ENABLING COMPONENT REUSE FROM EXISTING BUILDINGS THROUGH MACHINE LEARNING

Using Google Street View to Enhance Building Databases

DEEPIKA RAGHU¹, ARETI MARKOPOULOU², MATHILDE MARENGO³, IACOPO NERI⁴, ANGELOS CHRONIS⁵ and CATHERINE DE WOLF⁶

¹,⁶ETH Zurich, Dept. of Civil, Environmental and Geomatic Engineering (D-BAUG), Stefano-Franscini-Platz, 8093 Zurich, Switzerland.
²,³,⁴,⁵Institute for Advanced Architecture of Catalonia, Carrer de Pujades, 102, 08005, Barcelona, Spain

¹draghu@ethz.ch, 0000-0003-0725-6415
²areti@iaac.net, 0000-0002-5400-0837
³mathilde.marengo@iaac.net, 0000-0002-6249-0960
⁴iacopo.neri@iaac.net, 0000-0002-9246-6494
⁵angelos.chronis@iaac.net, 0000-0002-6961-2975
⁶cdewolf@ethz.ch, 0000-0003-2130-0590

Abstract. Intense urbanization has led us to rethink construction and demolition practices on a global scale. There is an opportunity to respond to the climate crisis by moving towards a circular built environment. Such a paradigm shift can be achieved by critically examining the possibility of reusing components from existing buildings. This study investigates approaches and tools needed to analyse the existing building stock and methods to enable component reuse. Ocular observations were conducted in Google Street View to analyse two building-specific characteristics: (1) façade material and (2) reusable components (window, doors, and shutters) found on building facades in two cities: Barcelona and Zurich. Not all products are equally suitable for reuse and require an evaluation metric to understand which components can be reused effectively. Consequently, tailored reuse strategies that are defined by a priority order of waste prevention are put forth. Machine learning shows promising potential to visually collect building-specific characteristics that are relevant for component reuse. The data collected is used to create classification maps that can help define protocols and for urban planning. This research can upscale limited information in countries where available data about the existing building stock is insufficient.

Keywords. Machine Learning; Component Reuse; Google Street View; Material Banks; Building Databases; SDG 11; SDG 12.
1. Introduction

Billions of tonnes of construction and demolition waste (CDW) is annually disposed of in landfills, causing severe environmental damage (Duan et al., 2019). Mismanagement of CDW significantly contributes to climate change as it causes water, ground, and air pollution which consequently damages ecosystems and affects human health. Studies show that up to 95% of non-hazardous CDW is either reusable or recyclable (Ma et al., 2020). Yet, insufficient access to information on the CDW stream hinder reuse opportunities, resulting in a linear “take-make-waste” economy. The concept of a “circular economy” (CE) is quickly gaining momentum in the construction industry. In a CE, the value resources are maintained for as long as possible to minimise the generation of waste. Recent studies for a CE in the construction industry view ‘buildings as material banks’ (BAMB; Debacker et al., 2016). The focus is on new buildings which are being developed with building information modelling (BIM) to retain high value components for future use. However, these studies do little to address the fundamental barriers of reusing materials and components in already existing buildings which lack digital records. This has led to the reframing of BAMB to ‘existing buildings as material banks’ (E-BAMB; Rose and Stegemann, 2018). In order to successfully reuse components from existing buildings, detailed information about the building stock is required.

The contribution of this research is to investigate the prospect of estimating component reuse from residential buildings that are currently facing renovation or demolition needs in two European cities: (i) Barcelona and (ii) Zurich. The buildings chosen to be studied were built in the 20th century and are deficient of necessary data to enable reuse. Image processing techniques such as deep learning are employed for extracting information on the building components. The methods are applied on street view images of both Barcelona and Zurich to show its applicability despite regional differences. To the authors’ knowledge, this is the first study to identify materials on existing building facades using Google Street View. Prospective reuse measures are then tailored to each building component based on their as-built characteristics. This can enable long-term strategic planning for effective reuse of the residential building stock built in the 20th century.

2. 'As-built' documentation of existing buildings

Public building registers generally collect information on general building statistics (built up area, height, number of floors, etc.) but still lack key indicators on the components of the building. In recent years, laser scanning (LS) technology that can be used to produce BIM models is being increasingly explored to identify component attributes in existing buildings. Although the technology is quite promising, some of the barriers to LS include high costs for the sophisticated equipment, the time consuming and laborious processing of the complex data and the inconvenient and large file types (Uotila et al., 2021). Due to the high expenses associated with the technology, it is only viable in large projects when scheduled for immediate demolition. Another source of information on reusable building components are reused material marketplaces (RMMs). RMMs are supply-led interfaces where building components or products are sold online. Usually, contractors lack incentive to use
RMMs and don’t put up their unwanted items for sale. For this reason, RMMs have still not gained enough traction for effective sales. Furthermore, the items are put up on the platform at the very moment they arise as waste, making it difficult to incorporate them into the design of new buildings.

Architects require detailed information about the reusable elements early in the design process. The Pareto principle outlines the importance of time in its ability to influence a project (Figure 1a). The earlier in time decisions are taken in a project lifecycle, the greater the potential to affect the outcome of the project, when compared to decisions taken at a later stage (Chini and Balachandran, 2012). Hence, generating data for an inventory of reusable building components at an urban scale requires quick, economically feasible and simple methods that can complement existing computer vision methods of data collection. Such a ‘material database’ of reusable materials can allow designers to check the forthcoming availability of components and assess their suitability for reuse in new building projects ahead of time (Figure 1b).

Untapped data sources that could help enable E-BAMB through early identification are open, unstructured image datasets of buildings. In this paper, Google Street View, which provides 360-degree panoramic imagery of streets and their surroundings in urban and rural areas across the globe is studied. Many applications in urban planning have been implemented using Google Street view such as estimating the demographic makeup of the cities (Gebru et al., 2017), estimating the building age (Li et al., 2018) and studying the relationships between city appearance and the health of its residents (Dubey et al., 2016). However, utilizing Google Street View for predictions of building-specific suitability for reuse remains an unexplored area of research.

3. Methods

In this section, the methods of collecting data, developing ML models, and applying the algorithms to the residential building stock built between 1925 and 1975 that are facing renovation or demolition needs in Barcelona and Zurich are described. First, public registers in both cities are examined to identify gaps in information required for component reuse. Second, ocular observations are conducted in Google Street View to obtain a comprehensive image dataset. For processing this data, ML models are trained
to predict building-specific characteristics that are hitherto unknown. Third, tailored strategies that help assess the feasibility of reusing the building components are presented. Finally, the potential of using the generated data in building classification maps for urban planning is depicted.

To investigate the data availability in public registers that could be relevant for building component reuse, open government databases in Barcelona and Zurich were queried (Table 1).Registrations of construction year, built-up area or building use, provide insights on the distribution of building types in cities (Figure 2). Furthermore, notifications on ongoing construction, renovation and demolition projects in the cities support studies for estimating the supply and demand of reusable components. Yet, these databases are still limited by the lack of information on reusable components from existing buildings. Based on the gap between the available data in the public registers and the data needed to assess the feasibility of reusing building components, two building-specific characteristics: (i) façade materials and (ii) reusable components (window, doors, and shutters) were chosen to be studied.

Table 1. An overview of the relevant data available to enable reuse in Barcelona and Zurich.

<table>
<thead>
<tr>
<th>Data source</th>
<th>Building characteristics</th>
<th>Measurement type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjuntament de Barcelona</td>
<td>Year of construction</td>
<td>Scale variable (year)</td>
</tr>
<tr>
<td></td>
<td>Number of floors</td>
<td>Ordinal</td>
</tr>
<tr>
<td></td>
<td>Construction condition</td>
<td>Ordinal (ruins, bad, okay, good)</td>
</tr>
<tr>
<td></td>
<td>Building permit</td>
<td>Nominal (construction, renovation, demolition)</td>
</tr>
<tr>
<td>Stadt Zurich</td>
<td>Year of construction</td>
<td>Scale variable (year)</td>
</tr>
<tr>
<td></td>
<td>Number of floors</td>
<td>Ordinal</td>
</tr>
<tr>
<td></td>
<td>Building volume</td>
<td>Scale variable (m²)</td>
</tr>
<tr>
<td></td>
<td>Building permit</td>
<td>Nominal (construction, renovation, demolition)</td>
</tr>
</tbody>
</table>

Figure 2a. Map of Barcelona building stock distributed by built-up area (m²); Figure 2b. Map of Zurich building stock distributed by date of construction (year).
4. Scheme 1: Façade material

Historically, the construction industry has made use of structured data (3D models, databases, spreadsheets) but tended to ignore all other sources of data (text documents, site images, project schedules) as it was too overwhelming to draw conclusions from due to the challenges of data preparation, representation and analysis (Soibelman et al., 2008). Today, emerging technologies like ML allow the processing and management of vast sources of unstructured data. Material classifications via ML are popular among the research community in almost every domain. In the construction industry, most existing literature on material recognition focuses on construction activities in site images for monitoring purposes (Akinosho et al., 2020). This section explores the use of a ML algorithm to detect three building façade materials, namely brick, stone, and mixed material facades. To generate a dataset, 2160 images of the residential stock in Barcelona and 1780 images of the residential stock in Zurich that were constructed between 1945 and 1975, facing renovation or demolition needs were observed in Google Street View. The geo-tagged images are downloaded through the Google Street View API, with the associated metadata, i.e., the image size and pitch value which were set to be 610×610 pixels and 0 degrees, respectively. In some cases, observations were not possible due to lack of coverage in Google Street View. However, the dataset was still sufficient for iterative testing. Samples of the captured images are illustrated in Figure 3a. Images of each façade material type were manually classified by placing them in folders with corresponding IDs. The data was divided into three sections: 70% was used for training, 10% for validation and 20% for testing. Since the training data set was limited in size, data augmentation was employed to extend the dataset for training by manipulating images in the dataset artificially. The data was augmented by changing the contrast, gamma, and saturation of the image to prevent overfitting. This helped validate the robustness of the model to recognize not just different colours and appearances but also the texture of the material (Figure 3b).

![Figure 3a](image)

**Figure 3a.** Samples of the ocular observations conducted in Google Street View; **Figure 3b.** Data augmentation by changing illumination of image; **Figure 3c.** Results of the façade material prediction.

A convolutional neural network (CNN) is a deep learning approach that has been widely used for analysing visual images. CNNs can learn highly abstract features and
can identify objects efficiently. To optimize computing power, the lightweight neural network MobileNetV2 (Sandler et al., 2019) with transfer learning was employed. Transfer learning is the process of training and predicting on a new dataset using a pre-trained model that has been learned on a previous dataset. This has been proven to be an efficient way for adaptation of the CNN to a new training task for datasets that are not large enough to train a CNN with many parameters from scratch. The model was created from the MobileNetV2 model pre-trained by Google on the ImageNet dataset utilizing dropout and batch normalization. An accuracy of 67% was achieved with 150 epochs per run. With an increase of epochs per run to 300, 74% accuracy was attained. Thus, with further data augmentation and training, more accurate and reliable results can be achieved. The system yields sufficient results on facade images of both Barcelona and Zurich, despite regional and architectural variation in the facades', showing general applicability of the technique (Figure 3c).

The predicted characteristics can then help estimate the reuse potential of the building facades (Figure 4). It is advantageous to know whether a building has a brick façade or not, as brick facades must often be preserved due to their cultural and historical value (Šekularac et al., 2020). Tailored reuse strategies for the facades are put forth based on a priority order of waste prevention.

Figure 4. Tailored reuse strategies for facades from buildings marked for demolition

The deep learning model used to identify facade materials can be replicated to predict other characteristics such as if a building has an eave overhang. Buildings marked for demolition with a brick façade and an eave overhang may be directly reused as the elements are close to their original status and for their original purpose, needing almost no further processing. Lendager Group’s Resource Rows project in Copenhagen demonstrates such brick façade reuse as seen on the left in Figure 4. Brick modules from historical breweries are cut out, installed in steel frames, and stacked on a new building (Lendager Group, The Resource Rows in 2021). Brick façades with no eave overhang may require operations of cleaning, renovating, or repairing. This method of reuse alters the brick to serve a new function. The reclaimable materials from mixed material façades are more complex. In this case, rethinking the function of the material is necessary. In combination with other substances, the waste material can be utilized to produce secondary goods with a new utility.

The dataset of buildings due for renovation or demolition were identified from building permits in the year 2020. Due to the relatively short time span of the building permit, some permits may not have been issued at the time of study or the work on the
building may already have been completed. Given the time-sensitivity of building permits, accounting for their temporality is critical. Methods to augment existing databases dynamically by building owners or contractors are essential. Supplementarily, anomaly detection techniques on facades as investigated by Barahona et al., 2021, can help automate the identification of buildings requiring retrofits.

5. Scheme 2: Reusable components (windows/doors/shutters) on building facades

Public datasets such as the Ecole Centrale Paris (ECP) have been used as a benchmark for facade parsing previously (Teboul et al., 2010). The ECP dataset only contains facades that are manually rectified and viewed in a front-parallel direction. Hence, the same dataset generated in Scheme 1 was used with the intention of encompassing a large variety of window types from different camera positions. This scheme explores the use of the state-of-art neural network Mask R-CNN (Girshick et al., 2014) for window detection. The VGG Image Annotator (VIA) was used to manually annotate the images. VIA is a single HTML file that can be opened in a browser. The tool saves the annotations in a JSON file with each mask as a set of polygon points. A collection of 150 images with 1478 annotated windows was generated. These images were then split into a training (120 images) and validation dataset (30 images). Transfer learning was applied with weights pre-trained from the COCO dataset. The implementation was hosted on Google Colaboratory, for free GPU usage. Some predicted results were imprecise, and the window boundary was incorrectly annotated due to the presence of open shutters (Figure 5). Overall, experimental detection results showed that with only transfer learning, the suggested approach can produce instance segmentation of windows on façade images of different countries despite variation in the façade and window appearance.

![Figure 5. Detection of windows for reuse on the Barcelona (left) and Zurich (right) image datasets](image-url)
These results obtained can be combined with public register data to identify tailored reuse strategies (Figure 6). For example, windows in new buildings or buildings which have undergone recent renovation and have an energy performance certificate (EPC) rating of ‘A’ to ‘C’ may be directly reused in new construction. This is because of the high thermal quality of the windows. For older buildings with an EPC rating of ‘D’, improving the insulation of the window glazing or changing the window frames may be necessary. For very old buildings with an EPC rating of ‘E’ to ‘G’, alternate design strategies for using the glazing material will need to be considered because of the poor thermal quality of the windows.

Figure 6. Tailored reuse strategies for windows from buildings marked for demolition

This section is limited to identifying the quantity of reusable components from street view images. Additional information on the surface area of each identified component can be retrieved with a reference of the height of the building. This calculation can be performed with traditional computer vision techniques as shown by Kolenbrander et al., 2017.

6. Data navigation and impact

Enhanced E-BAMB databases with component information could generate more long-term reuse strategies that can be used for decision-making. Most components that are picked out for reuse are most often just about to be consigned as waste. This makes it difficult to incorporate reuse in new design workflows. In this regard, building classification maps, wherein a building address is provided to visually collect building specific information can be useful (Figure 7a). The time buffer generated with this approach increases the likelihood of components being reused when they are available for sale (Figure 7b). The process also expands the prospect of finding a substitute for common components in advance. This methodology will not eliminate the need for intermediate storage altogether, but it may help in partial reduction of storage time, thereby also reducing warehouse costs. The introduction of a material database that can be accessed during the planning stage of projects allows identified reusable components to be incorporated into new design development schemes. Detailed specifications can be updated after the procurement of the components is complete.
ENABLING COMPONENT REUSE FROM EXISTING BUILDINGS THROUGH MACHINE LEARNING

Figure 7a. Urban-scale E_BAMB classification map showing the location of brick, stone, and mixed material facades in the city of Barcelona; Figure 7b. Project-scale understanding of E-BAMB data

7. Conclusion and future work

This paper explores the opportunity to make better estimations about component reuse from existing buildings. To do so, machine learning methods are used to enhance building databases with information relevant for reuse. This is especially important as the review of public registers showed no data on materials or components. Data can be expensive to obtain, hence, it is essential to clarify for what purpose the register is established so that the data collection methods are feasible. ML methods can be used to find time-dependent patterns which makes them quite suitable for obtaining building-specific information. The methodology proposed in this paper can be used for data collection on materials and components available for reuse in existing buildings at national or regional levels. The contribution of this paper is the addition of façade material and window count as building-specific characteristics. Furthermore, the purpose of this study is to showcase how existing buildings stock data can be easily accessed to make decisions during early planning stages for matching supply and demand of reusable components in a quick and inexpensive way that can enable a wide range of analysis.

The scope of this paper is limited to the building specific predictions from street-view images. However, the automated methods described in this study can be used on other non-centralized, geo-located data, such as social media images to collect data on the building interior. Clarity on how much uncertainty can be allowed concerning the data quality needs to be established. For greater data accuracy or for filling critical data gaps that cannot be met with the automated methods described in this paper, pre-demolition audits with physical inspections can be carried out. Future work could explore other techniques to enrich building databases with more building characteristics relevant for energy retrofitting (von Platten et al., 2020).

References


