Component-level Residential Building Material Stock Characterisation Using Computer Vision Techniques

Menglin Dai,*,† Jakub Jurczyk,‡ Hadi Arbabi,‡ Ruichang Mao,¶ Wil Ward,‡ Martin Mayfield,‡ Gang Liu,† and Danielle Densley Tingley‡

†*College of Urban and Environmental Sciences, Peking University, 100871 Beijing, China*

‡*Department of Civil and Structural Engineering, The University of Sheffield, Sheffield, S1 3JD, United Kingdom*

¶*Department of Environmental and Resource Engineering, Technical University of Denmark, Kgs Lyngby 2800, Denmark*

E-mail: menglin.dai@pku.edu.cn

Phone: +44 (0)7446885153

Abstract

The residential building material stock constitutes a significant part of the built environment, providing crucial shelter and habitat services. The hypothesis concerning stock mass and composition has garnered considerable attention over the past decade. While previous research has mainly focused on the spatial analysis of building masses, it often neglects the component-level stock analysis or requires heavy labour cost for onsite survey. This paper presents a novel approach for efficient component-level residential building stock accounting in the UK, utilising drive-by street view images and building footprint data. We assessed four major construction materials: brick, stone, mortar, and glass. Compared to traditional approaches that utilise surveyed material intensity data, the developed method employs automatically extracted physical dimensions of building components incorporating predicted material types to calculate material mass. This not only improves efficiency but also enhances accuracy in managing the heterogeneity of building structures. The results revealed an error rate of 5% and 22% for mortar and glass mass estimations, and 8% and 7% for brick and stone mass estimations, with known wall types. These findings represent significant advancements in building material stock characterisation and suggest that our approach has considerable potential for further research and practical applications. Especially, our method establishes a basis for evaluating the potential of component-level material reuse, serving the objectives of a circular economy.

Keywords: building material stocks, urban sustainability, circular economy, deep learning, computer vision, building facade, street view imagery

Synopsis: This study introduces a computer vision-based method for precise component-level quantification of building materials, advancing circular economy efforts.

Introduction

Residential buildings, integral to human habitation and contributing 62.2% of total building carbon emissions[1], play a pivotal role in the built environment achieving the United Nations' Sustainable Development Goals[2]. These structures rely on construction materials which have formed a major portion of the anthropogenic mass of approximately 1.1 teratonnes, which has exceeded living biomass since 2020.[3,4] Knowledge, e.g. type and quantity, of these materials is crucial to facilitating a circular economy,[5] aiding in building decarbonisation efforts by reducing demand for new materials through urban mining[6] and facilitating related policy-making.[7] Existing research largely studies the residential building stock from a geographical viewpoint, focusing on spatial analysis of material stock,[8,9] but often overlooks individual building analysis. This gap warrants further investigation for a more geographically specific, nuanced understanding.

The state-of-the-art techniques for acquiring building stock information can be aptly classified into three distinct approaches:[8,10,11] top-down, bottom-up, and remote sensing. The top-down approach involves viewing target objects as a complete system, with the material stock being equivalent to the mass balance of inflow and outflow within the system. The utilisation of socioeconomic statistics data is a prevalent practice in the implementation of this approach.[12–15] As a result, it is often employed for material stock simulation at the level of a nation due to the data availability.[8,16] Conversely, the bottom-up approach begins at the end-use object inventory stage, where the number of buildings is collated to derive the building material stock for a specific time period, with material intensity coefficients employed to calculate the material stock.[17–22] The implemented bottom-up approaches have shown their capability of achieving considerably finer resolution than top-down approaches. Recently, the application of remote sensing data for material stock accounting has garnered considerable attention.[23–25]

To accurately evaluate building material stock at the individual building-level resolution, component-level understanding of each structure is imperative.[26,27] Consequently, top-down and satellite image-based remote sensing methods are insufficient for this task. Frequently utilised bottom-up techniques rely on publicly available data for material intensity, necessitating meticulous and labour-intensive preparation. In scenarios where no such data is accessible, collaboration with domain experts may prove a feasible alternative, but it is still highly labour intensive.[28] An alternative is to develop an automated method for estimating building material stock to facilitate a more efficient procedure.

Street view imagery serves as a potent data source, encapsulating substantial building-related information with promising utility to estimate building material stock. Over the past decade, navigation companies, notably Google[29], have significantly enhanced the availability of such imagery. While previous research has leveraged Street view imagery across various domains[30], including specific building attributes such as type[31], age[32], and window-to-wall ratio[33], the application of this data type for comprehensive building stock evaluation remains relatively uncharted and poses distinct challenges.

One salient challenge is a lack of data containing facade imagery and registered building information. Machine learning techniques, particularly deep learning, offer promising avenues for building attribute estimation using street view imagery. However, these methods necessitate substantial, diverse, and accurately labelled training data, which presents significant challenges in terms of data collection, annotation, and quality control.

Another challenge lies in the limited availability of data capture locations; Street View services may not consistently provide frontal, complete views of each building, thus resulting in partial or oblique images. Moreover, discerning intricate building details necessitates high-resolution imagery. For instance, distinguishing between non-rendered cavity walls and solid brick walls requires the visualisation of brick patterns and, consequently, the discernible mortar joints, typically 1 cm thick in masonry structures. To attain the requisite clarity, imagery should ideally possess a minimum of 1 pixel per centimetre, translating to a 600-pixel image size for a standard two-storey, 6-meter tall building. Nevertheless, the Google Street View API[29] currently limits downloadable image size to 640 pixels, which significantly constrains the effective use of this service for detailed analysis.

This paper presents a novel method that integrates high-quality geo-referenced street view facade images with machine learning-based computer vision techniques to enable component-level building material stock characterisation. The study typically focuses on houses lower than three storeys which make up 94% of UK households.[34] This proposed approach facilitates the characterisation of individual buildings by capturing essential features such as component quantity, composition, built form, and age while requiring fewer assumptions than conventional bottom-up models. To develop this approach, we have compiled a comprehensive dataset of 2,292 houses, enriched with high-resolution facade images and detailed attributes, and created specialised datasets for interior wall length estimation using facade features (300 UK houses) and construction material recognition (13,562 labelled image patches). Additionally, a new building age detection dataset (9,278 images) has been introduced, along with an innovative multi-task deep learning model for simultaneous building age and built form recognition.

Materials and Methods

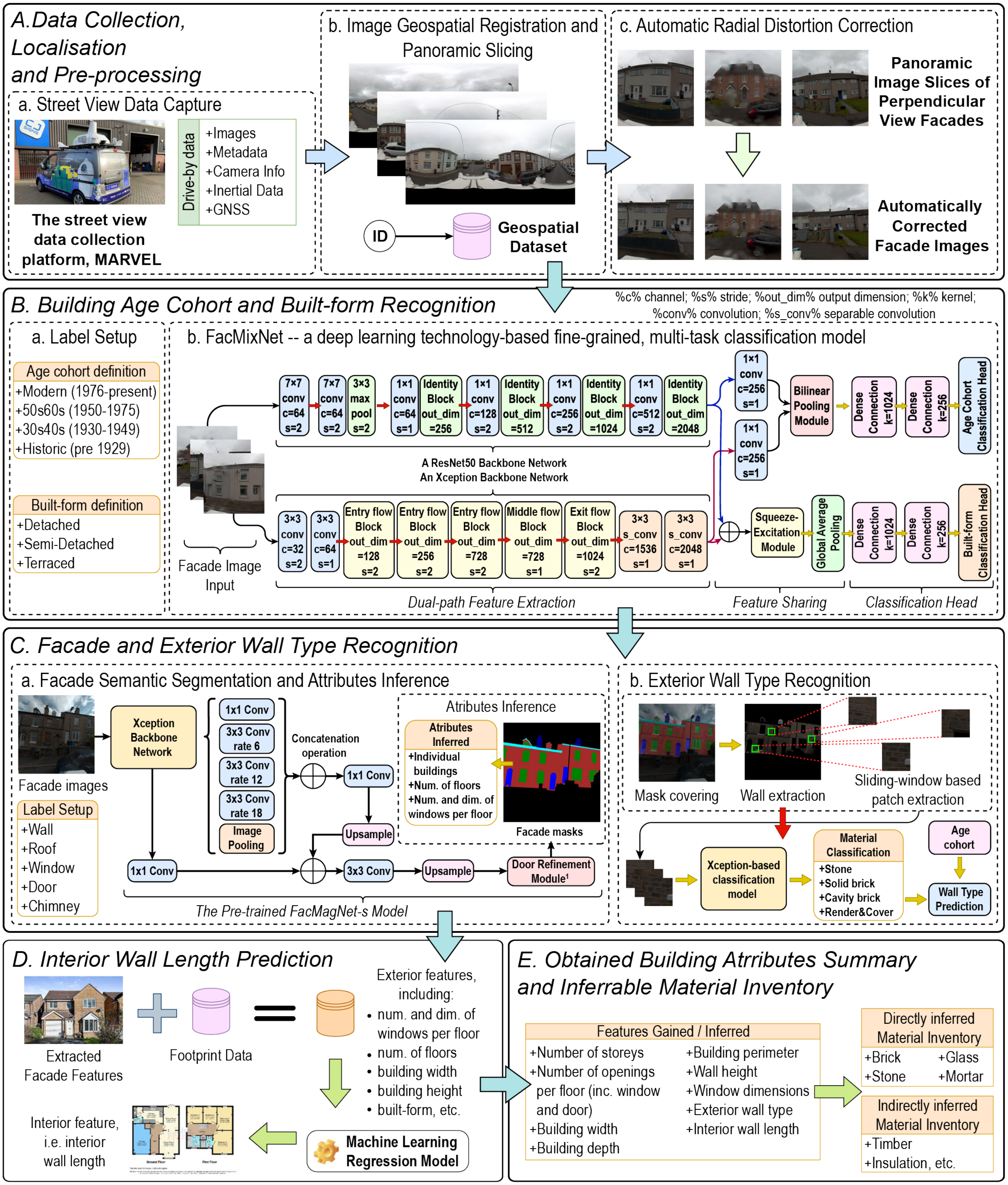
Workflow Overview

Figure 1 illustrates the workflow of the developed approach, which comprises of five individual modules, from data collection to material stock calculation. The data collection process aims to establish a matched facade image & building footprint data. Subsequently, the facade images are utilised to predict houses' age cohort, built form, i.e., detached, terraced, and semi-detached, and component locations, as illustrated in panels B and C. The predicted facade masks facilitate the estimation of the number of floors and openings i.e. windows and doors per floor. Moreover, the predicted masks are employed to extract patches from wall areas, which are subsequently used to predict the wall material type. The estimated building attributes are then fed into a regression model to predict the inner wall length, followed by the computation of the volume of inner walls. In the final step, all the achieved building attributes are used to calculate the mass of each designated material type.

Overall, our approach uses visible facade features and publicly available building footprint data to derive an individual house's material stock. The material stock of a given area can then be determined by summing the material stock of all individual buildings. The selected visible features include facade wall type, age, built form, and the spatial distribution of building openings, e.g. windows and doors. Wall type directly correlates to wall materials and in some cases may indicate insulation conditions while building age and built form link to pre-defined archetype databases such as the TABULA dataset[36]. These databases aid in the estimation of invisible materials such as timber from floor decks and roof structures and insulation. The spatial distribution of windows can be used to estimate the mass of glass. By incorporating dimension information obtained from the building footprint database, the mass of brick, stone, mortar and glass can then be directly obtained. This approach represents a significant contribution to the field of material stock analysis, as it provides a comprehensive and automated workflow for the estimation of building material stock from facade images.

Drive-by Data Capture

A built vehicle-mounted data capture platform[37] was employed to collect street view facade images in this study. The platform contains an advanced multi-camera rig and an onboard inertial measurement unit (IMU) and a global navigation satellite system (GNSS) unit for



[1] Referring to Dai et, al. (2021) for the refinement module structure

Figure 1: The developed material stock inventory characterisation pipeline. The data capture platform figure in Panel A-a is adapted from Dai et al. (2022).[35]

capturing raw image data and synchronised orientation/location data. The camera rig comprises six separate 2/3-inch Sony IMX264 CMOS sensors with 2048×2448 effective pixels and wide-angle lenses with a frequency of 30 frames per second (FPS). The cameras are oriented with one on the top pointing upwards and the other five positioned horizontally along the sides forming a regular pentagon. The combined capture has a field-of-view of 90% of the full sphere.

During the data collection process, the capture platform traverses the designated area, which for demonstration in this paper is Merthyr Tydfil town, Wales, the UK. The platform travels at a speed of approximately 4.5 meters per second and captures data with a frequency of 10 FPS, around 12 images for every meter travelled. At a distance of 10 meters from the sensing vehicle, each pixel in the image is representative of an area of approximately 2.5 cm^2 on the target surface which exceeds the 1cm per pixel requisite stated previously. The onboard IMU/GNSS unit provides an orientation accuracy of $0.1°$ and a location accuracy of up to 0.1m, which leads to a 0.25m frame position accuracy at the 4.5m/s driving speed. The position data from the IMU/GNSS unit is recorded in the World Geodetic System (WGS84) format, consisting of longitude and latitude coordinates. For applications within the UK, these coordinates are reprojected to the Ordnance Survey National Grid reference system (OSGB 1936).

The previously proposed algorithm, designed to extract perpendicular views of specified houses and register these with footprint data, is employed to obtain the 'face-on' views of the designated houses[38,39]. The method first reconstructs the panoramic image using frames from all five sensors and then slices the captured panorama based on vehicle orientations. Then by adopting the Ordnance Survey topographic identifier (TOID), the extracted perpendicular view slices are linked to the property footprint data[40]. The extraction of perpendicular-view facade images from captured panoramas inherently introduces radial distortion. Given that these images are sliced from panoramic views, traditional chessboard-based image rectification methods[41] prove inapplicable. To address this, an automatic radial distortion correction approach[42] is implemented on the building-registered images prior to further processing.

Datasets

This section introduces all datasets built or used in this study. More details of these datasets such as annotation protocols, data distributions and comparison studies are available in the supporting information.

The Housing Attributes Dataset This dataset comprises 2,292 houses, featuring facade images, footprint data, age cohorts, built forms, wall types, and opening information. It is based on the footprint-registered street view facade images obtained through the data collection procedure in Merthy Tydfil, Wales, the UK. The facade data is then aligned with the UK Energy Performance Certificate (EPC) records,[43] employing property identifiers to obtain wall type, building age and built form data. Subsequently, visual wall materials, building dimensions, and opening sizes are manually obtained.

The Age-Built form Dataset A dataset comprising 9,278 annotated images was constructed using Google Street View data by following the subsequent steps: Initially, 21,207 EPC records were collected. Duplicate and invalid records were filtered out, and the remaining location data was used to query and download corresponding building images via the Google Street View API. Each downloaded image was manually inspected for address matching, appropriate camera orientation, and suitable resolution, retaining only images with perpendicular views and acceptable resolution. The finalised dataset was randomly divided into training (80%), validation (10%), and test sets (10%).

The EPC categorises household age cohorts based on energy performance, but direct utilisation of these labels is challenging as narrow age cohorts may result in significant confusion as shown in previous research.[32] This study harmonises EPC age cohort classifications by integrating the TABULA Building Archetype project[36] and the BRE Housing Survey[44], resulting in four simplified groups: pre-1930s, 1930-1949, 1950-1975, and 1976-present. Additionally, built-form labels are streamlined, merging end-terraced and mid-terrace categories into a single terrace label.

The Facade Recognition Dataset The Sheffield Crookesmoor facade recognition dataset which is fine-labelled for building facade semantic segmentation is adopted in this study.[45] This dataset consisted of 997 images annotated to five categories: window, door, wall, roof and chimney and specifically focusing on UK housing, and the data was obtained from the same data capture platform as in this study. The dataset has been randomly split into 80%, 10%, and 10% training, validation, and test sets, respectively.

The Material Patch Dataset A material patch dataset comprising 13,562 images was constructed for this study, labelled into four categories: solid brick, cavity brick, stone, and render. This dataset was created using data from the previous captures with the same data collection platform as this study.[35,39,45] The dataset was randomly split using the same ratio as in previous datasets, i.e., 80% for training, 10% for validation, and 10% for testing.

The Inner Wall Regression Dataset The dataset contains 300 houses with floorplans and annotations across the UK. The dataset is sourced from real estate websites Zoopla (`https://www.zoopla.co.uk/`) and Savills (`https://www.savills.com/`). Co-authors with architecture expertise are responsible for collecting raw data to ensure its high quality. The samples were meticulously labelled—visually and through AutoCAD—enabling precise identification and measurement of features, including building types, interior wall length, perimeters, width, depth and quantities and dimensions of windows and doors.

Building Age Cohort and Built Form Recognition

In recent years, automatic recognition of building ages using facade images has garnered significant attention in the context of deep learning.[32,39,46,47] Zeppelzauer et al. (2018)[32] pioneered a patch-based approach, wherein building images are segmented into patches containing potential age-relevant features. These patches are input to a deep learning model, and aggregated predictions inform the building's age estimate.

However, this patch-slicing technique, initially designed for fine-grained classification tasks, is non-differentiable and introduces extra computational costs due to its two-stage nature of localisation and classification.[48] Bilinear pooling, a fully differentiable technique, has emerged as a compelling alternative, achieving comparable performance with reduced computational overhead and enabling end-to-end training.[49]

The proposed model in this study, named FacMixNet, integrates a shared feature extraction architecture for the dual prediction of building age and built form—two attributes that are conventionally classified separately.[46,50] The novel multitask learning framework posits the potential interrelation of features used for both age and built form recognition, as depicted in Figure 1 Panel B.

FacMixNet adopts a dual-path architecture, using ResNet50[51] and Xception[52] as distinct backbone networks for feature extraction. For age prediction, a bilinear pooling module fuses features, but given its substantial memory demands, 1×1 convolution kernels are applied to reduce feature map dimensions. In contrast, for built-form prediction, FacMixNet employs a straightforward concatenation operation, premised on the assumption of more distinct feature categories. A channel-wise attention module is introduced post-concatenation to accentuate key features.

In summary, this study introduces FacMixNet, a multi-task deep learning model designed for the dual prediction of building age and built form from facade images. It strategically employs bilinear pooling and concatenation methods, reflecting the nuanced demands of these distinct yet related classification tasks, and provides an efficient, streamlined solution for building attribute recognition. Training details are available in the supplementary file.

Facade Segmentation and Attributes Estimation

Facade semantic segmentation seeks to identify building components, such as windows and doors, in images at the pixel level. While rectified and cropped facade images have been prevalent for 3D procedural modelling,[53,54] recent work has shifted towards using street view images, which are more abundant and preserve building context.[45,55]

Door recognition, a critical aspect of segmentation, remains challenging due to its typically lower intersection-over-union (IOU) metrics compared to overall dataset mean IOU.[45,55,56] Doors are pivotal for discerning building attributes, such as the number of floors and individual units in terraced structures.

In this study, a more efficient version of the previously developed FacMagNet model[45] termed FacMagNet-s is proposed. FacMagNet-s leverages a Deeplabv3+[57] model as a multi-class classifier for predicting all classes. Instead of employing an object detection model to refine predicted components, FacMagNet-s directly utilises the door predictions to calculate bounding boxes, followed by the previously designed magnifier module. The model structure is shown in Figure 1 Panel C-a and training details are available in the supplementary file.

To isolate the spatial distribution of facade openings of individual houses, this work presents a pixel intensity-based algorithm. It applies a morphological open operation to a binary facade mask, followed by a vertical pixel intensity projection. This method identifies the primary building structure by detecting the longest span between sharp declines in pixel intensity, effectively excluding neighbouring structures from the analysis. For terraced and semi-detached houses, the location of doors is utilised to extract separate residences.

In estimating the distribution of windows on facades, we leverage the positions of doors, roofs, and wall trisection lines. The methodology is founded on specific architectural regularities, the details of which are outlined as assumptions in the supplementary file. By incorporating these assumptions, we develop a rule-based algorithm with K-means clustering to estimate the window arrangements in various house types, including terraced, semi-detached, and detached houses. Implementation particulars of this algorithm are detailed in the pseudo-code provided in the supplementary material. Once the vertical distribution of windows is ascertained, the achieved primary building location is then applied to obtain the window distribution of the designated house.

Exterior Wall Type Recognition

Wall construction types are crucial for estimating material requirements. In the UK, brick is the dominant construction material, with stone also being used. The British Energy Performance Certificates (EPCs) categorise walls into four types: cavity walls, solid brick walls, sedimentary rock walls, and igneous rock walls.[58] Stone and brick walls have distinct visual features, and brick layouts differ between solid and cavity walls.

A significant challenge is the rendering of outer walls, which can obscure the wall's texture—a key identifying feature—making visual data unreliable as shown in Figure 2. Moreover, different materials can be used within the same building for decorative purposes. An Xception model is first trained on the built material recognition dataset. Training details of the Xception model are available in the supplementary file.

Noting the correlation between wall types and building ages—solid walls in Victorian buildings and cavity walls become prevalent from the 1920s[59]—this study then introduces an age-assisted wall identification approach. This integrates visual characteristics, experiential insights, and building age data to overcome the ambiguities introduced by rendering and intra-building material variations, thereby enhancing wall identification accuracy.

During inference, depicted in Figure 1 Panel C-b, the predicted wall mask is used to extract wall area samples from raw images, which are then analysed by the trained Xception model. Subsequently, we employed a sliding box to randomly sample 50 patches from the wall area. These samples were then fed into the trained Xception model for inference. Acknowledging that diverse materials may be present on a single wall, we propose a material-ranking approach for filtering predictions. In this scheme, solid brick holds the highest priority; rendering wall has the lowest, and cavity brick and stone are of equal priority. For fully rendered buildings constructed post-1929, cavity brick is the default classification. For those built pre-1929, they are classified as solid brick. While buildings with stone features built after the 1930s are still classified as cavity brick.

Interior Wall Regression

To accurately estimate the material stock, the interior walls play a crucial role and acquiring such information is a tremendous challenge. The interior wall information can be retrieved from architectural layout plans or via onsite surveys. However, the former option is often hampered by severe data scarcity, while the latter is significantly labour-intensive. Meanwhile, to calculate the material stock, a detailed floor layout is not necessary but the total length of the inner wall is adequate. With known inner wall length and height from footprint data, the weight of the inner walls can be calculated if we know its material type which has been inferred in the previous section.

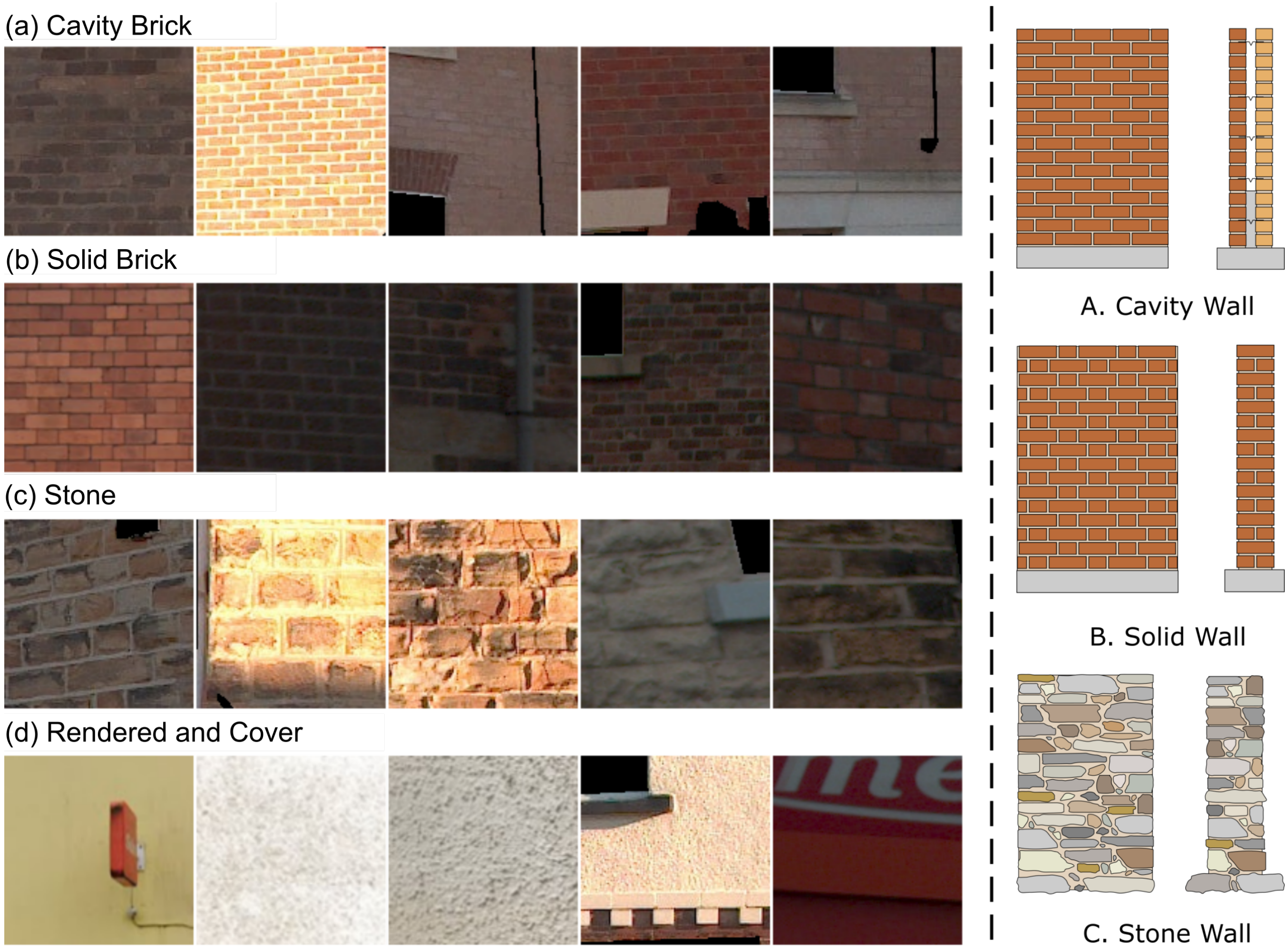


Figure 2: Typical Masonry Wall Patterns and Wall Type Structures. The three wall types are approximated to their archetypal forms: cavity walls consist of two layers of bricks, solid walls exhibit a cross-laid pattern, commonly referred to as Flemish bond, and stone walls are constructed using amorphous shaped stone chunks, which are adhered together using mortar.

We posit that there exists a correlation between the total inner wall length of a building and its external features, including windows, building length, building depth, and overall architectural form. Consequently, estimating the inner wall length presents a regression problem. In a similar vein, Yuan et al.[60] investigated the use of building exterior features to estimate the weight of buildings based on waste management data from Hong Kong. Their study involved the construction of a dataset comprising 78 building samples, utilising multiple data sources. Taking inspiration from their work, we built the inner wall regression dataset. We adhere to the same process as their work for our regression analysis.[60]

In choosing regression models, the Multi-Layer Perceptron (MLP) is adopted. The model structure is determined using a grid search strategy with the maximum number of hidden layers being 2 and the number of kernels ranging from 5 to 10. The activation function is determined to be Relu and the optimiser is chosen to be Adam. The number of epochs is determined to be 5000 and the early stopping is enabled. The model and evaluation are performed using the Python scikit-learn package.[61]

Material Stock Estimation

As our approach is vision-based, we primarily focus on calculating materials which are visible in a building in this paper including brick which is the major English housing construction material[62], stone, glass and mortar. By establishing connections between the derived building attributes and supplementary data sources, such as historical construction standards, we can deduce the use of other types of construction materials, including insulation, wood and metals. It is important to note, however, that as this paper serves as a framework, we do not delve into the estimation of the whole material masses at this stage.

In the UK, bricks typically adhere to a standard size of $215 \times 102.5 \times 65$mm, while mortar maintains a standard thickness of 10 mm.[63] Utilising these dimensions, a double-skin brick wall, whether a cavity wall or solid brick wall, necessitates an average of 118 bricks per square meter. Consequently, for a single-skin brick wall, the brick requirement would be reduced to 59 bricks per square meter. The brick density is assumed to be 2000 kg/m^3 and the mortar density is assumed to be 2300 kg/m^3.

A notable aspect of this study is the consideration of structural elements in buildings. First, the exterior walls are all assumed to be load-bearing. Second, when a building's long side exceeds 6 meters, the typical maximum span of a timber structural beam, a calculation is introduced: the building's long side length is divided by 6 meters, and the resulting quotient is multiplied by the building's short side length. This product is then assumed to represent the length of load-bearing walls within the building. Then the rest of the interior walls are assumed to be single-skin brick walls.

Despite the irregular patterns of stone walls, the approach assumes that stone used for construction can be treated similarly to regular bricks for weight calculations. In other ways, the calculation of the stone house is assumed to be the same as the brick houses. The density of stone is assumed to be 2500 kg/m^3.

To estimate glass weight, the resolution of the image is inferred by assuming each floor is 3m tall and obtaining the number of pixels of the wall height in the image. The glass area is then calculated from window dimensions, with an assumption of a 70-millimetre-wide window frame and a 4-millimetre glass thickness. Besides, buildings built after 1970 are assumed to be double glazing and otherwise to be single glazing. Using this area, the density of the glass and the assumed glazing type, the weight of the glass is computed.

Results and Discussion

Material Mass Estimation Performance

Figure 3 demonstrates the prediction performances of the computer vision-driven material stock estimation approach. EPC records and manually measured attributes provide reference values, while the proposed method furnishes predicted values. Each subplot includes a line with a gradient of one, denoting the expectation function, and the mass distributions for reference and prediction are displayed on the top and right margins, respectively.

The data reveals exemplary performance in mortar mass prediction, with a 5% error rate among the 2,292 samples, evidenced by the close alignment of the fitting function to the expectation function.

The prediction of glass mass has achieved the second-best result with an overall error rate of 22%. However, The variance of glass prediction is higher than the mortar.

Brick and stone mass predictions showcase three distinct clusters. One cluster aligns well with the expectation function, pointing to accurate wall type predictions. The other two clusters, residing along the x and y axes, represent misclassified wall types.

Individual Models' Performances

Figure 4 shows the trained classification models' deployment performances using confusion matrices. All confusion matrices have been normalised by dividing their number of true labels. All evaluation measures used in this study, including the confusion matrix and various evaluation metrics, are detailed in the supporting information.

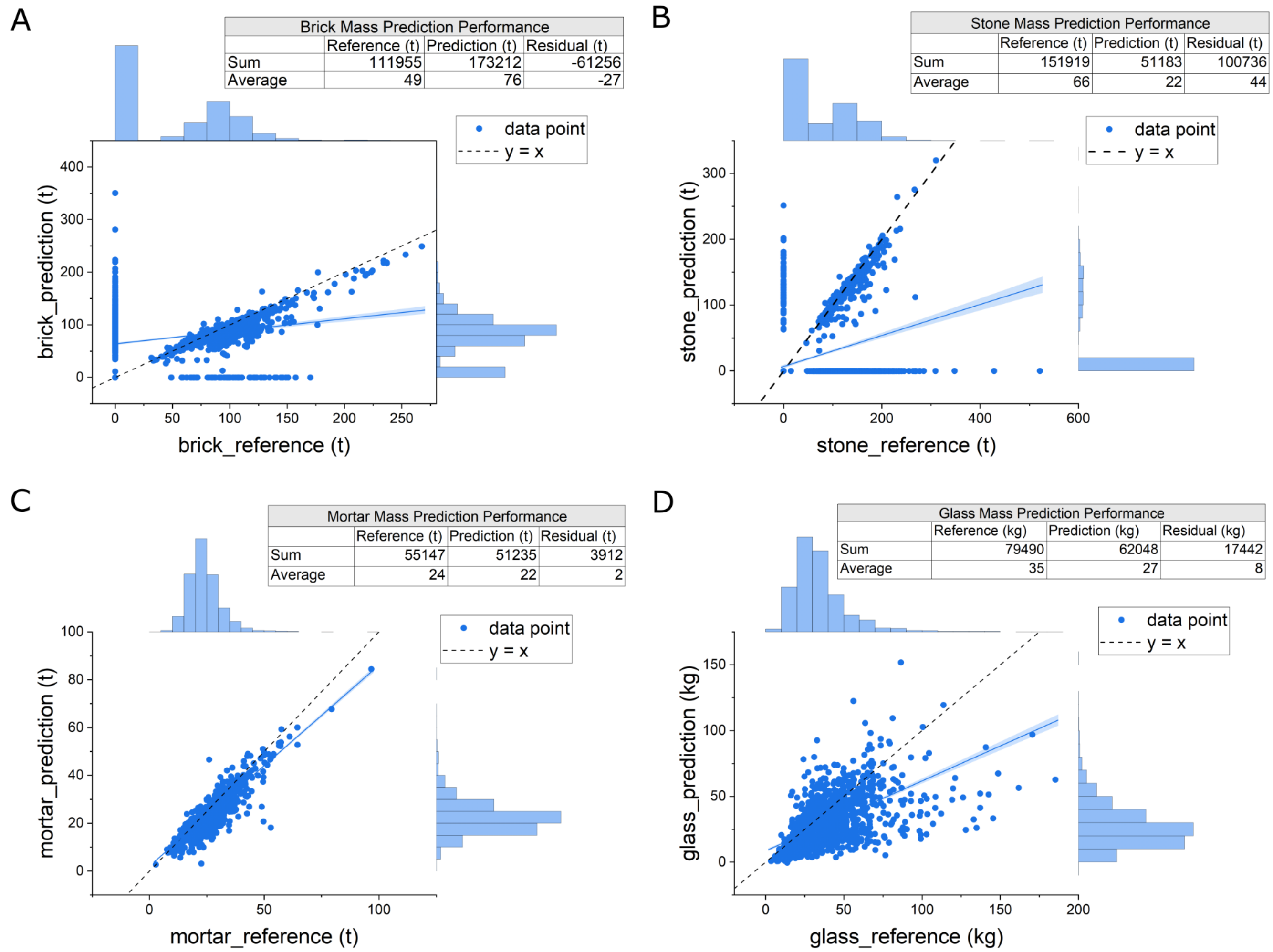


Figure 3: The mass prediction performance results for each material type. The reference values indicate the masses calculated by using EPC data and manually measured variables. The prediction values are produced by using the proposed computer vision-based mass estimation approach. For each material type, their linear regression fitting with a 95% confidence interval is calculated and another line with gradient being one is also presented. The total and average masses of each material and residuals of reference and prediction values are recorded in the top-right table of each figure.

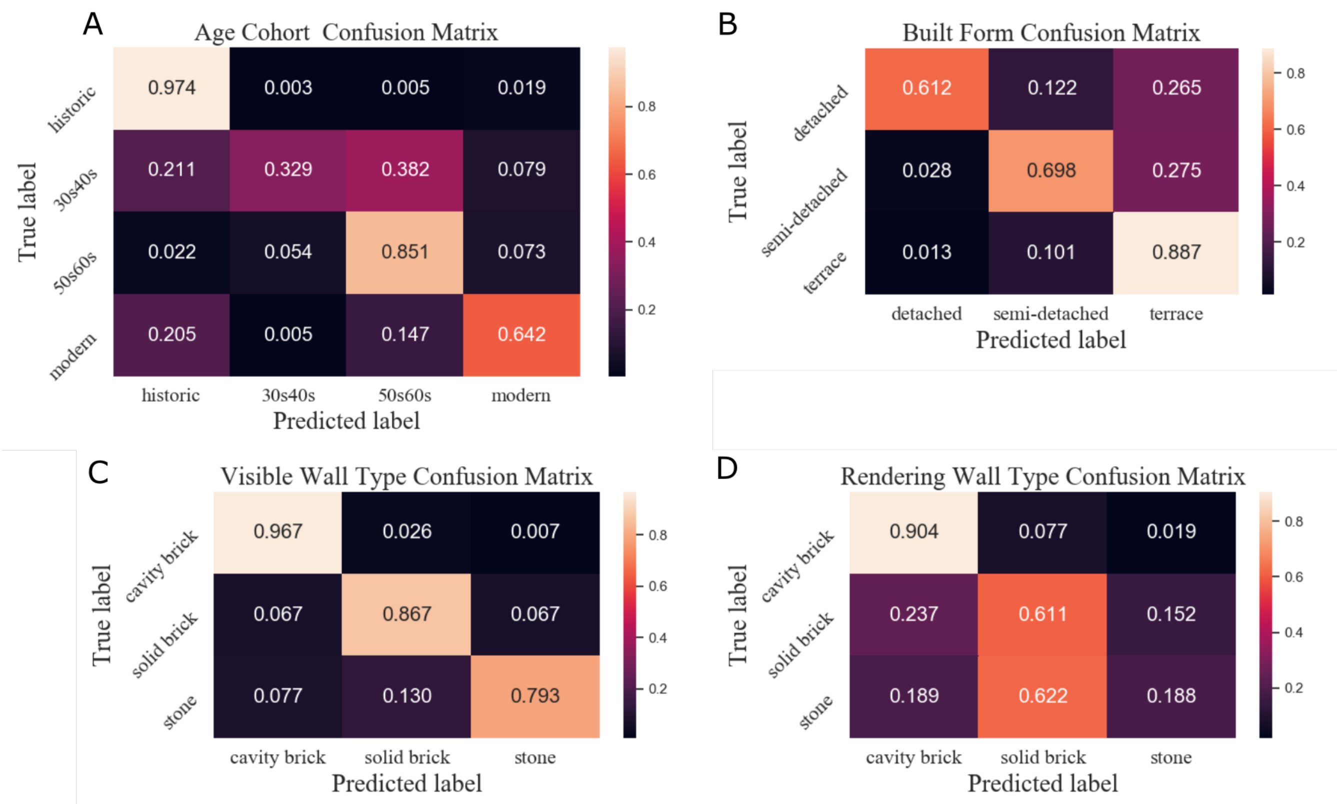


Figure 4: The results of the deployment of the age cohort (as seen in Subfigure A), built form (Subfigure B), and wall type predictions are presented. Due to the inability to directly predict the wall types of rendering walls, the prediction performance is assessed separately. Subfigure C pertains to buildings with visible wall materials, while Subfigure D relates to those with rendering walls.

Age cohort prediction The model achieved a 90% accuracy on the deployment dataset. As illustrated in Figure 4-A, the model exhibited superior performance for historic (97.4%) and 50s60s (85.1%) buildings but was less effective in identifying modern (64.2%) and 30s40s (32.9%) structures. Although the overall accuracy on the deployment data surpassed the validation results (90% versus 86%), the model underperformed in classifying modern (64.2% compared to 76.5%) and 30s40s (32.9% versus 49.1%) buildings. Notably, the deployment data indicated confusion between modern and historic buildings, a trend absent in the validation dataset. The 30s40s category consistently displayed significant misclassifications towards neighbouring age groups in both datasets.

Built form prediction On the deployment dataset, the built form prediction attained an overall accuracy of 85% similar to the validation set (88%). As depicted in Figure 4-B, the model demonstrated reduced efficacy in classifying semi-detached (69.8%) and detached (61.2%) buildings. This trend was also observed in the validation dataset, though the performance was superior—77.4% for semi-detached and 70.6% for detached buildings.

Wall type prediction Figure 4-C and D present the performances of wall type predictions. Distinct evaluations are conducted for visible walls and rendered walls, given the latter's predictions hinge on assumptions and age cohort determinations. For visible walls, the model excelled with an overall accuracy of 90%, achieving 96.7% for cavity walls, 86.7% for solid brick walls, and 79.3% for stone walls. Conversely, rendered wall predictions in Figure 4-D indicate that solid walls are often misclassified as cavity brick or stone. Notably, stone predictions frequently coincide with solid bricks due to the assignment of rendered walls in historic buildings to solid brick types, leading to the potential overlooking of rendered stone walls.

Segmentation and Inner Wall Regression Models Figure 5 demonstrates the quantitative performance of the proposed FacMagNet-s model. The wall and window predictions

Table 1: The inner wall regression results. Ten different random states have been tested on the built dataset. Four evaluation metrics are selected to evaluate the model performance.

Seed	**Avg.**	**1**	**6**	**22**	**31**	**38**	**48**	**72**	**93**	**98**
R^2	**0.81**	0.82	0.81	0.80	0.85	0.83	0.81	0.81	0.80	0.79
MAE	**2.67**	2.48	2.68	2.97	2.69	2.49	2.51	2.89	2.73	2.75
MSE	**12.00**	10.09	13.10	14.07	12.19	10.91	10.55	13.58	12.86	12.64
RMSE	**3.46**	3.18	3.62	3.75	3.49	3.30	3.25	3.68	3.59	3.56

have achieved 91.4% and 91.1% in pixel accuracy, respectively which lays a robust foundation for window distribution and glass mass estimation.

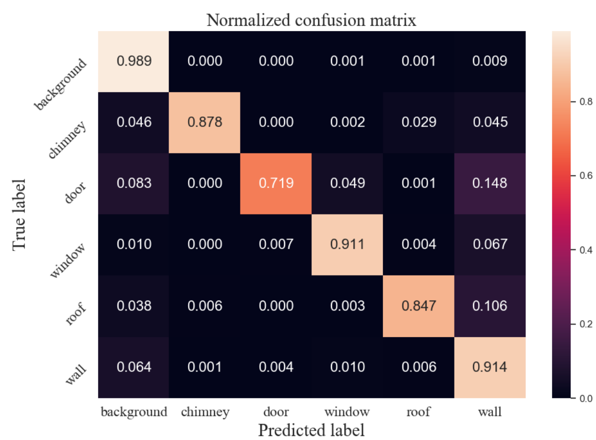


Figure 5: The confusion matrix of the proposed facade segmentation model.

Meanwhile, Table 1 presents an evaluation of the Multi-Layer Perceptron (MLP) algorithm's performance for the task of inner wall regression. To ascertain model robustness, ten distinct random state seeds are assessed. Overall, the MLP model attains an R^2 score of 0.81, signifying a pronounced correlation between the chosen independent variables and the dependent variable, namely, the inner wall length. With a mean inner wall length computed at 30.75m, the RMSE suggests an average prediction error margin of 11%.

Facade Information-based Building Material Stock Estimation

Factors influencing mass prediction accuracy fall into two categories: dimensional and classification. Dimensional inaccuracies encompass errors in predicting building footprint and window size, while classification errors arise from the performance of trained models. Notably, errors in predicting wall types have a significant bearing on estimating brick and stone masses, as illustrated in Figure 3.

Figure 4-C and D reveal that visible walls exhibit markedly superior accuracy for solid brick and stone than rendered walls. This underscores that the presence of rendered walls primarily constrains precise estimations. When samples with incorrect wall type predictions are excluded, brick and stone mass prediction errors stand at 8% and 7% respectively—substantially lower than the overarching error rates of 55% and 67%. Such findings, along with the mortar mass prediction results, suggest that dimensional errors have minimal impact. This supports the efficacy of employing the Douglas-Peuker algorithm and bounding boxes for estimating floorplan dimensions.

Glass mass predictions hinge on the accuracy of window dimension predictions, which in turn are greatly influenced by facade segmentation quality. The number of wall pixels dictates image resolution, and the precision of window recognition further affects these estimations. Additionally, inconsistencies in age cohort predictions sway glass mass predictions given their role in specifying glazing types.

In summation, the data affirms the viability of utilising facade images and computer vision methods to gauge building material stock, particularly for visible wall types. The methodology, however, demonstrates limitations in predicting the mass of rendered wall buildings, especially for stone.

The Reference Level Uncertainties and Approach Limitations

In this research, the reference level is derived from EPC records coupled with manually measured dimensions. However, given that EPC records are based on onsite evaluations, they inherently possess inaccuracies[64]. This makes the difference between the calculated reference level and the ground truth uncertain. A study by Hardy et al. (2019)[64] assessed the precision of EPC records and found that 27% of them contained discrepancies, with approximately 11% of buildings exhibiting wall types inconsistent with their records.

Our team manually labelled the housing attribute dataset based on their visual wall types. Findings reveal that the accuracy stands at 90.8% for cavity brick, 86.7% for solid brick, and 91.3% for stone walls. Consequently, the exact error rate for rendered walls is indeterminate. Moreover, from our observations, a subset of rendered wall buildings appear more akin to brick than stone walls, intensifying the challenge of accurately ascertaining the true wall types for rendered walls.

The results for the inner wall length regression demonstrate inherent uncertainties. The benchmark for inner wall length derives from EPC records complemented by manual measurements. Nonetheless, the model imparts an error margin of $\pm 11\%$, thus making the benchmark somewhat indeterminate. Additionally, while the predicted inner walls are presumed to be constructed from structural materials, i.e. brick or stone, the prevalence of plasterboard as partition walls in contemporary homes cannot be overlooked. Though efforts were made to exclude evident plasterboard partition walls during dataset annotation, the exact nature of the inner walls remains ambiguous without a comprehensive onsite survey.

Additionally, our assumptions consider only brick and stone as construction materials. Yet, some houses may be constructed using concrete blocks. Given that EPC records lack detailed descriptions concerning concrete blocks, and visually determining a building's construction as concrete-based is inherently challenging, the incorporation of concrete remains ambiguous.

The Path Towards an Efficient Component-Level Building Material Stock Future

A fundamental limitation of contemporary methods for estimating building material stock is resolution. Typical strategies, such as employing optical remote sensing or nighttime light images, fail to achieve component-level material estimation. Conversely, traditional bottom-up accounting necessitates labour-intensive onsite surveys.

Drive-by facade images offer a cost-effective and efficient avenue for obtaining building attributes. Our research underscores that, with the aid of computer vision techniques, building attributes crucial for stock estimation can be reliably discerned, except in the case of buildings with rendering walls. While this approach mandates the use of building footprint data, the increasing ubiquity of built environment research renders this data readily accessible.[65,66] As such, the method developed and presented in this paper signifies a step towards a streamlined component-level mass estimation.

Moving forward, our primary objectives are to address the current method's limitations: particularly concrete estimation, rendered wall construction material prediction, and the exclusion of plasterboard. Given the constraints of computer vision techniques, only visible attributes can be captured. Nevertheless, buildings, as rigorously regulated constructs, possess attributes that might be deduced using age cohort data. For instance, the popularity of concrete buildings and plasterboard surged post-World War II, attributed to their cost-effectiveness and ease of installation. Encoding building regulations into our approach promises a more holistic and precise trajectory for component-level building mass estimation.

Overall, building stock, a critical component of the built environment, serves as a repository of readily available and recoverable materials, effectively acting as an "above-ground mine". The foundation of a circular economy lies in the perpetuation of a materials loop, which strives to diminish and ultimately negate the need for extracting virgin resources. Precise knowledge of building materials at the component level allows for an accurate assessment of secondary resources and the forecasting of material demand. This granularity in accounting for building stock is indispensable, not only for tapping into the vast potential of material reuse but also for propelling the full scope of product recovery necessary for a thriving circular economy.

Data and Code Availability

Due to privacy constraints, image data from the vehicle-mounted capture platform will remain inaccessible to the public. However, the Google Street View data employed for model training is retrievable through the Street View API. Upon this paper's publication, the query locations—encompassing both location details and labels from their EPC records—the inner wall regression dataset, the building attributes dataset, and developed software packages will be made available on the designated GitHub repository: `https://github.com/MerlinDai/MARVEL_StockQuantification`.

Author Contributions

Menglin Dai: Conceptualisation, Methodology, Software, Validation, Investigation, Data Curation, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualisation **Jakub Jurczyk:** Methodology, Validation, Data Curation **Hadi Arbabi:** Writing - Review & Editing **Ruichang Mao:** Writing - Review & Editing **Wil Ward:** Data Curation **Martin Mayfield:** Resources, Funding acquisition **Gang Liu:** Supervision, Writing - Review & Editing **Danielle Densley Tingley:** Conceptualisation, Resources, Project administration, Funding acquisition, Supervision, Writing - Review & Editing

Acknowledgement

This work was supported by EPSRC Active Building Centre Research Programme, United Kingdom [EP/V012053/1] and EPSRC Multi-Scale, Circular Economic Potential of Non-Residential Building Scale [EP/S029273/1].

Supporting Information Available

The following files are available free of charge.

- supplementary_file.pdf: This file contains additional details of built datasets and developed machine learning models.

References

(1) International Energy Agency *Tracking Clean Energy Progress 2023*; 2023; License: CC BY 4.0.

(2) United Nations Sustainable Development Goals. 2015; `https://www.un.org/sustainabledevelopment/sustainable-development-goals/` (accessed 2023-06-06).

(3) Elhacham, E.; Ben-Uri, L.; Grozovski, J.; Bar-On, Y. M.; Milo, R. Global human-made mass exceeds all living biomass. *Nature* **2020**, *588*, 442–444.

(4) Krausmann, F.; Wiedenhofer, D.; Lauk, C.; Haas, W.; Tanikawa, H.; Fishman, T.; Miatto, A.; Schandl, H.; Haberl, H. Global socioeconomic material stocks rise 23-fold over the 20th century and require half of annual resource use. *Proceedings of the National Academy of Sciences* **2017**, *114*, 1880–1885.

(5) Densley Tingley, D. Embed circular economy thinking into building retrofit. *Communications Engineering* **2022**, *1*, 1–4.

(6) Koutamanis, A.; van Reijn, B.; van Bueren, E. Urban mining and buildings: A review of possibilities and limitations. *Resources, Conservation and Recycling* **2018**, *138*, 32–39.

(7) Wiedenhofer, D.; Steinberger, J. K.; Eisenmenger, N.; Haas, W. Maintenance and expansion: modeling material stocks and flows for residential buildings and transportation networks in the EU25. *Journal of Industrial Ecology* **2015**, *19*, 538–551.

(8) Lanau, M.; Liu, G.; Kral, U.; Wiedenhofer, D.; Keijzer, E.; Yu, C.; Ehlert, C. Taking stock of built environment stock studies: Progress and prospects. *Environmental science & technology* **2019**, *53*, 8499–8515.

(9) Li, Q.; Gummidi, S. R. B.; Lanau, M.; Yu, B.; Liu, G. Spatiotemporally Explicit Mapping of Built Environment Stocks Reveals Two Centuries of Urban Development

in a Fairytale City, Odense, Denmark. *Environmental Science & Technology* **2022**, *56*, 16369–16381.

(10) Augiseau, V.; Barles, S. Studying construction materials flows and stock: A review. *Resources, Conservation and Recycling* **2017**, *123*, 153–164.

(11) Mohammadiziazi, R.; Bilec, M. Building material stock analysis is critical for effective circular economy strategies: a comprehensive review. *Environmental Research: Infrastructure and Sustainability* **2022**, *2*, 032001.

(12) Shi, F.; Huang, T.; Tanikawa, H.; Han, J.; Hashimoto, S.; Moriguchi, Y. Toward a low carbon–dematerialization society: Measuring the materials demand and CO2 emissions of building and transport infrastructure construction in China. *Journal of Industrial Ecology* **2012**, *16*, 493–505.

(13) Huang, T.; Shi, F.; Tanikawa, H.; Fei, J.; Han, J. Materials demand and environmental impact of buildings construction and demolition in China based on dynamic material flow analysis. *Resources, Conservation and Recycling* **2013**, *72*, 91–101.

(14) Fishman, T.; Schandl, H.; Tanikawa, H.; Walker, P.; Krausmann, F. Accounting for the material stock of nations. *Journal of Industrial Ecology* **2014**, *18*, 407–420.

(15) Gao, X.; Nakatani, J.; Zhang, Q.; Huang, B.; Wang, T.; Moriguchi, Y. Dynamic material flow and stock analysis of residential buildings by integrating rural–urban land transition: A case of Shanghai. *Journal of Cleaner Production* **2020**, *253*, 119941.

(16) Nasir, U.; Chang, R.; Omrany, H. Calculation methods for construction material stocks: A systematic review. *Applied Sciences* **2021**, *11*, 6612.

(17) Tanikawa, H.; Hashimoto, S. Urban stock over time: spatial material stock analysis using 4d-GIS. *Building research & information* **2009**, *37*, 483–502.

(18) Ortlepp, R.; Gruhler, K.; Schiller, G. Material stocks in Germany's non-domestic buildings: a new quantification method. *Building Research & Information* **2016**, *44*, 840–862.

(19) Heeren, N.; Fishman, T. A database seed for a community-driven material intensity research platform. *Scientific data* **2019**, *6*, 23.

(20) Mao, R.; Bao, Y.; Huang, Z.; Liu, Q.; Liu, G. High-resolution mapping of the urban built environment stocks in Beijing. *Environmental science & technology* **2020**, *54*, 5345–5355.

(21) Lanau, M.; Liu, G. Developing an urban resource cadaster for circular economy: A case of Odense, Denmark. *Environmental science & technology* **2020**, *54*, 4675–4685.

(22) Sprecher, B.; Verhagen, T. J.; Sauer, M. L.; Baars, M.; Heintz, J.; Fishman, T. Material intensity database for the Dutch building stock: Towards Big Data in material stock analysis. *Journal of Industrial Ecology* **2022**, *26*, 272–280.

(23) Haberl, H.; Wiedenhofer, D.; Schug, F.; Frantz, D.; Virağ, D.; Plutzar, C.; Gruhler, K.; Lederer, J.; Schiller, G.; Fishman, T.; others High-resolution maps of material stocks in buildings and infrastructures in Austria and Germany. *Environmental science & technology* **2021**, *55*, 3368–3379.

(24) Bao, Y.; Huang, Z.; Wang, H.; Yin, G.; Zhou, X.; Gao, Y. High-resolution quantification of building stock using multi-source remote sensing imagery and deep learning. *Journal of Industrial Ecology* **2023**, *27*, 350–361.

(25) Liu, Z.; Saito, R.; Guo, J.; Hirai, C.; Haga, C.; Matsui, T.; Shirakawa, H.; Tanikawa, H. Does Deep Learning Enhance the Estimation for Spatially Explicit Built Environment Stocks through Nighttime Light Data Set? Evidence from Japanese Metropolitans. *Environmental Science & Technology* **2023**, *57*, 3971–3979.

(26) Arbabi, H.; Lanau, M.; Li, X.; Meyers, G.; Dai, M.; Mayfield, M.; Densley Tingley, D. A scalable data collection, characterization, and accounting framework for urban material stocks. *Journal of Industrial Ecology* **2022**, *26*, 58–71.

(27) Francart, N.; Gummidi, S.; Hoxha, E.; Birgisdottir, H. A Danish model of building macro-components to promote circularity. Journal of Physics: Conference Series. 2023; p 192001.

(28) Shen, L.; Yang, Q.; Yan, H. Spatial Characterization Analysis of Residential Material Stock and its Driving Factors: A Case Study of Xi'an. *Buildings* **2023**, *13*, 581.

(29) Anguelov, D.; Dulong, C.; Filip, D.; Frueh, C.; Lafon, S.; Lyon, R.; Ogale, A.; Vincent, L.; Weaver, J. Google Street View: Capturing the World at Street Level. *Computer* **2010**, *43*, 32–38.

(30) Biljecki, F.; Ito, K. Street view imagery in urban analytics and GIS: A review. *Landscape and Urban Planning* **2021**, *215*, 104217.

(31) Kang, J.; Körner, M.; Wang, Y.; Taubenböck, H.; Zhu, X. X. Building instance classification using street view images. *ISPRS journal of photogrammetry and remote sensing* **2018**, *145*, 44–59.

(32) Zeppelzauer, M.; Despotovic, M.; Sakeena, M.; Koch, D.; Döller, M. Automatic prediction of building age from photographs. Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval. 2018; pp 126–134.

(33) Szcześniak, J. T.; Ang, Y. Q.; Letellier-Duchesne, S.; Reinhart, C. F. A method for using street view imagery to auto-extract window-to-wall ratios and its relevance for urban-level daylighting and energy simulations. *Building and Environment* **2022**, *207*, 108108.

(34) Ministry of Housing, Communities & Local Government English Housing Survey, Households Report, 2017-18. 2018; `https://www.gov.uk/government/statistics/english-housing-survey-2017-to-2018-households` (accessed 2023-06-06).

(35) Dai, M.; Ward, W. O.; Arbabi, H.; Densley Tingley, D.; Mayfield, M. Scalable Residential Building Geometry Characterisation Using Vehicle-Mounted Camera System. *Energies* **2022**, *15*, 6090.

(36) Loga, T.; Stein, B.; Diefenbach, N. TABULA building typologies in 20 European countries—Making energy-related features of residential building stocks comparable. *Energy and Buildings* **2016**, *132*, 4–12.

(37) Meyers, G.; Zhu, C.; Mayfield, M.; Tingley, D. D.; Willmott, J.; Coca, D. Designing a Vehicle Mounted High Resolution Multi-Spectral 3D Scanner: Concept Design. Proceedings of the 2nd Workshop on Data Acquisition to Analysis. 2019; pp 16–21.

(38) Ward, W.; Dai, M.; Arbabi, H.; Sun, Y.; Tingley, D.; Mayfield, M. Measuring the Cityscape: A Pipeline from Street-Level Capture to Urban Quantification. *IOP Conference Series: Earth and Environmental Science* **2022**, *1078*, 012036.

(39) Ward, W.; Li, X.; Sun, Y.; Dai, M.; Arbabi, H.; Tingley, D. D.; Mayfield, M. Estimating energy consumption of residential buildings at scale with drive-by image capture. *Building and Environment* **2023**, *234*, 110188.

(40) Ordnance Survey MasterMap topography layer. 2023; `https://www.ordnancesurvey.co.uk/business-government/products/mastermap-topography` (accessed 2023-05-15).

(41) A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence* **2000**, *22*, 1330–1334.

(42) Cucchiara, R.; Grana, C.; Prati, A.; Vezzani, R. A Hough transform-based method for radial lens distortion correction. 12th International Conference on Image Analysis and Processing, 2003. Proceedings. 2003; pp 182–187.

(43) BPIE Energy Performance Certificates across the EU. 2015.

(44) Piddington, J.; Nicol, S.; Garrett, H.; Custard, M. The Housing Stock of the United Kingdom. *BRE Trust: Watford, UK* **2020**, `https://files.bregroup.com/bretrust/The-Housing-Stock-of-the-United-Kingdom_Report_BRE-Trust.pdf` (accessed 2023-06-06).

(45) Dai, M.; Ward, W. O.; Meyers, G.; Tingley, D. D.; Mayfield, M. Residential building facade segmentation in the urban environment. *Building and Environment* **2021**, *199*, 107921.

(46) Sun, M.; Zhang, F.; Duarte, F.; Ratti, C. Understanding architecture age and style through deep learning. *Cities* **2022**, *128*, 103787.

(47) Benz, A.; Voelker, C.; Daubert, S.; Rodehorst, V. Towards an automated image-based estimation of building age as input for Building Energy Modeling (BEM). *Energy and Buildings* **2023**, *292*, 113166.

(48) Wei, X.-S.; Song, Y.-Z.; Mac Aodha, O.; Wu, J.; Peng, Y.; Tang, J.; Yang, J.; Belongie, S. Fine-grained image analysis with deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence* **2021**, *44*, 8927–8948.

(49) Lin, T.-Y.; RoyChowdhury, A.; Maji, S. Bilinear CNN models for fine-grained visual recognition. Proceedings of the IEEE international conference on computer vision. 2015; pp 1449–1457.

(50) Meng, C.; Song, Y.; Ji, J.; Jia, Z.; Zhou, Z.; Gao, P.; Liu, S. Automatic classification

of rural building characteristics using deep learning methods on oblique photography. Building Simulation. 2022; pp 1–14.

(51) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016; pp 770–778.

(52) Chollet, F. Xception: Deep learning with depthwise separable convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017; pp 1251–1258.

(53) Martinovic, A.; Knopp, J.; Riemenschneider, H.; Van Gool, L. 3d all the way: Semantic segmentation of urban scenes from start to end in 3d. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015; pp 4456–4465.

(54) Gadde, R.; Jampani, V.; Marlet, R.; Gehler, P. V. Efficient 2D and 3D Facade Segmentation Using Auto-Context. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2017**, *40*, 1273–1280.

(55) Kong, G.; Fan, H. Enhanced Facade Parsing for Street-Level Images Using Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing* **2020**, *59*, 10519–10531.

(56) Zhang, G.; Pan, Y.; Zhang, L. Deep learning for detecting building façade elements from images considering prior knowledge. *Automation in Construction* **2022**, *133*, 104016.

(57) Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. Proceedings of the European Conference on Computer Vision (ECCV). 2018; pp 833–851.

(58) Watson, P. An introduction to UK energy performance certificates (EPCs). *Journal of Building Appraisal* **2010**, *5*, 241–250.

(59) Fewins, C. The pros and cons of different construction systems. *Home Building & Renovating* **2004**, 1–11.

(60) Yuan, L.; Lu, W.; Xue, F.; Li, M. Building feature-based machine learning regression to quantify urban material stocks: A Hong Kong study. *Journal of Industrial Ecology* **2023**, *27*, 336–349.

(61) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; others Scikit-learn: Machine learning in Python. *the Journal of Machine Learning research* **2011**, *12*, 2825–2830.

(62) Summers, C.; Hulme, J. *BRE Client Report for EPISCOPE Pilot Project: England, Intelligent Energy Europe Programme for the European Union*; Client Report 297809, 2016; `https://episcope.eu/fileadmin/episcope/public/docs/pilot_actions/GB_EPISCOPE_NationalCaseStudy_BRE.pdf` (accessed 2023-06-06).

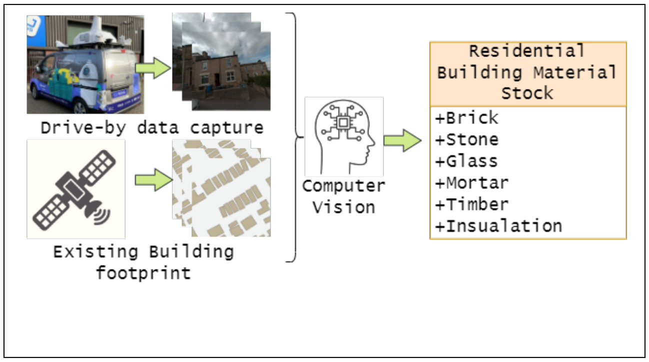
(63) Institution, B. S. *BS EN 771-1 Specification for Masonry Units. Part 1: Clay Masonry Units*; 2020.

(64) Hardy, A.; Glew, D. An analysis of errors in the Energy Performance certificate database. *Energy policy* **2019**, *129*, 1168–1178.

(65) Sirko, W.; Kashubin, S.; Ritter, M.; Annkah, A.; Bouchareb, Y. S. E.; Dauphin, Y.; Keysers, D.; Neumann, M.; Cisse, M.; Quinn, J. Continental-scale building detection from high resolution satellite imagery. **2021**, *arXiv preprint arXiv:2107.12283*, `https://arxiv.org/abs/2107.12283` (accessed 2023-06-06).

(66) Milojevic-Dupont, N.; Wagner, F.; Nachtigall, F.; Hu, J.; Brüser, G. B.; Zumwald, M.; Biljecki, F.; Heeren, N.; Kaack, L. H.; Pichler, P.-P.; others EUBUCCO v0. 1: European building stock characteristics in a common and open database for 200+ million individual buildings. *Scientific Data* **2023**, *10*, 147.

TOC Graphic


Drive-by data capture
Existing Building footprint
Computer Vision
Residential Building Material Stock
+Brick
+Stone
+Glass
+Mortar
+Timber
+Insualation