Data augmentation approach in detecting roof pathologies with UASs images

L Staffa[1], J Nogueira[2], M Lima[1], A Ottoni[3], D B Costa[1] and M Novo[2]

[1]Department of Construction and Structures, Federal University of Bahia, Salvador, BA, Brazil
[2]Department of Electrical and Computer Engineering, Federal University of Bahia, Salvador, BA, Brazil
[3]Technologic and Exact Center, Federal University of Recôncavo da Bahia, Cruz das Almas, BA, Brazil

Abstract. Machine learning and computer vision techniques contribute to the automation roof pathologies identification from images collected with Unmanned Aerial System (UASs). However, one of the challenges for practical machine learning model tuning is the small-data problem. One strategy is to adopt data augmentation for generating more training data from existing images. This paper evaluates data augmentation in detecting pathologies in roof inspections with UASs images. The study adopted data augmentation for training two models in an image processing system. The training and tests using data augmentation images obtained superior results in accuracy, precision, recall, F-score, negative precision, and specificity metrics compared to the study using only original photos. These results indicate that data augmentation improves the adopted system's performance in identifying roof pathologies in UAS images. This inspection system proposed with such integrated technologies would make it possible to increase transparency, simplify steps and reduce the time to perform roof inspections, streamlining the preparation of reports and application of corrective actions.

1. Introduction

The roof inspection is a periodical activity aiming to ensure that needed repairs are made on time, thus avoiding severe damage to the roof structure and other building parts. However, these inspections are often neglected due to the time required for assessment, labor cost, difficulty accessing these structures, and safety risks [1].

One solution for buildings roofs inspection is Unmanned Aerial System (UAS) for image capture [2]. UASs can provide a large amount of data in a short period, allowing the visualization of pathologies in hard-to-reach areas, safely and at low cost, when compared to manual visual inspection methods [3].

Associating the large number of images collected with the UAS, machine learning, and computer vision can contribute significantly to image processing for the automatic identification of pathologies on roofs [4]. However, one of the main challenges in applying Deep Learning is using datasets with few image examples [5, 6]. A few training examples can lead the ML model to low accuracy rates in the classification process [5].

The application of Deep Learning with Data Augmentation for image processing is an area with significant results in recent literature [6]. The data augmentation approach makes it possible to generate artificial images from the original data [5]. The dataset of new photos is added to the training process and improves Deep Learning methods' generalization ability. Thus, the original images are transformed by applying rotation, zoom, and shear features.

Therefore, this study proposes a data augmentation approach to increase the database for identifying pathologies in roof images obtained with UASs. The artificial images tend to increase the generalization capacity of the trained algorithm, thus improving the system's performance in identifying non-conformities in new photos.

2. Literature Review

Razali et al. [7] and Kaamin et al. [8] used UAS to inspect historic buildings (facade and roof) in Malaysia. In the inspections performed in these two studies, the auto-capture and photogrammetry methods were used on the images, allowing defects to be identified and referenced. Razali et al. [7] and Kaamin et al. [8] identified pathologies such as corrosion, vegetation growth, and broken tiles.

Silveira et al. [3] evaluated the use of UAS for roof inspection on post-occupation residential buildings, aiming to assist in the decision-making process. The authors assessed items that can be viewed with UAS and developed a manual flight mode protocol for flight planning and data collection with UAS. Bown e Miller [2] recommends the use of individual still images process for sloped roof inspection with UAS, for being faster than other options (image stitching, video recording, streaming), also providing evidence in a manageable format (small individual photos). According to Silveira et al. [3], UAS use allows to increase transparency, simplify steps, decrease the time to carry out roof inspections, speed up the preparation of reports, and apply corrective actions.

Staffa et al. [4] tested and validated Custom Vision software, from Microsoft, as an image processing tool to identify roof pathologies from images captured with UAS. Microsoft Custom Vision uses Artificial Intelligence and Deep Learning techniques in the computer vision task. Classes of roof pathologies were trained and tested with the image processing algorithms in a database of 1661 images collected with UAS from 61 roofs of buildings in use. The results indicated 72% identification of the non-conformities analyzed, with 11,5% error.

Some recent work has already applied Data Augmentation to increase the number of images in the training database [9, 6]. Bang et al. [6] used three Data Augmentation techniques to train a Convolutional Neural Network (R-CNN type) to detect construction elements, such as workers, crane trucks, and concrete trucks, in 656 original images collected with a UAS. According to the authors, the application of Data Augmentation techniques improved the recall, precision, and f-measure indexes. Dung et al. [9] also used Data Augmentation in the crack detection task using Deep Learning. The increased number of training images improved the accuracy up to 5% in detecting cracks in steel bridges.

According to Chollet and Allaire [5], one way to apply Data Augmentation is to use a framework for automatically generating artificial images, such as R software in conjunction with the Keras library. The Keras library has implemented the image_data_generator function for Data Augmentation processing. This method makes it possible to apply more than ten types of random transformations to the images, thus generating new artificial data.

3. Research Methodology

Figure 1 shows the main research steps adopted in this research.

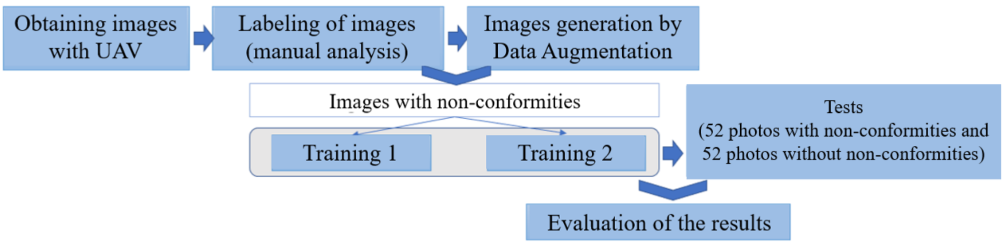

Figure 1. Research design.

The database adopted includes database with 1661 images and 1743 new images collected with UAS, totalizing 3404 photos [4]. These photos were from 173 sloped roofs with fiber-cement tiles from 11 residential buildings in use, located in the Metropolitan Area of Salvador in Brazil. A DJI Phantom 4 drone and the inspection protocol defined by Silveira et al. [3] were used for image collection.

The inspection protocol defined by Silveira et al. [3] consists of flying over and obtaining photos of the entire roof surface at the height of approximately 6 meters above the roof to ensure the quality of the resolution of the images. In addition, the UAS has to fly a 360º from the center of the top and continue taking the photos. This technique allows obtaining a large number of images in a short time. For the flight with UAS, it is necessary to prepare a flight plan, identify the best trajectory to be followed by the UAS, manage the flight time, plan the replacement of the UAS batteries, and finally, carry out the flight get the roof shots.

In the next step, the authors labeled the images from the database with the identification of non-conformities. A total of 2673 non-conformities were identified in 913 photos, and 52 photos were separated for testing. It is worth noting that finding more than one type of non-conformity per image is possible.

For the classes of non-conformities with less than 120 images, Data Augmentation was used to increase the available database. This technique was used in 6 of the 11 categories, presented in bold in Table 1. The Data Augmentation application was developed in R software, adopting the Keras library [5]. Algorithm 1 presents the proposed code for generating artificial images based on Chollet and Allaire [5].

Algorithm 1 Data Augmentation approach using Keras library and R Software. Based on Chollet and Allaire [5]

```
1   #Stage 1 – Data Augmentation Settings
2   library(keras)
3   base_dir <- "C:/Data/"
4   train_datagen<- image_data_generator(
5           rescale = 1/255,
6           rotation_range = 40,
7           zoom_range = 0.2,
8           horizontal_flip = TRUE,
9           vertical_flip = TRUE,
10   )
11   nrep <- 5
12   nimg <- 20
13  #Stage 2 – Loop of Image Generator
14   cont<-0
15   for (j in 1:nrep)
16     for (i in 1:nimg) {
17       cont <- cont +1
18       fnames <- list.files(base_dir, full.names = TRUE)
19       img_path <- fnames[[i]]
20       img <- image_load(img_path, target_size = c(1200, 1900))
21       img_array <- image_to_array(img)
22       img_array <- array_reshape(img_array, c(1, 1200, 1900, 3))
23       augmentation_generator <- flow_images_from_data(
24          img_array,
25          generator = train_datagen,
26          batch_size = 1
27       )
28       batch <- generator_next(augmentation_generator)
29       t <-paste(cont,".jpeg")
30       jpeg(filename = t,width = 1200, height = 900, units = "px")
31       plot(as.raster(batch[1,,,]))
32       dev.off()
33     }
34   }
```

Algorithm 1 consists of two stages: Data Augmentation Settings and Loop of Image Generator. In the first phase (lines 1 to 12), the necessary settings for the generation of new images are performed, such as the declaration of the Keras library, an indication of the folder for reading the images, the number of repetitions (nrep), and the number of original images (nimg). In addition, the image_data_generator() function is also declared, responsible for defining the adopted random transformations. In this sense, five transformations were used in the images: (1) rescale - normalizes the pixel values of the images between 0 and 1; (2) rotation_range - defines a random rotation of up to 40 degrees; (3) zoom_range - zooms the image up to 20%; (4) horizontal_flip - whether to randomly flip images horizontally; (5) vertical_flip: whether to randomly flip images vertically.

Then, in the second stage (lines 13 to 34), the artificial image generation loop's commands are executed. In this aspect, the function flow_images_from_data stands out, responsible for reading the data and applying the image generator. Thus, it was possible to generate 733 artificial images of roofs with non-conformities. Figure 2 presents examples of images generated by Data Augmentation.

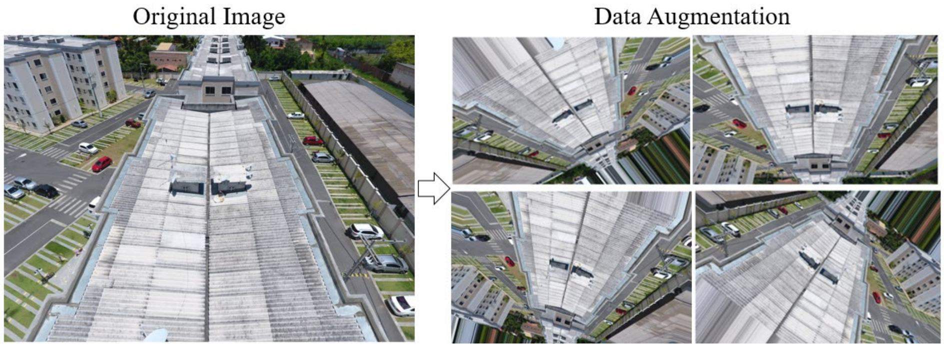

Figure 1. Images Generation by Data Augmentation.

In the fourth stage, two Custom Vision training sessions were performed using the original images and generated. The Custom Vision is an artificial intelligence service that allows customized image classifier models using pre-trained neural networks (ResNet and AlexNet). The object detector feature was used, in which the user labels classes to train the algorithm then. Table 1 presents the number of photos used in each training and the equivalent percentage of original and artificially generated images. Trained the algorithm only using authentic images [4], the photos of the classes "proper arrangement of antennas and wires" and "integrity of tiles (broken)" were not counted because the Custom Vision requires a minimum of 16 photos to perform the class training and there was not this minimum amount.

Table 1. Number of photos used in training per classes of non-conformities

Classes	**Staffa (2020) [4]**	**Training 1**			**Training 2**		
	No. Photos (Total)	No. Photos (Total)	% Original images	% Generated images	No. Photos (Total)	% Original images	% Generated images
1. Gutter cleanliness	182	303	69%	31%	656	41%	59%
2. Accumulation of algae, lichens, and mosses	157	157	100%	0%	401	50%	50%
3. Presence of residues on the roof	117	144	90%	10%	423	73%	27%
4. Flashing's integrity	93	93	100%	0%	233	92%	8%
5. Gutter integrity	49	55	100%	0%	90	59%	41%
6. Sealing the meeting between flashings	47	137	34%	66%	228	22%	78%
7. Presence of extra tile on the roof	45	51	90%	10%	118	96%	4%
8. Trapdoor cover open	33	33	100%	0%	141	27%	73%
9. Poor fastening of the flashing	16	73	26%	74%	136	15%	85%
10. Proper arrangement of antennas and wires	13	13	-	-	120	0%	100%
11. Integrity of the tiles (broken)	5	5	-	-	127	5%	95%
TOTAL	757	1064	-	-	2673	-	-
AVERAGE	-	-	79%	21%	-	44%	56%

Therefore, the models generated in the three training sessions were tested using 104 images, 52 of which have no problems (negative class), and the other 52 have non-conformities (positive class). The randomness criterion was adopted for the selection of these images.

The evaluation of the results includes the analysis of training and test images. The training data were evaluated to identify (1) the capacity of the trained models to identify non-conformities and (2) the influence of the images generated by Data Augmentation in-class identification. The analysis of the test data aims to: (1) verify the behavior of trained models in a situation where there were no pathologies in the image (negative class) and, (2) evaluate the trained models' ability to recognize at least one pathology in a positive class photograph, adopting a minimum detection rate of 30%.

For this analysis, six evaluation metrics were used (Silva et al., 2017): accuracy, precision, recall, F-score, negative accuracy, and specificity, according to Equations:

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

$$Precision = \frac{TP}{TP + FP} (2)$$

$$Recall = \frac{TP}{TP + FN} (3)$$

$$F - score = \frac{TP}{TP + FN} (4)$$

$$Negative\, precision = \frac{TN}{TN + FN} (5)$$

$$Specificity = \frac{TN}{TN + FP} (6)$$

Where, TP (true positives) refers to the identification of at least one non-conformity in a positive class image; FN (false negatives) refers to not identifying any problems when the roof has any pathology; TN (true negatives) is the indication that there are no pathologies on a roof without problems (negative class); FP (false positives) is the detection of non-conformities in negative class roof images.

Among these measures, the most general metric is accuracy. Accuracy calculates the percentage of correct predictions by dividing the number of accurate predictions by the total number of observations. Other measures also have the number of true positives identified, such as recall, precision, and F-Score. The recall measures the capacity to identify true positives (non-conformities) concerning the number of false negatives (unidentified non-conformities). The precision measures the percentage of hits for true positives (non-conformities) concerning the number of false positives. The F-score is a metric that relates recall to precision. Finally, the specificity and negative precision measures evaluate the classifications with the factor of the number of true negatives (roofs without pathologies) and the number of false positives and false negatives, respectively.

4. Findings and Discussion

Table 2 presents the results of the training performed in Custom Vision. The values in bold represent the Precision and Recall values that were improved after applying the Data Augmentation (Training 1 and Training 2), compared to the training performed only with original images [4].

Table 2. Results of the Custom Vision model training (precision and recall).

Classes	Staffa (2020) [4]		Training 1		Training 2	
	Precision	Recall	Precision	Recall	Precision	Recall
Gutter cleanliness	62.10%	29.50%	**87.90%**	**39.70%**	**80.50%**	22.90%
Accumulation of algae, lichens, and mosses	73.50%	56.80%	**81.80%**	28.10%	70.80%	19.80%
Presence of residues on the roof	82.10%	51.10%	**84.20%**	**56.10%**	**86.20%**	40.30%
Flashing’s integrity	58.10%	38.30%	**83.30%**	37.00%	**87.50%**	25.00%
Gutter integrity	77.80%	43.80%	50.00%	17.60%	0.00%	0.00%
Sealing the meeting between flashings	42.90%	6.80%	**50.00%**	**9.80%**	**66.70%**	4.90%
Presence of extra tile on the roof	75.00%	66.70%	**100.00%**	**80.00%**	**90.00%**	**75.00%**
Trapdoor cover open	100.00%	57.10%	83.30%	**83.30%**	91.30%	**75.00%**
Poor fastening of the flashing	0.00%	0.00%	**80.00%**	**25.00%**	0.00%	0.00%
Proper arrangement of antennas and wires	-	-	-	-	**100.00%**	**41.20%**
Integrity of the tiles (broken)	-	-	-	-	**87.50%**	**53.80%**
Average	**63.50%**	**38.90%**	**77.83%**	**41.84%**	**69.14%**	**32.54%**

After Training 1, the accuracy improved to 7 classes of non-conformities compared to Staffa et al. (2020). The detection of "Presence of extra tile on the roof" stands out achieved 100% precision, increasing 25% to the non-use of artificial images. For Training 2, the precision was higher by five classes than Staffa et al. (2020). Moreover, the generation of new images allowed the training of two new classes ("Proper arrangement of antennas and wires" and "Integrity of shingles"). A higher precision value was achieved in both classes, being 100% and 87.5%, respectively.

Concerning the recall metric, comparing the models trained only with original images (Staffa et al., 2020), six classes for Training 1 and two classes for Training 2 achieved better performance. The recall results stand out for the "Trapdoor cover open " class, 83.3% in Training 1 and 75% in Training 2, against only 57.1% in training without using images generated by Data Augmentation.

Table 3 presents the six metrics results (Equations 1 to 6) adopted to evaluate the test phase. For all evaluation metrics, the models trained with images generated by Data Augmentation (Training 1 and Training 2) presented superior results concerning the results obtained at Staffa et al. (2020). Training 2 achieved the best Accuracy (75.00%), Precision (77.08%), and Specificity (78.85%) scores. Training 1 presented better values for recall (80.77%), F-score (74.34%), and Negative Precision (76.74%). Considering the average of all the indicators, Training 2, which was trained with the highest proportion of generated images (56%), performed better than both Training 1, which was performed with only 21% of rendered images, and Staffa et al. (2020) trained with only original photos.

Table 3. Results of the indicators after testing the models.

Metric	Staffa (2020) [4]	Training 1	Training 2
Accuracy	72.12%	72.12%	**75.00%**
Precision	69.49%	68.85%	**77.08%**
Recall	78.85%	**80.77%**	71.15%
F-score	73.87%	**74.34%**	74.00%
Negative Precision	75.56%	**76.74%**	73.21%
Specificity	65.38%	63.46%	**78.85%**
Average	72.54%	72.71%	**74.88%**

Table 3 highlights the results of two metrics: accuracy in training 2 (77.08%) and recall in training 1 (80.07%). The increase in accuracy in training 2 implies a greater tendency to identify pathologies in non-conforming roofs. In addition, it is a result of the decrease in false positives, that is, incorrectly detecting pathologies in negative class images (without problems). On the other hand, the increase in recall in training 1 indicates a decrease in false negatives, that is, no longer indicating pathologies in images of the positive class (roofs with problems).

5. Conclusions and Further Research

This study evaluated data augmentation to detect pathologies in roof inspections with UASs images. The Data Augmentation technique was applied to expand an image processing system's training database. After rendering images, two models were trained, and the results were compared with a model trained only with original images [4]. Training 2 performed with a higher proportion of generated images (56% generated images and 44% original images), and the results indicated improvements in precision for 7 of the 11 non-conformities classes used in training. The results found in the testing stage suggest that the application of Data Augmentation allowed improved accuracy, precision, recall, F-score, negative precision, and specificity in the tests of pathology identification in images of roofs captured with UASs.

The application of data augmentation presents practical contributions to automate the analysis of roof images captured with UAS. The central aspect is the increase in the number of database images when the number of original photos is small. Consequently, the Deep Learning model trained with the artificial images increases its generalization capacity, thus improving the system's performance in identifying non-conformities in new photos. Therefore, the data augmentation approach allows, for example, to generate images for classes of pathologies with low frequency of occurrence and, consequently, few original photographs. Moreover, data augmentation can also have financial and operational benefits, such as reduced travel and labor costs to collect images for model training.

This work has some limitations, so it expects the development of future work on the following points. Evaluate other combinations and transformations' values in generating artificial images, such as differential rotation and zoom values. Therefore, it would be possible to create new Data Augmentation models, seeking to optimize the precision and recall metrics in the same training. Besides, there is a need to establish a standard for manually marking the non-conformities in the photos for training. It is still subjective and different ways of marking can influence the system's performance.

6. Acknowledgement

This work was supported by the Ministry of Education of Brazil through CAPES (Agency for Support and Evaluation of Graduate Education), and by the Ministry of Science, Technology, Innovation, and Communication of Brazil through CNPq (National Council for Scientific and Technological Development) (Grant Number Grant Number 421262/2018-4). We would also like to thank the company for the opportunity to develop this investigation.

References

[1] Conceição J, Poça B, Brito J, Flores-Colen, I and Castelo A 2017 Inspection, diagnosis, and rehabilitation system for flat roofs *Journal of Performance of Constructed Facilities* **31** 04017100

[2] Bown M and Miller K 2018, The use of unmanned aerial vehicles for sloped roof inspections considerations and constraints *Journal of Facility Management and Research* **2** 12-18

[3] Silveira B, Melo R and Costa D Using uas for roofs structure inspections at post-occupational residential buildings In: Toledo Santos E,, Scheer S, (eds) *Proc. of the 18th International Conference on Computing in Civil and Building Engineering* ICCCBE 2020 Lecture Notes in Civil Engineering vol 98 Springer Cham

[4] Staffa L, Sa L, Lima M and Costa D 2020 Use of image processing techniques for inspection of building roof structures for technical assistance purposes (in Portuguese) *Proc. of Encontro Nacional de Tecnologia do Ambiente Construído* ENTAC

[5] Chollet F and Allaire J 2018 Deep learning with R *Manning Publications*

[6] Bang S, Baek F, Park S, Kim W and Kim 2020 Image augmentation to improve construction resource detection using generative adversarial networks, cut-and-paste, and image transformation techniques *Automation in Construction* **115** 103198

[7] Razali S, Kaamin M, Razak S, Hamid N, Ahmad N, Mokhtar M, Ngadiman N and Sahat 2019 Application of uav and csp1 matrix for building inspection at muzium negeri, seremban *International Journal of Innovative Technology and Exploring Engineering* **8** 1366-1372

[8] Kaamin M, Ahmad N, Razali S, Mokhtar M, Ngadiman N, Masri D, Hussin I and Asri L 2020 Visual inspection of heritage mosques using unmanned aerial vehicle (uav) and condition survey protocol (csp) 1 matrix: a case study of tengkera mosque and kampung kling mosque, melaka *Journal Of Physics: Conference Series* **1529** 032107

[9] Dung C, Sekiya H, Hirano S, Okatani T and Miki C 2019 A vision-based method for crack detection in gusset plate welded joints of steel bridges using deep convolutional neural networks *Automation in Construction* **102** 217-229