Estimating Urban Heat Island Effect through Building Façade System and Form Detection on Street View and Aerial Images

Po-Chien Yu, Pin-Chun Cheng, Jacob J. Lin and Lufan Wang
National Taiwan University, Taiwan
d09521001@ntu.edu.tw

Abstract. Reducing urban energy consumption is critical to achieving the 2050 Net Zero Emission Goal. In the past fewer years, researchers and experts have focused on improving energy efficiency; however, not only does the building energy efficiency affect energy consumption, but the urban scale factors also play an essential role in it, like the urban heat island effect. To predict urban electricity consumption, many researchers use different urban planning parameters such as road width, green belt, building parameters such as building use type, and other public parameters for analysis, but the calculation and statistical steps require a lot of manpower, in addition, many researchers use some confidential parameters in the research such as single-user power data, which involve privacy issues and are difficult to obtain in different cities or countries, causing their results are limited in a specific area. Using computer vision and deep learning, this research can quickly and effectively identify and calculate 10 urban and building characteristic parameters such as urban green ratio, average road width, window ratio, and building height from aerial images. The above characteristic parameters are combined with the government's public information that can be obtained on the Internet. This research uses regression analysis to quickly predict urban power consumption, which can be used for analysis in different cities. In advance, it can provide an objective prediction result to government offices while making an urban (renewal) plan.

1. Introduction

Reducing urban energy consumption is critical to achieving the 2050 Net Zero Emission (International Energy Agency, 2021). Recent studies have significantly improved energy efficiency through simulation for interior spaces. However, on a larger scale, such as city blocks or districts, other factors, for example, building form and facade system, also play an essential role in the urban heat island effect (UHI) (Ma *et al.*, 2020). To estimate the UHI effect and electricity consumption, many researchers use urban planning parameters such as road width, green belt, building parameters such as building type, and other general metrics for analysis. Nevertheless, the prediction model is complicated to understand, and the recent deep learning model's reasoning is behind the black box. Moreover, many parameters are closed to the public, making most approaches hard to scale and limited to certain areas.

This research aims to develop a method that quickly and effectively detects and calculates urban and building parameters affecting energy consumption and the UHI effect. These parameters are combined with the public information obtained on the internet to simulate and predict energy consumption and the UHI effect. This research uses regression analysis to predict urban power consumption and further simulates the UHI, which can be used for analysis in different cities. Furthermore, it can provide an objective prediction result to government offices while making an urban (renewal) plan.

2. Related Research

UHI effect is critical to the urban energy consumption. Urban heat island (UHI) can have significant impacts on building energy consumption by increasing space cooling demand and decreasing space heating demand. However, the impacts of UHI on building energy consumption were understudied due to challenges associated with quantifying UHI-induced temperature change and evaluating building energy consumption. The correlations between the urban island effect and distinct urban attributes exist. The heating load indexes of the buildings located in downtown are 1.5-5% less than those in the suburbs. (Li *et al.*, 2019) As the heat island intensity raises by 1℃, the average heating energy consumption will decrease 5.04%.(Zhou *et al.*, 2017) Urban heat island effect increases energy consumption of buildings by increasing space cooling demand and decreasing space heating demand. Urban heat island effect can be evaluated by using remote sensing data to measure surface temperatures and impervious surface area, or using a holistic evaluation methodology that takes into account factors such as land use, vegetation cover, and building density (Deilami, Kamruzzaman and Liu, 2018). Other study also takes the floor area ratio into consideration (Zhou *et al.*, 2017).

Data-driven energy analysis can work on urban energy analysis and prediction. Many energy analysis studies started by modeling and characterizing the building performance through developing simulation-based and data-driven methods. In recent years, simulation-based techniques usually use commercial software such as EnergyPlus (Crawley *et al.*, 2001), eQUEST, etc. with manual adjustment and input of building energy model data, like material type, air condition, amount of lighting, to quantify and predict building energy performance from various energy conservation measures (ECMs). Urban and building energy simulation required plenty effort on variables collection (Chen, Han and de Vries, 2020). Some studies pointed out the task of creating a reliable building energy model of a new or existing neighborhood can be broken into the following subtasks: simulation input organization (data input), thermal model generation and execution (thermal modeling) as well as result validation (validation) (Reinhart and Cerezo Davila, 2016). On the other hand, data-driven energy analysis is a method of analyzing energy data using statistical and machine learning techniques to identify patterns and trends in energy consumption. It is used to optimize energy use and reduce costs by identifying areas where energy can be saved. It is also used to predict future energy consumption and to identify areas where new sources of energy can be developed (Amasyali and El-Gohary, 2018; Lin, Zhang and Zuo, 2022).

Street view and aerial images can help to extract building features. In recent years, due to the rapid development of computer vision technology, related studies have been tried to obtain the characteristics of residential buildings through the use of high -altitude photos, and used for energy analysis. A study use the overhead image and regression model to estimating resdential building energy consumption (Streltsov *et al.*, 2020). Also, some researchers us aerial and street view image to predict residential building (Rosenfelder *et al.*, 2021). These studies use street view and aerial images to classify building type (residential or commercial), and extract features (e.g. area, perimeter, building density, pool size, number of stories, vegetation) to regress and predict the energy usage. However, these studies mainly are all focused on individual residential building's energy prediction without considering their surroundings, or choose countryside or suburban as their experiment field. One report suggests that building energy consumption is not only related to the systems and people in it, but also interact with exterior systems like the street, tree, other building ('using satellite images, scholars develop a model to quantify buildings' energy use.pdf', no date). this research will use street view and serial images to extract both building and urban features, and use regression method to analyze and predict energy consumption.

3. Methodology

This study uses computer vision technology to analyze street view and aerial images to establish relationships between them. A dataset is created from Manhattan Island in New York City, USA, which contains a total of 39 blocks with 2D building elevation views reconstructed in this study.

This study presents a framework for automated prediction of electricity consumption from city-scale aerial images and 3D models. The work of establishing this framework in this research could be divided into four phases: data collection, deep learning-based image segmentation model, segmentation data post-processing, and data-driven regression model. Figure 1 demonstrates the flowchart of the whole procedure of this framework.

In the first phase of this research, we sought data that could be quickly accessed and had enough characteristics to represent a city's energy consumption. To predict electricity consumption, we chose aerial images and 3D building models as the beginning data. Aerial images were mainly used to identify the city's composition so that we could know the surrounding situation of each building. On the other hand, 3D building models were more likely to present visual features of buildings themselves. We specifically extracted orthographic façade images from 3D models using an automated procedure. We chose these two sources of data because they both obtain a large amount of potential information to develop and their massive scale is very easy to obtain due to advanced development of photogrammetry technology nowadays.

The second phase of this research involves building a deep learning-based image segmentation model. As shown in the flowchart, we developed two different semantic segmentation models to deal with our data from two sources: aerial images and orthographic images extracted from 3D building models. The main reason for using deep learning models here is to automate the process of identifying the features we want from our image data. Aerial images are used to identify the composition of the city (around the buildings), so we use a semantic segmentation model to identify buildings, roads, and green areas in a large scope from an aerial view.

In this research, we used a semantic segmentation model to identify every window on the building façade from orthographic façade images extracted from 3D building models. The façade image is about the appearance of the building, and we are especially interested in the window rate of the building.

In the third phase of this research, segmentation data post-processing played a character of bridge to connect the segmentation models to the regression model ultimately. Overall, we did some summarization and extraction to segmentation data generated from two deep learning models, then created a list of critical parameters we needed to build the regression model.

In the fourth phase, we built a data-driven regression model to predict our target: electricity consumption. Combining parameters of building and its surrounding, like building area, green ratio, road with …and so on, then we can do regression analysis with the energy consumption data.

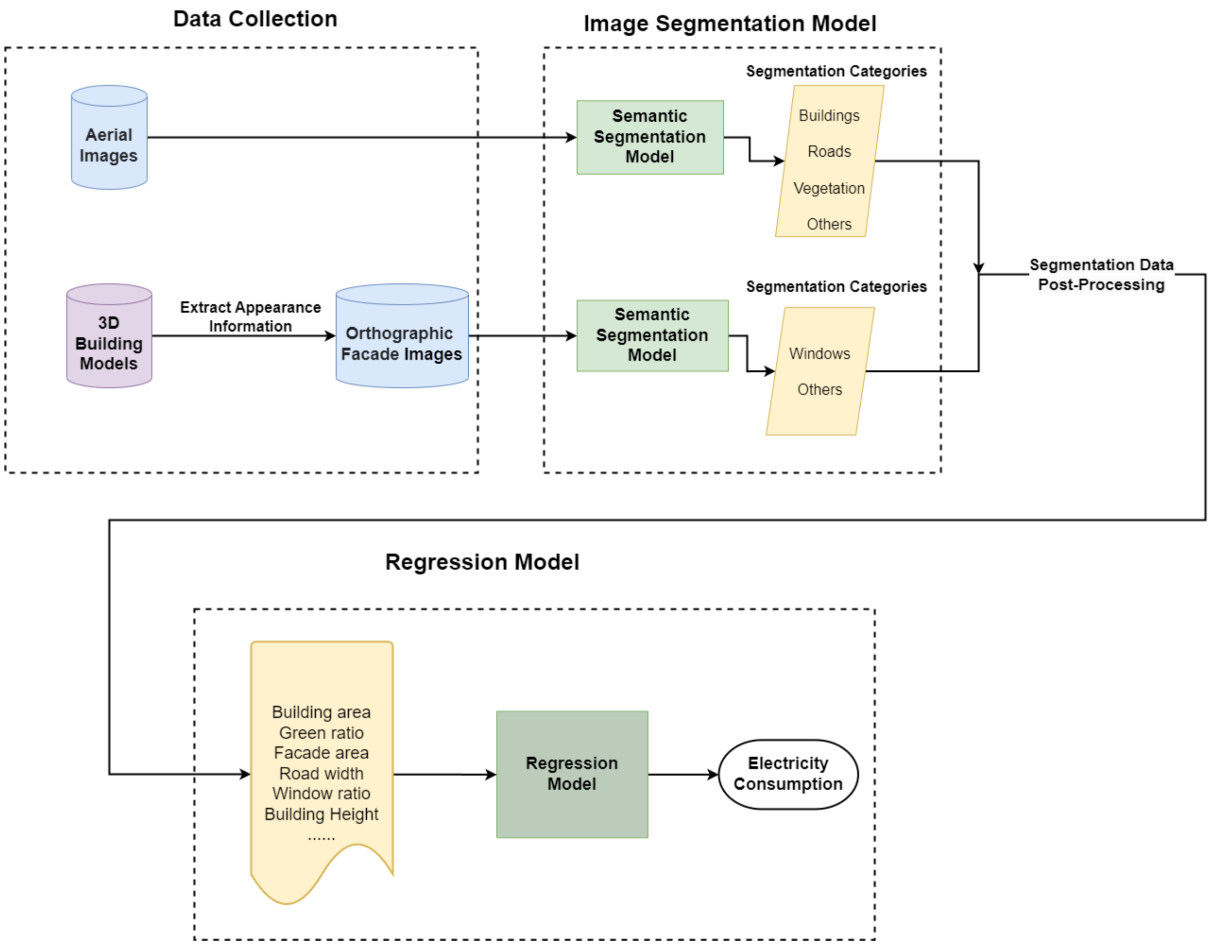

Figure 1. Flowchart of working framework.

3.1 Data Collection

Aerial Image. In this study, we selected Manhattan's financial district as the demonstration area, which is the core area of the city and mainly contains residential, office, skyscraper, landmarks, and so on. First, we used orthographic aerial images as one of the main data sources. The orthographic aerial images were generally generated from camera drones over the city and finished after orthorectification. This kind of image data has a great advantage in resolution compared with the satellite aerial images usually used before. The data used in this study were obtained from open-source maps on the Internet. The raw image could be downloaded by 19200 x 19200 resolution per image in PNG format. We captured the area of Manhattan's financial district, with a range about 1 kilometer by 1 kilometer, and the resolution is 0.011 square meter per pixel. The aerial images from many different open-source maps are convenient for anyone to access and utilize quickly. However, unlike satellite images, aerial images cannot be generated once in such a big scope. Therefore, the map would be divided into many parts with different time and light conditions, which make an obvious sense of stitching in the overall view. Figure 2a and Figure 2b show the approximate scope and demonstrate the level of resolution of the data.

3D Building Model. The 3D building model used here is to collect all the potential parameters that can represent energy consumption in building appearance information. The 3D building appearance model of most cities is also accessible in many online maps now. We use Blend API to transform the models into 3D mesh in Blender (Blender Development Team, 2022), then split them into individual buildings. Since we want to use image recognition to extract characteristics from the model, we developed an automated script in Blender to generate four-direction orthographic façade images and top-view images for subsequent usage. The resolution of each image generated here adopts 2160 x 2160 pixels to ensure the minimization of the loss from conversion. Figure 3a shows the view of the 3D mesh model in Blender, and Figure 3b is one generated orthographic façade image.

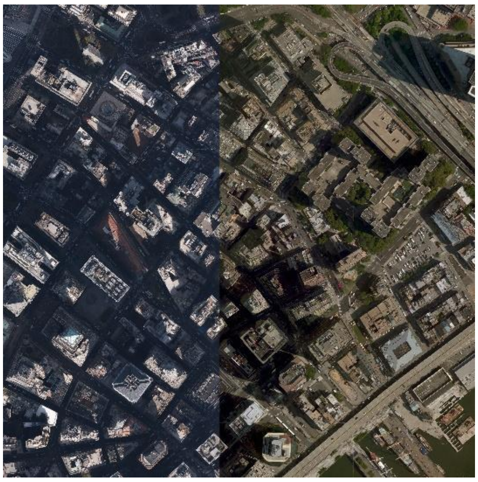


Figure 2a. Aerial image of specific area and Figure 2b. Larger scale aerial image


Figure 3a. Building Mesh Model and Figure 3b. Orthographic façade

3.2 Deep Learning.

Aerial Image Segmentation. We built a deep learning-based semantic segmentation model to extract characteristics by recognizing city composition through aerial images. For this model, we used the structure of MMSegmenation (MMSegmentation Contributors, 2020), which is based on PyTorch (Zhu *et al.*, 2019) and is a semantic segmentation library of OpenMMLab (Contributors, M. M. C. V., 2018). Pre-trained weights and configs of PSPNet (Zhao *et al.*, 2017) trained with iSAID dataset (Zamir *et al.*, 2019) were utilized in this model for initial setup.

To train the model, many famous aerial image datasets were tried to use in the beginning, but they didn't show good enough performances for most datasets consisting of satellite images. Some of the other datasets closer to our data miss the class, vegetation, we are particularly interested in. Ultimately, the model was trained with data we annotated on the same map aerial image ourselves.

The training data was annotated in the online annotation tool, Supervisely, and segmented into four classes: Building, street, vegetation, and all others. Each image annotated here has a resolution of 2400 x 2400 pixels, and a total of 64 images were annotated.

About the training result, Figure 1 presents the evaluations of each class and the whole model. Using IoU and accuracy for classes separately; Overall accuracy, mean IoU, and mean accuracy were adopted for evaluating overall model. Figure 4 demonstrates the prediction of trained segmentation model.

Based on the three main classes extracted, we can generate the parameters we need in later regression models.

Table 1. Evaluation indicators of aerial image model

Evaluation indicators					
Indicators for classes			Indicators for whole model		
class	IoU	Acc	aAcc	mIoU	mAcc
Other	75.35	89.32	87.26	74.43	83.98
Road	76.22	85.03			
Vegetation	61.49	72.03			
Building	84.67	89.54			

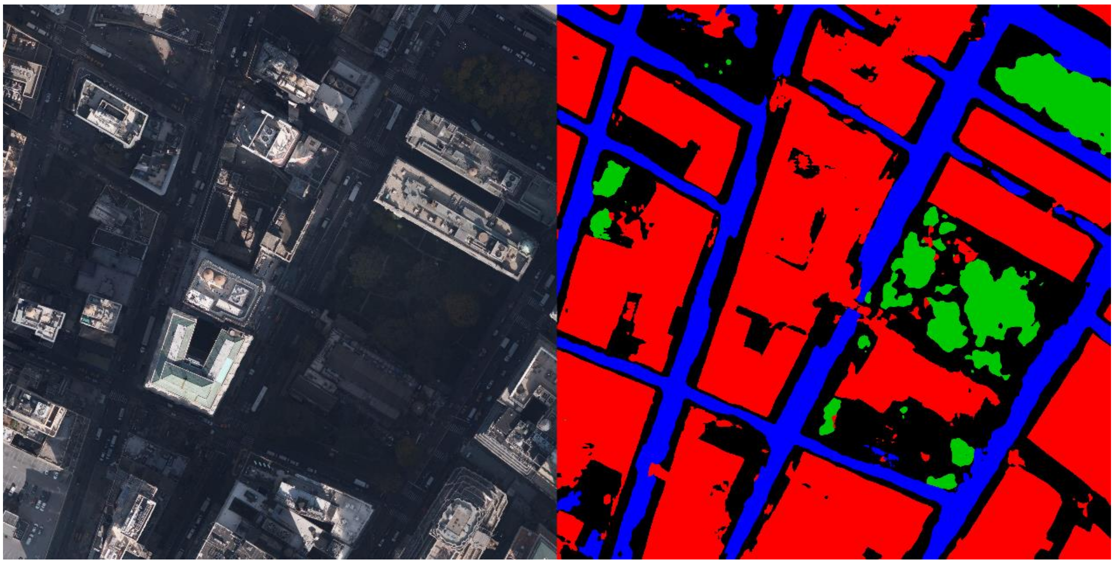

Figure 4. Segmentation and original aerial image

Orthographic Façade Image Segmentation. Previously, an orthographic façade image was generated from a 3D mesh model. We are concerned with the energy performance of building façades and specifically with the window ratio. Similar to the aerial image model, we used the same structure of semantic segmentation model to calculate windows area. A Mask R-CNN (He *et al*., 2020) model was tried in the early phase to recognize windows, but in the end, a semantic segmentation model did a more precise job in calculating windows area.

Pre-trained weights and configs of PSPNet (Zhu *et al*., 2019) trained with Cityscapes dataset (Cordts *et al*., 2016) were utilized in this model for initial setup. The training dataset was also annotated by ourselves. Most previous façade image datasets are limited to residential houses, but there are skyscrapers and glassy towers in Manhattan where glass curtain walls are not recognized well by models trained on those datasets. In the end, our training dataset consisted of 156 façade images annotating windows.

Since other characteristics would be obtained using computer vision methods, windows are the only class we focus on in the segmentation model. Table 2 shows evaluations of each class and the whole model using IoU and accuracy; overall accuracy, mean IoU and mean accuracy were also used for evaluation here. Figure 5 is one of the prediction results.

So far, we have built a window recognizing model for façade images and a city composition segmentation model for aerial images.

Table 2. Evaluation indicators of façade image model.

Evaluation indicators					
Indicators for classes			Indicators for whole model		
class	IoU	Acc	aAcc	mIoU	mAcc
Background	93.64	95.47	94.22	77.55	89.78
Window	61.47	84.09			

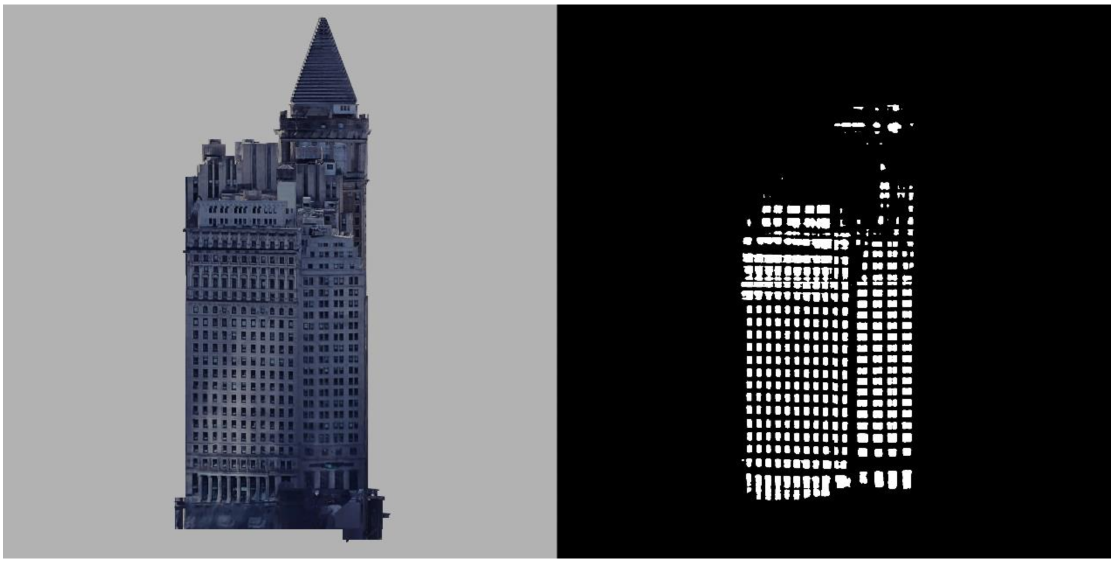

Figure 5. Predicting and original façade image

3.3 Segmentation Data Post-processing and Regression Analysis

For the later step, using NYC energy performance map data to build a regression model, the demonstration area has been divided into 19 subdivisions, each with several buildings. The results of two segmentation models are applied here for the computational transformation into parameters. Building ratio, green ratio, road ratio, window ratio …and so on, simple computer vision algorithms are used to complete the generation of parameters.

After obtaining the characteristics of urban buildings by the method mentioned above, this study also obtained the city's public electricity consumption (EUI, energy use intensity) and building types, and other public information from the Internet in regression analysis.

In the regression model, we set EUI as a dependent variable, windows ratio as an independent variable, and control variables include gross area, green ratio, and road ratio. We expect the result to support the following hypothesis: EUI increases while the window ratio increases. The results are Table 3.

Table 3. Regression result (R-squared is 0.05)

Log(EUI)=C(1)+C(2)*Window ratio+C(3)*Log(gross area)+C(4)*(green ratio + road ratio)				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	5.5496	0.5692	9.7491	0
C(2)	0.0081	0.0054	1.4903	0.1413
C(3)	-0.0505	0.0481	-1.0514	0.2972
C(4)	0.0059	0.0085	0.6922	0.4915

The return results show that the current parameters are less than 20% of the EUI interpretation rate (R-Squared). The variable part shows that the rise of the window ratio will increase EUI, but its impact is not significant. In addition, past research believes that the increase in road width and green space area helps alleviate UHI, but the return result positively affects EUI.

In the 19 regions divided in this study, there are a total of 65 buildings, classified as 27 residential buildings, 29 office buildings, 8 hotels, and 1 school according to building type. We hope to understand the energy behavior of different types of buildings and their relationship with the aforementioned parameters. The regression results are as Table 4 and Table 5.

Table 4. Residential building regression result(R-squared is 0.11)

Log(EUI)=C(1)+C(2)*Window ratio+C(3)*Log(gross area)+C(4)*(green ratio + road ratio)				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	4.8042	0.651	7.3802	0
C(2)	-0.014	0.0083	-1.6843	0.1056
C(3)	0.0329	0.0596	0.5517	0.5865
C(4)	0.0073	0.0118	0.6225	0.5397

Table 5. Office building regression result(R-squared is 0.12)

Log(EUI)=C(1)+C(2)*Window ratio+C(3)*Log(gross area)+C(4)*(green ratio + road ratio)				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	5.542	1.0387	5.3356	0
C(2)	0.0115	0.0065	1.7743	0.0882
C(3)	-0.0517	0.0799	-0.6466	0.5238
C(4)	0.0037	0.012	0.3103	0.7589

In the analysis of residential buildings, we can find that when the window ratio increases, it will help reduce the use of power. On the contrary, the window ratio will increase power consumption at office buildings. The result of the two shows considerable prominences.

4. Conclusions and Future Work

The thesis aims to explore the possibility of using publicly available information on the internet to assist in capturing the external features of urban buildings by computer vision method and to use regression analysis to explore the possibility of using them for urban energy prediction. Through our research, we have completed the identification of building features (window ratio, green ratio, road ratio) for 65 buildings in Manhattan, New York City, using street view and aerial images and combined with publicly available building energy consumption data (energy use intensity, building using type, gross area), we have explored the possibility of energy prediction. The current regression model accuracy is still low, but it does show that window ratio has different effects on energy use for residential and office buildings.

This study also found that government-publicized network information was less than expected. For example, in terms of the window ratio identified in this study, it is difficult to obtain credible public information on the Internet, making it difficult to verify its correctness as a result. At the same time, EUI data is considered confidential information in many countries or regions, making obtaining correct and credible data in public areas difficult.

Future research will focus on four directions:

1. Improve aerial and orthographic façade image segmentation, and recognize the building height for further regression.
2. Continue to find more relevant public information can be sought to validate the research's identification results, increase data reliability.
3. Expand the research scope and increase sample size to improve the reliability of the regression model.
4. Will continue to seek public information from other countries, and provide suggestions for urban planning by comparing energy consumption between cities in different climatic regions.

References

Amasyali, K. and El-Gohary, N.M. (2018) 'A review of data-driven building energy consumption prediction studies', Renewable and Sustainable Energy Reviews. Elsevier Ltd, pp. 1192–1205. Available at: https://doi.org/10.1016/j.rser.2017.04.095.

Chen, H.C., Han, Q. and de Vries, B. (2020) 'Urban morphology indicator analyzes for urban energy modeling', Sustainable Cities and Society, 52. Available at: https://doi.org/10.1016/j.scs.2019.101863.

Cordts, M. et al. (2016) 'The Cityscapes Dataset for Semantic Urban Scene Understanding', Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem, pp. 3213–3223. Available at: https://doi.org/10.1109/CVPR.2016.350.

Crawley, D.B. et al. (2001) 'EnergyPlus: Creating a new-generation building energy simulation program', Energy and Buildings, 33(4), pp. 319–331. Available at: https://doi.org/10.1016/S0378-7788(00)00114-6.

Deilami, K., Kamruzzaman, M. and Liu, Y. (2018) 'Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures', International Journal of Applied Earth Observation and Geoinformation. Elsevier B.V., pp. 30–42. Available at: https://doi.org/10.1016/j.jag.2017.12.009.

He, K. et al. (2020) 'Mask R-CNN', IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(2), pp. 386–397. Available at: https://doi.org/10.1109/TPAMI.2018.2844175.

International Energy Agency (2021) Net Zero by 2050 - A Roadmap for the Global Energy Sector. Available at: www.iea.org/t&c/.

Lin, P., Zhang, L. and Zuo, J. (2022) 'Data-driven prediction of building energy consumption using an adaptive multi-model fusion approach', Applied Soft Computing, 129. Available at: https://doi.org/10.1016/j.asoc.2022.109616.

Li, X. et al. (2019) 'Urban heat island impacts on building energy consumption: A review of approaches and findings', Energy, 174, pp. 407–419. Available at: https://doi.org/10.1016/j.energy.2019.02.183.

Ma, R. et al. (2020) 'Modeling urban energy dynamics under clustered urban heat island effect with local-weather extended distributed adjacency blocks', Sustainable Cities and Society, 56(November 2019), p. 102099. Available at: https://doi.org/10.1016/j.scs.2020.102099.

Reinhart, C.F. and Cerezo Davila, C. (2016) 'Urban building energy modeling - A review of a nascent field', Building and Environment, 97, pp. 196–202. Available at: https://doi.org/10.1016/j.buildenv.2015.12.001.

Rosenfelder, M. et al. (2021) 'Predicting residential electricity consumption using aerial and street view images', Applied Energy, 301. Available at: https://doi.org/10.1016/j.apenergy.2021.117407.

Streltsov, A. et al. (2020) 'Estimating residential building energy consumption using overhead imagery', Applied Energy, 280. Available at: https://doi.org/10.1016/j.apenergy.2020.116018.

'using satellite images, scholars develop a model to quantify buildings' energy use.pdf' (no date).

Blender: Blender Development Team. (2022) Free and open source 3D creation suite. [Computer software]. https://www.blender.org/

MMSegmentation: MMSegmentation Contributors.(2020) Openmmlab semantic segmentation toolbox and benchmark. https://github.com/open-mmlab/mmsegmentation

Zamir, S.W. et al. (2019) 'iSAID: A Large-scale Dataset for Instance Segmentation in Aerial Images'. Available at: http://arxiv.org/abs/1905.12886.

Contributors, M. M. C. V. (2018). MMCV: OpenMMLab computer vision foundation.

Zhao, H. et al. (2017) 'Pyramid scene parsing network', Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua, pp. 6230–6239. Available at: https://doi.org/10.1109/CVPR.2017.660.

Zhou, Y. et al. (2017) 'Urban morphology on heat island and building energy consumption', Procedia Engineering, 205, pp. 2401–2406. Available at: https://doi.org/10.1016/j.proeng.2017.09.862.

Zhu, Z. et al. (2019) 'Asymmetric non-local neural networks for semantic segmentation', Proceedings of the IEEE International Conference on Computer Vision, 2019-Octob, pp. 593–602. Available at: https://doi.org/10.1109/ICCV.2019.00068.

NYC Office of Climate and Environmental Justice. NYC Energy & Water Performance Map. https://benchmarking.cityofnewyork.us/#/