Building and Environment 177 (2020) 106805

Contents lists available at ScienceDirect

Building and Environment

journal homepage: http://www.elsevier.com/locate/buildenv

Automatic detection of building typology using deep learning methods on street level images

Daniela Gonzalez [a], Diego Rueda-Plata [b], Ana B. Acevedo [a], Juan C. Duque [c,*], Raúl Ramos-Pollán [d], Alejandro Betancourt [c], Sebastian García [c]

[a] Department of Civil Engineering, Universidad EAFIT, Medellin, Colombia
[b] Universidad Industrial de Santander, Bucaramanga, Colombia
[c] Research in Spatial Economics (RiSE) Group, Universidad EAFIT, Medellin, Colombia
[d] Universidad de Antioquia, Medellín, Colombia

ARTICLE INFO

Keywords:
Seismic risk assessment
Convolutional neural networks
Exposure model
Street-view data
SDG 11

ABSTRACT

An exposure model is a key component for assessing potential human and economic losses from natural disasters. An exposure model consists of a spatially disaggregated description of the infrastructure and population of a region under study. Depending on the size of the settlement area, developing such models can be a costly and time-consuming task. In this paper we use a manually annotated dataset consisting of approximately 10,000 photos acquired at street level in the urban area of Medellín to explore the potential for using a convolutional neural network (CNN) to automatically detect building materials and types of lateral-load resisting systems, which are attributes that define a building's structural typology (which is a key issue in exposure models for seismic risk assessment). The results of the developed model achieved a precision of 93% and a recall of 95% when identifying nonductile buildings, which are the buildings most likely to be damaged in an earthquake. Identifying fine-grained material typology is more difficult, because many visual clues are physically hidden, but our model matches expert level performances, achieving a recall of 85% and accuracy scores ranging from 60% to 82% on the three most common building typologies, which account for 91% of the total building population in Medellín. Overall, this study shows that a CNN can make a substantial contribution to developing cost-effective exposure models.

1. Introduction

Disaster risk management is a worldwide concern, and earthquakes have been a main cause of casualties and economic losses from natural disasters over the past several decades [1]. The numbers of people affected, as well as the economic losses for a given earthquake, depend on three factors [2,3]: the frequency and size of the earthquakes in the study region (seismic hazard), the inventory of people and infrastructure (exposure), and the ability of buildings to sustain earthquake loading (vulnerability). As described in the Sendai Framework for Disaster Risk Reduction [4] the exposure of people and infrastructure has increased worldwide faster than their vulnerabilities have decreased. This trend has generated new risks and caused a steady rise in disaster losses.

This paper focuses on the exposure model, which is a description of all the assets in the region under study. The development of an exposure model is challenging, especially in places such as emerging countries where information is not routinely collected by government agencies. When small settlements are considered, information on building characteristics may be obtained from in-situ surveys that carry only moderate costs for human and economic resources. However, as the size of the modeled settlements increases, the cost and time required to conduct these surveys also increases [5–8].

Currently, some of the valuable information used to develop exposure models can be obtained from data already available online, such as images from Google Street View (GSV). This article explores applying convolutional neural networks (CNN) to such imagery to automatically determine building materials and the types of lateral load-resisting systems. These attributes define building structural typology, which is a key issue in exposure and vulnerability models for seismic risk assessment. For this purpose, we compiled and manually annotated a dataset containing approximately 10,000 GSV images of buildings within the urban area of Medellín, which is the second-largest city in

* Corresponding author. Carrera 49, 7 Sur-50, Department of Mathematical Sciences, EAFIT University, 050022, Medellín, Antioquia, Colombia.
E-mail address: jduquec1@eafit.edu.co (J.C. Duque).

https://doi.org/10.1016/j.buildenv.2020.106805
Received 18 January 2020; Received in revised form 7 March 2020; Accepted 10 March 2020
Available online 20 March 2020
0360-1323/© 2020 Elsevier Ltd. All rights reserved.

Colombia, with a population of 2.5 million inhabitants. We used the compiled dataset to fine-tune state-of-the-art publicly available pre-trained CNNs to predict the building material and the lateral load-resisting system type and achieved promising levels of accuracy when predicting the material and ductility of the lateral load-resisting systems. The fine-grained typology prediction performance was not as accurate; nonetheless, the overall results can be effectively used by modelers as complementary additional data that can significantly decrease the resources needed for developing exposure models. Although CNN and deep learning methods have been previously used for building classification (see Section 2.4), individual building classification based on its structural typology as approached in this work has not, to the knowledge of the authors, been previously addressed.

In the context of Latin American cities, which are some of the most urbanized regions in the world, using a tool like the one presented in this paper, will give local governments the possibility to conduct seismic risk assessments in less time and at a lower cost, thereby facilitating to some extent the achievement of the sustainable development goal (SDG) 11, "Make cities and human settlements inclusive, safe, resilient and sustainable." Latin American cities must face the challenge of revisiting their built environments and increasing their quality to improve the safety of their cities. In this context, exposure models are a key component of seismic risk assessments, which are essential for understanding disaster risks, one of the four action priorities of the Sendai Framework for Disaster Risk Reduction.

The remainder of this paper is organized as follows. Section 2 describes previous works conducted in this area. Section 3 describes the data collection process and the dataset construction. Section 4 describes the experiments. The results and subsequent discussion are presented in Section 5. Finally, Section 6 presents the main conclusions of this study.

2. Related works

2.1. Exposure models

Knowledge of building typology in a given region is a key input in developing an exposure model for seismic risk assessment [9]. The exposure model contains a detailed description of the assets in a region, including properties, infrastructure, population, and economic activities [10]. An exposure model, along with seismic hazard and vulnerability models, is used to estimate the probability of losses if an earthquake should occur [2,3]; in other words, exposure models are used to assess seismic risk. A seismic risk assessment for the building stock in a region requires classifying its buildings based on their structural typology, a parameter that defines the building behavior under seismic load. Building typology is a function of the building's ability to resist lateral loads and involves the building materials, building height, date of construction and the shape of the building plan, among other factors [11–13]. In this paper, we follow the building taxonomy for earthquake assessment developed by the Global Earthquake Model (GEM) Foundation [14]. This work considers only residential buildings whose structural typology is based on three of the attributes included in the GEM taxonomy: the material of the lateral load-resisting system, the lateral load-resisting system type, and the building height. The lateral load-resisting system refers to the horizontal and vertical elements that transfer lateral seismic forces to the building's foundations; it can be a system of walls only, beams/columns, or beams/columns/walls. The lateral load-resisting system can also be constructed of different materials, including reinforced concrete, masonry, steel, earth, stone, etc. We also included building ductility in the typology definition. Ductility refers to the building's capacity to sustain deformation before collapse. This parameter can be inferred from the lateral load-resisting system and the date of building construction, because the latter parameter is closely related to seismic building code enforcement.

2.2. Building typologies in Medellín

The building stock used in lateral load-resisting systems varies among locations due to differences in construction practices, material availability, building age, weather, etc. In the city of Medellín, the high levels of socio-economic inequality are reflected in high urban heterogeneity. According to Ref. [15], "Medellín has a portfolio of buildings mainly comprised of low- and intermediate-rise masonry structures and reinforced concrete structures, both medium- and high-rise, in developing residential areas." Reference [16] identified nine building typologies for Medellín using the GEM taxonomy. The building typology "others" was included to consider buildings built with unconventional materials; however, this type constitutes an insignificant percentage of Medellín's building stock and is thus excluded from the analysis in the present work.

The lateral load-resisting systems for the eight considered typologies for Medellín are wall (LWAL) systems, infilled frame (LFINF) systems, or dual frame-wall systems (LDUAL). A system can be made of masonry or reinforced concrete (CR); in masonry buildings the lateral load-resisting system is composed of masonry walls (vertical planar elements), which can be unreinforced (MUR: masonry without any form of reinforcement), confined (MCF: construction in which masonry walls are first laid and then the horizontally and vertically reinforced confining elements are cast), or reinforced (MR: masonry wall construction in which reinforcement is embedded in such a manner that both materials act together to resist forces). The lateral load-resisting system of CR buildings can be frames, walls or a combination of both. A moment-resisting frame is a beams and columns structure with strong and rigid beam-to-column connections. In some cases bays of frames are infilled with masonry walls, in which case the system is termed an infilled frame. In a wall system, the walls resist both gravity and horizontal forces; CR walls are monolithic, which is not the case with masonry walls. In a dual CR frame-wall system, both the frame and CR walls resist lateral loads. The third attribute for building classification, ductility, refers to the building's capacity to sustain deformation without failure. Building typologies are classified as either ductile (DUC) or nonductile (DNO). Fig. 1 shows examples of each building typology.

2.3. Determining structural building typologies

The lateral load-resisting system and its material can be identified only from the blueprints or by direct expert observations. In most cases, access to structural blueprints is limited, and for structures arising from self-construction processes, this information does not exist. Consequently, expert surveys seem to be the best option for compiling building inventories. However, in large cities it is not possible to survey each asset; therefore, several assumptions need to be made to establish an exposure model. Some authors have gathered building attributes from census data [9], while others acquire statistics from samples of surveyed structures and use them to allocate the load-resisting system of non-surveyed buildings [15,17,18]. Expert opinions are necessary to gather details not included in the sources of information such as national housing databases, census data, etc. Since the early 2000s, the development of remote sensing technologies and their high degree of penetration in urban environments worldwide have created opportunities for gathering data used in air exposure models [5,19–21]. This technology allows for the measurement of variables such as plan-built areas, building height, type of roof, building classification (in terms of in-plan area, number of stories, detached/attached dwelling, etc.), and building age (by comparing images taken at different years). More recently, some authors have started to use GSV, replacing fieldwork with virtual tours to visualize façades and gather lateral load resisting systems and materials remotely [16,22]. Reference [8] used GSV images to automatically retrieve and map the number of stories of urban buildings. Either by expert observation or by automatic means, the identification of a building typology solely based on its façade can be a challenging task.

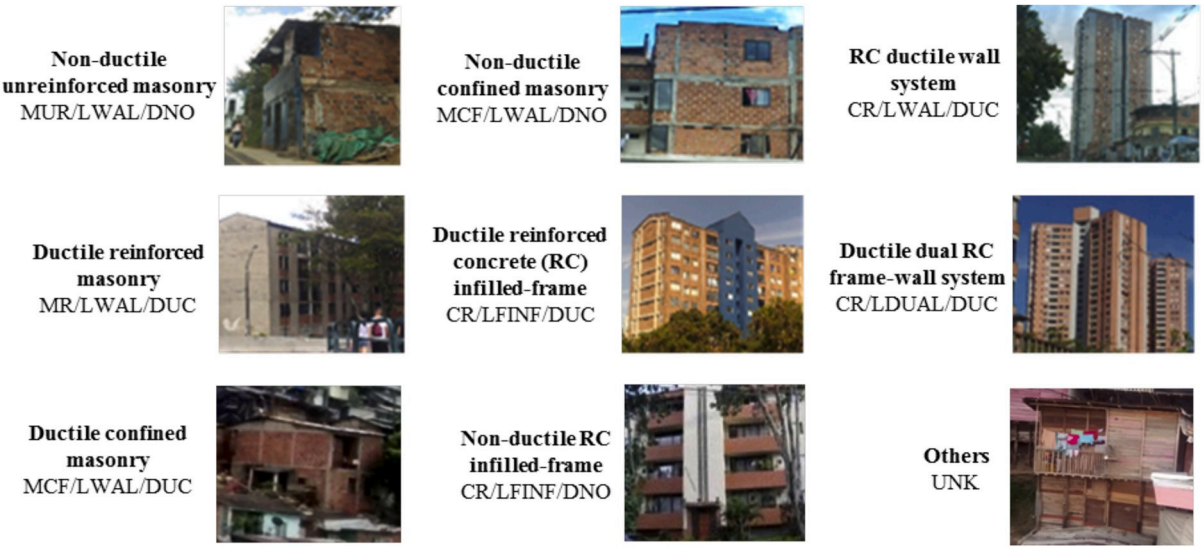

Fig. 1. Building typologies for the residential building stock of Medellín (names according to the GEM taxonomy). Images were sourced from photos taken by the authors in the city of Medellín in Jan 2018.

Therefore, uncertainty in building classification is always present. The possibility of using automated methods to increase the sample size at low cost helps to reduce such uncertainty.

In this paper, we aim to take advantage of recent developments in artificial intelligence, particularly deep learning, to develop a system capable of learning structural building typology from experts and automatically classifying large numbers of façades captured from GSV. From the best of our knowledge, this is the first work that uses CNN for the automatic identification of the structural building typology.

2.4. Deep learning

In recent years deep learning methods have achieved wide success on diverse of perceptual tasks [23]. These models avoid the need to manually design specific image feature detectors by looking for a set of transformations directly from the data, and have achieved remarkable results, particularly for computer vision problems such as natural scene classification and object detection [24].

The CNN [25] is a specialized artificial neural network architecture intended for signal processing in general and two-dimensional images in particular. In a CNN, images pass through successive layers arranged in a hierarchical pattern recognition system that simultaneously learns the features that are best suited to classify an image, recognize objects, etc.

A typical CNN architecture (Fig. 2) consists of a succession of convolutional layers accompanied by other supporting layers (such as pooling, dropout, etc.). The convolutional layers learn a set of filters or kernels that are activated when a specific feature pattern is present somewhere in the input image. Their convolutional and hierarchical nature provides certain scale and location invariance capabilities. Pooling layers progressively reduce the spatial size of the representation, diminishing the number of parameters and computation in the network. Finally, the CNN is completed with fully-connected (FC) layers that act as a regular multilayer perceptron to provide the networkâ€™s output (class probabilities, etc.).

CNNs learn hierarchies of visual features (Fig. 3)—that is, compositional patterns of visual structures through which the current dataset can be better understood. For instance, the first level set of visual primitives of a building dataset is probably composed of various color gradients and mostly vertical and horizontal border-like structures. Then, richer visual structures are composed by using the previous structures as building blocks. Each level corresponds to a CNN layer. The corresponding filters are not prebuilt by the experimenter, they are learned from the data.

Other than simulation-based approaches (see Ref. [6,26]) and classical statistical or modelling methods [27–29], there has been little use of machine learning methods for seismic risk assessment. Certain machine learning approaches have been based on applying standard algorithms to data acquired from sensors or databases (public registries, etc.) such as the works in Ref. [30,31] or [32]. However, various image processing approaches have been proposed in related areas, including earthquake damage estimation [33,34], estimation of seismic building structural types [35], building detection [36], crack damage detection [37], exposure estimation [38,39], landslide risk assessment [40], etc. Nevertheless, the vast majority of these approaches are based on satellite imagery.

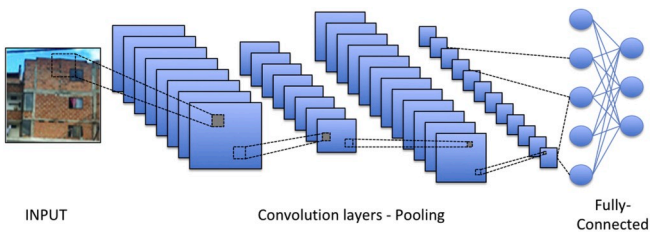

Fig. 2. Example of a CNN architecture.

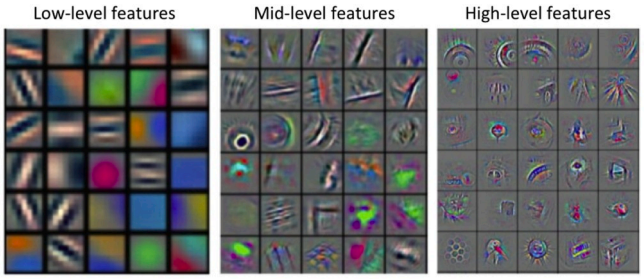

Fig. 3. Example of hierarchical features.

CNN's are starting to be used (although timidly) in this field. For instance, in Ref. [41], a CNN was used to process ground velocity records to detect earthquakes. However, CNN's have been used to extract information and classify GSV images, including street number recognition [42], traffic sign recognition [43], number of building stories [8], building view factors estimation [44], estimation of street-level solar irradiance [45], estimation of sky, tree and building view factors [46], classification and mapping of urban canyon geometry [47], classification of building instances [48], description of the form and composition of cities from a human-centric perspective [49], and estimation of neighborhood demographic composition [50].

3. Data

3.1. Study area

Between the early 1950s and early 1970s, Medellínâ's population grew from 358,189 to 1,077,252 inhabitants, mainly due to forced displacement from rural areas to informal urban settlements on the valley slopes. After the urban war against the drug cartels, from the mid-1980s to the early 1990s, Medellín's annual population growth reached 2.6%. Today, Medellín is the second most populous city in Colombia with an estimated population of 2.5 million living in an area of 1152 km^2, divided into 271 neighborhoods grouped in 16 urban districts (comunas).

Medellín (Fig. 4) is located in an intermountain valley at 1460 m above mean sea level, in a medium seismic hazard zone [51]. The possibility of earthquakes and the city's size has made seismic risk a concern in recent decades ([9,15,17,18,52]). The construction quality of Medellín's building portfolio is closely related to the inhabitants' economic levels and building age: the best construction practices are found in medium-high and high-income zones. Seismic design has been mandatory only since 1984 in Colombia; therefore, few buildings built before that date can withstand seismic loads. The bulk of Medellín building stock consists of informal construction in low-income zones that do not meet code requirements [17] and thus can be expected to exhibit poor performance under seismic loads.

3.2. The dataset

This paper relies on a dataset for Medellín compiled by Ref. [16], who used GSV to sample over ten thousand buildings distributed throughout the entire city. An image of the façade of each building was collected using GSV; then, the images were stored (along with building information) using the Inventory Data Capture Tool, IDCT [53]. There are three aspects to keep in mind when using GSV: (1) in residential areas, images are usually updated every two to three years, (2) GSV captures the images from the perspective of the street and about 3 m apart, and (3) currently, not all cities have imagery available for all streets. Therefore, in those cases in which there are areas of the city not covered by GSV imagery or when the available images do not reflect recent changes, it is possible to gather the information by own means and process it with our method.

A set of attributes were measured and linked to each image, including the type and material of the lateral load-resisting system, number of stories and ductility level. This virtual survey required 342 working hours over a four-month period by a master student and two final-year civil engineering students. To guarantee good performance when assigning the lateral load-resisting type, the students were trained by a guide published by the Federal Emergency Management Agency, FEMA [54], to perform rapid visual screening of buildings for potential seismic hazards. The surveyed buildings were selected such that different socio-economic levels, construction practices and building heights were included.

A total of 9989 buildings from those surveyed by Ref. [16] were considered for the machine learning process. The excluded buildings were those that possess identical characteristics, such as buildings that belong to the same residential complex. Fig. 5 presents the geographical distribution of the buildings used for machine prediction. Table 1 presents the building typology distribution of the dataset defined as follows: material of the lateral load-resisting system/lateral load-resisting system/ductility level. As expected, the majority of the surveyed buildings (60.96%) are unreinforced masonry structures (MUR/LWAL/DNO), which is the most common building typology in Medellín. The material

Fig. 4. Location of Medellin.

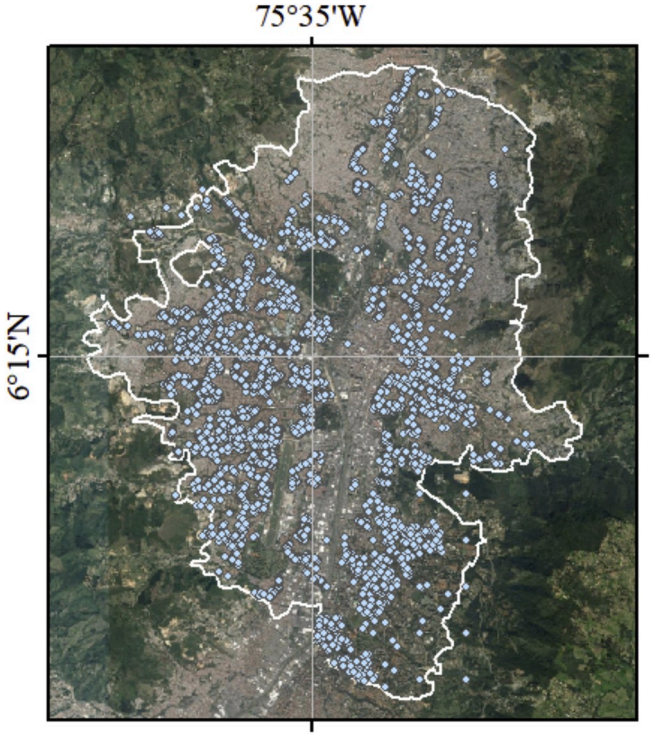

Fig. 5. Surveyed buildings.

Table 1
Dataset description.

Building ID	Building typology	No. of buildings	Percentage (%)
1	CR/LDUAL/DUC	128	1.28
2	CR/LFINF/DNO	1921	19.23
3	CR/LFINF/DUC	1081	10.82
4	CR/LWAL/DUC	128	1.28
5	MCF/LWAL/DUC	167	1.67
6	MCF/LWAL/DNO	231	2.31
7	MR/LWAL/DUC	195	1.95
8	MUR/LWAL/DNO	6089	60.96

of the lateral load-resisting system is either concrete (CR) or masonry (MUR, MR, MCF). In our dataset, 66.89% of the surveyed buildings are masonry structures, while the remaining 33.11% are concrete structures. Regarding building ductility, 17.49% of the buildings in the dataset are ductile (DUC) and (82.51%) are nonductile (DNO).

Fig. 6A describes the spatial distribution of the predominant number of stories and the socio-economic stratum at the neighborhood scale. Since 1994 Colombia uses socioeconomic strata to classify its population in groups with similar socioeconomic characteristics. There exist six strata ranging from lower-low (1) to high (6). The variables utilized for the socioeconomic characterization are housing characteristics and quality of the neighborhood. The spatial patterns show a clear predominance of 1–2-story buildings throughout the city. Fig. 6B presents the composition of our dataset in terms of the number of stories and the socio-economic stratum. Most of our dataset is concentrated in strata 1–2 (74.5%) and 3–4 (24.6%) and 82% of the buildings are between 1 and 4 stories. Finally, Fig. 6C, also based on our database, shows that a significative portion of nonductile buildings is concentrated in socio-economic strata 1 and 2.

4. Methodology

A general outline of the study methodology is shown in Fig. 7. To comprehensively understand the deep learning methods applied to this problem, the experimental methodology involved executing each experiment (1) with five network architectures of different complexities; and (2) resampling the original dataset three times to obtain three independent distributions, which were each split into training, testing and validation subsets. We averaged the performance metrics across distributions to ensure the robustness of our selection.

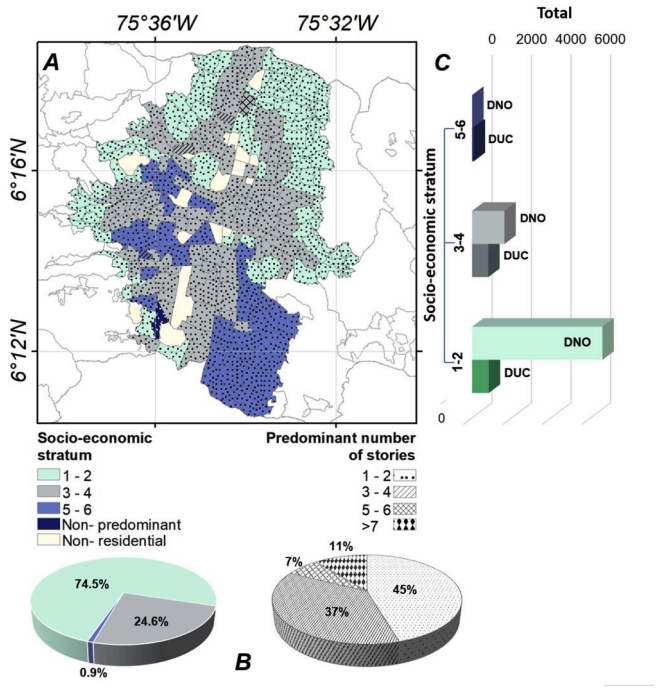

Fig. 6. Building characteristics according to predominant number of stories, socio-economic stratum, and ductility.

Our methodology also applied data augmentation techniques to enhance model performance (we generated new images by mirroring, rotating and transforming the original images to increase the variability of the visual structures), and we tested for the impact of including additional information (the number of stories) into the network architectures as shown in Fig. 8.

At this stage, we decided not to add any object removal algorithm (such as for trees or cars occluding the buildings) since our aim with this work is to build a solid simple baseline against which we can later compare further improvements (including object removal). As our research progresses beyond this work, we want to be able to make a cost/benefit analysis of each potential improvement. For instance, if some additional stage involves a large increase in compute time, or memory footprint, etc., but yields modest improvement over the baseline, its applicability could be limited. Right now, we are focused on establishing a simple enough CNN based baseline.

In total, we conducted 30 experiments with the three resampled distributions, five network architectures, and two information modes (only images and images + number of stories). The training times required between 6 and 13 h and were executed on a computer equipped with the Keras/Tensorflow deep learning framework and a K80 NVIDIA GPU.

The following subsections describe each of these aspects in detail.

4.1. Resampled distributions

In a typical machine learning experimental workflow, data is used for three purposes: (1) training models; (2) tuning model parameters and selecting the best-performing model; and (3) measuring the performance of the selected models with data not used in any of the previous stages. Through these operations, we can obtain an unbiased estimate of the model performance, which has a greater chance of being representative of the performance that would be observed when the models are applied to real-world environments.

Therefore, the dataset is randomly split into three subsets: training, validation and testing, which are used in each of the stages mentioned above. The split should be performed carefully so that class proportions are preserved and each split remains representative of the full dataset. Our splits involved using 60% of the data for training, 20% for validation, and 20% for testing.

Additionally, we repeated each complete experimental workflow three times: each time we performed a random train/test/validation split, augmented the data and trained the different network architectures using both image-only and multimodal settings (see the sections below). This process is depicted in Fig. 7.

To ensure consistency, we fixed three random train/test/validation splits at the beginning of our experimentation. We therefore created three data distributions, each consisting of a different 60/20/20 split of the full dataset.

4.2. Data augmentation

Data augmentation is an important step to reduce overfitting in a Deep Learning workflow. It consists of adding noise to the training set to avoid the model learning specific patterns and affecting its generalizability to unseen data. Noise is added to the training set by including new images derived from the originals after applying geometric transformations (e.g., horizontal mirroring, small rotations, and shear transformations), cropping a small percentages of the images, or applying Gaussian blur.

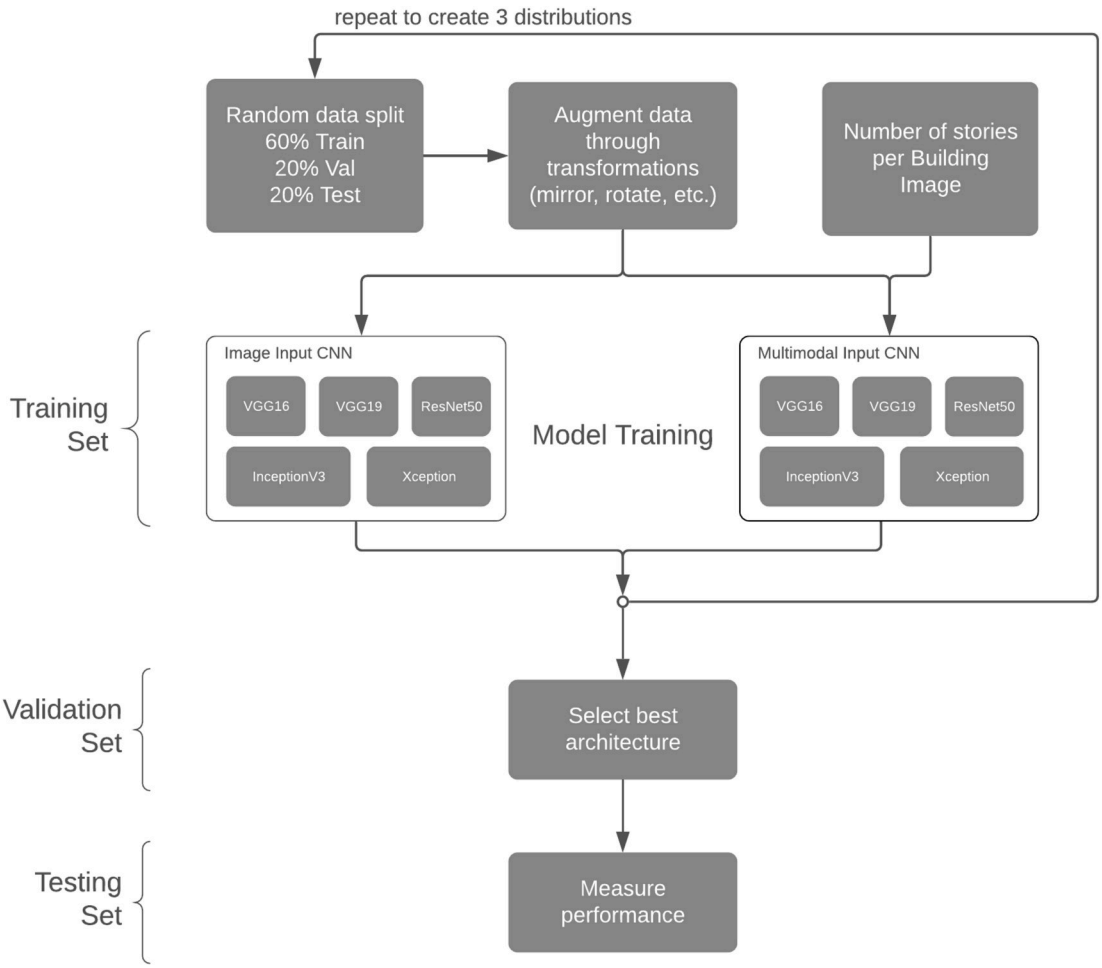

Fig. 7. Methodology workflow.

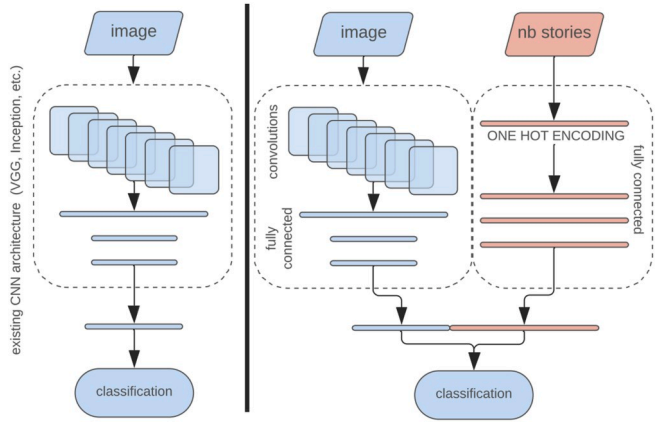

Fig. 8. Image only classification architecture (left) and multimodal architecture (right).

4.3. Network architectures and model training

We selected five state-of-the-art CNN networks that have shown remarkable success on ImageNet, a large scale classification challenge with over 1.5 Million images, 1000 classes, and many other image analytic tasks [55]: VGG16, VGG19, InceptionV3, ResNet50, and Xception.

VGG19 introduces multiple stacks of smaller convolutional filters to reduce the number of parameters [56]. InceptionV3 introduces the inception module, which is a block of filters of different sizes that capture meaningful features at different scales [57]. ResNet50 manages the vanishing gradient problem that occurs in truly deep networks by using shortcut connections between blocks of convolutional layers to allow gradient information to be directly transmitted further down the network [58]. Finally, Xception builds on the previous networks and introduces a two-step convolution in which a spatial convolution is performed independently for each channel in the image, followed by a 1x1 convolution across all the channels [59].

These architectures and their respective pretrained weights on ImageNet allow us to use a technique called fine-tuning, in which the earlier layers of the network are preserved, retaining the low level visual features they learned on the ImageNet dataset, while the later layers are retrained to learn features specific to a target domain (our problem in this case). Fine-tuning has been shown to improve generalization performance on many classification tasks [60].

4.4. Multimodal architecture

Our dataset also includes additional attributes that associate each building by its number of stories. Including additional information can potentially improve the performance of the classification models, but too many attributes affect the possibility of applying the models in contexts where similar information is unavailable.

To understand the value of injecting information about the number of stories into the model, we devised two sets of experiments: (1) using only images as input; (2) using images plus the number of stories as input. The latter approach is denoted as *multimodal learning* because the input information is provided in complementing modalities (images and

numeric data).

The image-only models need to be adapted to accept multimodal information. To achieve this, we devised a simple three layer multilayer perceptron architecture whose input is the number of stories in a one-hot encoding with 36 elements (one for each story). In a one-hot encoding of n possible values, an integer i is mapped to a n-length zero vector, whose i-th position is set to one. For example, the integer values {0, 1, 2, 3} are mapped to the vectors {1000, 0100, 0010, and 0001}, respectively. Then, we simultaneously inject the image into whichever CNN architecture we are using, and input the number of stories into our perceptron. Finally, we collate both outputs, which together produce the final classification. This procedure is shown in Fig. 8.

4.5. Performance metrics

Various performance metrics can be used in classification tasks, but no individual metric is fully informative in a general manner. Accuracy (the percentage of correct classifications) is arguably the metric most often used but is somewhat misleading when the input classes are unbalanced, or the profits and costs from correct and incorrect classifications are not symmetric. In these cases, precision and recall are also commonly used metrics. Starting from a confusion matrix like the one presented in Table 2, precision is a measure of exactness calculated as the percentage of positive predictions that were correct (true positives) divided by the number of predicted positives. Recall is a measure of completeness calculated as the percentage of positives that were correctly identified as such (true positives) divided by the number of actual positives).

In this context, we propose analyzing the results using a two stage performance assessment process. First, we use precision and recall with respect to the building class typologies that are most likely to be damaged in an earthquake(i.e., non-ductile buildings), which includes the typologies MUR/LWAL/DNO, MCF/LWAL/DNO, and CR/LFINF/DNO. Therefore, we expect our models to perform well at differentiating these typologies from the rest. This aspect measures the classification architecture performances based on their capacity to successfully identify fragile structures (i.e., to avoid misclassifying a non-ductile structure as ductile) and minimizes the possibility of putting people at risk by misclassifying their dwellings. Then, we use accuracy to gain a finer understanding of misclassifications on each of the eight types of buildings.

The results are shown in Table 3, in which our best-performing model classified all the buildings in each typology, grouped by ductility on the first validation distribution. From this table we obtain the ductility-based confusion matrix shown in Table 4 by aggregating the buildings in each quadrant. Then, we use that confusion matrix to compute precision and recall with respect to ductility, as indicated in Equations (1) and (2). After completing this process, the recall results indicates that 94% of the total relevant results (i.e., non-ductile buildings) are correctly classified by ResNet50, and the precision indicates that 93% of the buildings classified as relevant by ResNet50 (i.e., the non-ductile buildings) are actually relevant.

Then, we use accuracy to inspect each individual class to understand the nature and impact of specific misclassifications. For instance, according to Table 3, out of the 276 buildings classified by ResNet50 as class 2, 169 were indeed class 2 buildings—an accuracy of 61%. However, 55 of these (20%) were classified as class 8, but those misclassifications have no impact on the ductility precision/recall because both classes are DNO. Nevertheless, 35 (13%) were misclassified into class 3, which does change the ductility assessment.

Table 2
Elements of a confusion table.

Result	Actual Class 1	Actual Class 0
Predicted Class 1	*True Positive*	*False Positive*
Predicted Class 0	*False Negative*	*True Negative*

Table 3
Confusion matrix using ResNet50 (image input only) on the first distribution of the validation subset.

Predicted class		Actual class								
		DNO			DUC					
		2	6	8	1	3	4	5	7	Total
DNO	2	169	4	55	0	35	1	6	6	276
	6	10	7	12	0	0	0	5	0	34
	8	159	29	1104	0	18	0	28	24	1362
DUC	1	1	0	0	11	31	3	0	0	46
	3	27	0	0	17	105	5	0	0	154
	4	2	0	1	7	17	18	0	1	46
	5	7	5	18	0	0	0	1	0	31
	7	17	0	17	0	2	2	1	10	49
Total		392	45	1207	35	208	29	41	41	1998

1:CR/LDUAL/DUC, 2:CR/LFINF/DNO, 3:CR/LFINF/DUC, 4:CR/LWAL/DUC, 5:MCF/LWAL/DUC, 6:MCF/LWAL/DNO, 7:MR/LWAL/DUC, 8:MUR/LWAL/DNO.

Table 4
Ductility-based confusion matrix using ResNet50 (image input only) on the first distribution of the validation subset.

$$recall_{Ductility} = \frac{1,549}{1,549 + 95} = 0.94 \quad (1)$$

$$precision_{Ductility} = \frac{1,549}{1,549 + 123} = 0.93 \quad (2)$$

	DNO	DUC	Total
DNO	1549	123	1672
DUC	95	231	326
Total	1644	354	1998

4.6. Architecture selection

Recall and precision are complementary metrics. On the one hand, a confusion matrix that produces high recall and low precision indicates that the model tends to classify many buildings as positive, but many of those predictions are incorrect when compared to the true classifications. On the other hand, a confusion matrix that produces low recall and high precision indicates a classification model that correctly identifies true positives, but correctly classifies only a few instances compared to the actual number of buildings within the positive or relevant category. Therefore, an ideal classification model with both high recall and high precision will classify many buildings as positive predictions, and those predictions will largely be correct when compared against the actual classification.

Algorithm 1 presents the procedure that we followed to select the best classification models. The first part of the algorithm calculates the ductility-based recall and precision for each architecture (VGG16, VGG19, InceptionV3, ResNet50, and Xception) and input option (image only and multimodal). Since we have three independent validation distributions (v_dist1, v_dist2, v_dist3), we calculate the average recall and precision for each model. Then, we select those cases whose average recall is greater than or equal to 0.95 and whose average precision is greater than or equal to 0.90. Finally, in the second part of the algorithm we apply the selected architectures and input options to the three independent test subsets (t_dist1, t_dist2, t_dist3) and average them. This second operation ensures unbiased recall and precision estimates.

Algorithm 1. Ductility-based Architecture Selection

$Selected = \{\}$
for architecture a in {VGG16, VGG19, InceptionV3, ResNet50, Xception} **do**
  for input i in {image only, multimodal} **do**
    for validation distribution v in {v_dist1, v_dist2, v_dist3} **do**
      calculate $recall_{a,i,v}$
      calculate $precision_{a,i,v}$
    end for
    calculate $avg\_rec\_v_{a,i} = \frac{\sum_v recall_{a,i,v}}{3}$
    calculate $avg\_prec\_v_{a,i} = \frac{\sum_v precision_{a,i,v}}{3}$
    if $avg\_rec_{a,i} \geq 0.95$ and $avg\_prec_{a,i} \geq 0.90$ **then**
      $Selected \cup \{[a, i]\}$
    end if
  end for
end for
for a, i in $Selected$ **do**
  for test distribution t in {t_dist1, t_dist2, t_dist3} **do**
    calculate $recall_{a,i,t}$
    calculate $precision_{a,i,t}$
  end for
  calculate $avg\_rec\_t_{a,i} = \frac{\sum_t recall_{a,i,t}}{3}$
  calculate $avg\_prec\_t_{a,i} = \frac{\sum_t precision_{a,i,t}}{3}$
end for
RETURN $avg\_rec\_v_{a,i}$, $avg\_prec\_v_{a,i}$, $avg\_rec\_t_{a,i}$, $avg\_prec\_t_{a,i}$

5. Results

Following the architecture selection process described above, Table 5 lists the validation (top table) and test (bottom table) metrics for all the considered architectures and distributions. In particular, note (1) the similar behavior of the test and validation splits across the different distributions, indicating a desirable representativity of the data; and (2) that no distinguishability exists between models regarding precision; thus, we use recall to select the best model. The three best architectures whose *avg_rec* values are above 95% are Resnet50$_i$, Xception$_i$, and Resnet50$_m$. Although these three models achieve the same validation performance, we selected Resnet50$_i$ because it uses less information than does Resnet$_m$, and its computational training time was approximately half that of the other algorithms because its optimization process converged faster. Fig. 9 presents the mean aggregated confusion matrix for the three distributions on the Resnet50$_i$ architecture. The obtained results are encouraging because, on average, 89.8% of the buildings are classified in their correct ductility categories. However, 4.4% of the buildings are classified as ductile when in fact they are not (this is precisely the type of error that we seek to minimize) while 5.8% of the buildings are classified as non-ductile when in fact they are. While the first type of error underestimates risk because it fails to capture

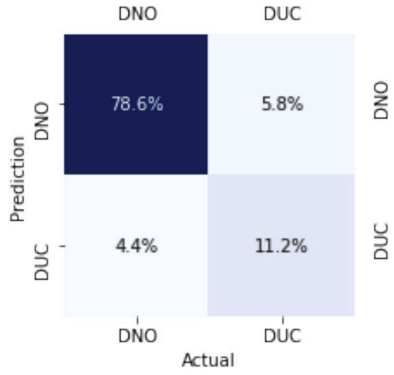

Fig. 9. Confusion matrix for *avg_rec* $= \frac{0.786}{0.786+0.44} = 0.95$ and *avg_prec* $= \frac{0.786}{0.786+0.58} = 0.93$ (resnet50_i_test mean of the three distributions). The entire matrix adds up to 100% focusing on the overall magnitude of correct and incorrect predictions with respect to the total population.

Table 5
Results from Algorithm 1.

	Recall$_v$				*Precision$_v$*			
	dist1	dist2	dist3	*avg*	dist1	dist2	dist3	*avg*
Vgg16[a]	0.86	0.88	0.91	0.88	0.94	0.93	0.93	0.93
Vgg16[b]	0.89	0.85	0.85	0.86	0.94	0.92	0.93	0.93
Vgg19[a]	0.91	0.93	0.91	0.92	0.94	0.92	0.93	0.93
Vgg19[b]	0.93	0.89	0.88	0.90	0.94	0.92	0.93	0.93
InceptionV3[a]	0.93	0.92	0.91	0.92	0.93	0.92	0.92	0.93
InceptionV3[b]	0.93	0.92	0.93	0.93	0.93	0.93	0.92	0.93
Xception[a]	0.96	0.94	0.95	**0.95**	0.93	0.93	0.92	**0.93**
Xception[b]	0.89	0.94	0.95	0.93	0.95	0.93	0.93	0.93
Resnet50[a]	0.95	0.95	0.95	**0.95**	0.94	0.92	0.92	**0.93**
Resnet50[b]	0.96	0.95	0.95	**0.95**	0.95	0.92	0.91	**0.93**
	Recall$_t$				*Precision$_t$*			
	dist1	dist2	dist3	*avg*	dist1	dist2	dist3	*avg*
Vgg16[a]	0.85	0.88	0.91	0.88	0.93	0.93	0.94	0.94
Vgg16[b]	0.87	0.85	0.84	0.85	0.93	0.93	0.94	0.94
Vgg19[a]	0.88	0.94	0.94	0.90	0.92	0.93	0.93	0.93
Vgg19[b]	0.91	0.90	0.88	0.90	0.93	0.94	0.94	0.94
InceptionV3[a]	0.93	0.93	0.92	0.93	0.92	0.94	0.94	0.93
InceptionV3[b]	0.94	0.92	0.93	0.93	0.92	0.94	0.93	0.93
Xception[a]	0.96	0.95	0.96	**0.96**	0.93	0.94	0.94	**0.93**
Xception[b]	0.88	0.93	0.95	0.92	0.93	0.94	0.94	0.94
Resnet50[a]	0.94	0.94	0.96	**0.95**	0.93	0.93	0.93	**0.93**
Resnet50[b]	0.95	0.96	0.96	**0.95**	0.93	0.93	0.93	**0.93**

[a] Image only.
[b] Multimodal.

buildings at potential risk, the second type of error overestimates the number of structures and humans at risk.

Next we evaluate the performance of the selected architecture when classifying each of the 8 building typologies. Fig. 10 reports the accuracy for each individual building typology. The best performance is obtained for typologies 2, 8 and 3, with accuracy levels above 60%. Observe that many misclassifications occur within the diagonal quadrants, which has little impact from the ductility perspective. For instance, 70% of typology 1 predictions are actually typology 3, but the model still correctly predicts the ductility nature of those misclassified buildings. Note also that 86% of the typology 5 predictions are incorrectly classified: their actual typologies are either 2 or 8. However, these classes are much less frequent overall. According to Fig. 9 all the DNO buildings classified by the model as DUC amount to only 4.4% of the total. Remarkably, however, this matches the stability of results reported in Table 5; we can obtain significant recall improvements only by using more sophisticated architectures. This result reflects the fact that a nominal recall level (about 0.88) is relatively easy to obtain, but improving the recall further is considerably more difficult.

We used ductility to select the best model because we consider it important to minimize the errors in classifying non-ductile structures as ductile (i.e., we want to avoid errors that could led to underestimation of the risk). Furthermore, it is important to understand the performance of other parameters, such as material of the lateral load-resisting system (LLRS). The dataset includes two types of materials: reinforced concrete (CR-buildings with ids of 1–4) and masonry (M-buildings with ids of 5–8). Fig. 11 shows the accuracy of material prediction, from which it can be seen that the accuracy levels exceed 85% for both types of material. Fig. 10 shows that misclassifications greater than or equal to 10% on the material of the LLRS occur for only one type of CR building, CR/LFINF/DNO (id 2), which are misclassified as all types of masonry (ids 5 to 8) and one type of M building, MUR/LWAL/DNO (id 8), that is classified as CR/LFINF/DNO (id 2). These errors are concentrated in one building class: CR/LFINF/DNO (id 2), which is an expected mistake because this building typology consists of CR beams and columns infilled with masonry walls; therefore, in this typology, the material of the LLRS can easily be mistaken for masonry.

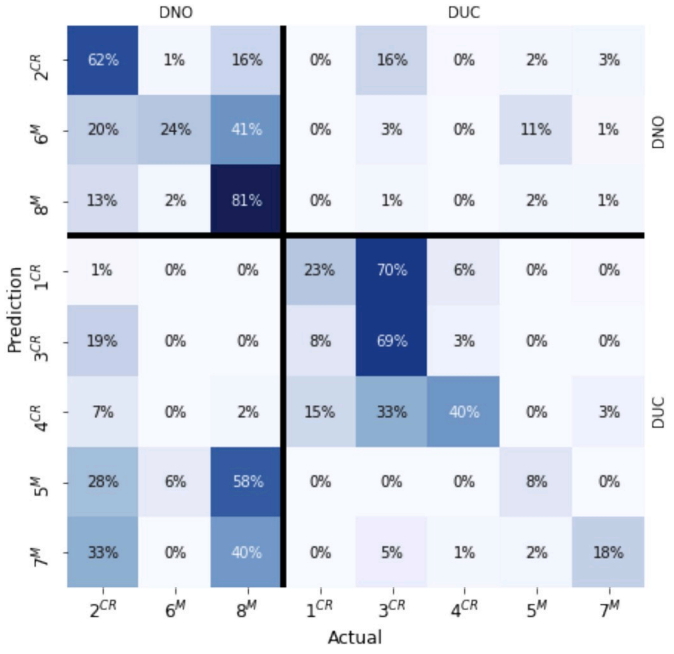

Fig. 10. Confusion matrix 8 classes for **accuracy** (resnet50_i_test mean of the three distributions). Each row adds up to 100%. We focus on how the predictions of each typology are distributed with respect to actual values. (**CR**: concrete; **M**: masonry).

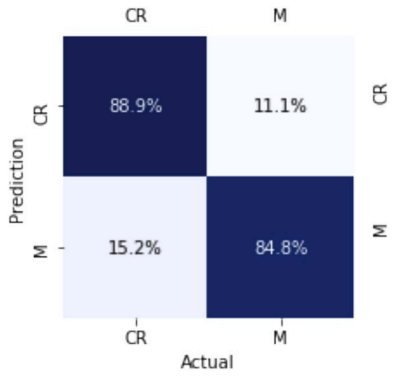

Fig. 11. Confusion matrix for material accuracy (resnet50_i_test mean of the three distributions). Observe that each row adds up to 100% so that we focus on accuracy per class (i.e. 88.9% of the times the model predicted **CR** it was correct, where are 84.8% of **M** predictions were correct.

When both building ductility and material of the LLRS are considered, the accuracy levels exceed 60% for all the typologies except MCF/LWAL/DUC (id 5) and MR/LWAL/DUC (id 7). For this analysis, the values from each row of Fig. 10 with same ductility and material as the predicted typology must be added, e.g. for building typology 4, with DUC-ductility and CR-material, accurate predictions took place for typologies 1, 3 and 4, achieving a total accuracy level of 88% (15% + 33% + 40%). The accuracy levels for both ductility and material are as follows: 62% for CR/LFINF/DNO (id 2); 65% for MCF/LWAL/DNO (id 6); 83% for MUR/LWAL/DNO (id 8); 99% for CR/LDUAL/DUC (id 1); 80% for CR/LFINF/DUC (id 3); 88% for CR/LWAL/DUC (id 4); 8% for MCF/LWAL/DUC (id 5); and 20% for MR/LWAL/DUC (id 7). Clearly, the model has problems correctly identifying typologies 5 and 7. However, these two typologies are difficult to identify from images because MCF/LWAL/DUC (id 5) has masonry walls with horizontal and vertical reinforced concrete elements, making it easy to misclassify as MUR/LWAL/DNO (in which masonry walls are predominant) and CR/LFINF/DNO (a typology that also has masonry walls with RC elements) can easily be misclassified as MCF/LWAL/DUC. On the other hand, MR/LWAL/DUC (id 7) typology corresponds to a masonry wall construction in which reinforcement is embedded; therefore, the reinforcement is not exposed in the structure. Consequently, identification of this typology from only an image is not an easy task, and buildings in which concrete elements are covered or do not exist may be misclassified as MR/LWAL/DUC buildings.

Although the ductility and material accuracy levels are high, certain misclassifications occur at the typology level. Thus, the next question is: Are those mistakes plausible even to an expert's eye? Therefore, we summarize the opinions of expert engineers regarding the most common errors found in our classification model below.

- 70% of the buildings classified as CR/LDUAL/DUC (id 1) are actually CR/LFINF/DUC (id 3) buildings: both typologies are usually tall buildings with beam-and-column systems. The difference between those typologies is the presence of reinforced walls in the CR/LDUAL/DUC system; sometimes, however, those reinforced walls are not easily discernable from an image because they are covered by the façade or exist inside the building. Even a trained engineer may be hesitant to classify this typology from an image.
- 58% of the buildings classified as MCF/LWAL/DUC (id 5) and 41% of the buildings classified as MCF/LWAL/DNO (id 6) are actually MUR/LWAL/DNO (id 8) buildings: From Fig. 1 it can be observed that the three typologies have a lateral load-resisting system made of masonry that all show similar patterns and colors in the images. The subtle difference between these three typologies is the presence of confining elements such as thin columns and beams in the MCF/LWAL typologies (5 and 6).

- 40% of buildings classified as MR/LWAL/DUC (id 7) are actually MUR/LWAL/DNO (id 8) buildings: Differentiating between these two typologies is difficult even for experts because the differences are not visible in the façade: reinforcement in MR/LWAL/DUC (id 7) buildings is embedded in the masonry. Thus, both typologies have similar external appearances.
- 33% of the buildings classified as CR/LWAL/DUC (id 4) are actually CR/LFINF/DUC (id 3) buildings: Both typologies refer to tall buildings with similar external shapes. As addressed earlier, it is common to cover the building façade with masonry elements, which prevents correctly identify the wall material—in this case, concrete for id 4 and masonry for id 3.
- 33% of buildings classified as MR/LWAL/DUC (id 7) are actually CR/LFINF/DNO (id 2) buildings: MR/LWAL/DUC (id 7) is a building typology difficult to identify from an image because it consists of walls with embedded reinforcement that cannot be seen in the façade; therefore, the masonry wall is the most prevalent feature of the image. On the other hand, buildings with CR/LFINF/DNO (id 2) typology have reinforced concrete elements (beams and columns) with bays infilled with masonry walls. Once again, the masonry wall is the more distinctive feature in the image.
- 28% of buildings classified as MCF/LWAL/DUC (id 5) are actually CR/LFINF/DNO (id 2) buildings: both these typologies have reinforced concrete columns and beams with the presence of masonry walls. Although the dimensions of the concrete elements are smaller in the MCF/LWAL/DUC (id 5) images, it is difficult to extract this particular characteristic from the image.

Finally, we present the confusion matrix for the best multimodal architecture in Fig. 12 to evaluate the impact from including the number of stories. The results show a slight reduction in misclassifications such as CR/LFINF/DNO is observed for MCF/LWAL/DNO (−4%), MCF/LWAL/DUC (−7%) and MR/LWAL/DUC (−5%). We noticed minor improvements on some of the more difficult classes to predict, such as CR/LDUAL/DUC (+4%), CR/LWAL/DUC (+4%), and MCF/LWAL/DUC (+5%). Including the number of stories in the architectures provides some contribution to improving the predictions for a subset of classes; however, when considering its operational implications in terms of both data collection and application of these models in other contexts, these gains become less significant.

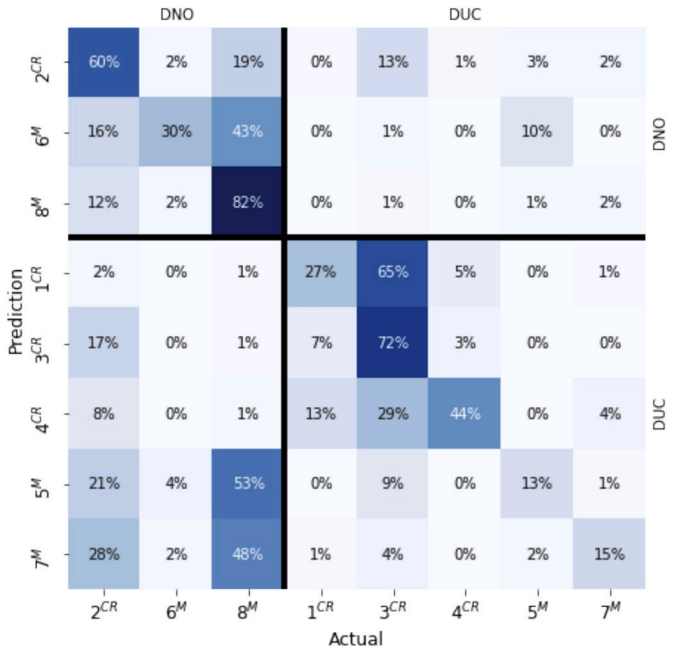

Fig. 12. Confusion matrix 8 classes for **accuracy** (resnet50_m_test mean of the three distributions) when using the multimodal input (i.e., images plus the number of stories). Each row adds up to 100%. We focus on how the predictions of each typology are distributed with respect to the actual values. (**CR**: concrete; **M**: masonry).

6. Conclusions

This study used a manually annotated dataset of approximately 10,000 photos at the street level within the urban area of Medellín to explore the potential of using a CNN to classify buildings according to their lateral load-resisting system. Among the five network architectures trained in this study, ResNet50 showed the best performance because it classified fewer non-ductile buildings as ductile (4.4%). Ductility was the parameter chosen to define the best architecture because we considered it important to minimize the errors of classifying non-ductile buildings as ductile (i.e., we wanted to avoid errors that could underestimate risk). Comparisons of the metrics among the architectures when using only image input (*i*) and architectures using multimodal input (*m*) indicates only a limited impact of including the number of stories as additional input.

The selected architecture had a non-ductile recall accuracy of approximately 95% on the material type—85% for masonry and 89% for concrete—and a typology accuracy exceeding 60% on three of the eight building typologies. Although the typology accuracy was considered successfully for only three of the eight considered typologies (CR/LFINF/DNO, CR/LFINF/DUC and MUR/LWAL/DNO), the accuracy levels for both ductility and material type indicate the reliability of the results for predicting such characteristics and constitutes useful information that can be used by exposure modelers. Notably, the buildings of the three typologies with the highest accuracy constitute 91% of the buildings in the dataset. Future research should concentrate on two directions. First, the performance of this approach in cities that include building typologies not covered in this work should be explored. Second, because unreinforced masonry is the most common typology in the dataset, the potential of CNN to classify floor diaphragm types, which can also be critical for improving the classification performance of unreinforced masonry buildings, should be explored.

The outputs of our model are intended to be used by an exposure model developer aiming to reduce the implementation cost. The good performance of our model in detecting ductility and building material can be used either to classify exhaustively all the buildings in small cities or in big cities to process a spatially-distributed sample of buildings that can be used to interpolate ductility and material for the rest of the city [9,17]. Predictions on building typologies should be used with caution as for some building classes the CNN performance is not good enough for a direct building classification. However, the information provided by the CNN could be used together with additional information, such as census data and expert judgment, to produce a more accurate building classification.

A risk assessment requires three input models: seismic hazard (the probability of earthquake occurrence), vulnerability (building ability to sustain earthquake loads), and exposure (inventory of exposed assets); all of them should be suitable for the location where the risk is assessed. Over the last decades, efforts have been made to produce seismic hazard assessments, which are now available at different scales worldwide. As a matter of fact, seismic design codes use results from seismic hazard assessments to define seismic actions for structural design. Regarding the second input, the vulnerability model, there is a vast number of vulnerability studies that can be used, or adapted if required, to generate the vulnerability model. Finally, contrary to seismic hazard assessments that are available at different scales and vulnerability models that are either available or may be adapted from existing ones, the exposure models must necessarily be developed specifically for the studied location. Unfortunately, exposure models have not received as much attention as the hazard and vulnerability models, especially in emerging countries where the information required for its development is difficult

to access. Also, the development of an exposure model is a time and resource-consuming activity as it must have a detailed description of the exposed buildings. It is in this context in which our methodology becomes a significant contribution to seismic risk assessment.

Author contributions

Conceptualization, A.B.A., J.C.D., and R.R.; methodology, R.R. and J. C.D.; software, D.R., R.R., A.B., and S.G.; validation, R.R. and A.B.A.; formal analysis, A.B.A., J.C.D., and R.R.; data curation, D.G. and A.B.A.; writing–original draft preparation, A.B.A., J.C.D., D.G., and R.R.; writing–review and editing, J.C.D., A.B.A., D.G., and R.R.; visualization, D. G., D.R., J.C.D., and R.R.; supervision, J.C.D. and A.B.A.; project administration, A.B.A. and J.C.D.; funding acquisition, J.C.D. and A.B.A.

Funding

This document was completed with support from the PEAK Urban programme, supported by UKRI's Global Challenge Research Fund, Grant Ref: ES/P011055/1. Financial support for the database collection was provided by EAFIT University, Grant Ref: 690–000026.

Data availability statement

All data files required to reproduce the analysis in the paper are hosted at the Github repository "building_typology", available at the url: https://github.com/Rise-group/building_typology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are grateful for the support with the Apolo supercomputer resources and the team at EAFIT. The usual disclaimer applies.

References

[1] H. Ritchie, M. Roser, Natural disasters, Our World in Data, https://ourworldindata.org/natural-disasters, 2020.
[2] Hazus–mh Mr5, Multi-Hazard Loss Estimation Methodology–Earthquake Model, Department of Homeland Security, Federal Emergency Managment Agency, Mitigation Division, Washington, DC, USA, 2010. Tech. Rep.
[3] V. Silva, H. Crowley, M. Pagani, D. Monelli, R. Pinho, Development of the Openquake Engine, the Global Earthquake Model's Open-Source Software for Seismic Risk Assessment, vol. 72, 07 2014.
[4] U.G. Assembly, The sendai framework for disaster risk reduction 2015–2030, Resolution A/Res/69/283, http://www.unisdr.org/files/resolutions.N, 2015, 1516716.
[5] F. Yamazaki, H. Mitomi, M. Matsuoka, K. Honda, Inventory development for natural and built environments-remote sensing technologies for inventory development and risk assessment-characteristics of satellite images in bangkok, Thailand, Tech. Rep. (2000). NIED, Miki, Hyogo, Japan, Tech. Rep. on The Development of Earthquake and Tsunami Mitigation Technologies and their Integration for the Asia-Pacific Region.
[6] A.C. Caputo, A. Vigna, Numerical simulation of seismic risk and loss propagation effects in process plants: an oil refinery case study, in: ASME 2017 Pressure Vessels and Piping Conference, American Society of Mechanical Engineers, 2017. V008T08A024–V008T08A024.
[7] M. Pittore, M. Wieland, "Toward a rapid probabilistic seismic vulnerability assessment using satellite and ground–based remote sensing 68 (2013) 115–145.
[8] G.C. Iannelli, F. Dell'Acqua, Extensive exposure mapping in urban areas through deep analysis of street-level pictures for floor count determination 1 (2017), 16.
[9] C. Yepes-Estrada, V. Silva, J. Valcárcel, A.B. Acevedo, N. Tarque, M.A. Hube, G. Coronel, H. Santa Maria, Modeling the residential building inventory in south America for seismic risk assessment, Earthq. Spectra 33 (1) (2017) 299–322.
[10] C. Geiß, H. Taubenböck, Remote sensing contributing to assess earthquake risk: from a literature review towards a roadmap, Nat. Hazards 68 (1) (2013) 7–48.
[11] K. Jaiswal, D.J. Wald, "Creating a global building inventory for earthquake loss assessment and risk management," U.S. Geological Survey Open-File Report 2008-1160, Tech. Rep. (2008).
[12] C. Rojahn, R.L. Sharpe, Earthquake Damage Evaluation Data for california, Atc-13, Tech. Rep., Redwood City, CA, USA, 1985. Applied Technology Council.
[13] Hazus–mh Mr4 Technical Manual, Tech. Rep., Washington, DC, USA, 2003. Federal Emergency Management Agency, Mitigation Division.
[14] S. Brzev, C. Scawthorn, A.W. Charleson, L. Alle, M. Greene, K. Jaiswal, V. Silva, "Gem building taxonomy version 2.0," GEM Foundation, Pavia, Italy, Tech. Rep. (2013), 2013-02 V1.0.0, Tech. Rep..
[15] M.A. Salgado-Gálvez, D. Zuloaga-Romero, G.A. Bernal, M.G. Mora, O.-D. Cardona, Fully probabilistic seismic risk assessment considering local site effects for the portfolio of buildings in medellín, Colombia, Bull. Earthq. Eng. 12 (2) (2014) 671–695.
[16] D. Gonzalez, A.B. Acevedo, "Actualización del modelo de exposición sísmica para viviendas en medellín (colombia) y su aplicación en la evaluación del riesgo sísmico para viviendas de mampostería no reforzada," in *VIII Congreso Nacional de Ingeniería Sísmica*, Barranquilla, Colombia, 2017.
[17] A.B. Acevedo, J.D. Jaramillo, C. Yepes-Estrada, V. Silva, F.A. Osorio, M. Villar-Vega, Evaluation of the seismic risk of the unreinforced masonry building stock in antioquia, Colombia, Nat. Hazards 86 (1) (2017) 31–54.
[18] L.F. Restrepo, M.R. Villarraga, J.D. Jaramillo, Y. Farbiarz, A.F. Vélez, D.A. Rendón, F.P. Ángel, Microzonificación y evaluación del riesgo sísmico del Valle de Aburrá, Área Metropolitana del Valle de Aburrá, 2007.
[19] D. Dutta, K. Serker, Urban building inventory development using very high-resolution remote sensing data for urban risk analysis, Int. J. Geoinf. 1 (1) (2005), 190–116.
[20] H. Taubenböck, A. Roth, S. Dech, Vulnerability assessment using remote sensing: the earthquake prone megacity istanbul, Turkey, in: Proceedings of ISRSE, 2007, pp. 1–5, 2007.
[21] S. Valero, J. Chanussot, P. Gueguen, Classification of basic roof types based on vhr optical data and digital elevation model, in: *IGARSS 2008 - 2008* IEEE International Geoscience and Remote Sensing Symposium, vol. 4, 2008. IV – 149–IV – 152.
[22] M. Pittore, M. Wieland, K. Fleming, Perspectives on global dynamic exposure modelling for geo-risk assessment, Nat. Hazards 86 (1) (2017) 7–30.
[23] Y. Bengio, A. Courville, P. Vincent, Representation learning: a review and new perspectives, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2013) 1798–1828.
[24] J. Schmidhuber, Deep learning in neural networks: an overview, Neural Network. 61 (2015) 85–117.
[25] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[26] S. Esposito, I. Iervolino, A. d'Onofrio, A. Santo, F. Cavalieri, P. Franchin, Simulation-based seismic risk assessment of gas distribution networks, Comput. Aided Civ. Infrastruct. Eng. 30 (7) (2015) 508–523.
[27] D. Cook, K. Fitzgerald, T. Chrupalo, C. Haselton, Comparison of Fema P-58 with Other Building Seismic Risk Assessment Methods, 2017.
[28] N. Frolova, V. Larionov, J. Bonnin, S. Sushchev, A. Ugarov, M. Kozlov, Seismic risk assessment and mapping at different levels, Nat. Hazards 88 (1) (2017) 43–62.
[29] S. Mangalathu, J.-S. Jeon, J.E. Padgett, R. DesRoches, Performance-based Grouping Methods of Bridge Classes for Regional Seismic Risk Assessment: Application of Anova, Ancova, and Non-parametric Approaches, *Earthquake Engineering & Structural Dynamics*, 2017.
[30] I. Riedel, P. Guéguen, M. Dalla Mura, E. Pathier, T. Leduc, J. Chanussot, Seismic vulnerability assessment of urban environments in moderate-to-low seismic hazard regions using association rule learning and support vector machine methods, Nat. Hazards 76 (2) (2015) 1111–1141.
[31] Y. Zhang, H.V. Burton, H. Sun, M. Shokrabadi, A machine learning framework for assessing post-earthquake structural safety, Struct. Saf. 72 (2018) 1–16.
[32] W.D. Fisher, T.K. Camp, V.V. Krzhizhanovskaya, Anomaly detection in earth dam and levee passive seismic data using support vector machines and automatic feature selection, J. Comput. Sci. 20 (2017) 143–153.
[33] Y. Bai, B. Adriano, E. Mas, H. Gokon, S. Koshimura, Object-based building damage assessment methodology using only post event alos-2/palsar-2 dual polarimetric sar intensity images, J. Disaster Res. 12 (2) (2017) 259.
[34] A.J. Cooner, Y. Shao, J.B. Campbell, Detection of urban damage using remote sensing and machine learning algorithms: revisiting the 2010 Haiti earthquake, Rem. Sens. 8 (10) (2016) 868.
[35] C. Geiß, P.A. Pelizari, M. Marconcini, W. Sengara, M. Edwards, T. Lakes, H. Taubenböck, Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques, ISPRS J. Photogrammetry Remote Sens. 104 (2015) 175–188.
[36] M. Vakalopoulou, K. Karantzalos, N. Komodakis, N. Paragios, Building detection in very high resolution multispectral data with deep learning features,, in: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, 2015, pp. 1873–1876.
[37] Y.-J. Cha, W. Choi, O. Büyüköztürk, Deep learning-based crack damage detection using convolutional neural networks, Comput. Aided Civ. Infrastruct. Eng. 32 (5) (2017) 361–378.
[38] M. Wieland, M. Pittore, S. Parolai, J. Zschau, Exposure estimation from multi-resolution optical satellite imagery for seismic risk assessment, ISPRS Int. J. Geo-Inf. 1 (1) (2012) 69–88.
[39] C. Geiß, A. Schauß, T. Riedlinger, S. Dech, C. Zelaya, N. Guzmán, M.A. Hube, J. J. Arsanjani, H. Taubenböck, Joint use of remote sensing data and volunteered geographic information for exposure estimation: evidence from valparaíso, Chile, Nat. Hazards 86 (1) (2017) 81–105.

[40] A. Raouf, Y. Peng, T.I. Shah, Integrated use of aerial photographs and lidar images for landslide and soil erosion analysis: a case study of wakamow valley, moose jaw, Canada, Urban Sci. 1 (2) (2017) 20.
[41] T. Perol, M. Gharbi, M. Denolle, Convolutional Neural Network for Earthquake Detection and Location, 2017 *arXiv preprint arXiv:1702.02073*.
[42] I. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnoud, V. Shet, Multi-digit Number Recognition from Street View Imagery Using Deep Convolutional Neural Networks, 2014 *arXiv preprint Arxiv:1312.6082*, 2015.
[43] J. Jin, K. Fu, C. Zhang, Traffic sign recognition with hinge loss trained convolutional neural networks, IEEE Trans. Intell. Transport. Syst. 15 (5) (2014) 1991–2000.
[44] F.-Y. Gong, Z.-C. Zeng, F. Zhang, X. Li, E. Ng, L.K. Norford, Mapping sky, tree, and building view factors of street canyons in a high-density urban environment, Build. Environ. 134 (2018) 155–167.
[45] F.-Y. Gong, Z.-C. Zeng, E. Ng, L.K. Norford, Spatiotemporal patterns of street-level solar radiation estimated using google street view in a high-density urban environment, Build. Environ. 148 (2019) 547–566.
[46] J. Liang, J. Gong, J. Zhang, Y. Li, D. Wu, G. Zhang, Gsv2svf-an interactive gis tool for sky, tree and building view factor estimation from street view photographs, Build. Environ. 168 (2020) 106475.
[47] C.-B. Hu, F. Zhang, F.-Y. Gong, C. Ratti, X. Li, Classification and mapping of urban canyon geometry using google street view images and deep multitask learning, Build. Environ. 167 (2020), 106424.
[48] J. Kang, M. Körner, Y. Wang, H. Taubenböck, X.X. Zhu, Building instance classification using street view images, ISPRS J. Photogram. Remote Sens. 145 (2018) 44–59.
[49] A. Middel, J. Lukasczyk, S. Zakrzewski, M. Arnold, R. Maciejewski, Urban form and composition of street canyons: a human-centric big data and deep learning approach, Landsc. Urban Plann. 183 (2019) 122–132.
[50] T. Gebru, J. Krause, Y. Wang, D. Chen, J. Deng, E.L. Aiden, L. Fei-Fei, Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the United States, Proc. Natl. Acad. Sci. Unit. States Am. (2017) 201700035.
[51] "Estudio general de amenaza sísmica de Colombia," Asociación Colombiana de Ingeniería Sísmica, Comité AIS–300, Bogotá, Colombia, Tech. Rep. (2009).
[52] "Instrumentación y microzonificación sísmica del área urbana de medellín," Sistema Municipal para la Prevención y Atención de Desastres, Medellín, Colombia, Tech. Rep. (1999).
[53] C.J. Jordan, K. Adlam, K. Laurie, W. Shelley, J. Bevington, User guide: windows tool for field data collection and managment, GEM Foundation, Pavia, Italy, Tech. Rep. (2014), 2014-04 V1.0.0, Tech. Rep..
[54] Rapid Visual Screening of Buildings for Potential Seismic Hazards: Supporting Documentation, vol. 154, FEMA, Washington, DC, USA, 2002. Standard.
[55] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (3) (2015) 211–252.
[56] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014 arXiv preprint arXiv:1409.1556.
[57] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826.
[58] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[59] F. Chollet, Xception: deep learning with depthwise separable convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1251–1258.
[60] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks?, in: Advances in Neural Information Processing Systems, 2014, pp. 3320–3328.