FINAL YEAR PROJECT 2021-2022

Building extraction and classification of aerial images of rooftop for estimating maximal PV panel installation

TEAM 18

INTRODUCTION

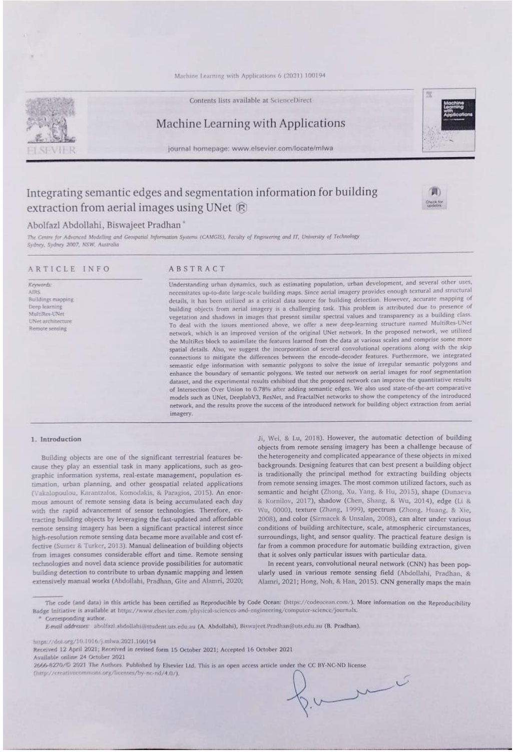
- ❑ A lot of fossil fuel gets wasted in generating electricity and it is highly anticipated that if this continues, there would be dearth of electricity in the next few decades. Potential tapping of renewable resources from solar, wind energy is of great interest.
- ❑ However, traditional approaches such as online assessment and utilities interconnection filings are time consuming and costly, and involve a lot of human effort where concerned people have to go to the site to look into the rooftop area to determine the installation of PV panels.
- ❑ As a first step, to eliminate and reduce the manual work of profiling rooftop, we use an automated approach to segment rooftops. Using aerial images to identify objects on the earth's surface has attracted great attention and aerial image segmentation has gained importance in recent years.
- ❑ However the variety and intricate look of the buildings in mixed backdrops has made it difficult for automatic detection of building objects from remote sensing data.
- ❑ In our project, we employ automated approach for the detection of the usable rooftop of a building with the input satellite imagery.
- ❑ We aim to segment the input rooftop images by using various deep learning models such using ensembling approach to identify the type of roof and use maximum fitting approach to output rooftop images with PV panels installed.

TEAM MEMBERS

Shruthi M - 2018103592
Gayathri M - 2018103535
Jayapriya M - 2018103029

PROJECT GUIDE

Prof Dr. P. Uma Maheshwari



Machine Learning with Applications 6 (2021) 100194

Contents lists available at ScienceDirect

Machine Learning with Applications

journal homepage: www.elsevier.com/locate/mlwa

Integrating semantic edges and segmentation information for building extraction from aerial images using UNet

Abolfazl Abdollahi, Biswajeet Pradhan *

The Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, Sydney 2007, NSW, Australia

ARTICLE INFO

Keywords:
AIRS
Buildings mapping
Deep learning
MultiRes-UNet
UNet architecture
Remote sensing

ABSTRACT

Understanding urban dynamics, such as estimating population, urban development, and several other uses, necessitates up-to-date large-scale building maps. Since aerial imagery provides enough textural and structural details, it has been utilized as a critical data source for building detection. However, accurate mapping of building objects from aerial imagery is a challenging task. This problem is attributed due to presence of vegetation and shadows in images that present similar spectral values and transparency as a building class. To deal with the issues mentioned above, we offer a new deep-learning structure named MultiRes-UNet network, which is an improved version of the original UNet network. In the proposed network, we utilized the MultiRes block to assimilate the features learned from the data at various scales and comprise some more spatial details. Also, we suggest the incorporation of several convolutional operations along with the skip connections to mitigate the differences between the encode–decoder features. Furthermore, we integrated semantic edge information with semantic polygons to solve the issue of irregular semantic polygons and enhance the boundary of semantic polygons. We tested our network on aerial images for roof segmentation dataset, and the experimental results exhibited that the proposed network can improve the quantitative results of Intersection Over Union to 0.78% after adding semantic edges. We also used state-of-the-art comparative models such as UNet, DeeplabV3, ResNet, and FractalNet networks to show the competency of the introduced network, and the results prove the success of the introduced network for building object extraction from aerial imagery.

1. Introduction

Building objects are one of the significant terrestrial features because they play an essential task in many applications, such as geographic information systems, real-estate management, population estimation, urban planning, and other geospatial related applications (Vakalopoulou, Karantzalos, Komodakis, & Paragios, 2015). An enormous amount of remote sensing data is being accumulated each day with the rapid advancement of sensor technologies. Therefore, extracting building objects by leveraging the fast-updated and affordable remote sensing imagery has been a significant practical interest since high-resolution remote sensing data became more available and cost effective (Sumer & Turker, 2013). Manual delineation of building objects from images consumes considerable effort and time. Remote sensing technologies and novel data science provide possibilities for automatic building detection to contribute to urban dynamic mapping and lessen extensively manual works (Abdollahi, Pradhan, Gite and Alamri, 2020; Ji, Wei, & Lu, 2018). However, the automatic detection of building objects from remote sensing imagery has been a challenge because of the heterogeneity and complicated appearance of these objects in mixed backgrounds. Designing features that can best present a building object is traditionally the principal method for extracting building objects from remote sensing images. The most common utilized factors, such as semantic and height (Zhong, Xu, Yang, & Hu, 2015), shape (Dunaeva & Kornilov, 2017), shadow (Chen, Shang, & Wu, 2014), edge (Li & Wu, 0000), texture (Zhang, 1999), spectrum (Zhong, Huang, & Xie, 2008), and color (Sirmacek & Unsalan, 2008), can alter under various conditions of building architecture, scale, atmospheric circumstances, surroundings, light, and sensor quality. The practical feature design is far from a common procedure for automatic building extraction, given that it solves only particular issues with particular data.

In recent years, convolutional neural network (CNN) has been popularly used in various remote sensing field (Abdollahi, Pradhan, & Alamri, 2021; Hong, Noh, & Han, 2015). CNN generally maps the main

The code (and data) in this article has been certified as Reproducible by Code Ocean: (https://codeocean.com/). More information on the Reproducibility Badge Initiative is available at https://www.elsevier.com/physical-sciences-and-engineering/computer-science/journals.

* Corresponding author.
E-mail addresses: abolfazl.abdollahi@student.uts.edu.au (A. Abdollahi), Biswajeet.Pradhan@uts.edu.au (B. Pradhan).

https://doi.org/10.1016/j.mlwa.2021.100194
Received 12 April 2021; Received in revised form 15 October 2021; Accepted 16 October 2021
Available online 24 October 2021

OVERALL OBJECTIVES

To estimate the placement of photovoltaic panels on the rooftop for an aerial satellite image, the following steps are to be performed:

- ❑ Detect the building in a given satellite image using U-Net model and perform background subtraction to extract the building.
- ❑ Annotate the extracted building rooftops using CVAT/VIA annotator tool.
- ❑ Perform object detection to segment roofs based on the type (flat/tilted).
- ❑ Use maximum fitting algorithm to find the optimized no of solar panels based on type of roof to maximize energy consumption.

LITERATURE SURVEY

S.NO	CITATION	METHODOLOGY	ADVANTAGES	LIMITATION
1.	**An aerial image segmentation approach based on enhanced multi-scale convolutional neural network, 2019** Xiang Li, Yuchen Jiang, Hu Peng and Shen Yin in 2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)	1. Segmentation model is performed using an encoder-decoder architecture. 2. A U-Net is constructed as the main network, and the bottom convolution layer of U-Net is replaced by a set of cascaded dilated convolution with different dilation rates. 3. Add an auxiliary loss function after the cascaded dilated convolution	1. From the aspect of design and training, the approach does not involve manual features and does not require specific preprocessing or post-processing, which can reduce the influence of subjective factors 2. The auxiliary loss function helps to make the network converge faster and optimize.	1. Segmentation of large buildings work well but boundaries and middle parts are misaligned. 2. The bulges on the boundaries are lost and edges are not detected properly. 3. The algorithm performs well in one of the subset (countryside and forest) but does not perform well when tested on a different subset.

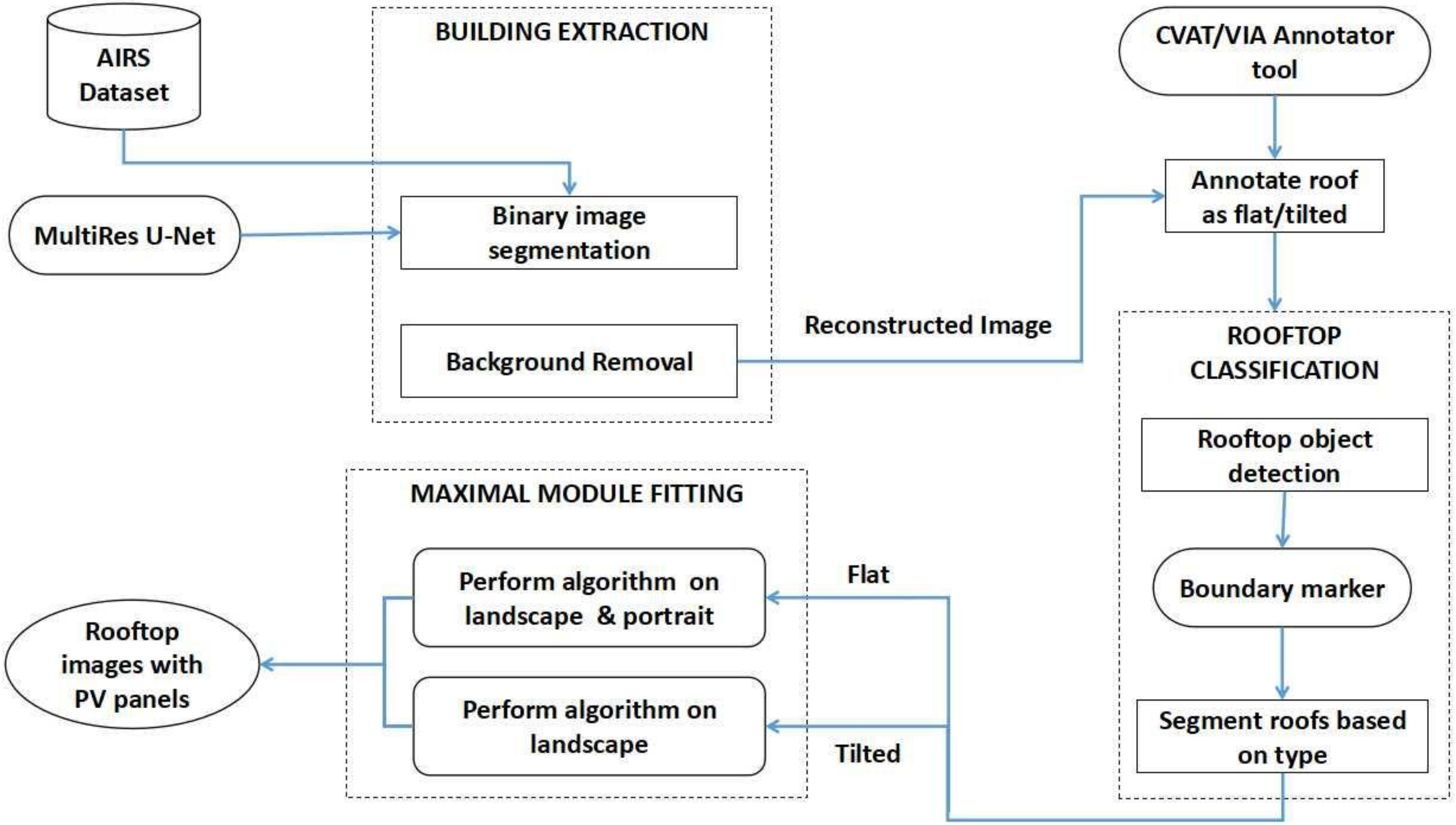
LITERATURE SURVEY

S.NO	CITATION	METHODOLOGY	ADVANTAGES	LIMITATION
2.	**Convolutional Neural Network Based Solar Photovoltaic Panel Detection in Satellite Photos, 2017** Vladimir Golovko, Sergei Bezobrazov, Alexander Kroshchanka and Anatoliy Sachenko in 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications	1. Collect data from Google Maps. 2. Perform pre-processing techniques like image resizing, image sharpening. 3. Train a 6 layer CNN model.	1. Here, the authors have used the low-quality satellite imagery (Google Maps photos), instead of the high resolution color satellite orthoimagery that enables decreasing the requirements for the approach. 2. Simple 6 layer CNN model.	1. Better segmentation could have been performed by training with other CNN models. 2. Bad quality satellite images have led to inaccurate classification. 3. No validation on the dataset as in some cases solar panels look similar to roof tops.

LITERATURE SURVEY

S.NO	CITATION	METHODOLOGY	ADVANTAGES	LIMITATION
3.	**Deep Convolutional Neural Network Application on Rooftop Detection for Aerial Imagery, 2019** Mengge Chen, Jonathan Li, in Journal of Computational Vision and Imaging Systems	1. It is primarily based on Mask R-CNN with 3 stages. 2. Feature extraction is based on existing deep learning model. 3. RPN (Regional Proposal Network) is used to find RoI. 4. Object classification is then performed.	1. Efficient and feasible approach to extract detached house from aerial images. 2. RoIAlign method is used instead of RoIPool for better feature extraction.	1. Edges of the building are not detected properly. 2. Training data was less and hence less accuray. 3. Comparatively less precision with other new state of art models.
4.	**Solar Potential Analysis Of Rooftops Using Satellite Imagery, 2019** Akash Kumar, Delhi Technology University, in *ArXiv* abs/1812.11606	1.Dataset is manually collected for India. 2.Adaptive Edge Detection and Contours are focused to segment out rooftop boundaries and obstacles present inside them along with polygon shape approximation.	1.Provides a comparative analysis of the solar potential of the building. 2. Several types of the rooftop are considered to learn the intra-class variations.	1.The image quality of satellite imagery is very deficient hence the edges are not detected properly. 2.There are some outliers that are plotting solar panels outside the building rooftop area.

ARCHITECTURE DIAGRAM



ARCHITECTURE DIAGRAM

The above block diagram gives a high level overview on the 3 modules.

The 1st module contains AIRS dataset which is fed to the MultiRes U-Net model to detect and extract the buildings. The background is removed, the segmented rooftop is reconstructed and fed to the next module.

We annotate the reconstructed image using CVAT/VIA software to perform rooftop object detection algorithm where boundaries are marked and rooftops are segmented based on type.

Based on the segmented roofs, we set a size for solar panels that gets overlaid on the rooftop image. The no of panels to be fitted needs to be maximized in order to get maximum efficiency and this is done with the help of maximum fitting algorithm.

LIST OF MODULES

Module I - Building Extraction.

Module II - Rooftop Classification.

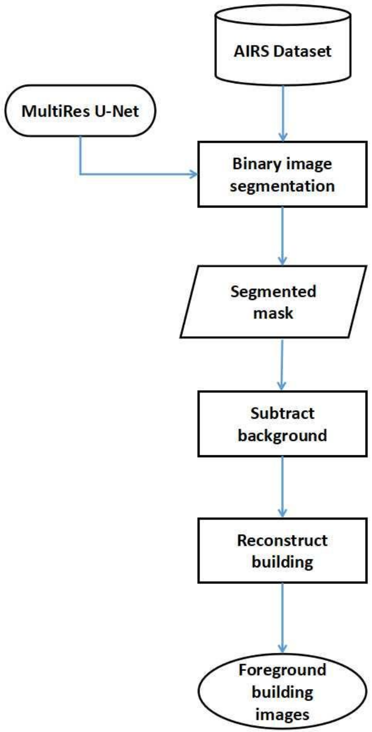
Module III - Maximal Module Fitting.

MODULE DESIGN

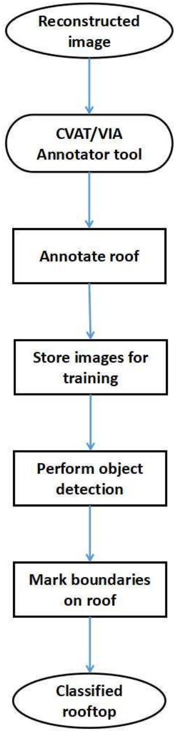
MODULE I - BUILDING EXTRACTION

INPUT: AIRS Dataset
OUTPUT: Reconstructed rooftops without background

- ➔ We use AIRS dataset and select around 1k images from the dataset.
- ➔ The input images are fed to the MultiRes U-Net model to perform binary segmentation.
- ➔ The background is subtracted from the image.
- ➔ The segmented rooftop is reconstructed and fed to the next module.



MODULE DESIGN



MODULE II - ROOFTOP CLASSIFICATION

INPUT: Reconstructed image
OUTPUT: Classified rooftop

➔ Using CVAT/VIA tool, we annotate each of the segmented , reconstructed rooftop images as flat or tilted.

➔ This is done by using the software and marking boundaries on the rooftop and assigning labels.

➔ The masked image along with input is stored for training.

➔ Object detection is performed on the masked image to mark boundaries on roof

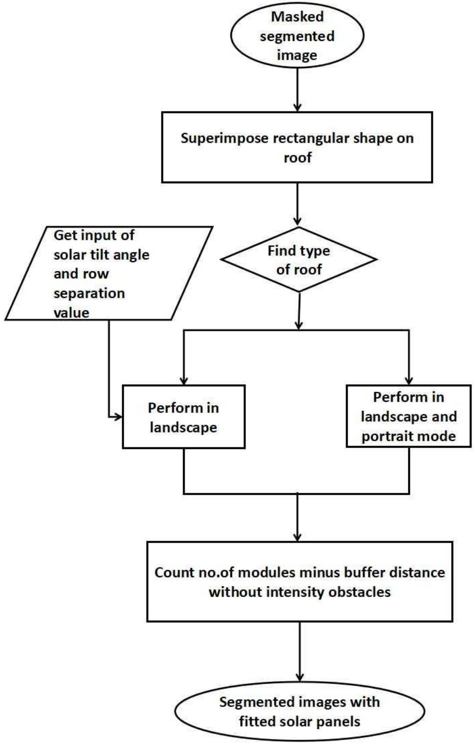
➔ The rooftops are classified and segmented based on type.

MODULE DESIGN

MODULE III - MAXIMAL MODULE FITTING

INPUT: Masked segmented image.
OUTPUT: Segmented image with solar panels of specific size fitted.

- ➔ The masked segmented image stored in the new database is taken and the roofs are superimposed with rectangular shape.
- ➔ Then the type of the roof is determined as flat or slope.
    - If the type of the roof is flat, we get user input(solar tilt angle and row separation value) and perform maximum fitting algorithm on landscape mode.
    - If the type of the roof is slope, we perform maximum fitting algorithm on landscape and portrait mode.
- ➔ The buffer distance without intensity obstacles is subtracted from the count of number of modules.
- ➔ The segmented images are fitted with the solar panels.



PERFORMANCE MEASURES

- **IoU - Intersection over Union /Jaccard Coefficient**

  To quantify the accuracy of our model to predict size for solar PV arrays, we use Jaccard coefficient which is widely used in prior work to measure the similarity between detected regions and ground truth regions. Jaccard Similarity Index(JSI) measures the similarity for the two sets of pixel data, with a range from 0% to 100%. The higher the percentage, the more precise prediction. It is defined as follows:

$$JSI = \frac{r_d \bigcap r_g}{r_d \bigcup r_g}$$

  where r_d denotes the detected region for a solar PV array, and r_g indicates the groundtruth region for a solar PV array

- **DICE Coefficient**

  We use DICE coefficient to compare the pixel-wise agreement between a predicted segmentation and its corresponding ground truth. DICE coefficient is 2 times the area of overlap divided by the total number of pixels in both the images.The formula is given by

$$\frac{2 * |X \cap Y|}{|X| + |Y|}$$

  where X is the predicted set of pixels and Y is the ground truth.

PERFORMANCE MEASURES

- **MCC - Matthews Correlation Coefficient**

  We use the MCC , a standard measure of a binary classifier's performance, where values are in the range −1.0 to 1.0, with 1.0 being perfect roof segmentation, 0.0 being random roof segmentation, and −1.0 indicating roof segmentation is always wrong. The expression for computing MCC is below, where TP is the fraction of true positives, FP is the fraction of false positives, TN is the fraction of true negatives, and FN is the fraction of false negatives, such that TP+FP+TN+FN= 1.

$$\frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

- **Accuracy**

  Accuracy is the percentage of correct predictions for the test data. It can be calculated easily by dividing the number of correct predictions by the number of total predictions.

$$\text{accuracy} = \frac{\text{correct predictions}}{\text{all predictions}}$$

DATASET

- **AIRS Dataset -** The AIRS (Aerial Imagery for Roof Segmentation) dataset provides a wide coverage of aerial imagery with 7.5 cm resolution and over 220k buildings.

- AIRS dataset covers almost the full area of Christchurch, the largest city in the South Island of New Zealand during the flying seasons of 2015 and 2016, and the supplied images are ortho-rectified DOMs with RGB channels.

- It contains around 800 images in training set, 90 each in testing and validation set.

- We select around 1000 images from residential areas and use the VIA tool to annotate the type of roofs and combine both the images to form an annotated dataset.

REFERENCES

[1] Q. Li, Y. Feng, Y. Leng and D. Chen, " SolarFinder: Automatic Detection of Solar Photovoltaic Arrays," 2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), 2020, pp. 193-204, doi: 10.1109/IPSN48710.2020.00024.

[2] Edun, Ayobami & Harley, Joel & Deline, Chris & Perry, Kirsten. (2021). Unsupervised azimuth estimation of solar arrays in low-resolution satellite imagery through semantic segmentation and Hough transform. Applied Energy. 298. 10.1016/j.apenergy.2021.117273.

[3] B. Chatterjee and C. Poullis, "On Building Classification from Remote Sensor Imagery Using Deep Neural Networks and the Relation Between Classification and Reconstruction Accuracy Using Border Localization as Proxy," 2019 16th Conference on Computer and Robot Vision (CRV), 2019, pp. 41-48, doi: 10.1109/CRV.2019.00014.

[4] V. Golovko, S. Bezobrazov, A. Kroshchanka, A. Sachenko, M. Komar and A. Karachka, "Convolutional neural network based solar photovoltaic panel detection in satellite photos," 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2017, pp. 14-19, doi: 10.1109/IDAACS.2017.8094501.

[5] Kumar, Akash & Sreedevi, Indu. (2018). Solar Potential Analysis of Rooftops Using Satellite Imagery. *ArXiv* abs/1812.11606.

REFERENCES

[6] Chen, Mengge and Jonathan Li. “Deep convolutional neural network application on rooftop detection for aerial image.” *ArXiv* abs/1910.13509 (2019): n. Pag.

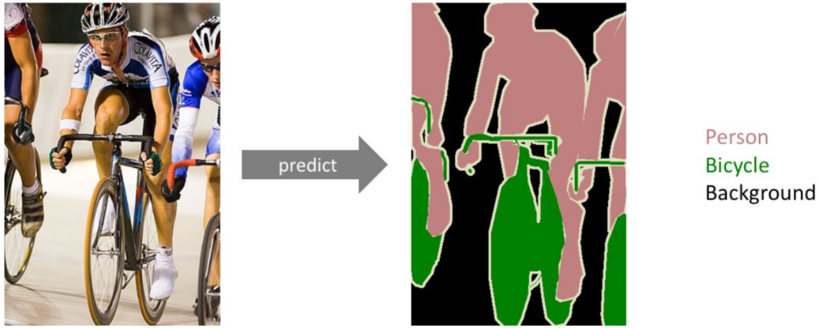
[7] Peiran Li, Haoran Zhang, Zhiling Guo, Suxing Lyu, Jinyu Chen, Wenjing Li, Xuan Song, Ryosuke Shibasaki, Jinyue Yan. (2021). Understanding rooftop PV panel semantic segmentation of satellite and aerial images for better using machine learning. Advances in Applied Energy, Elsevier. Volume 4, 100057, ISSN 2666-7924. doi: 10.1016/j.adapen.2021.100057.

[8] Qi, Chen & Wang, Lei & Wu, Yifan & Wu, Guangming & Guo, Zhiling & Waslander, Steven. (2018). Aerial Imagery for Roof Segmentation: A Large-Scale Dataset towards Automatic Mapping of Buildings. ISPRS Journal of Photogrammetry and Remote Sensing, Elsevier. Volume 147, pp. 42-55.

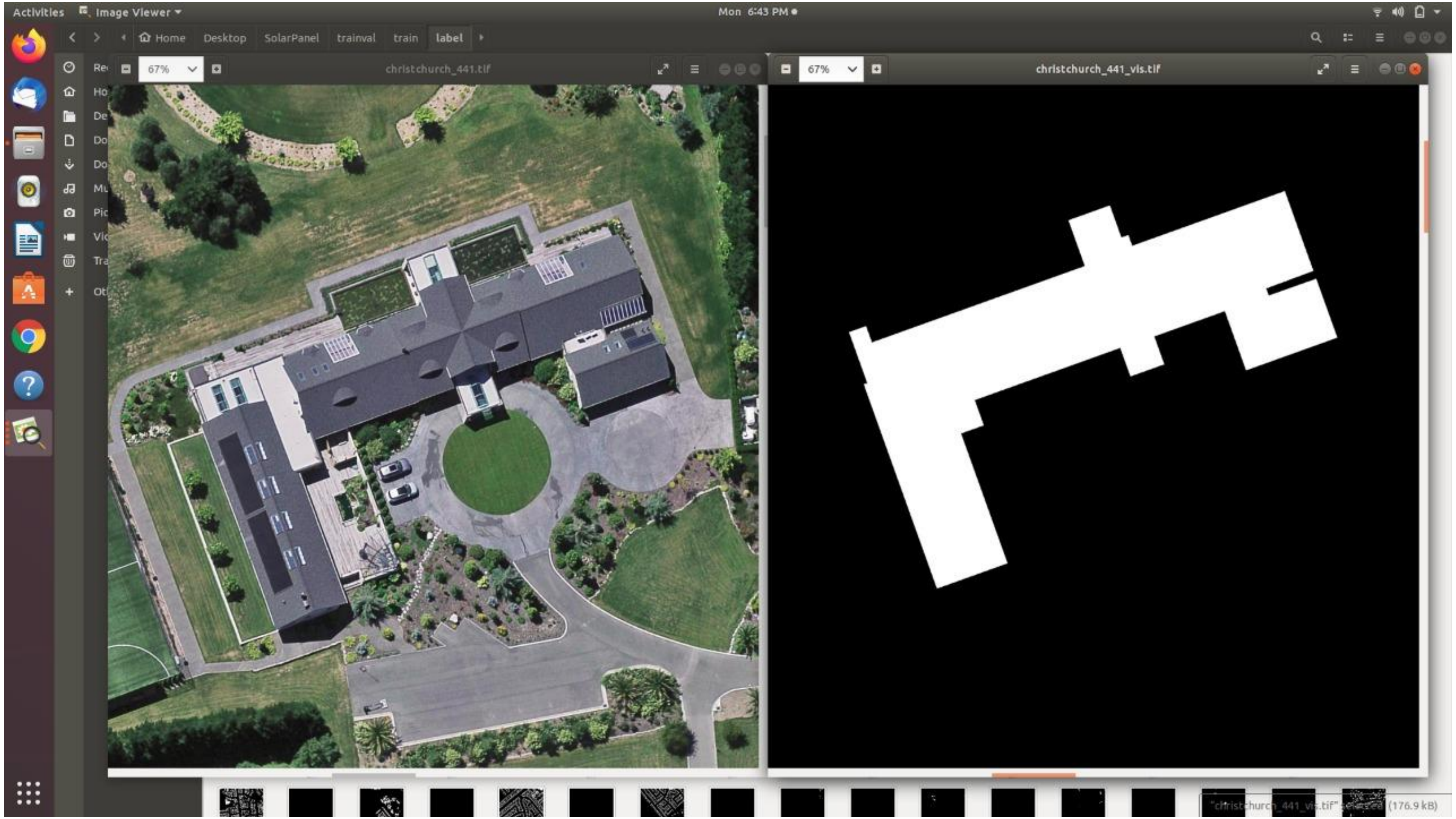
[9] Cai, Yuwei & He, Hongjie & Yang, Ke & Fatholahi, Sarah Narges & Ma, Lingfei & Xu, Linlin & Li, Jonathan. (2021). A Comparative Study of Deep Learning Approaches to Rooftop Detection in Aerial Images. Canadian Journal of Remote Sensing. 47. 1-19. 10.1080/07038992.2021.1915756.

[10] Srinivasan Iyengar, Stephen Lee, David Irwin, Prashant Shenoy, and Benjamin Weil. 2018. WattHome: A Data-driven Approach for Energy Efficiency Analytics at City-scale. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.

pixel of an image with a corresponding ***class*** of what is being represented. Because we're predicting for every pixel in the image, this task is commonly referred to as **dense prediction**.



An example of semantic segmentation, where the goal is to predict class labels for each pixel in the image. (Source)


Activities
Image Viewer
Mon 6:43 PM
Home
Desktop
SolarPanel
trainval
train
label
67%
christchurch_441.tif
67%
christchurch_441_vis.tif
"christchurch_441_vis.tif" selected (176.9 kB)