

# A training dataset for machine learning-based prediction of window opening position in a naturally ventilated building

Jeremy Wong<sup>\*1</sup>, Julian Donges<sup>2</sup>, Andrea Gasparella<sup>2</sup>, and Adam Rysanek<sup>1,3</sup>

<sup>1</sup>Department of Mechanical Engineering, University of British Columbia, Vancouver, Canada

<sup>2</sup>Faculty of Science and Technology, Free University of Bozen-Bolzano, Bolzano, Italy

<sup>3</sup>School of Architecture and Land Architecture, University of British Columbia, Vancouver, Canada

**Abstract.** Window operation is the main strategy used by building occupants to naturally ventilate buildings. However, common approaches to measure window operation for energy and comfort assessments are still technically complex or insufficient; typical window open/close sensors often provide only binary information about the opening state of a window, not the extent to which the window is open. This paper is the first outcome of a research project that seeks to use photo imagery and machine learning to predict the variable opening state of windows on a real multi-family residential passive house located in Vancouver, Canada. The employed windows are European-style in that they can be opened in tilt or turn mode. To eventually train the algorithm, a ground-truth dataset is constructed by manually changing the opening state of sixteen windows every minute over a 15-hour test period spanning three days and taking a photo of the windows at each instance, measuring the angle each time. This paper documents the first outcome of the overall project: the publication of the training dataset itself, with over 10,000+ images of a building façade taken, under variable-but-known window opening state, and under various light conditions. The paper presents the testing methodology undertaken for generation of the dataset and provides instructions for how to access the dataset. In the future, these images will be used to calibrate a machine learning model to estimate window opening/closing state of the tested building. The dataset can also be extended for semantic segmentation in support of other machine learning problems.

## 1 Introduction

As society becomes more reliant on energy-heavy heating and cooling methods, one age-old solution has potential to address the concern. In both residential and commercial sectors, the effect of natural ventilation on building performance, inclusive of energy efficiency and indoor air quality (IAQ), cannot be understated.

As technology has advanced, approaches to modelling have shifted from first-principles physics-based techniques to machine learning black-box methods. Validating energy and air quality models have increasingly relied on real-world empirical data. While building management systems do track and record data, publicly accessible data in this domain is far and few. Without it, the performance gap between simulation/modelling and the real world will never be closed[1].

This work seeks to create a publicly available dataset that will foster further research into energy and IAQ modelling, specifically in regards to prediction of window opening/closing cycles in real-world conditions.

## 2 Background

### 2.1 Natural Ventilation and Indoor Air Quality

It is well established that natural ventilation can play a major role on affecting people's comfort, health, wellness in buildings and their overall environmental performance[2]. Conversely, it is understood that buildings with low indoor ventilation rates and unmitigated latent heat gains can have negative impacts on human health[3], including decreased school performance amongst children[4].

Upon the onset of the the COVID-19 pandemic, discussions on the importance of natural ventilation rate took on a new meaning, particularly in regards to the effect of natural ventilation on disease transmission rates in buildings[5, 6]. Though the pandemic has waned, the discussion around mechanical and natural ventilation and disease transmission continues[7, 8]

\*Corresponding Author: [jeremwong@outlook.com](mailto:jeremwong@outlook.com)



Figure 1: Footprint of the building used in the experiment. Situated at the corner of Wesbrook Mall and Gray Avenue in Vancouver, Canada, the building is a multi-unit residential low-rise. Photos were taken from the south side facing north, indicated by the camera and light red shading. Windows are marked with four-point stars.

## 2.2 Energy modelling

Alongside air quality management, natural ventilation has long-been considered an important feature of energy efficient building design. Peng et al.'s review of energy usage of traditional HVAC systems combined with natural ventilation found that on average energy savings of 28% could be achieved in buildings featuring mixed-mode delivery ventilation systems versus 100% mechanically-ventilated systems[9].

During the COVID-19 pandemic, calls to improve the indoor air quality in buildings via higher ventilation rates were not exempt from concerns about the energy impact of such decisions. Aviv et al.'s work on combining radiant heating/cooling systems with natural ventilation demonstrated the viability of mixed-mode building control to simultaneously improve building ventilation rates and decrease building energy use[10].

## 2.3 Measuring occupant use of natural ventilation systems

In the prediction of building performance in buildings that feature natural ventilation systems, the uncertainty of how occupants choose to open and close windows (or similar systems) is known to be high[11, 12]. In multi-unit residential buildings, which is the context of this work, residents can encounter different thermal regimes in suites

across an entire building. One façade may be warmer than another due to solar loads, one floor cooler than another due to buoyancy, and seasonal variations can affect window opening behaviour altogether[13]. In addition, each occupant in a building may have a unique behavioural profile that is difficult to ascertain with surveys or traditional data-gathering methods. Some occupants may be more adaptable to warmer or cooler indoor air temperatures, some may be more active than others in using passive systems to control the indoor environment, some may be more or less informed about the principles of building physics[14].

Measuring the specific behaviour of occupants with respect to natural ventilation is found to rest on the capability to measure the operation of windows at time-scales appropriate for building performance evaluation. Magnetic sensors, connected to a central data acquisition system, can be used to measure the open/close status of individual windows in a real-time manner[15]. Another method, growing in familiarity, is the use of machine learning—specially computer vision models—to predict window open position directly from photo imagery.

## 2.4 Machine vision-based prediction of window open/close position

Extracting identifiable architectural features from images, such as windows, was first shown in work by Debevec et al. which generated 3D models from 2D images using depth maps and stereo images[16]. Similar work involving ground-based images utilized mutual information to detect repetitions in the façade to build high quality 3D models[17].

In subsequent years, algorithms for predicting façade features have become more intense and reliant on machine learning. Work in 2010 by Wendel et al. used scale invariant feature transform with Harris corners to extract repetitive façade patterns[18]. Martinovic et al. used recurrent neural networks in his three-layer approach to façade parsing[19]. Most recently, Sun and Zhai used the computer-vision model YOLO\_v3 for identifying and segmenting open windows on unobstructed façades[20].

In all of these works, and future works that may utilize these established models, calibration is important, yet there exists no uniformity in training data structure[21]. In particular, to date, there is no known standard dataset that can be used to calibrate computer vision-based algorithms that predict window open/close position.

While several computer vision datasets exist for general applications of computer vision algorithms, datasets specific to the problem of window open/close prediction do not exist. Imagenet, one of the largest open-source datasets, is composed of many classes of everyday items but exclude an explicit window class[22]. Microsoft's COCO dataset suffers the same shortfall[23]. Finally, both CIFAR-10 and CIFAR-100[24] contain 10 and 100 classes, respectively, none of which are windows.

Collectively, these dataset labels best serve other types of machine learning tasks. Yet for estimation of natural

ventilation performance, data labels need to identify windows and classify window states, i.e., open, closed, and ideally angle measurement. For this reason, Sun et al. constructed their own dataset by taking over 10000 photos, which at time of writing, is not publicly available[20].



Figure 2: Photo of camera setup during the experiment. To minimize movements between photos, the camera was mounted on a tripod and shielded from the direct sunlight to prevent overheating. The operator manually pressed the camera trigger for each photo.

## 2.5 Scope of paper

This paper is the first of a series of studies to evaluate machine vision-based approaches to predict window open/close behaviour in buildings. The objective of this work has been to generate and disseminate a standard training dataset for general machine vision solutions. In this paper, the methodology for generating the standard dataset is presented, the format and data architecture of the dataset is described, and information about general access to the dataset is provided.

## 3 Methods

### 3.1 Case study building

A case study building was provided for this work. The BCR8 "Evolve" buildings is a six-story, 105-unit multi-unit residential building located in Vancouver, British Columbia Canada. The building is Passive House-certified and was constructed between 2019 and 2022.

Access to the building was provided to this study's re-

searchers prior to tenancy of the building. Researchers were given temporary access to all suites of the building to operate building systems and take imagery. A site plan for the building is provided in figure 1.

### 3.2 Camera and camera location

Photos were taken from a camera located on the balcony of an east-facing suite located on the 5th-floor of the building. The suite was located at the "South" wing of the building shown in Fig. 1. The camera was pointed to face the opposing exposed "East" wing of the building, identified by the red highlighting of the camera's view in Fig. 1. Images taken by the camera could capture the view of sixteen manually-operated windows.

The windows installed in the building are Innotech Defender 88PH+ Pro with European-style tilt and turn capability. Each window features a custom-built movable exterior shade, but the shade was open during all experiments to ensure views from the windows were unobstructed. The camera used was an Olympus E-M1 MarkII. The lens was an Olympus M.14-42mm F3.5-5.6. It captured 20 megapixel images with a resolution of 3800 by 5100 pixels. Fig. 2 shows the camera setup; for stability, it was mounted on a tripod and placed in the shade to prevent direct sunlight exposure. Photos were taken over the course of three days in July 2022. Lighting conditions were sunny or slightly overcast. Camera settings were adjusted to capture 5 frames with varying exposure times (5F 2.0EV) using a technique called bracketing.

### 3.3 Experimental process

The windows were operated by several volunteers over a total of 15 hours spanning the experimental days. Each window was pre-configured to be openable to one of six discrete positions: closed, tilted, turned 5°, 15°, 45°, and 90°. Images were taken during repeated opening / closing cycles. Every two minutes, the window position of all windows captured by the camera would be changed. Each volunteer was allotted the first minute to change the positions of windows assigned to them. Subsequently the camera operator would capture multiple images of the façade over a 50 second period. The remaining 10 seconds served as a grace period between cycles.

The camera operator also managed the encoding of timestamps into captured images and ensuring that the time log of window opening/closing cycles was identically logged.

The purpose of capturing repeated cycles of photos for the same building and façade was to ensure a rich dataset could be captured under varying solar angles, ambient light levels, and resulting shadows and/or artificial light pollution.

### 3.4 High Dynamic Range image reconstruction

We merge the five differently-exposed, or bracketed, frames to create a single High Dynamic Range (HDR)



Figure 3: Five bracketed images and reconstructed HDR image (right). The left image shows the original exposure, followed by four increasing exposure times. The reconstructed image shows much better tonal information in shadows and washout regions compared to the original bracketed images.

image. In the direct sunlight, each photo is expected to have a high range of brightnesses leading to over- and under-exposed regions. Reconstructing HDR images from bracketed exposures retain tonal information from these areas[25]. This process is best suited for static or stationary photo scenes, under high- and low-light conditions.

Although end-users of cameras and smart phones may be accustomed to "HDR" options when taking photos using these devices, available methodologies to generate HDR images do actually vary. The first HDR image was made by Gustave Le Gray in 1857[26], which was made by combining negatives of the same photo taken with different exposures. Since then, many digital reconstruction techniques have been developed, most notably those of Robertson[27] and Debevec[28]. Mertens reconstruction technique is a more recent development[29]. The algorithm weights pixels based on contrast, saturation and exposure before smoothing with Gaussian and Laplacian pyramids.

In this work, we use OpenCV-Python to merge the bracketed images. For HDR reconstruction, we use the Mertens algorithm. After merging the frames, we crop the HDR images into individual window segments.

### 3.5 Segmenting and metadata tagging

Following HDR reconstruction, we segment each of the images into subimages focusing on individual windows. This follows the convention set by CIFAR, ImageNet, and COCO, where smaller images of individual objects are used in classification training tasks[24, 22, 23].

Rather than provide a separate file to notate each of the images, we include the label for each segmented image in the filename. Each follows the scheme of `<date>_ID_<label>_WIN_<window#>`, guaranteeing a unique filename for each image while providing important contextual information.

Since small variations of the camera positions were observed between experiments, we include the `<date>` as a prefix for optimal control of the dataset. The ID substring is the original filename assigned to the photo by the camera.

The `<label>` substring of the filename is one of the following classes: `Close`, `Tilt`, `5`, `15`, `45`, or `Turn`, as

referred to in Sec. 3.3. While the ultimate goal is to estimate continuous window position to extract ventilation rate, the experiment complexity and number of photos needed grows exponentially. In contrast, we have fewer classes with far more examples with greater variance.

The combination of the two substrings `WIN` and `<window#>` of the filename indicates which window the image is focused on, ordered from left to right, top to bottom. This allows subsets of windows to be chosen for user flexibility.

## 4 Results and discussion

During the experimental days, a total of 6500 raw images were captured of 6 discrete window positions. The time of day of all captured window operating cycles ranged from 9am to 9pm, but were pro-dominantly captured during the hours of 11 and 6pm. These hours coincided with peak solar loading on the captured façade. A total of 1300 unsegmented HDR images were produced and stored in the presented dataset, one for each observed window position.

### 4.1 Full HDR images

An example of five bracketed images and resulting Mertens-reconstructed HDR image for one instance are shown in Fig. 3. We see that the reconstructed image retains better contrast, saturation and exposure compared to the differently-exposed images.

Although the windows are visible in the bracketed images, the HDR image shows more resilience to shadows and reflections. Since the windows are surrounded by white shading forms, the frames in the original images are often under-exposed in some of the raw images. The HDR image extracts this information from otherwise over-exposed images, indicating promising performance for future detection problems.

While one could construct the dataset consisting of bracketed images, it is unlikely to be useful. The HDR reconstructed image preserves the valuable information from each exposure. This reduction of input size can reduce training times in machine learning applications.



Figure 4: Sample segmented HDR images showing windows under different lighting conditions at different angles.

## 4.2 Segmented HDR images

The segmented images are shown in Fig. 4. We see a sample of different windows at various angles under unique lighting conditions; these varying conditions across the dataset can aid generalization for supervised learning tasks. Since each window is at a different location, the camera perspective differs slightly between images. For detection tasks, these variations can improve the training process.

The reflections in the glazing introduce more variation in the dataset. Windows located on the fifth floor and higher show a uniform reflection of the sky. For lower floors, the reflections contain features from the surrounding area; foliage, buildings, and other structures from the built environment are all visible.

Figure 4 also shows how each window cropping differs in size. Since the camera was placed at an oblique angle to the facade, windows on the lower floors and further from the camera appear much smaller and less square. While this can be detrimental in some machine learning pipelines, it is easy to resize the images if necessary.

## 4.3 Data accessibility

At time of publication, there is no automated way of retrieving the dataset we have created. The full, labelled dataset can be accessed by contacting the authors.

## 5 Conclusion

In this work, a publicly-available dataset is described, consisting of HDR images of windows with labels describing window state built into the filename of each image. The dataset can be used in any supervised learning scenario where knowing window position is necessary. This is most likely to be used in the in the training and testing of machine vision approaches for window open/close prediction, and subsequently the calibration of IAQ and building energy modeling tools that derive predictive models from real-world monitoring. Such a step will form a follow-up work to this paper.

Some known limitations of the presented dataset include slight variations in image sizes between experimental days. Since the camera had to be removed between data collection sessions, the camera position was slightly changed. To account for this, image registration techniques could be used.

Another limitation of this dataset is the lack of variance in window model. Since all of the windows were the same, those wanting to train on different styles of windows will need to create a different dataset. Additionally, more data from winter months may be necessary for production-grade deployments.

## References

- [1] P. de Wilde. *Automation in Construction* **41** (2014)
- [2] K. W. Tham. *Energy and Buildings* **130** (2016)
- [3] J. Sundell. *Indoor air* **14** (2004)
- [4] M. J. Mendell, G. A. Heath. *Indoor air* **15** (2005)
- [5] Y. Li, G. M. Leung, J. Tang, X. Yang, C. Chao, J. Z. Lin, J. Lu, P. V. Nielsen, J. Niu, H. Qian, et al. *Indoor air* **17** (2007)
- [6] C. Sun, Z. Zhai. *Sustainable Cities and Society* **62** (2020)
- [7] J. Ding, C. W. Yu, S.-J. Cao. *Indoor and Built Environment* **29** (2020)
- [8] Z. Pang, X. Lu, Z. O'Neill. *Building Simulation* (2022)
- [9] Y. Peng, Y. Lei, Z. D. Tekler, N. Antanuri, S.-K. Lau, A. Chong. *Energy and Buildings* **276** (2022)
- [10] D. Aviv, K. W. Chen, E. Teitelbaum, D. Sheppard, J. Pantelic, A. Rysanek, F. Meggers. *Applied Energy* **292** (2021)
- [11] A. Roetzel, A. Tsangrassoulis, U. Dietrich, S. Busching. *Renewable and Sustainable Energy Reviews* **14** (2010)
- [12] Y. Zhang, P. Barrett. *Building and Environment* **50** (2012)
- [13] Z. Shi, H. Qian, X. Zheng, Z. Lv, Y. Li, L. Liu, P. V. Nielsen. *Build Environ* **130** (2017)
- [14] S. D'Oca, T. Hong. *Building and Environment* **82** (2014)
- [15] R. V. Jones, A. Fuertes, E. Gregori, A. Giretti. *Building and Environment* **118** (2017)

- [16] P. E. Debevec, C. J. Taylor, J. Malik. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '96, 11–20 (1996)
- [17] P. Müller, G. Zeng, P. Wonka, L. Van Gool. *ACM Trans. Graph.* **26** (2007)
- [18] A. Wendel, M. Donoser, H. Bischof. In M. Goesele, S. Roth, A. Kuijper, B. Schiele, K. Schindler (editors), *Pattern Recognition*, 51–60 (2010)
- [19] A. Martinović, M. Mathias, J. Weissenberg, L. Van Gool. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, C. Schmid (editors), *Computer Vision – ECCV 2012*, 416–429 (2012)
- [20] G. Sun, P. Lin, Y. Li. *Journal of Physics: Conference Series* **1769** (2021)
- [21] Z. Wang, J. Liu, Y. Zhang, H. Yuan, R. Zhang, R. S. Srinivasan. *Renewable and Sustainable Energy Reviews* **143** (2021)
- [22] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255 (2009)
- [23] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, P. Dollár. *Microsoft COCO: Common Objects in Context* (2014)
- [24] A. Krizhevsky, G. Hinton. *Learning multiple layers of features from tiny images*. Technical Report 0, University of Toronto, Toronto, Ontario (2009)
- [25] J. Howard. *HDR: an introduction to high dynamic range photography* (2007)
- [26] G. Le Gray. *Seascape* (1857)
- [27] M. Robertson, S. Borman, R. Stevenson. In *Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348)*, volume 3, 159–163 (1999)
- [28] P. E. Debevec, J. Malik. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH 1997, 369–378 (1997)
- [29] T. Mertens, J. Kautz, F. Van Reeth. In *15th Pacific Conference on Computer Graphics and Applications (PG'07)*, 382–390 (2007)