The overfitting problem could be minimized using tools such as early stopping and the dropout rate.

In this way, differences were observed between the training of the different groups. In Figure 33, the best accuracies of each group were plotted.

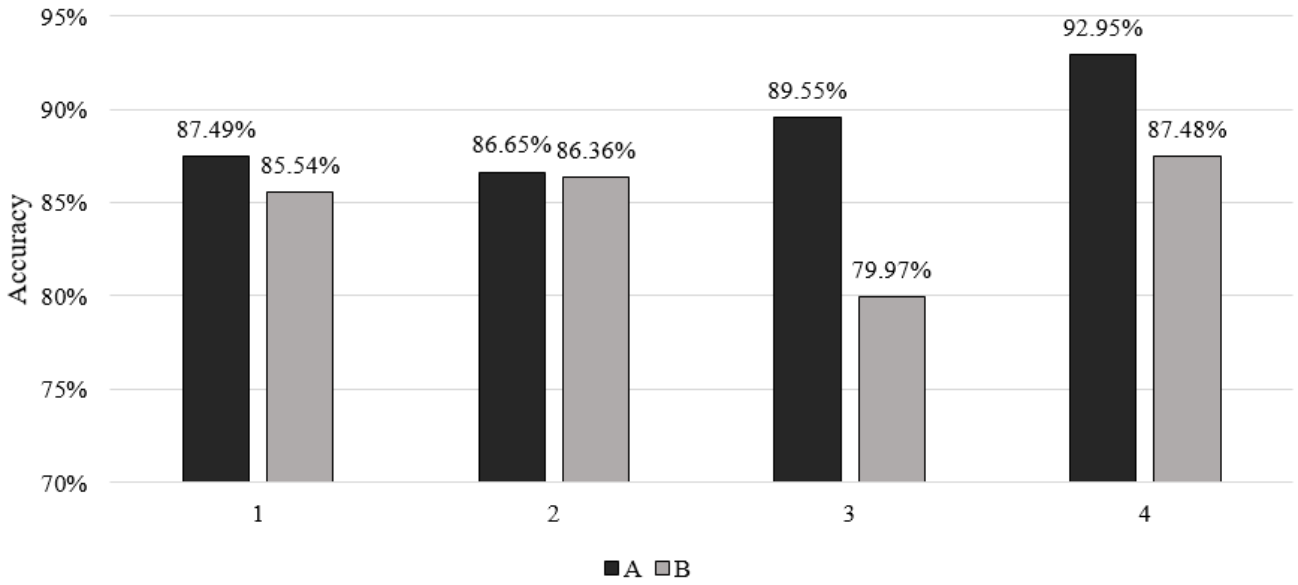


Figure 33. Best validation accuracies of each group. (1) Total bank; (2) Smooth; (3) Scrapped; (4) Rough.

In Figure 33, it is possible to evaluate all required situations more clearly. The first point is that almost all balanced groups had lower accuracy than the unbalanced ones, to a greater or lesser extent. Groups A2 and B2 had almost no variation in terms of better accuracy. Groups A1 and B1 had similar values to groups A2 and B2. Groups A3 and B3 had the greatest variation in accuracy between them, with A3 also corresponding to the most unbalanced database. Group A4 was the most accurate value, however, it was also unbalanced. In this case, the accuracy of group B4 was lower than that of group A4, but with a smaller variation than that which occurred between groups A3 and B3. The accuracy of group B4 was the best among the balanced groups and was also next to groups A1, B1, A2 and B2. This may indicate that the ability to identify cracks in a roughened surface finish is close to identifying in a smooth.

The scrapped type finishing was the one that presented the greatest differences between the training of groups A and B. In this case, it was a more unbalanced surface finishing type, which may explain the accuracy of 89.55% in group A3. This is because the network may be more accurate in the images of the class that do not contain a crack, which would increase the accuracy of the network, considering that the volume of images in this category is much greater than in the other. A tool that can help to evaluate this scenario would be the confusion matrix and the verification of the accuracy in each of the classes.

The accuracy of the B3 group was the lowest among all the trainings, 79.97%. This issue can be justified by the complexity of identifying cracks in the scrapped finish. The grooves for this type of finish are practically traces, which can be resemble to some types of fissures. This can be observed in the labeling of the data, in view of the large number of low visibility images within the group.

The A4 group of the rough class had the highest value among the accuracies—92.95%. However, it was also unbalanced, therefore, the same observations made for the surface finish of the scrapped type can also be applied to the rough cast finish, in this case. Group B4 had the best accuracy in relation to the balanced groupings, with 87.48%. This value was close to those obtained in the smooth type finish, which may indicate that the difficulty of identifying cracks is similar between these two types of finishes. In theory, the surface of the rough cast finish has a much rougher appearance than the smooth one; however, it does not have much information that can be confused with cracks, as is the case with the scrapped type. Another point is that, in the smooth finish, it is possible to verify that some textures also have a less uniform appearance, as seen in Figure 11.

The smooth type finish had similar results between the accuracy of groups A1 and B1. This can be explained by the fact that the A1 group was already well-balanced, with a variation of 501 images between the classes in a universe of 22,795. The surface of the smooth type would be the one that most resembled the surface of concrete; however, the results accuracies of A1 and B1 were 86.65% and 86.36%, respectively, which is below those of concrete studies like Chow et al. [25], Ali et al. [26], Islam et al. [27], and Chaiysarn et al. [28]. This may indicate that the application of computer vision in mortar coating images is actually more complex than in the case of concrete.

The groupings that contained all surface finishes had close values between unbalanced and balanced. A1 had an accuracy of 87.49% and B2 of 85.54%. These values were close to those of the A2 and B2 groups. As most of the images in the database were of the smooth type coating, it is possible that the network tended to hit this type of coating. This could be evaluated using the confusion matrix.

Figure 34 shows the training times in relation to each group.

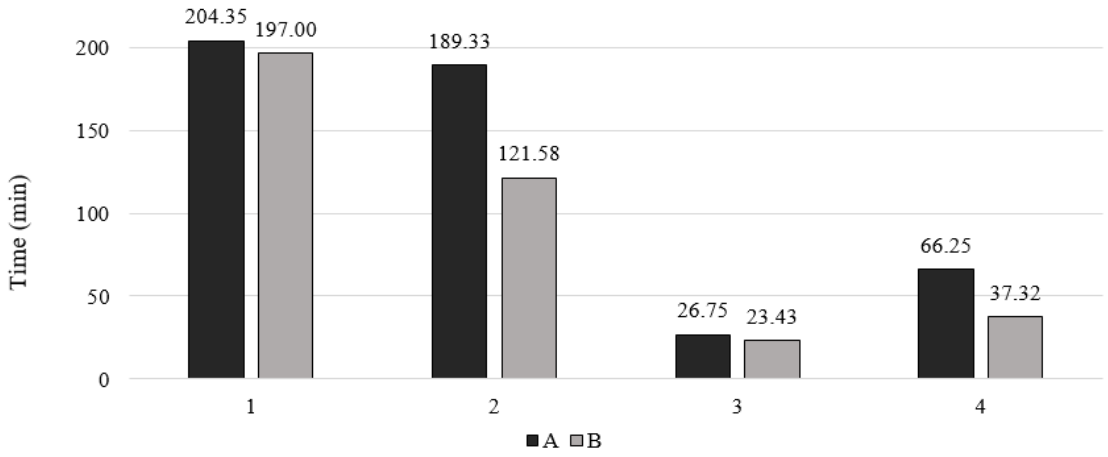


Figure 34. Training time of each group. (1) Total bank; (2) Smooth; (3) Scrapped; (4) Rough.

It is possible to observe that the training time varies between groups. This is justifiable because the groupings have different amounts of data. In view of this situation, the training time was plotted in relation to the number of images used to carry out the process. This can be seen in Figure 35.

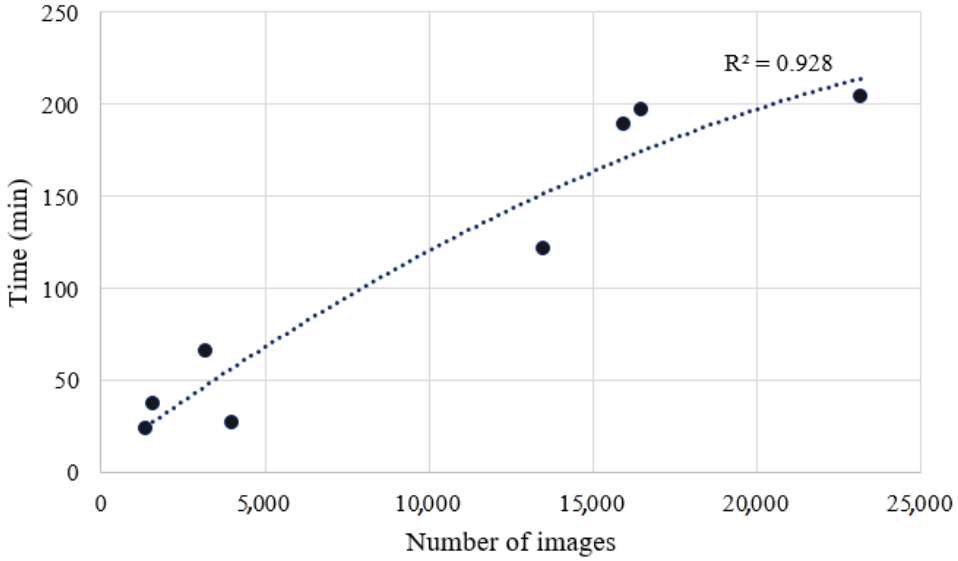


Figure 35. Training time by number of images.

It was verified that the training time increases according to the number of images. By adding a polynomial trend line, it was possible to obtain a coefficient of determination (R^2) of 0.9279.

Thus, it was possible to observe several relevant points to be investigated using the database proposed from AI approaches.

4. Conclusions

With this study, it was possible to verify issues that have not yet been addressed in the literature, such as the influence of the surface finish of a mortar coating on the detection of cracks. For this purpose, transfer learning was applied to the VGG16 network in order to train a classifier.

This study also presented a new public database with cracks images in mortar coating with different types of surface finishes. Several types of noise were observed in the database, such as the presence of background in the images, vegetation, encounters between elements, windows, doors, electrical and hydrosanitary installations, shadows, graffiti, and various objects. The surface finishes showed variability within each type as in relation to their textures and colors.

The images in the database present cracks with different types of configurations and levels of visibility. In addition, the database is unbalanced in all classes delimited in the labeling process. The scrapped group was the most unbalanced in relation to the presence of cracks and was also the one that presented more images in which it was difficult to identify the discontinuities. Because of this, overfitting was observed in all training sessions.

Groups A1, A2, B1, B2, and B4 showed close accuracy values ranging between 85.5% and 87.5%. The rough and scrapped finishes were the ones that showed the greatest difference between the accuracy values of the balanced and unbalanced groups. The group with the balanced scrapped coating, B3, had the lowest accuracy value, with 79.9%.

One of the main limitations of the research was that the training of the VGG16 was carried out without a test group and the analysis was carried out based on the training accuracy. The most appropriate approach was to use the trained network to classify images from a group that was not used during the training stage, which would be the test group. Another issue is that only the accuracy metric was analyzed; in order to have richer discussions, it would be necessary to evaluate other assertiveness metrics. However, the training served its purpose of raising possible points to be studied with the proposed database.

Thus, it is possible to conclude that the type of mortar coating influences the training performance of a CNN. The proposed database has 33,088 images and important characteristics in relation to the possible variability to be found in inspections of facades. Therefore, it presents potential for further civil and computer vision applications.

The database developed in this study can be accessed at the link available in the Data Availability Statement at the end of the article or in [35]. An explanation of how the database was divided into the folders within the link can be found in Appendix A.

Author Contributions: R.d.A.G.S. elaborated the research, developed the procedure, supervised and performed data labeling, conducted training, and wrote the text of the article; F.P.N. realized with data labeling and revised the final version of the article; H.F. guided the research, revised the article throughout the writing process and the final version. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001.

Data Availability Statement: The database created in this article is available in: https://doi.org/10.34740/KAGGLE/DSV/6091986. Accessed on 20 July 2023.

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

Appendix A

The database was divided into training, validation, and test groups in the proportions: 70% for training, 20% for validation, and 10% for testing. This configuration is the most appropriate, as it makes it possible to test the network on a database that was not used in training.

To obtain the proportions of 70% for training and 30% for validation used in this article, just group the validation and test folders.

In the database link, you can find two folders called: All_Database and Split_Database.

The All_Database folder contains the database with all the images, divided into training, validation, and test groups.

In the Split_Database folder, the database is divided into the following folders:

- CR—Which contains the images of the surface finish of the rough type with noise;
- CS—Which contains the images of the surface finish of the rough type without noise;
- GR—Which contains the images of the surface finish of the type scrapped with noises
- GS—Which contains the images of the surface finish of the type scrapped without noise;
- LR—Which contains the surface finish images of the smooth type with noise;
- LS—Which contains the images of the surface finish of the smooth type without noise.

In this way, it is possible to carry out different combinations by joining the respective folders, such as, for example, grouping the CS and CR folders to have a set that only has the slate type finish. All folders are divided between training, validation, and testing groups, in the specified proportions.

References

1. Schwab, K. *The Fourth Industrial Revolution*; World Economic Forum: Cologny, Switzerland, 2016.
2. Taulli, T. *Artificial Intelligence Basics: A Non-Technical Introduction*; Apress: Monrvia, CA, USA, 2019.
3. Vinod Chandra, S.S.; Anand Hareencran, S. *Artificial Intelligence and Machine Learning*; PHI Learning: New Delhi, India, 2014.
4. Tan, K. The framework combining artificial intelligence and construction 3D printing in civil engineering. *MATEC Web Conf.* **2018**, *206*, 01008. [CrossRef]
5. Zhang, Y. Safety management of civil engineering construction based on artificial intelligence and machine vision technology. *Adv. Civ. Eng.* **2021**, *2021*, 3769634. [CrossRef]
6. Palencia, S.; Díaz, S.Á.; Chicote, M.A. Machine learning for the improvement of deep renovation buildings projects using As-Built BIM models. *Sustainability* **2021**, *13*, 6576. [CrossRef]
7. Prieto, S.A.; Mengiste, E.T.; Soto, B.G. Investigation the Use of Chat GPT for the Scheduling of Construction Projects. *Buildings* **2023**, *13*, 857. [CrossRef]
8. Chang, D.W.M. Sustainable building maintenance for safer and healthier cites: Effectiveness strategies for implementing the Mandatory Building Inspection Scheme (MBIS) in Hong Kong. *J. Build. Eng.* **2019**, *24*, 100737. [CrossRef]
9. Silva, A.; Brito, J. Do we need a buildings' inspection, diagnosis and service life prediction software? *J. Build. Eng.* **2019**, *22*, 335–348. [CrossRef]
10. Garcia Sobrinho, R.A.; Motta, L.A.; Guterres, P.R.C. Analysis of the incidence of pathological manifestation in public buildings who have gone through rehabilitation processes. In Proceedings of the XVI Congreso Latinoamericano de Patología de la Construcción y XVII de Control de Calidad en la Construcción, Online, Brasil, 21 October 2021. [CrossRef]
11. ImageNet. Available online: https://www.image-net.org/index.php (accessed on 1 June 2023).
12. Silva, L.F.; Saade, D.C.M.; Sequeiros, A.C.; Silva, A.C.; Paiva, A.C.; Bravo, R.S.; Conci, A. A new database for breast research with infrared images. *J. Med Imaging Health Inform.* **2014**, *4*, 92–100. [CrossRef]
13. Katija, K.; Orestein, E.; Schining, B.; Lundsten, L.; Barnard, K.; Sainz, G.; Boulais, O.; Cromwell, M.; Woodward, B.; Bel, K.L.C. FathomNet: A global image database for enabling artificial intelligence in the ocean. *Sci. Rep.* **2022**, *12*, 15914. [CrossRef]
14. Leenhardt, R.; Li, C.; Mouel, J.; Rahmi, G.; Saurin, J.C.; Chotel, F.; Boureille, A.; Amiot, X.; Delvaux, M.; Duburque, C.; et al. CAD-CAP: A 25,000-image data-set serving the development of artificial intelligence for capsule endoscopy. *Endosc. Int. Open* **2020**, *8*, E415–E420. [CrossRef]
15. Li, B.; Wang, K.C.P.; Zhang, A.; Yang, E.; Wang, G. Automatic classification of pavement crack using deep convolutional neural network. *Int. J. Pavement Eng.* **2017**, *21*, 457–463. [CrossRef]
16. Zhang, C.; Nateghinia, E.; Miranda-Moreno, L.F.; Sun, L. Pavement distress detection using convolutional neural network (CNN): A case study in Montreal, Canada. *Int. J. Transp. Sci. Technol.* **2022**, *11*, 298–309. [CrossRef]
17. Jiang, Y.; Pang, D.; Li, C. A deep learning approach for fast detection and classification of concrete damage. *Autom. Constr.* **2021**, *128*, 103785. [CrossRef]

18. Dung, C.V.; Anh, L.D. Autonomous concrete crack detection using deep fully convolutional neural network. *Autom. Constr.* **2019**, *99*, 52–58. [CrossRef]
19. Sonntag, H.; Grunewald, J. Methodology for praxis-oriented development of a building refurbishment concept including consideration of potentially existing moisture related problems and façade restoration measures. *E3S Web Conf.* **2020**, *172*, 23008. [CrossRef]
20. Tomal, E.S. The problem of biological destruction of façades of insulated buildings—Causes and effects. *IOP Conf. Ser. Mater. Sci. Eng.* **2017**, *245*, 032012. [CrossRef]
21. Edis, E.; Colen, I.; Brito, J. Passive thermographic moisture problems in façades with adhered ceramic cladding. *Constr. Build. Mater.* **2014**, *51*, 187–197. [CrossRef]
22. Lee, K.; Hong, G.; Sael, L.; Lee, S.; Him, H.Y. MultiDefectNet: Multi-class defect detection of building façade based in deep learning convolutional network. *Sustainability* **2020**, *12*, 9785. [CrossRef]
23. Guo, J.; Wang, Q.; Li, Y.; Liu, P. Façade defects classification from imbalanced dataset using meta learning -based convolutional network. *Comput. Aided Civ. Infrastruct. Eng.* **2020**, *17*, 1403–1418. [CrossRef]
24. Guo, J.; Wang, Q.; Li, Y. Semi-supervised learning based on convolutional neural network and uncertainty filter for façade defects classification. *Comput. Civ. Infrastruct. Eng.* **2020**, *36*, 302–317. [CrossRef]
25. Chow, J.K.; Su, Z.; Wu, J.; Li, Z.; Tan, P.S.; Liu, K.; Mao, X.; Wang, Y. Artificial intelligence-empowered pipeline for image-based inspection of concrete structures. *Autom. Constr.* **2020**, *120*, 103372. [CrossRef]
26. Ali, L.; Alnajjar, F.; Jassmi, H.A.; Cocfo, M.; Khan, W.; Serhani, M.A. Performance evaluation on deep CNN-Based crack detection and localization techniques of concrete structures. *Sensors* **2021**, *21*, 1688. [CrossRef] [PubMed]
27. Islam, M.M.; Hossain, B.; Akhtar, N.; Moni, M.A.; Hasan, K.F. CNN based on transfer learning models using data argumentation and transformation fir detection of concrete crack. *Algorithms* **2022**, *15*, 287. [CrossRef]
28. Chaiysarn, K.; Buatik, A.; Moramad, H.; Zhou, M.; Kongsilp, S.; Poovarodom, N. Integrate pixel-level CNN-FCN crack detection via photogrammetric 3D texture mapping of concrete structures. *Autom. Constr.* **2022**, *140*, R713–R715. [CrossRef]
29. Weidman, S. *Deep Learning from Scratch: Building with Python from First Principles*; O'Reilly Media: Sebastopol, Russia, 2019.
30. Zeiler, M.D.; Fergus, R. Visualizing and Understanding Convolutional Networks. *arXiv* **2013**, arXiv:1311.2901. [CrossRef]
31. Chollet, F. *Deep Learning with Python*, 2nd ed.; Manning: New York, NY, USA, 2021.
32. Gonçalves, C.B.; Spuza, J.R.; Fernandes, H. CNN architecture optimization using bio-inspired algorithms for breast cancer detection in infrared images. *Comput. Biol. Med.* **2022**, *142*, 105205. [CrossRef]
33. Chaves, E.; Gonçalves, C.B.; Albertine, M.K.; Lee, S.; Jeon, G.; Fernandes, H.C. Evaluation of transfer learning of pre-trained CNNs applied to breast detection on infrared images. *Appl. Opt.* **2020**, *59*, E23–E28. [CrossRef]
34. *NBR 15575-4*; Residential Buildings—Performance, Part 4: Requirements for Internal and External Vertical Sealing Systems. *NBR 15575-4*; Brazilian Association of Technical Standards (ABNT): Rio de Janeiro, Brazil, 2021.
35. Garcia Sobrinho, R.A.; Piauhy Neto, F.; Fernandes, H. (Database) Cracks Images in Mortar Coating. Kaggle. 2023. Available online: https://www.kaggle.com/datasets/garciasobrinho/database-cracks-images-in-mortar-coating (accessed on 16 July 2023).

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.