



## Article

# Open Data-Driven 3D Building Models for Micro-Population Mapping in a Data-Limited Setting

Kittisak Maneepong <sup>1</sup>, Ryota Yamanotera <sup>1</sup>, Yuki Akiyama <sup>2,\*</sup>, Hiroyuki Miyazaki <sup>3</sup>, Satoshi Miyazawa <sup>4</sup> and Chiaki Mizutani Akiyama <sup>5</sup>

<sup>1</sup> Graduate School of Integrative Science and Engineering, Tokyo City University, Tokyo 158-0087, Japan; g2291605@tcu.ac.jp (K.M.); g2381644@tcu.ac.jp (R.Y.)

<sup>2</sup> Faculty of Architecture and Urban Design, Tokyo City University, Tokyo 158-0087, Japan

<sup>3</sup> GLODAL, Inc., Yokohama 231-0062, Japan; miyazaki@glodal-inc.com

<sup>4</sup> LocationMind Inc., Tokyo 101-0048, Japan; miyazawa@locationmind.com

<sup>5</sup> Reitaku University, Chiba 277-0065, Japan; chakiyam@reitaku-u.ac.jp

\* Correspondence: akiyamay@tcu.ac.jp; Tel.: +81-03-5707-0104 (ext. 3268)

**Abstract:** Urban planning and management increasingly depend on accurate building and population data. However, many regions lack sufficient resources to acquire and maintain these data, creating challenges in data availability. Our methodology integrates multiple data sources, including aerial imagery, Points of Interest (POIs), and digital elevation models, employing Light Gradient Boosting Machine (LightGBM) and Gradient Boosting Decision Tree (GBDT) to classify building uses and morphological filtration to estimate heights. This research contributes to bridging the gap between data needs and availability in resource-constrained urban environments, offering a scalable solution for global application in urban planning and population mapping.

**Keywords:** urban population mapping; building height estimation; building use classification; machine learning



**Citation:** Maneepong, K.; Yamanotera, R.; Akiyama, Y.; Miyazaki, H.; Miyazawa, S.; Akiyama, C.M. Open Data-Driven 3D Building Models for Micro-Population Mapping in a Data-Limited Setting. *Remote Sens.* **2024**, *16*, 3922. <https://doi.org/10.3390/rs16213922>

Academic Editors: Iván Puente-Luna and Xavier Núñez-Nieto

Received: 7 September 2024

Revised: 12 October 2024

Accepted: 21 October 2024

Published: 22 October 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Population maps serve as a crucial factor in understanding human settlement patterns and interactions with the environment. Traditional population maps, such as choropleth maps and dasymetric maps, over administrative boundaries [1–3], such as state or district levels, have provided valuable insights for broad-scale analysis. However, more refined population mapping, from gridded data [4,5] to micro-dasymetric mapping is becoming increasingly feasible using building information [6,7].

The realization of high-resolution population estimates at the building level is expected to make significant contributions to a wide range of fields that have traditionally relied on population data, such as waste generation estimation [8], urban energy consumption calculation [9], and disaster response planning [10].

Efforts have been made to generate population data at the building unit using high-resolution satellite images, focusing on building density and area [11]. The process requires high-resolution aerial images with a resolution of 5 m or better to clearly capture building features. However, such high-quality images are not yet widely available or openly accessible for global utilization. A number of studies focusing on open data have utilized Points of Interest (POIs) [12,13], land use data to classify building use [12], building height [14,15], and building footprints [12,15]. These studies demonstrate that population mapping is effective in areas with well-defined land use. Nevertheless, challenges persist in regions where information on land use is absent or incomplete [12]. Furthermore, detailed building features, such as stairs and entrances, are also being used in the mapping process [12,14].

Three-dimensional (3D) city models and building information are crucial for understanding urban settlement patterns [16], enabling multifaceted research on the built environment dynamics, for instance, disaster mitigation and risk assessment [17,18], environmental

analysis [19], demographic studies [15,20], and public health management [21,22]. The utilization of 3D building models, in particular, offers substantial potential for urban planning and management [23,24]. Urban built-up data, which include building footprints, is used in population mapping [12,14,25–27] as it indicates human settlement. These data come in various forms and resolutions. The Global Human Settlement Layer (GHSL) provides coarser-resolution datasets for urban analysis [28]. In contrast, OpenStreetMap offers more detailed, community-driven data on urban structures [29]. Tech companies like Microsoft [30] and Google [31] have recently applied machine learning techniques to improve mapping capabilities. These efforts have led to initiatives, such as the Overture Map project [32], which aims to create comprehensive global mapping data. However, these initiatives typically require substantial computational resources and high-resolution aerial imagery for both training and inference.

Complementing spatial usage data, such as land use or building use, are commonly employed as covariates in population mapping [12,14,25–27,33,34]. Recent advancements in machine learning and satellite imagery have led to the development of datasets related to spatial interactions, specifically, land use classification. Cheng et al. [35] developed a method to classify cities into 16 types based on satellite imagery, building characteristics, and access networks, with the aim of better understanding urban environments in developing countries. Similarly, Wang et al. [36] successfully identified 14 types of urban land use, including detailed classifications of urban green spaces, parking lots, and plazas, using high-resolution remote sensing images and the deep learning model. Furthermore, satellite imagery and deep learning models have increasingly been applied to tasks, such as slum mapping, with the accuracy of these methods improving year by year [37]. As demonstrated, advancements in machine learning and related approaches within the urban spatial domain are contributing to the development of various datasets beyond just building information. Understanding spatial usage, integrated with building data, allows for a more comprehensive understanding of spatial patterns, and human interactions.

### *1.1. Challenges in Population Mapping and Building Data Acquisition in Data-Limited Settings*

Creating and maintaining high-quality building data, as described above, poses a significant challenge, particularly in data-limited settings. In such settings, the confluence of these challenges in building data acquisition and micro-scale population mapping presents a significant barrier to advancement in micro-scale urban studies. Some regions lack the necessary data infrastructure, technical expertise, and financial resources to implement traditional methods of 3D city modeling and high-resolution population mapping. The gap between data requirements and data availability creates a significant barrier to the effective execution of urban planning, resource allocation, and disaster preparedness initiatives. This challenge is particularly pronounced in urban areas in developing countries, where the pace of urbanization and population growth necessitates access to accurate and timely data. Consequently, there is a high demand for capturing up-to-date population distributions at the highest possible resolution and with high frequency [38,39].

The micro-population mapping approach presented in this study utilizes building uses (e.g., residential, commercial, industrial) and volume to estimate the population at a building level. These two attributes were selected to maintain the generality of the dataset, ensuring its applicability while capturing enough nuance of the building–human relationship. Building use indicates occupancy patterns, while volume correlates with potential occupant capacity. A significant challenge posed during the recreation of this approach was that some regions lacked the necessary attributes in the building data [40]. This gap further complicates effective urban planning, resource allocation, and disaster preparedness in these areas.

Meanwhile, the most widely used method for understanding the spatial distribution of populations is the use of population statistics. While population statistics are commonly used to understand population distribution in developed countries, there are numerous challenges associated with such data in developing countries. These challenges include

unreliable data collection systems, incomplete spatial and attribute coverage, insufficient technical infrastructure and funding for statistical development, lack of government transparency and commitment to statistical improvement, and the issue of populations not captured in statistics, such as those in the informal sector [41–43]. As a result, it is difficult to establish and maintain accurate population statistics, particularly in remote areas and slums where data quality tends to be lower. These challenges affect policymaking and the provision of social services and hinder progress toward achieving the Sustainable Development Goals (SDGs) [44,45].

### 1.2. Objective of This Study

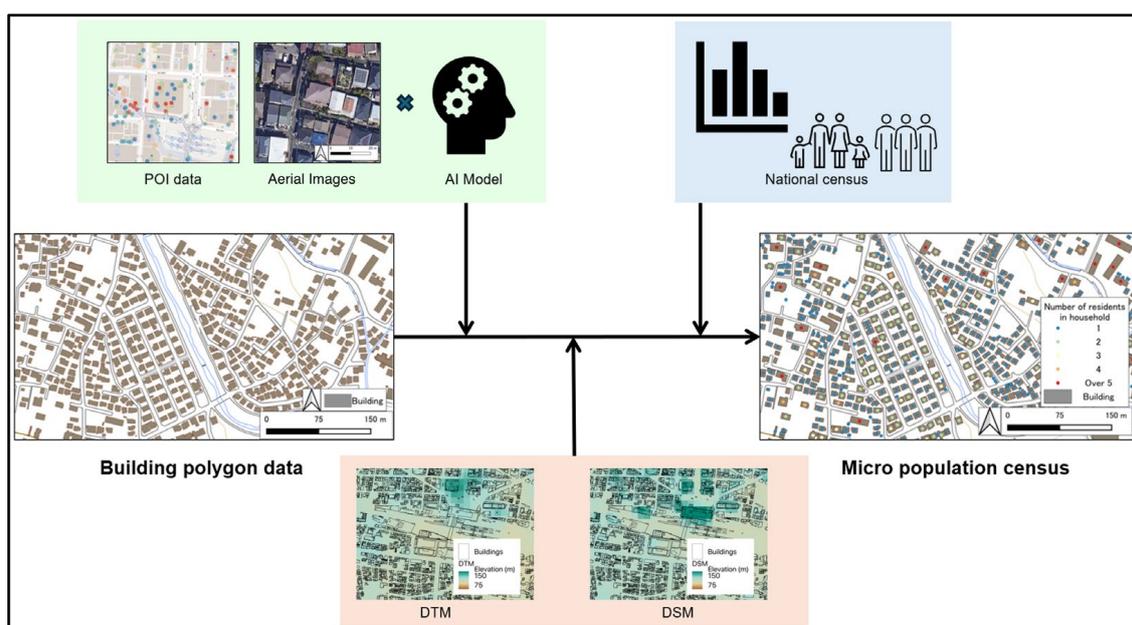
To address these challenges, particularly in developing countries and underdeveloped regions, our research aims to

1. Develop methods for micro-population estimation at the building unit scale, including techniques for estimating building height and usage;
2. Ensure the applicability and scalability of our framework across diverse regions by utilizing generally available data.

Our novel approach combines multiple source data with advanced population mapping techniques to bridge the critical gap between data needs and availability in resource-constrained settings. The model's efficacy is evaluated using Japan as a case study, where reference data are readily available for validation purposes.

## 2. Methods

This study aims to develop an innovative method for estimating individual building attributes (building use and number of floors) and population distribution using machine learning techniques based on aerial images and existing statistical data. Figure 1 illustrates the comprehensive forecasting workflow. The proposed approach is expected to enable accurate, high-frequency estimation of population distribution in regions where comprehensive building data are lacking. This method has the potential to address critical challenges in urban planning, resource allocation, and disaster management in areas with limited infrastructure data.



**Figure 1.** Workflow of the developing method for estimating individual population distribution.

## 2.1. Building Use Classification

We utilized machine learning (ML) techniques to classify building use for individual structures using satellite imagery and existing statistical data. Unlike previous methods that utilized street-level perspective images to classify buildings [46,47], our approach eliminates human intervention for image capturing, facilitating frequent data updates. Our study introduces a novel classification method that enables swift updates with enhanced accuracy [48]. Figure 2 illustrates the comprehensive forecasting workflow of building data building use classification. The methodology involves the fusion of a machine learning model by utilizing Points of Interest (POI) data and aerial imagery to classify building usage [49]. This approach overcomes the limitations of traditional methods, offering a more efficient and scalable solution for urban building use classification.

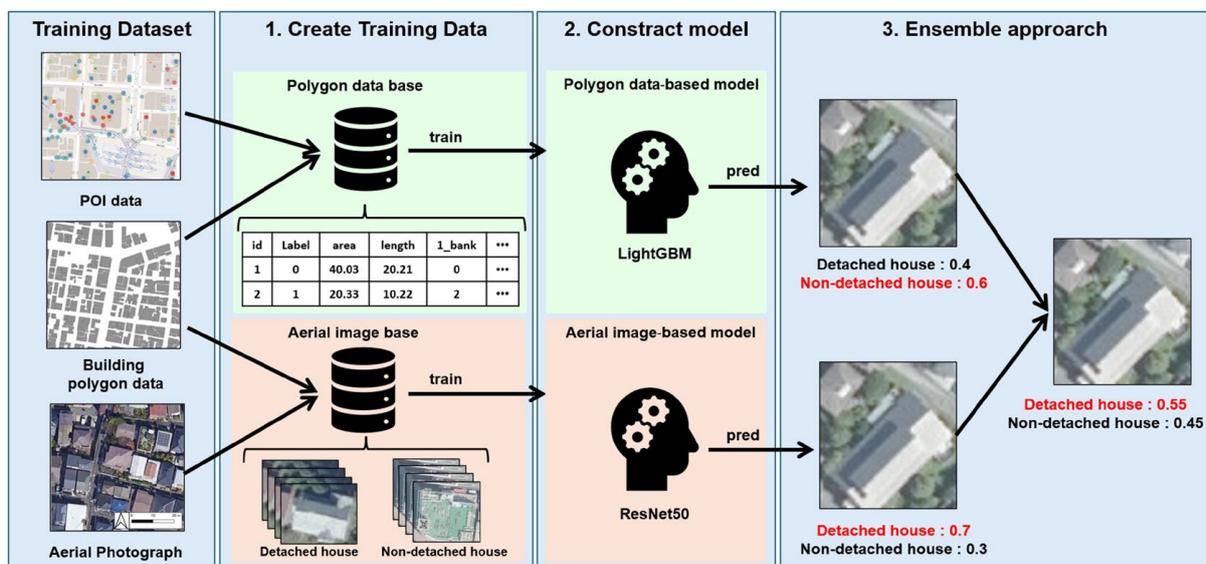


Figure 2. Comprehensive forecasting workflow of building use classification.

We developed and validated our models using Hachioji City, Tokyo prefecture, Japan (Hereafter, “Hachioji City”), as the study area, which was selected for its diverse and unbiased representation of building uses. Hachioji City, located in the western part of Tokyo within the Tama region, is home to approximately 570,000 residents. The city features diverse topography and urban functions, with mountainous and hilly terrain extending across the northern and western areas, offering abundant natural landscapes. In contrast, the southern and eastern parts consist of flatlands, where residential and commercial areas have developed. Due to its diverse geographical and topographical conditions and the variety of building types, Hachioji City was deemed an ideal study area for our study.

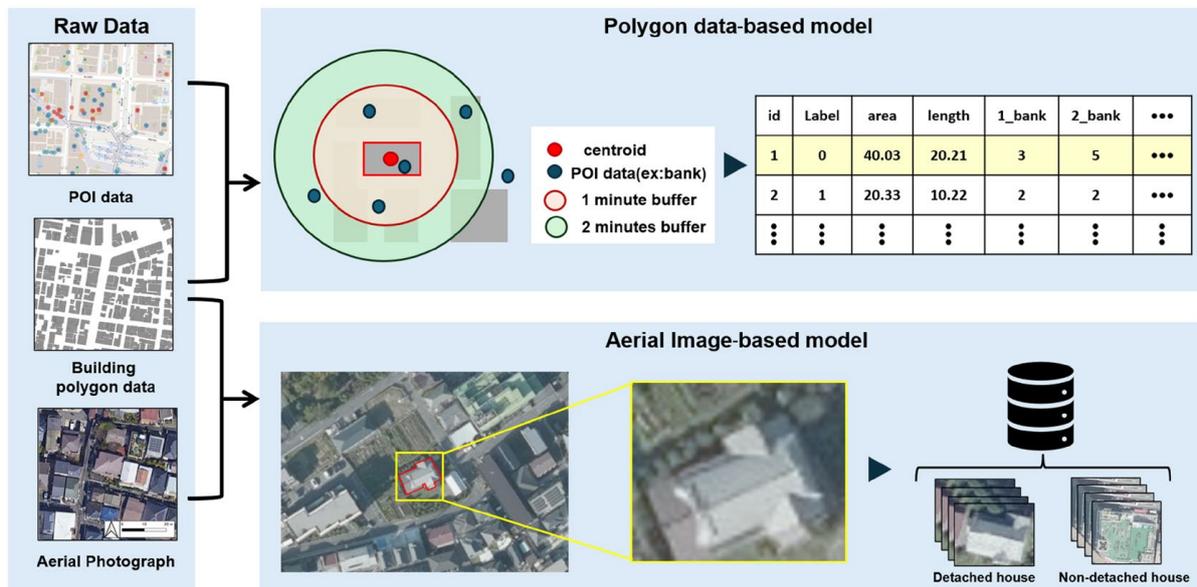
Our classification system employs three primary building use categories: multi-family housings, detached houses, and detached offices. It is derived from the data known as “Residential Maps” in Japan that provide detailed information on the distribution, shape, usage, and number of floors for individual buildings.

While the focus of this study is to provide the framework for urban areas in developing countries, we deliberately chose to study cities in Japan. This decision was made due to Japan’s highly reliable and comprehensive spatial data and statistics that cover the entire country, which provide a favorable environment for implementing this study and verifying the results.

### 2.1.1. Training Data Creation

The training datasets are created for two types of models, polygon data based and aerial image based, to classify building uses. The polygon data-based model was created using building polygon data and POI data, while the aerial image-based model was created

using building polygon data and aerial image. In addition, following the methodology of Akiyama et al. [15], in this study, multi-family housing, detached houses, and detached offices were selected as target variables for each classification model (detached house classification model and multi-family house classification model). The methodology used to create the respective training data is shown in Figure 3.



**Figure 3.** Method used to create the training data.

For the polygon data-based model, features were extracted from building polygon data, from the fundamental geospatial information provided as open data by the Geospatial Information Authority of Japan (GSI), and POI data, obtained from OpenStreetMap. Building features included area, circumference, and shape complexity. POI features were derived by quantifying the number of POIs within walking distance (1–5 min) of the center of each building for that category. For example, in the case of the “bank” POI category, we counted the number of banks within a 1 min walk (approximately 80 m), a 3 min walk (approximately 240 m), and a 5 min walk (approximately 400 m) from each building’s center. This approach allows us to capture the density of banking services in the immediate vicinity of a building, which could be indicative of its use or the character of its neighborhood. The list of variables for the polygon data-based model is presented in Table 1, including these POI-based features for various categories such as banks, restaurants, schools, and other relevant urban amenities. The list of variables for the polygon data-based model is presented in Table 1.

For the aerial image-based model, a bounding box of polygon data for each selected building and extracting the corresponding section from the aerial image was created. These cropped images were labeled with each building’s use to form a training dataset.

To ensure fair comparison and consistency in training conditions, we used the same set of buildings as training data for both models. This approach allowed us to minimize potential biases that could arise from differences in the training data selection, enabling a more accurate comparison of the performance of machine learning models in building use classification.

**Table 1.** List of variables in the machine learning model.

Categories	Variables	Data Source
Building Features	Area, complexity, length, number of vertices	Geospatial Information Authority of Japan
POI (Amenity)	Restaurant, pub, cafe, telephone booth, medical service, convenience store, fast food, theatre, hospital, school, bank, pharmacy, police station, post office, library	OpenStreetMap
POI (Shop)	Car repair shop, fabric shop, funeral directors, supermarket, convenience store, bakery, clothes shop, electronics shop, hairdresser, pharmacy	OpenStreetMap
POI (Tourism)	Hotel, museum, tourism information, tourism attraction spot, tourism viewpoint, tourism apartment, tourism hostel	OpenStreetMap

### 2.1.2. Classification Model Construction

This study employed two advanced tools for building use classification, LightGBM (Light Gradient Boosting Machine) for the polygon data-based model and ResNet50 (Residual Network with 50 layers) for the aerial image-based model, leveraging the previously described training data [50,51].

LightGBM, an enhancement of the Gradient Boosting Decision Tree (GBDT) algorithm, implements a leaf-wise tree growth strategy. This approach significantly accelerates learning speed and efficiently handles large datasets while minimizing memory usage—a crucial advantage when processing extensive building data.

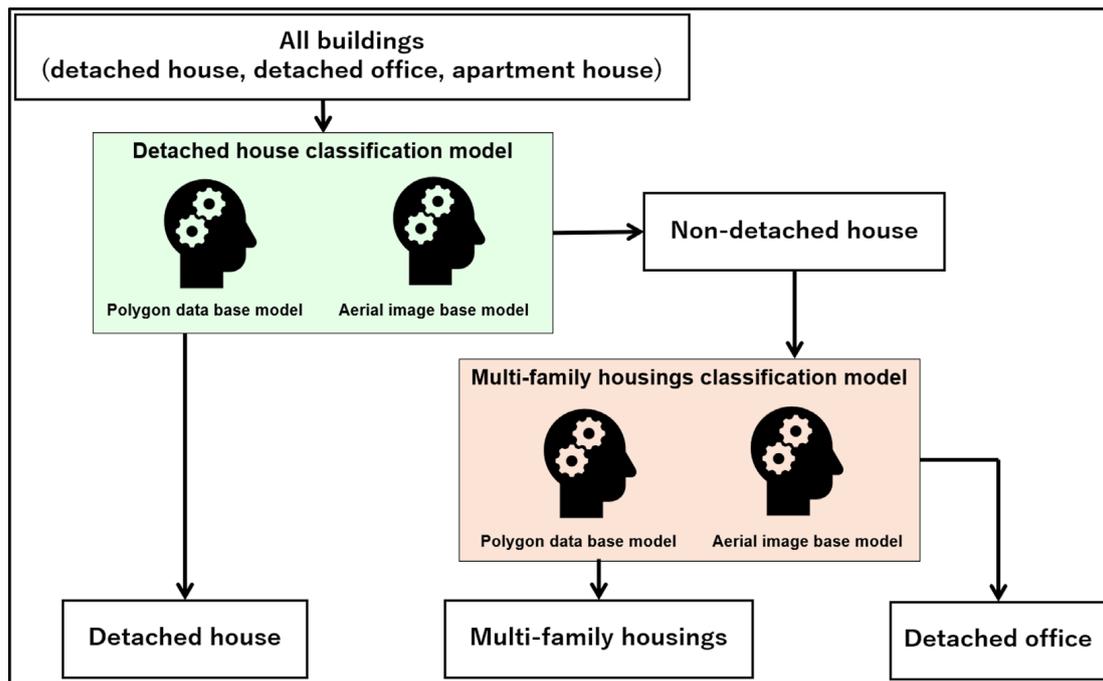
ResNet50 addresses the gradient vanishing problem through skip connections (residual blocks), enabling a very deep network structure. This architecture is particularly suited for classifying building images with complex visual features, as it can extract a wide range of low- and high-level features, effectively capturing building appearance and structural characteristics.

### 2.1.3. Ensemble Approach

This study implemented an innovative ensemble approach, combining machine learning models to predict building use. This approach aims to enhance classification accuracy and address the complexity of urban environments by integrating complementary information from diverse data types.

We implemented a multi-stage forecasting strategy using sequential binary classification models rather than a single multi-class classification approach. This method addresses challenges encountered in multi-class model training, particularly the difficulty in efficiently learning distinctive features across numerous building use categories such as detached houses, multi-family housing, and detached offices. Our multi-stage approach allows each model to focus on specific, distinguishing characteristics of building types at each stage, thereby enhancing overall classification performance. Moreover, given our primary focus on residential building classification, this stepwise approach is better aligned with subsequent population allocation processes. It allows for a more granular and accurate differentiation between residential building types, which is crucial for precise population estimation.

Figure 4 illustrates the comprehensive forecasting workflow. This methodology enables the construction of a hierarchical, highly accurate building use classification model. By addressing data imbalance issues and focusing on our primary objective of residential classification, we anticipate more robust and applicable results for urban population estimation and analysis.



**Figure 4.** Workflow of building use classification: two-stage forecasting flow.

## 2.2. Building Height Estimation

The creation of building height data relies on two key elements: the shape of the buildings and their respective heights. These factors dictate the overall structure of the city model and provide the depth and dimensionality necessary for rendering. In this study, methods for identifying surface object height, including the surface height model (SHM), utilize the Digital Surface Model (DSM) and the Digital Terrain Model (DTM) to accurately determine building heights in a grid format. Using the variance in elevation between the DTM and DSM, Equation (1) can be used to extract the height of the ground surface object.

$$SHM = DSM - DTM \quad (1)$$

### DEM Local Extrema Filtration

Filtration is employed to normalize the DSM by adapting the minimum and maximum values within a sliding window, as proposed by Huang et al. [52]. This method involves applying a filter that replaces each pixel value with the value derived from its neighboring vicinity. In our study, we decided to use the DSM as is for the maximum value in elevation to prevent oversmoothing.

The DSM is represented as an array of elevations, where each  $x_{ij}$  represents the pixel value at position  $(i, j)$  in the DSM matrix as follows:

$$DSM = \begin{bmatrix} x_{00} & \dots & x_{0n} \\ \vdots & \ddots & \vdots \\ x_{m0} & \dots & x_{mn} \end{bmatrix} \quad (2)$$

The neighboring values of  $x_{ij}$  can be identified as follows:

$$DSM_{ij} = \begin{bmatrix} x_{(i-1)(j-1)} & x_{(i-1)j} & x_{(i-1)(j+1)} \\ x_{i(j-1)} & x_{ij} & x_{i(j+1)} \\ x_{(i+1)(j-1)} & x_{(i+1)j} & x_{(i+1)(j+1)} \end{bmatrix} \quad (3)$$

For indices outside the matrix bounds, the mirror boundary condition is applied to prevent negative or out-of-range indices.

To create minimum filtration models, we apply minimum filters. The  $\min(DSM)$  matrix stores the minimum value defined as follows:

$$DTM \approx \min(DSM) = \begin{bmatrix} \min(DSM_{(0,0)}) & \dots & \min(DSM_{(0,n)}) \\ \vdots & \ddots & \vdots \\ \min(DSM_{(m,0)}) & \dots & \min(DSM_{(m,n)}) \end{bmatrix} \quad (4)$$

These matrices have the same dimensions as the initial DSM.

To normalize the DSM, the difference between the local maximum and minimum values is calculated. This is performed by subtracting the  $\min(DSM)$  matrix from the DSM matrix. The result represents the normalized DSM. The isolated height of objects above the ground level, denoted as  $SHM$ , is as follows:

$$SHM = DSM - \min(DSM) \quad (5)$$

Terrain correction is performed as described in Equation (6) by evaluating slopes greater than 10 degrees [52] using Horn's slope algorithms [53] applied to the reference DTM. This process identifies the areas with significant terrain variation, and in steep terrain, the model may overestimate building heights. The height of the slope to be corrected ( $slope_{corij}$ ) is calculated by determining the minimum and the maximum filtered model of the  $\min(DSM)$ . Specifically, the minimum value represents the "bottom of the slope", while the maximum value represents the "top of the slope". The slope height is then computed as the difference between these two values.

$$cSHM_{ij} = \begin{cases} SHM_{ij} - slope_{corij}, & \text{if } slope_{ij} \geq x\% \\ SHM_{ij}, & \text{if } slope_{ij} < x\% \end{cases} \quad (6)$$

where

- $cSHM_{ij}$  is the corrected SHM at position  $(i, j)$ ;
- $SHM_{ij}$  is the SHM at position  $(i, j)$ ;
- $slope_{ij}$  is the slope value at position  $(i, j)$ ;
- $slope_{corij}$  is the slope correction based on terrain data;
- $x\%$  is the threshold slope percentage.

The estimation of building height considers both usable net height  $H_{net}$  and total gross height  $H_{gross}$ , which are calculated using Equations (7) and (8), respectively.

$$H_{gross}(B_i) = \max\{cSHM(x, y) \mid (x, y) \in B_i\} \quad (7)$$

$$H_{net}(B_i) = \frac{1}{N} \sum_{(x,y) \in B_i} cSHM(x, y) \quad (8)$$

where

- $SHM(x, y)$  is the elevation value at a given pixel location  $(x, y)$ ;
- $B_i$  is a building polygon  $i$ ;
- $N$  is the total number of pixels within the building polygon.

### 2.3. Building Population Estimation

In this section, we integrate the building attribute classification and estimation methods with the population estimation approach developed by Akiyama et al. [15] to estimate the population per building unit. This method begins with the classification of building use followed by the identification of residential buildings, the calculation of building volumes,

and the proportional distribution of the population. The accuracy of these estimates is then validated against census data for each subregion. In this study, “subregion” refers to the areas known as “Cho-Cho-Aza,” which serve as the fundamental units of aggregation for statistical data in Japan. It refers to small administrative units used to organize areas within cities, towns, or villages. “Cho” typically represents a district or neighborhood, while “Aza” is a smaller subdivision within that area. These units are commonly used in official records to group and manage data for specific locations. Similar to city blocks or precincts in other countries, they facilitate the handling of demographic and geographic information at a detailed level, aiding in administrative and statistical processes.

For the verification process, we selected Setagaya Ward, Shibuya Ward, and Hachioji City in Tokyo prefecture, as well as Nagoya City in Aichi prefecture, Osaka City in Osaka prefecture, and Fukuoka City in Fukuoka prefecture. The selected areas were chosen to assess the versatility of the proposed method, as they encompass a wide range of urban characteristics. Each area exhibits distinct features differentiating them from the other cities, as outlined below. This approach enables a comprehensive evaluation of the method’s applicability across various urban contexts, ensuring its robustness in diverse settings.

- Setagaya Ward. Characterized by its vast residential areas and numerous upscale neighborhoods, Setagaya is known for its greenery and tranquil atmosphere. Unlike the business-centric areas of Shibuya or Shinjuku, Setagaya is predominantly residential.
- Shibuya Ward. As a commercial and business hub, Shibuya is particularly noted for its role as a center of youth culture and trendsetting. It is more urban and vibrant compared to the quieter Setagaya.
- Hachioji City. Situated in western Tokyo, Hachioji balances nature and urbanity. While it has a variety of commercial facilities, it retains a suburban feel, with less development compared to Setagaya or Shibuya.
- Nagoya City. This is the largest city in the Chubu region and the economic and industrial center of the area. It is a major urban hub, though its business district lacks the vibrant commercial culture seen in Shibuya.
- Osaka City. As the largest city in western Japan, Osaka is strongly characterized by its commercial activity. It is lively and historical, with popular tourist spots such as Dotonbori and Umeda. Compared to Nagoya, it has a more dynamic cultural scene, especially in terms of food and local businesses.
- Fukuoka City. As the central city of Kyushu, Fukuoka is known for its balanced integration of urban functionality and natural beauty. While it is smaller in scale compared to Osaka or Nagoya, it thrives as a compact city with significant cultural and economic exchanges with Asian countries.

### 2.3.1. Integration of Building Use Classification Results

The entirety of Tokyo was chosen as the target region for the building use classification to prevent overfitting to the characteristics of the specific area and to enhance the model’s versatility. This comprehensive approach is expected to yield more equitable and reliable results in subsequent accuracy validation.

### 2.3.2. Building Volume Calculation

After building use was assigned to the buildings in the study area, the calculation of building volume was performed. Firstly, the calculation of building volume was performed by multiplying the area of each building by the estimated number of floors. The number of floors was determined by dividing the estimated building height, obtained in the previous section, by an average floor height of 3 m. This approach allows for a more discrete height estimation, as the critical attribute for residential use is the number of floors [54]. To address this complexity, we developed a model to estimate the residential use of apartment buildings and determine the proportion of each building used for residential purposes.

To achieve this, we first created labels indicating whether each apartment building was used for residential purposes by leveraging nameplate data (containing household or

business information) and data on building use from residential maps. We then applied a machine learning method to construct a residential use estimation model, which produces predictions on a scale from 0 to 1. For example, a prediction of 0.7 indicates that 70% of the building is estimated to be used for residential purposes.

Finally, we calculated the volume used for residential purposes by multiplying this predicted value by the total building volume. For instance, if a building has a total volume of 1000 m<sup>3</sup> and the model predicts a residential use value of 0.6, the volume used for residential purposes is calculated as 1000 m<sup>3</sup> × 0.6 = 600 m<sup>3</sup>. This method enabled a more precise estimation of the portion of apartment buildings actually occupied by residents, leading to improved accuracy in subsequent population estimates. Apartment buildings with a predicted residential use value of less than 0.5 were classified as non-residential and excluded from the population assignment.

### 2.3.3. Population Allocation Algorithm

The population is prorated for each building using the constructed framework. In this study, population statistics and the number of households by municipality are prorated to generate population statistics on a building-by-building basis. The prorating method follows the influential study by Akiyama et al. [15].

The initial step involves calculating the volume of each building ( $v_{ij}$ ) by multiplying the building's area ( $s_{ij}$ ) by its number of floors ( $f_{ij}$ ), as expressed in Equation (9) as follows:

$$v_{ij} = s_{ij} \times f_{ij} \quad (9)$$

where

- $v_{ij}$  is the volume of the building  $j$  in subarea  $i$ ;
- $s_{ij}$  is the area of the building  $j$  in subarea  $i$ ;
- $f_{ij}$  is the number of floors of building  $j$  in subarea  $i$ .

Subsequently, the number of households is allocated to each building based on its relative volume within the subarea. The number of households assigned to building ( $h_{ij}$ ) is calculated using Equation (10) as follows:

$$h_{ij} = H_i \left( \frac{v_{ij}}{\sum_{k=1}^m v_{ik}} \right) \quad (10)$$

where

- $h_{ij}$  is the number of households assigned to building  $j$  in subarea  $i$ ;
- $H_i$  is the total number of households in subarea  $i$ ;
- $m$  is the number of buildings in the subarea.

To refine our population distribution estimate, the estimated area of each household ( $hs_{ij}$ ) is made from the building volume and the number of allocated households, as shown in Equation (11) as follows:

$$hs_{ij} = \frac{v_{ij}}{h_{ij}} \quad (11)$$

where

$hs_{ij}$  is the estimated area of each household.

Finally, the number of residents allocated to each household ( $r_{ij}$ ) is determined based on the estimated household area, as expressed in Equation (12) as follows:

$$r_{ij} = R_i \left( \frac{hs_{ij}}{\sum_{k=1}^m hs_{ik}} \right) \quad (12)$$

where  $R_i$  is the total number of residents in subarea  $i$ .

This methodological approach allows for an accurate estimation of population distribution, as the relative sizes of buildings and households within each subarea are taken into account.

### 3. Experiment and Results

#### 3.1. Datasets

This study utilized a combination of geospatial datasets, primarily from open data, to construct necessary data and estimate the micro-level population.

##### 3.1.1. Aerial Imagery

This study utilized aerial photographs provided by the Geospatial Information Authority of Japan (GSI) to construct a deep learning model. These images cover approximately half of Japan's land area, including plains and remote islands, which are crucial for effective land management, conservation, and regional development. The images were captured under optimal weather conditions using an aerial camera mounted on a survey aircraft, ensuring high-quality imagery with a spatial resolution of 0.5 m, free from cloud and fog interference.

The GSI aerial images were selected for this research due to their high accuracy and consistency with the base maps used to generate the training data. This alignment enhances the reliability and applicability of the deep learning model in analyzing land use patterns and urban structures across diverse Japanese landscapes.

##### 3.1.2. Building Polygon Data

This study integrated two primary data sources: the 2020 Fundamental Geospatial Data (FGD) from the Geospatial Information Authority of Japan (GSI) and the Zmap TOWN II (2020) commercial residential map by Zenrin, Inc. The FGD provides comprehensive geospatial data for the entirety of Japan, including 13 key elements, such as building perimeters, building shape, and elevation data, presented as highly accurate polygon data that align precisely with aerial imagery. While these maps offer extensive topographical information, they lack detailed building use attributes. To address this limitation, we incorporated the Zmap TOWN II data. Although the Zmap TOWN II is not open data, this integration with FGD enabled us to create a robust training dataset by leveraging the strengths of both sources to develop a more comprehensive and accurate model for building use classification. This approach paves the way for the potential open building use data in the future.

##### 3.1.3. Point of Interest (POI) Data

This study incorporated Point of Interest (POI) data from OpenStreetMap [29] to enhance the feature set for our machine learning model. POI data typically represent locations of public interest on maps, encompassing a diverse range of attributes, including public facilities, tourist attractions, and commercial establishments, such as convenience stores, restaurants, and gas stations. By leveraging this comprehensive POI data, we aim to capture the complex relationships between building functions and their surrounding urban context, potentially leading to more nuanced and accurate predictions in our AI-driven approach to urban analytics.

##### 3.1.4. DEM

One of the key data sources in building height estimation is the digital elevation model (DEM), which plays a crucial role in explaining the terrain either in the ground elevation model, such as in the Digital Terrain Model (DTM), or in portraying surface objects, such as in the Digital Surface Model (DSM). Using the openly available elevation model, the options have been narrowed down to the DEM, including AW3D30 [55] and NASADEM [56], as shown in Table 2. AW3D30 is considered one of the most recent global DSMs, as it is open,

has a high resolution, and has excellent accuracy. NASADEM is served as the DTM for the terrain references [52].

**Table 2.** Detailed information on the selected DEMs.

Characteristic	AW3D30 V4.1	NASADEM
Spatial resolution (m)	30	30
Vertical accuracy (m)	<5	3.5
Datum	ITRF97 and GRS80, using EGM96	WGS84/EGM96
Methodology	Photogrammetry	Interferometric SAR
Temporal coverage starts	January 2006	February 2000
Temporal coverage end	March 2016	February 2000
Last update	April 2024	February 2020
Data source	ALOS PRISM	SRTM, ASTER GDEM, ICESat
Availability	Free of charge	Free of charge

When considering the temporal coverage of these data sources, AW3D30 stands out due to its regular maintenance and updates, with coverage up to 2016. In contrast, NASADEM provides terrain input data up to the year 2000. However, the ground surfaces have not undergone significant changes, particularly in the urban areas where the main focus of this study lies. These data have been used for urban and regional planning and support [57,58], disaster management [59], environmental monitoring [60], resource management [61], and various other studies and have provided reliable and proven data. The ALOS global DSM is available in two products: AW3D with a 5 m spatial resolution and AW3D30 with a 30 m spatial resolution. AW3D30 is offered at no cost and is valuable for scientific research, education, and geospatial services. Its accessibility and up-to-date status make AW3D30 the primary DSM for estimating building height.

### 3.1.5. Building Height Reference

Project PLATEAU, an open and accessible 3D building data initiative by Japan's Ministry of Land, Infrastructure, Transport, and Tourism (MLIT), was employed to verify the accuracy of building heights. This project represents a significant advancement in digital twins technology [62,63] and adheres to the Open Geospatial Consortium (OGC) CityGML standards, ensuring interoperability and consistency.

### 3.1.6. Population Census

Population data from the Japan Statistics Bureau's Population Census are utilized to create and verify population statistics in this study. This comprehensive survey is conducted quinquennially, encompassing all individuals and households in Japan to assess the demographic and household composition, with the most recent one carried out in 2020 [64]. The survey collects a broad range of data, including household-level information, such as the building and the number of children in the household, as well as individual-level data, like age and occupation. These datasets are publicly available as open data and are compiled by prefecture, municipality, and subregion.

## 3.2. Results

### 3.2.1. Building Use Classification Results

This study evaluates the performance of our multi-stage building use classification model using four metrics: accuracy, precision, recall, and F-measure. The results for each classification stage are presented in Tables 3–5.

**Table 3.** Results of prediction for a detached house.

Building Use	Accuracy	Precision	Recall	F-Measure
Detached house	0.90	0.88	0.92	0.90
Non-detached house	0.90	0.92	0.90	0.90
				(N = 700)

**Table 4.** Results of prediction for multi-family housing.

Building Use	Accuracy	Precision	Recall	F-Measure
Multi-family housing	0.78	0.74	0.88	0.80
Non-multi-family housing	0.90	0.85	0.69	0.76

(N = 700)

**Table 5.** Results of two-stage prediction.

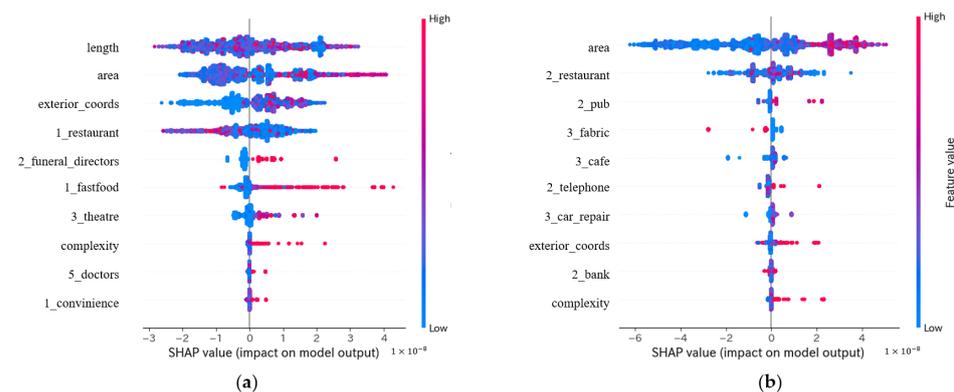
Building Use	Accuracy	Precision	Recall	F-Measure
Detached house	0.71	0.62	0.97	0.76
Detached office	0.71	0.70	0.46	0.55
Multi-family housing	0.71	0.90	0.71	0.80

(N = 1050)

The initial classification between detached and non-detached houses, as shown in Table 3, achieved high performance with 90% accuracy and all metrics exceeding 0.88, suggesting a clear distinction between these categories. Building on the first stage, the second stage, differentiating multi-family from non-multi-family housing, as detailed in Table 4, maintained good performance with 78% accuracy. Notably, multi-family housing showed a high recall at 0.88, indicating strong detection capability for apartment buildings. However, non-multi-family housing recall at 0.69 suggests room for improvement.

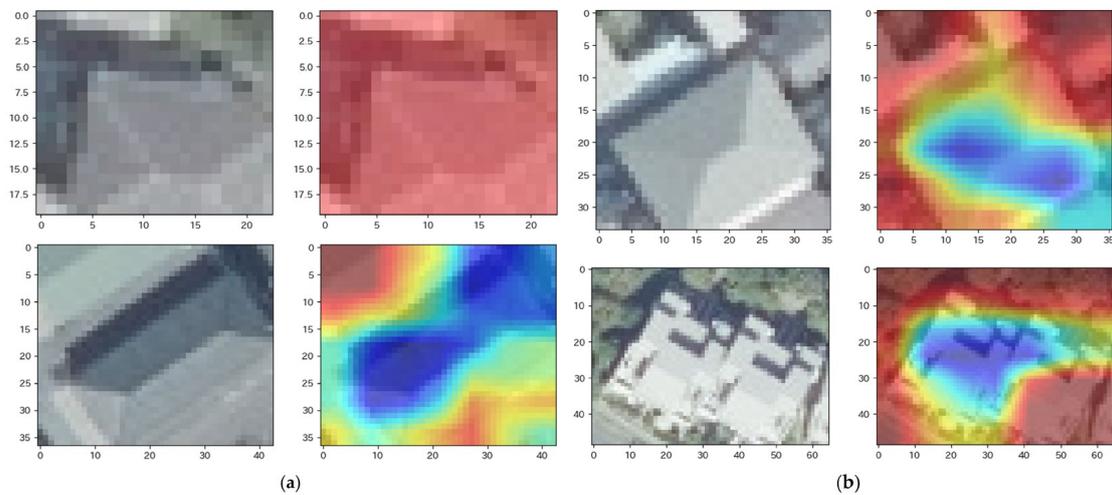
The two-stage forecasting flow, which includes detached houses, detached offices, and multi-family housing, as shown in Table 5, achieved 71% overall accuracy. Detached houses showed excellent recall at 0.97, while multi-family housing demonstrated high precision at 0.90. Detached offices, however, had a lower recall at 0.46 and were often misclassified as detached houses or apartment buildings.

To further understand the results, SHAP (SHapley Additive exPlanations) analysis, which was chosen due to its model-agnostic nature, versatility, and capacity to concurrently evaluate local and global feature importance, reveals that physical attributes derived from building polygon data are crucial determinants in classifying detached houses. Specifically, building perimeter (length) and area play significant roles, as shown in Figure 5.

**Figure 5.** (a) SHAP results of the detached house classification; (b) SHAP results of the multi-family housing classification.

Complementing the analysis, Grad-CAM, a method for visualizing the decision process of deep learning models, particularly convolutional neural networks (CNNs), visually highlights the parts of a building that significantly influence the model's decisions, as shown in Figure 6. Analysis of correctly classified cases confirms that the classification of detached houses focuses on the entire roof, indicating that the unique roof shape and structure are important distinguishing features. In the classification of apartment buildings, a strong recognition of the building outline is observed, suggesting that the large scale and

distinctive external shape are key to classification. However, analysis of misclassified cases reveals the model's limitations and its room for improvement.



**Figure 6.** (a) Grad-CAM results of true classification (top: detached house, bottom: multi-family housing); (b) Grad-CAM results of false classification (top: detached house, bottom: multi-family housing).

### 3.2.2. Building Height Estimation Results

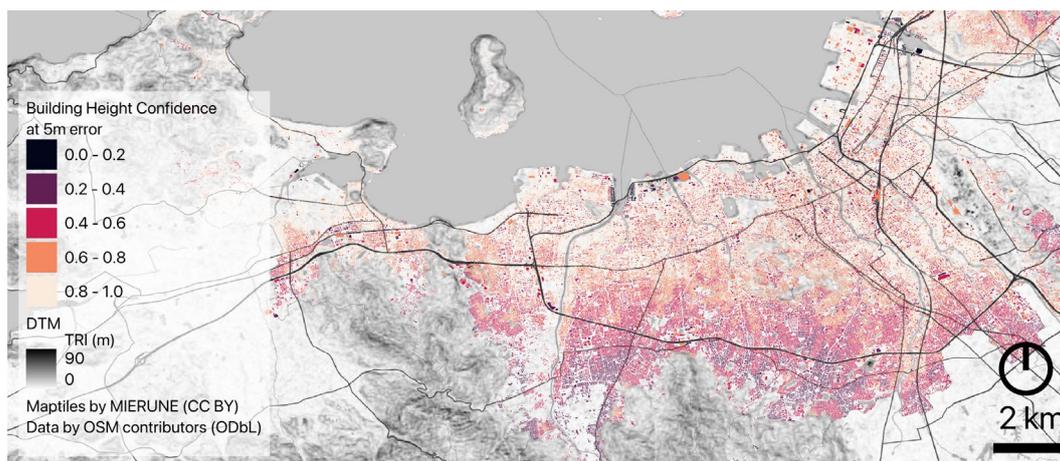
The results demonstrated that the accuracy varies by region, as shown in Table 6. This comparative analysis covers our entire study area. We defined the 5 m confidence level as the percentage of buildings measured with an error margin within 5 m. This threshold aligns with the precision of our input DSM data, which has an accuracy of 5 m or better, ensuring consistency between our estimations and the inherent limitations of the input data.

**Table 6.** Result of the building height estimation.

Metric	Tokyo	Hachioji	Nagoya	Osaka	Fukuoka
Total number of buildings	1,768,107	180,940	678,220	560,198	355,454
5m confidence	90.35%	65.85%	86.19%	93.55%	67.21%
90% confidence error (m)	4.9	19.6	6.4	3.8	15.0
RMSE	6.88	13.37	7.18	6.69	26.7
MAE	4.64	8.56	5.1	3.97	16.11
Mean TRI (m)	5.09	16.64	5.04	4.03	6.19
SD TRI (m)	4.75	14.23	3.30	3.60	10.11

Tokyo Special Sards, including Setagaya Ward and Shibuya Ward, as well as Nagoya City and Osaka City, exhibit a promising building height estimation of over 85%. In contrast, Hachioji City and Fukuoka City exhibit relatively lower accuracy. We identified that a factor influencing this variability is the terrain complexity. We measured it with the terrain ruggedness index (TRI), calculated from the DTM data of each area. The TRI represents the standard deviation of elevation differences between a center cell and its neighbor.

The results reveal a clear correlation between terrain complexity and building height estimation accuracy, as shown in the TRI and 5 m confidence. The example of the results over Fukuoka City, displayed in Figure 7, shows that lower estimation accuracy is clustered near the mountains in the southern region.



**Figure 7.** Result example over Fukuoka City showing 5 m error confidence and the terrain ruggedness index.

### 3.2.3. Population Allocation Results

Using the results of building attribute estimation, we generated building-by-building demographic data, as shown in Figure 8. The accuracy of these estimates was verified by aggregating the results by subregion and conducting a regression analysis, with population census data for each subregion serving as the true values. In this study, three metrics were used to assess accuracy: the coefficient of determination ( $R^2$ ), the RMSE, and the MAE. The MAE directly reflects the overall error size and is less sensitive to outliers, while the RMSE is more responsive to large errors. By employing these two indices alongside the coefficient of determination, we were able to account for the impact of small areas that represent significant outliers in the predictions.

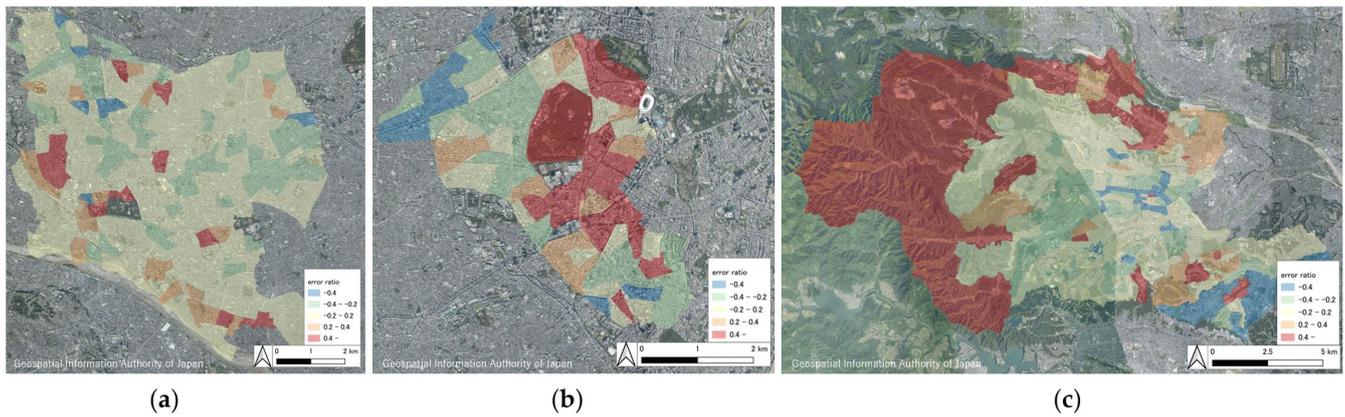


**Figure 8.** Estimated population per building in a part of Hachioji City, Tokyo.

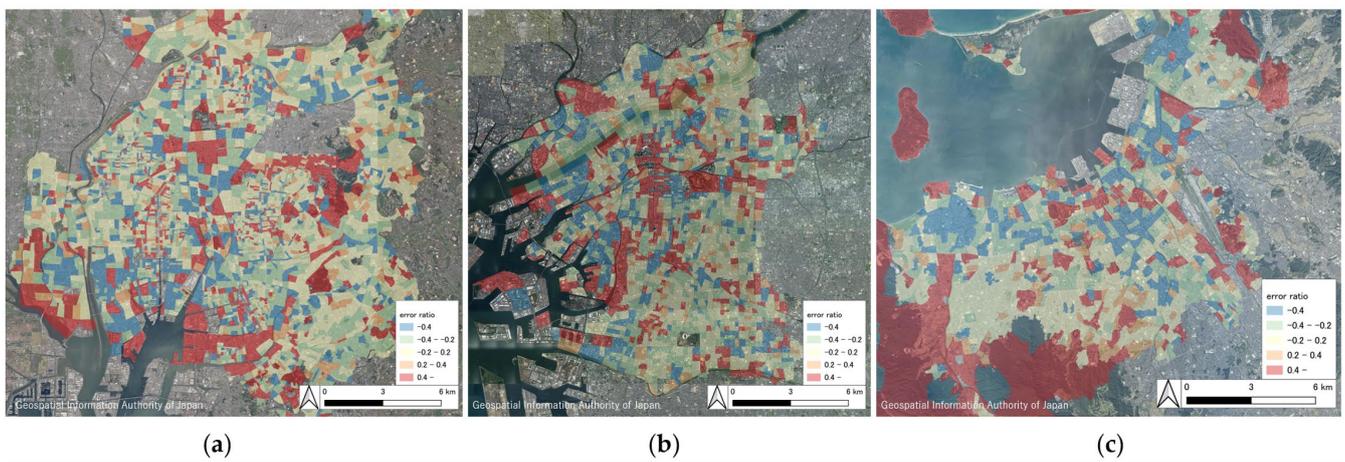
Table 7 presents the results of the accuracy validation, Figures 9 and 10 illustrate the number of visibility data calculated based on the error rate for each subregion, and Figures 11 and 12 display the regression analysis outcomes.

**Table 7.** Results of regression validation.

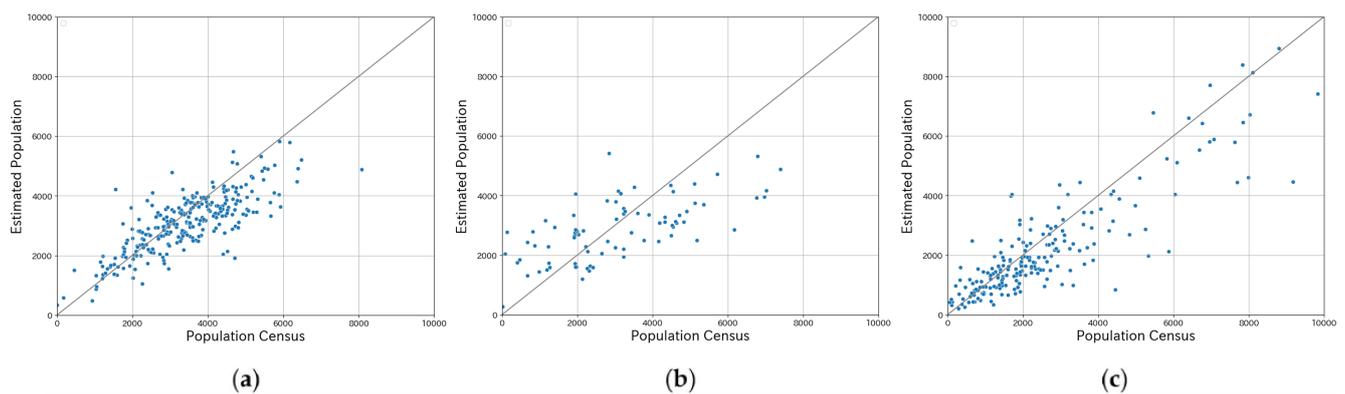
	Setagaya	Shibuya	Hachioji	Nagoya	Osaka	Fukuoka
$R^2$	0.56	0.41	0.82	0.44	0.35	0.14
RMSE	842	1371	1163	443	869	913
MAE	634	1137	769	222	578	6100



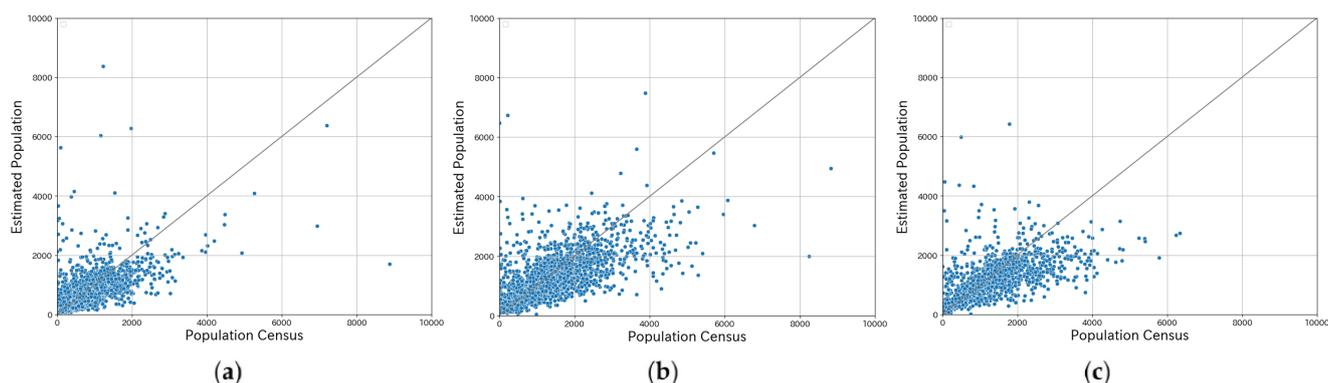
**Figure 9.** Result of error rate for each subregion in Tokyo prefecture: (a) case of Setagaya Ward; (b) case of Shibuya Ward; (c) case of Hachioji City.



**Figure 10.** Result of error rate for each subregion outside of Tokyo prefecture: (a) case of Nagoya City, Aichi prefecture; (b) case of Osaka City, Osaka prefecture; (c) case of Fukuoka City, Fukuoka prefecture.



**Figure 11.** Result of regression analysis in Tokyo prefecture: (a) case of Setagaya Ward; (b) case of Shibuya Ward; (c) case of Hachioji City.



**Figure 12.** Result of regression analysis in another Tokyo prefecture: (a) case of Nagoya City, Aichi prefecture; (b) case of Osaka City, Osaka prefecture; (c) case of Fukuoka City, Fukuoka prefecture.

As shown in Table 7, the coefficient of determination was particularly high in Hachioji City, reaching 0.82, but was below 0.5 in Shibuya Ward. Additionally, the MAE was significantly lower than the RMSE in Hachioji City and Setagaya Ward, indicating that while population estimation accuracy was relatively high for many buildings in these areas, there were substantial errors in certain buildings and regions.

#### 4. Discussion

This study aimed to develop a method for micro-population estimation at the building unit scale, with a particular focus on estimating building height and usage. Additionally, we sought to ensure the applicability and scalability of our framework across regions by emphasizing the integration of generally available data, preferably open data. The model was tested in Japan, where the reference data are available for validation. Our research has yielded several findings.

##### 4.1. Evaluation from Building Use Classification

First, the building use classification by the machine learning approach using two-stage forecasting flow showed the model's strength in distinguishing detached from non-detached houses and identifying apartment buildings. The decision-making process of the machine learning models was analyzed using SHAP and Grad-CAM, revealing the key features that each model prioritizes in classifying building use and verifying the complementarity between the polygon data-based and aerial image-based model.

SHAP analysis showed that the relative uniformity in size and shape is a characteristic feature influencing the identification of detached houses. Additionally, features extracted from POI data, such as the proximity to restaurants and the spatial co-location patterns of various POIs, play a significant role in identifying multi-family housing (apartment buildings). These POI features capture the semantic context surrounding buildings, reflecting urban development patterns, and real-world spatial relationships [65,66]. However, it is important to note that urban developments and zoning regulations can vary by region, which should be taken into account when adopting the approach. The classification of detached offices remains a challenge requiring further refinement.

Grad-CAM analysis reveals that misclassification of detached houses often occurs when the model fails to adequately capture roof characteristics, particularly when recognizing roofs of attached buildings, such as warehouses and garages. Misclassification of apartment buildings often arises due to difficulty in recognizing the contours of buildings with complex shapes. To mitigate this problem, using higher resolution image data and improvement in model architecture to properly capture the complex structure of buildings are recommended.

The results of both SHAP and Grad-CAM analysis confirm that different machine learning models, polygon data based and aerial image based, utilize different features for

building use classification, providing insights into the strengths and limitations of each of the models. The implementation of the two-stage forecasting approach enhances the efficacy of the proposed ensemble method, resulting in a more comprehensive and accurate building use classification.

The utilization of POI data presents two significant challenges. Firstly, the distribution characteristics of POI data may considerably vary across different regions. Even within Japan, the applicability of these characteristics is expected to differ between urban centers and suburban areas. Given that this study aims for international implementation, it is crucial to conduct a detailed examination of the variations in these characteristics across different countries and regions. The second challenge pertains to the completeness of POI data. Compared to Japan, developing countries, which this study intends to investigate in the further study, are likely to have limited availability of such data. Consequently, there is a concern that the model may become applicable only to data-rich areas, thus restricting its generalizability.

To address these challenges, we propose two potential improvements. Firstly, the construction of a model using representative data extracted from diverse regions through appropriate sampling techniques could be considered. This approach would enable the creation of a more versatile model that is not dependent on any specific region. Secondly, the creation of distinct models for different areas or land use zones could be effective. By selecting and applying the most suitable model based on regional characteristics, it is anticipated that accurate classification can be performed across diverse geographical contexts.

Future research should focus on implementing these proposed improvements and quantitatively evaluating their efficacy. Furthermore, it is essential to explore alternative spatial data sources that could potentially substitute for POI data, as well as develop new features that can be applied in data-scarce regions. This approach will enhance the model's adaptability to various urban environments globally, thereby increasing its practical applicability in diverse international settings.

#### *4.2. Terrain Impact on Building Height Estimation Accuracy*

The results of our building height estimation analysis reveal that the accuracy of the estimations varies significantly across different regions. Areas with relatively flat terrain, such as Tokyo, Nagoya, and Osaka, yield higher estimation accuracies of over 85%. In contrast, to better understand the phenomenon of underperforming estimations, our preliminary analysis shows that errors are more prevalent in terrains with approximately 10% or greater elevation variation.

Several factors may contribute to the increased error margins in mountainous regions. Firstly, the techniques used for height estimation may be insufficient for the topographic complexity, which is characterized by substantial variability in elevation. This complexity complicates the estimation process, as the interplay between terrain and building heights may lead to increased uncertainty. Secondly, data resolution may be insufficient to capture the fine details in rugged terrains, as the DEM data used in this study are 30 m. This limitation can result in inaccuracies when estimating building heights, as the terrain variations may not be adequately represented in the input data. Model calibration is also one significant factor. Although our estimation model accounts for terrain correction, further adjustments or the development of new frameworks tailored specifically to mountainous environments may be necessary. The unique characteristics of these regions may require specialized approaches to mitigate the impact of terrain complexity on building height estimation.

When utilizing the number of floors derived from estimated height, it is important to consider that both metrics represent similar building characteristics. However, exploring the margin of differences between building height and floor count could provide additional insights. Additionally, factors such as zoning should be considered when determining average floor height [67].

### 4.3. Population Allocation

The population allocation methodology has proven its performance in this study where the data are well maintained and commercialized [15]. However, our micro-population estimation model revealed several areas for improvement. These areas can be categorized into three main trends: (1) overestimation in sparsely populated areas, (2) inaccuracies in areas with high concentrations of non-residential buildings, and (3) misclassification of industrial structures in port and harbor areas.

The first trend observed was overestimation in sparsely populated areas, such as mountainous regions. In this study, Hachioji City and Fukuoka City exhibited lower accuracy due to this issue. In such areas, even small differences in the number of residents can lead to significant errors due to the small population size. Additionally, since this method proportionally allocates population statistics by building units within a municipality, some areas may receive a disproportionately small allocation compared to the actual distribution. To address this, future research should explore methods that allow for population prorating on a more granular, underlying unit basis.

The second observed trend was lower accuracy in areas with a high concentration of non-residential buildings, particularly commercial facilities, such as downtown areas. In these regions, there was a tendency to misclassify non-residential buildings, especially commercial properties, as residential buildings. This misclassification led to an over-allocation of the population. To improve accuracy in these areas, it is necessary to refine the threshold for determining residential use, enhance the model, and develop methods for accurately classifying areas not designated for residential purposes.

Finally, the misclassification of warehouses and other facilities located in ports and harbors emerged as a significant issue. In Osaka and Nagoya, the population of warehouses in port areas was misclassified, resulting in decreased accuracy. This problem combines elements of the previous two issues, and it is characterized by a lack of former residents and the area not being used for residential purposes. In such areas, it is essential first to develop methods to exclude these regions from the target population and to improve the model accordingly.

These improvement measures are expected to lead to the development of a more accurate building-by-building population estimation method. Moving forward, addressing these challenges is essential for developing a versatile population estimation model that can be applied across a broader range of urban environments.

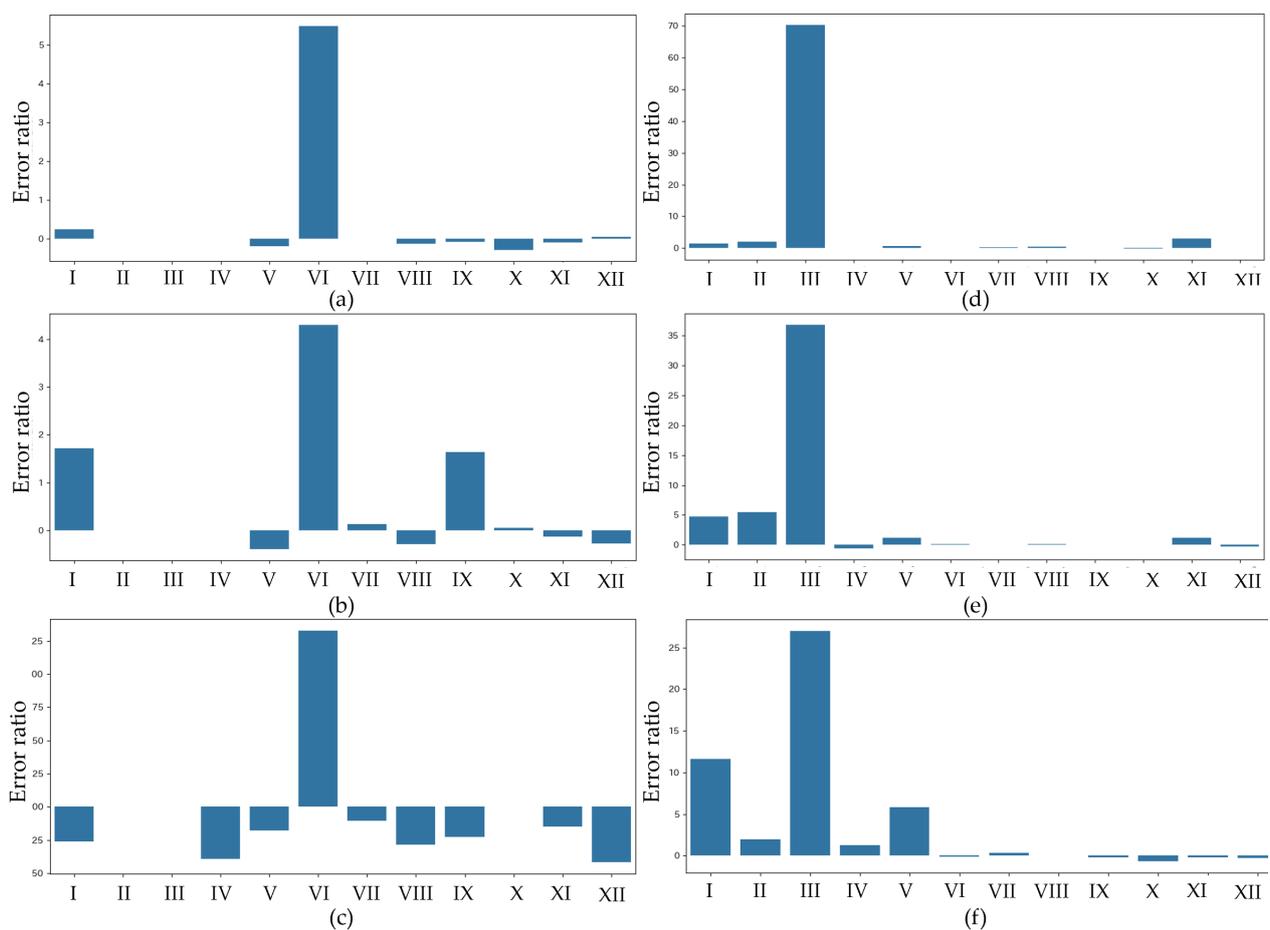
To further investigate these three trends, we conducted a comparative analysis of the relationship between error rates and land use zones by aligning Japanese land use zoning data with small-area boundaries. We utilized land use zoning data from the National Land Numerical Information database [68]. Table 8 defines each land use zone, while Figure 13 presents the average aggregated results of land use zones and error rates for each small area.

**Table 8.** Land use zoning classification and descriptions of Japan.

	Land Use Zoning	Descriptions
I	Commercial Zone	A zone for larger commercial facilities, such as banks, cinemas, restaurants, and department stores, as well as housing and small factories.
II	Industrial Zone	A zone where various industrial factories are allowed, but it also includes schools, offices, and some residential buildings.
III	Exclusive Industrial Zone	A zone designated exclusively for industrial factories. In this zone, no residential buildings, schools, offices, or hotels are allowed.
IV	Quasi-Industrial Zone	A zone for light industrial activities and small-scale factories that do not produce significant hazards or nuisances to the surrounding environment.

Table 8. Cont.

	Land Use Zoning	Descriptions
V	Quasi-Residential Zone	A zone located along roadsides that allows for auto repair shops, gas stations, and small industries while maintaining the residential environment.
VI	Category II: Mid-/High-Rise-Oriented Residential Zone	A zone for medium- to high-rise residential buildings. Shops, offices, and hospitals up to 1500 square meters are permitted.
VII	Category I: Mid-/High-Rise-Oriented Residential Zone	A zone for medium- to high-rise residential buildings. Shops, offices, and hospitals up to 500 square meters are permitted.
VIII	Category I: Exclusively Low-Rise Residential Zone	A zone designated for low-rise residential buildings, where small shops and offices are allowed near housing, as well as elementary and middle schools.
IX	Category II: Exclusively Low-Rise Residential Zone	A zone designated for low-rise residential buildings, where small shops and offices up to 150 square meters are allowed, in addition to housing.
X	Category II: Residential Zone	A zone primarily for residential use but also permits shops, offices, hotels, and karaoke boxes.
XI	Category I: Residential Zone	A zone intended for residential purposes. In this zone, larger shops, offices, hotels, and schools up to 3000 square meters are also permitted.
XII	Neighborhood Commercial Zone	A zone for small-scale commercial facilities that serve the daily needs of residents, such as small shops and grocery stores.



**Figure 13.** The average aggregated results of land use zones and error rates for each small area: (a) case of Setagaya Ward, Tokyo prefecture; (b) case of Shibuya Ward, Tokyo prefecture; (c) case of Hachioji City, Tokyo prefecture; (d) case of Nagoya City, Aichi prefecture; (e) case of Osaka City, Osaka prefecture; (f) case of Fukuoka City, Fukuoka prefecture.

Notably, we observed differences in trends between Tokyo, where the model was developed, and other cities. In Tokyo, larger errors were more common in areas, such as

Category I and Category II, Mid-/High-Rise-Oriented Residential Zone, where hospitals, universities, and stores can coexist with residences. The higher error rates in these areas are likely due to misclassifications between hospitals or certain-sized stores and apartment buildings or detached office buildings, and there is a tendency for taller buildings to exhibit greater estimation errors.

In contrast, predictions for areas outside Tokyo frequently showed misclassifications in exclusive industrial zones. This likely occurred because the Tokyo-based model did not include sufficient examples from these zones, leading to their classification as apartment buildings and the proportional allocation of population estimates. This finding corroborates previous observations of misclassifications in port areas.

Furthermore, both inside and outside Tokyo, commercial zones exhibited a tendency for overestimation. This overestimation likely stems from the misclassification of commercial facilities as apartment buildings combined with the general tendency for taller structures to produce larger estimation errors.

To address these issues, one approach is to appropriately sample training data from each land use zone during model development. In this study, large errors occurred in industrial areas not represented in the training data. By ensuring that training data include a diverse range of land use zones, we can mitigate such biases. Another strategy is to cluster areas based on their characteristics and develop separate models for each cluster or region. Our analysis revealed distinct error patterns between Tokyo, where the model was trained, and other regions. By clustering or grouping similar areas and constructing models tailored to these clusters, we can potentially improve prediction accuracy.

By adopting these methodologies and modeling approaches that account for regional characteristics, we aim to develop more effective population estimation techniques.

Our ultimate goal is to establish a robust methodology for generating reliable building-level demographic statistics in developing countries, which is essential for urban planning, resource allocation, and disaster management. Access to detailed data on both buildings and population enables us to conduct comprehensive analyses at the building unit level. An exemplar of this approach is illustrated in Figure 14, which represents a prototype for integrating population data at the building unit level with vulnerability factors in risk assessment [69]. This prototype classifies vulnerability levels for population distribution in each building, incorporating data from neighboring structures within a defined buffer zone.



**Figure 14.** Prototype of the application in fire risk assessment (Tokyo, Japan).

## 5. Conclusions

In this study, we developed a method for micro-population estimation at the building unit scale, focusing on estimating building height and usage while ensuring the applicability

and scalability of our framework across regions. The model was tested in Japan, where reference data were available for validation.

Our building use classification approach, utilizing a two-stage forecasting flow with machine learning, demonstrated strengths in distinguishing detached from non-detached houses and identifying apartment buildings. SHAP and Grad-CAM analyses provided insights into the decision-making processes of the models, revealing the key features prioritized in classifying building use and confirming the complementarity between polygon data-based and aerial image-based approaches. However, challenges remain in classifying detached offices and capturing complex building structures, which can be addressed through higher-resolution image data and improved model architectures.

The building height estimation analysis revealed that accuracy varies significantly across different regions, with higher accuracies in areas with relatively flat terrain. Errors were more prevalent in terrains with approximately 10% or greater elevation variation. Factors such as topographic complexity, data resolution, and model calibration contribute to increased error margins in mountainous regions. Future research should focus on developing specialized approaches tailored to these environments.

The population allocation methodology has three main areas for improvement to be considered: overestimation in sparsely populated areas, inaccuracies in areas with high concentrations of non-residential buildings, and misclassification of industrial structures in port and harbor areas. Addressing these challenges is essential for developing a versatile population estimation model applicable across a broader range of urban environments. Comparative analysis of error rates and land use zones revealed distinct trends between Tokyo and other cities, highlighting the need for appropriate sampling of training data from each land use zone and the development of separate models for different clusters or regions.

Overall, potential improvements include refining the building information and further specifying its usage and physical characteristics. Taking into account regional characteristics and utilizing multimodal data is also preferable. Additionally, setting region-specific thresholds and introducing pre-filtering processes could contribute to further accuracy gains. The resolution of the population data is also significant in micro-scale population mapping [12].

To facilitate the practical application of the proposed methodology, users of open data should consider focusing on several key points. These include identifying relevant open data sources, deriving building use information from alternative datasets where the data are absent, applying the methodology to map the micro-population of interest, and validating the results using ground truth data.

This study presented a novel approach to micro-population estimation at the building unit scale, demonstrating the application of machine learning for building use classification, morphology-based height estimation, and statistical population mapping. The results show the potential for high-resolution estimations in areas with uniform residential buildings while identifying challenges in mountainous regions and commercial areas. By addressing these limitations and extending the applicability of the method to diverse regions, particularly in developing countries, we aim to contribute to the collection of accurate building-level demographic data for urban planning, resource allocation, and disaster management. Future research should focus on refining the models to better capture complex building structures, developing specialized approaches for mountainous environments, and improving population allocation accuracy in non-residential areas. Furthermore, the exploration of alternative spatial data sources and the development of new features applicable to data-scarce regions will enhance the model's adaptability to various urban environments globally, thereby increasing its practical applicability in diverse international settings.

**Author Contributions:** Conceptualization, K.M. and Y.A.; methodology, K.M. and R.Y.; software, K.M. and R.Y.; validation, R.Y. and H.M.; formal analysis, K.M. and R.Y.; investigation, R.Y.; data curation, K.M., R.Y. and H.M.; writing—original draft preparation, K.M.; writing—review and editing, R.Y., Y.A., H.M. and C.M.A.; visualization, K.M. and R.Y.; supervision, H.M. and S.M.; project administration, Y.A.; funding acquisition, Y.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by JSPS KAKENHI, Grant Number JP24K00243.

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** Author Hiroyuki Miyazaki was employed by the company GLODAL, Inc.; Author Satoshi Miyazawa was employed by the company LocationMind Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Mennis, J. Dasymetric Mapping for Estimating Population in Small Areas. *Geogr. Compass* **2009**, *3*, 727–745. [CrossRef]
- Mennis, J. Generating Surface Models of Population Using Dasymetric Mapping. *Null* **2003**, *55*, 31–42. [CrossRef]
- Zandbergen, P.A.; Ignizio, D.A. Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates. *Cartogr. Geogr. Inf. Sci.* **2010**, *37*, 199–214. [CrossRef]
- WorldPop Gridded Population Estimate Datasets and Tools. Available online: <https://www.worldpop.org/methods/populations/> (accessed on 10 September 2024).
- European Commission, Joint Research Centre. *GHSL Data Package 2019: Public Release GHS P2019*; Publications Office: Luxembourg, 2019.
- Reiter, D.; Jehling, M.; Hecht, R. Benefits of Using Address-Based Dasymetric Mapping in Micro-Level Census Disaggregation. *AGILE GIScience Ser.* **2023**, *4*, 38. [CrossRef]
- Pirowski, T.; Szypuła, B. Dasymetric Population Mapping Using Building Data. *Ann. Am. Assoc. Geogr.* **2024**, *114*, 1001–1019. [CrossRef]
- Sakti, A.D.; Rinasti, A.N.; Agustina, E.; Diastomo, H.; Muhammad, F.; Anna, Z.; Wikantika, K. Multi-Scenario Model of Plastic Waste Accumulation Potential in Indonesia Using Integrated Remote Sensing, Statistic and Socio-Demographic Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 481. [CrossRef]
- Wang, Y.; Sun, G.; Wu, Y.; Rosenberg, M.W. Urban 3D Building Morphology and Energy Consumption: Empirical Evidence from 53 Cities in China. *Sci. Rep.* **2024**, *14*, 12887. [CrossRef]
- Katada, T.; Kuwasawa, N.; Shida, S.; Kojima, M. Scenario Analysis for Evacuation Strategies for Residents in Big Cities During Large-Scale Flooding. *J. JSCE* **2015**, *3*, 209–223. [CrossRef]
- Tiecke, T.G.; Liu, X.; Zhang, A.; Gros, A.; Li, N.; Yetman, G.; Kilic, T.; Murray, S.; Blankespoor, B.; Prydz, E.B.; et al. Mapping the World Population One Building at a Time 2017. *arXiv* **2017**, arXiv:1712.05839.
- Pajares, E.; Muñoz Nieto, R.; Meng, L.; Wulforst, G. Population Disaggregation on the Building Level Based on Outdated Census Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 662. [CrossRef]
- Bakillah, M.; Liang, S.; Mobasheri, A.; Jokar Arsanjani, J.; Zipf, A. Fine-Resolution Population Mapping Using OpenStreetMap Points-of-Interest. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 1940–1963. [CrossRef]
- Calka, B.; Nowak Da Costa, J.; Bielecka, E. Fine Scale Population Density Data and Its Application in Risk Assessment. *Geomat. Nat. Hazards Risk* **2017**, *8*, 1440–1455. [CrossRef]
- Akiyama, Y.; Miyazaki, H.; Sirikanjanaanan, S. Development of Micro Population Data for Each Building: Case Study in Tokyo and Bangkok. In Proceedings of the 2019 First International Conference on Smart Technology Urban Development (STUD), Chiang Mai, Thailand, 13–14 December 2019; pp. 1–6.
- Zhao, L.; Liu, X.; Xu, X.; Liu, C.; Chen, K. Three-Dimensional Simulation Model for Synergistically Simulating Urban Horizontal Expansion and Vertical Growth. *Remote Sens.* **2022**, *14*, 1503. [CrossRef]
- Gevaert, C.M.; Buunk, T.; Van Den Homberg, M.J.C. Auditing Geospatial Datasets for Biases: Using Global Building Datasets for Disaster Risk Management. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 12579–12590. [CrossRef]
- Ghaffarian, S.; Roy, D.; Filatova, T.; Kerle, N. Agent-Based Modelling of Post-Disaster Recovery with Remote Sensing Data. *Int. J. Disaster Risk Reduct.* **2021**, *60*, 102285. [CrossRef]
- Magnaye, A.M.T.; Kusaka, H. Potential Effect of Urbanization on Extreme Heat Events in Metro Manila Philippines Using WRF-UCM. *Sustain. Cities Soc.* **2024**, *110*, 105584. [CrossRef]
- Kajiwara, K.; Ma, J.; Seto, T.; Sekimoto, Y.; Ogawa, Y.; Omata, H. Development of Current Estimated Household Data and Agent-Based Simulation of the Future Population Distribution of Households in Japan. *Comput. Environ. Urban Syst.* **2022**, *98*, 101873. [CrossRef]
- Goniewicz, K.; Burkle, F.M.; Hall, T.F.; Goniewicz, M.; Khorram-Manesh, A. Global Public Health Leadership: The Vital Element in Managing Global Health Crises. *J. Glob. Health* **2022**, *12*, 03003. [CrossRef]

22. Irandoost, K.; Alizadeh, H.; Yousefi, Z.; Shahmoradi, B. Spatial Analysis of Population Density and Its Effects during the COVID-19 Pandemic in Sanandaj, Iran. *J. Asian Archit. Build. Eng.* **2023**, *22*, 635–642. [[CrossRef](#)]
23. Zhang, Z. Research on Urban Building Planning and Construction System under Computer 3D Digitization. In Proceedings of the 2023 IEEE International Conference on Sensors, Electronics and Computer Engineering (ICSECE), Jinzhou, China, 18–20 August 2023; pp. 1102–1106.
24. Agius, T.; Sabri, S.; Kalantari, M. Three-Dimensional Rule-Based City Modelling to Support Urban Redevelopment Process. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 413. [[CrossRef](#)]
25. Chen, H.; Wu, B.; Yu, B.; Chen, Z.; Wu, Q.; Lian, T.; Wang, C.; Li, Q.; Wu, J. A New Method for Building-Level Population Estimation by Integrating LiDAR, Nighttime Light, and POI Data. *J. Remote Sens.* **2021**, *2021*, 9803796. [[CrossRef](#)]
26. Schug, F.; Frantz, D.; van der Linden, S.; Hostert, P. Gridded Population Mapping for Germany Based on Building Density, Height and Type from Earth Observation Data Using Census Disaggregation and Bottom-up Estimates. *PLoS ONE* **2021**, *16*, e0249044. [[CrossRef](#)] [[PubMed](#)]
27. Wang, S.; Li, R.; Jiang, J.; Meng, Y. Fine-Scale Population Estimation Based on Building Classifications: A Case Study in Wuhan. *Future Internet* **2021**, *13*, 251. [[CrossRef](#)]
28. Corbane, C.; Florczyk, A.; Pesaresi, M.; Politis, P.; Syrris, V. GHS Built-up Grid, Derived from Landsat, Multitemporal (1975-1990-2000-2014), R2018A. 2018. Available online: <http://data.europa.eu/89h/jrc-ghsl-10007> (accessed on 27 June 2022).
29. Open Street Map. Available online: <https://www.openstreetmap.org/> (accessed on 4 September 2024).
30. Microsoft Global ML Building Footprints. Available online: <https://github.com/microsoft/GlobalMLBuildingFootprints> (accessed on 4 September 2024).
31. Google Open Buildings. Available online: <https://sites.research.google/open-buildings/> (accessed on 4 September 2024).
32. Overture Maps Foundation Overture Maps. Available online: <https://overturemaps.org/> (accessed on 4 September 2024).
33. Yao, Y.; Liu, X.; Li, X.; Zhang, J.; Liang, Z.; Mai, K.; Zhang, Y. Mapping Fine-Scale Population Distributions at the Building Level by Integrating Multisource Geospatial Big Data. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 1220–1244. [[CrossRef](#)]
34. Shang, S.; Du, S.; Du, S.; Zhu, S. Estimating Building-Scale Population Using Multi-Source Spatial Data. *Cities* **2021**, *111*, 103002. [[CrossRef](#)]
35. Cheng, Q.; Zaber, M.; Rahman, A.M.; Zhang, H.; Guo, Z.; Okabe, A.; Shibasaki, R. Understanding the Urban Environment from Satellite Images with New Classification Method—Focusing on Formality and Informality. *Sustainability* **2022**, *14*, 4336. [[CrossRef](#)]
36. Wang, Z.; Liang, Y.; He, Y.; Cui, Y.; Zhang, X. Refined Land Use Classification for Urban Core Area from Remote Sensing Imagery by the EfficientNetV2 Model. *Appl. Sci.* **2024**, *14*, 7235. [[CrossRef](#)]
37. Fisher, T.; Gibson, H.; Liu, Y.; Abdar, M.; Posa, M.; Salimi-Khorshidi, G.; Hassaine, A.; Cai, Y.; Rahimi, K.; Mamouei, M. Uncertainty-Aware Interpretable Deep Learning for Slum Mapping and Monitoring. *Remote Sens.* **2022**, *14*, 3072. [[CrossRef](#)]
38. Rajput, T.S.; Singhal, A.; Routroy, S.; Dhadse, K.; Tyagi, G. Urban Policymaking for a Developing City Using a Hybridized Technique Based on SWOT, AHP, and GIS. *J. Urban Plann. Dev.* **2021**, *147*, 04021018. [[CrossRef](#)]
39. Sisodia, P.S.; Tiwari, V.; Dahiya, A.K. Measuring and Monitoring Urban Sprawl of Jaipur City using Remote Sensing and GIS. *Int. J. Inf. Syst. Soc. Chang.* **2015**, *6*, 46–65. [[CrossRef](#)]
40. Biljecki, F. Exploration of Open Data in Southeast Asia to Generate 3d Building Models. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *VI-4/W1-2020*, 37–44. [[CrossRef](#)]
41. Bongaarts, J. United Nations Department of Economic and Social Affairs, Population Division World Family Planning 2020: Highlights, United Nations Publications, 2020. 46 p. *Popul. Dev. Rev* **2020**, *46*, 857–858. [[CrossRef](#)]
42. Walker, R.J. Population Growth and Its Implications for Global Security. *Am. J. Econ. Sociol.* **2016**, *75*, 980–1004. [[CrossRef](#)]
43. De La Croix, D.; Gobbi, P.E. Population Density, Fertility, and Demographic Convergence in Developing Countries. *J. Dev. Econ.* **2017**, *127*, 13–24. [[CrossRef](#)]
44. Abel, G.J.; Barakat, B.; Kc, S.; Lutz, W. Meeting the Sustainable Development Goals Leads to Lower World Population Growth. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 14294–14299. [[CrossRef](#)]
45. Leasure, D.R.; Jochem, W.C.; Weber, E.M.; Seaman, V.; Tatem, A.J. National Population Mapping from Sparse Survey Data: A Hierarchical Bayesian Modeling Framework to Account for Uncertainty. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 24173–24179. [[CrossRef](#)]
46. Kang, J.; Körner, M.; Wang, Y.; Taubenböck, H.; Zhu, X.X. Building Instance Classification Using Street View Images. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 44–59. [[CrossRef](#)]
47. Laupheimer, D.; Tutzauer, P.; Haala, N.; Spicker, M. Neural Networks for the Classification of Building Use from Street-View Imagery. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *IV-2*, 177–184. [[CrossRef](#)]
48. Okada, K.; Nishiyama, N.; Akiyama, Y.; Miyazaki, H.; Miyazawa, S. Development of Detailed Building Distribution Map to Support Smart City Promotion -an Approach Using Satellite Image and Deep Learning. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *10*, 189–196. [[CrossRef](#)]
49. Wei, Y.; Luo, G.; Yu, L.; Huang, Z. Identification of Urban Building Functions Based on Points of Interest and Spatial Relationships between Geographic Entities. *Appl. Sci.* **2024**, *14*, 4544. [[CrossRef](#)]
50. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Proceedings of the Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 3149–3157.

51. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition 2015. *arXiv* **2015**, arXiv:1512.03385.
52. Huang, H.; Chen, P.; Xu, X.; Liu, C.; Wang, J.; Liu, C.; Clinton, N.; Gong, P. Estimating Building Height in China from ALOS AW3D30. *ISPRS J. Photogramm. Remote Sens.* **2022**, *185*, 146–157. [[CrossRef](#)]
53. Horn, B.K.P. Hill Shading and the Reflectance Map. *Proc. IEEE* **1981**, *69*, 14–47. [[CrossRef](#)]
54. Alahmadi, M.; Atkinson, P.M.; Martin, D. A Comparison of Small-Area Population Estimation Techniques Using Built-Area and Height Data, Riyadh, Saudi Arabia. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1959–1969. [[CrossRef](#)]
55. Earth Observation Research Center, Japan Aerospace Exploration Agency (JAXA EORC). ALOS Global Digital Surface Model “ALOS World 3D—30m (AW3D30)”. 2016. Available online: [https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/aw3d30\\_e.htm](https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/aw3d30_e.htm) (accessed on 24 June 2022).
56. NASA JPL NASADEM Merged DEM Global 1 Arc Second V001. 2020. Available online: [https://lpdaac.usgs.gov/products/nasadem\\_hgtv001/](https://lpdaac.usgs.gov/products/nasadem_hgtv001/) (accessed on 6 June 2024).
57. Ridha, R.M.; Alwan, I.A.; Ismael, H.S. Accuracy Assessment of UAV Automated 3D City Model for Urban Planning. *AIP Conf. Proc.* **2023**, *2793*, 020004. [[CrossRef](#)]
58. Zhang, Z.H.; Liu, H. Research of Urban Digital Planning Model Based on GIS. *Appl. Mech. Mater.* **2014**, *543–547*, 4129–4132. [[CrossRef](#)]
59. Tadono, T.; Takaku, J.; Shimada, M. Global Digital Surface Model Generation by PRISM Onboard ALOS “Daichi” to Contribute Geo-Disaster Studies. *Int. J. Landslide Environ.* **2013**, *1*, 97–98.
60. Spasova, T.; Avetisyan, D. A Synchronized Remote Sensing Monitoring Approach in the Livingstone Island Region of Antarctica. In Proceedings of the Ninth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2023), Ayia Napa, Cyprus, 21 September 2023; Themistocleous, K., Michaelides, S., Hadjimitsis, D.G., Papadavid, G., Eds.; SPIE: Pune, Maharashtra; p. 63.
61. Gairabekov, I.G.; Hamzatov, A.I.; Mishieva, A.T.; Ibragimova, E.I.; Gairabekov, M.-B.I.; Gayrabekova, A.I. Development of a Digital Surface Model and a Digital Terrain Model Based on ERS Data. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *905*, 012025. [[CrossRef](#)]
62. Seto, T.; Furuhashi, T.; Uchiyama, Y. Role Of 3D City Model Data As Open Digital Commons: A Case Study Of Openness In Japan’s Digital Twin “Project Plateau”. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *48*, 201–208. [[CrossRef](#)]
63. Open geospatial. CityGML-3.0 Encodings, Encodings for the CityGML 3.0 Conceptual Model. Available online: <https://github.com/opengeospatial/CityGML-3.0Encodings> (accessed on 9 September 2024).
64. Statistics Bureau Population Census. Available online: <https://www.stat.go.jp/english/data/kokusei/index.html> (accessed on 28 September 2024).
65. Bivoltsis, A.; Trapp, G.; Knuiman, M.; Hooper, P.; Ambrosini, G.L. The Evolution of Local Food Environments within Established Neighbourhoods and New Developments in Perth, Western Australia. *Health Place* **2019**, *57*, 204–217. [[CrossRef](#)]
66. Zhao, Y.; Tang, X.; Liao, Z.; Liu, Y.; Liu, M.; Lin, J. Multi-Type Features Embedded Deep Learning Framework for Residential Building Prediction. *IJGI* **2023**, *12*, 356. [[CrossRef](#)]
67. Usui, H. Comparison of Precise and Approximated Building Height: Estimation from Number of Building Storeys and Spatial Variations in the Tokyo Metropolitan Region. *Environ. Plan. B Urban Anal. City Sci.* **2022**, *50*, 487–499. [[CrossRef](#)]
68. Bureau of Urban Development Tokyo Metropolitan Government Designation Policy and Standards Concerning Land-Use Specification. Available online: [https://www.toshiseibi.metro.tokyo.lg.jp/kanko/area\\_ree/](https://www.toshiseibi.metro.tokyo.lg.jp/kanko/area_ree/) (accessed on 4 October 2024).
69. Hermawan, Y.A.; Warlina, L.; Mohd, M. GIS-Based Urban Village Regional Fire Risk Assessment and Mapping. *INJIISCOM* **2021**, *2*, 31–43. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.