# Predicting Levels of Household Electricity Consumption in Low-Access Settings

Simone Fobi *[1], Joel Mugyenyi[2], Nathaniel J. Williams[3], Vijay Modi[1], and Jay Taneja[2]

[1]Department of Mechanical Engineering, Columbia University
[2]Department of Electrical & Computer Engineering, University of Massachusetts Amherst
[3]Golisano Institute for Sustainability, Rochester Institute of Technology
{sf2786,modi}@columbia.edu, {jmugyenyi,jtaneja}@umass, njwgis@rit.edu

## A. Supplementary Material

### A.1. Multi-Layer Perception (MLP) Architecture

Figure 1 shows the MLP architecture used to train *Model B* (Non-visual Data) and *Model D* (Building Characteristics Only). The MLP consists of 3 dense layers with 64, 32, and 16 filters respectively, all with ReLU activations. The last dense layer consists of a softmax activation. 25 % dropout was applied to minimize overfitting.

### A.2. Performance of building segmentation

Additional evaluation of the building segmentation task is done by observing how the classifier performs at varying training data sample sizes. Figure 2 shows the F1-score at different training data sample sizes when random subsets of the data are selected and either random weights or building segmentation weights are used to initialize model training. At each sample size increment, samples from the previous sample size are included. E.g. the 20 % dataset contains all the samples from the 5 % dataset. Initializing with building segmentation weights offers performance gains especially at smaller sample sizes. The improved performances with building segmentation weights suggests that underlying characteristics about buildings (rooftop type, color, size) provides relevant features for consumption prediction. This is inline with our initial findings that building characteristics are relevant in predicting consumption levels. In addition to improved model performance, building segmentation weights make the classifier less susceptible to label quality. Specifically when random weights are used for initialization, it is observed that the randomly selected subsample at 60 % of the full dataset, performed the best and performance dropped as more samples were added. This suggests that the ease | difficulty of the sub-sample significantly affects performance. Building segmentation weights initializes the model in a suitable learning space and has a regularizing effect even as harder labels may be introduced, allowing only additional useful information to be extracted. Obtaining large amounts of useful samples to appropriately

*We thank The Rockefeller Foundation for their generous support of this work under Grant 2018 POW 004.

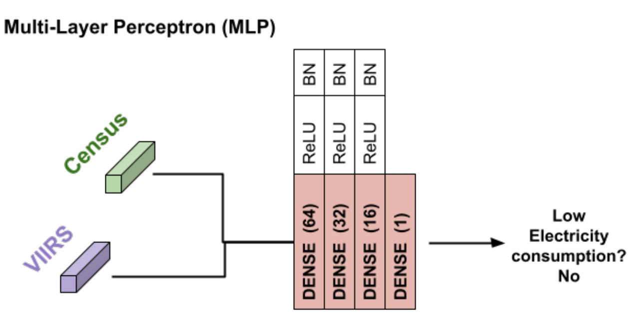


Figure 1. MLP architecture used to train *Model B* and *Model D*

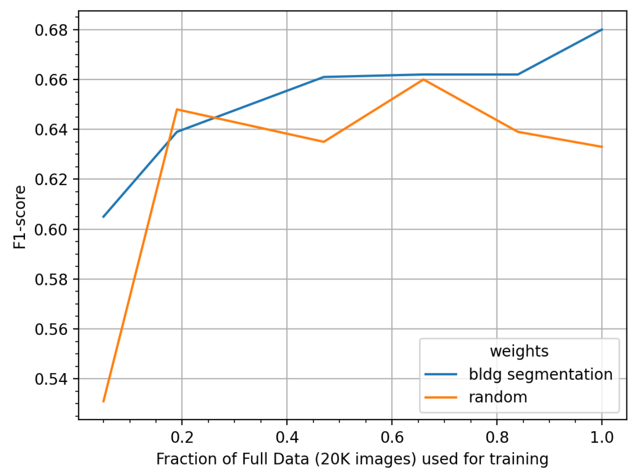


Figure 2. Comparison of prediction performance when the classifier is initialized with random weights versus building segmentation weights. Learning about building segmentation improves performance in low-data regimes and makes performance less susceptible to harder labels thereby offering a regularizing effect.

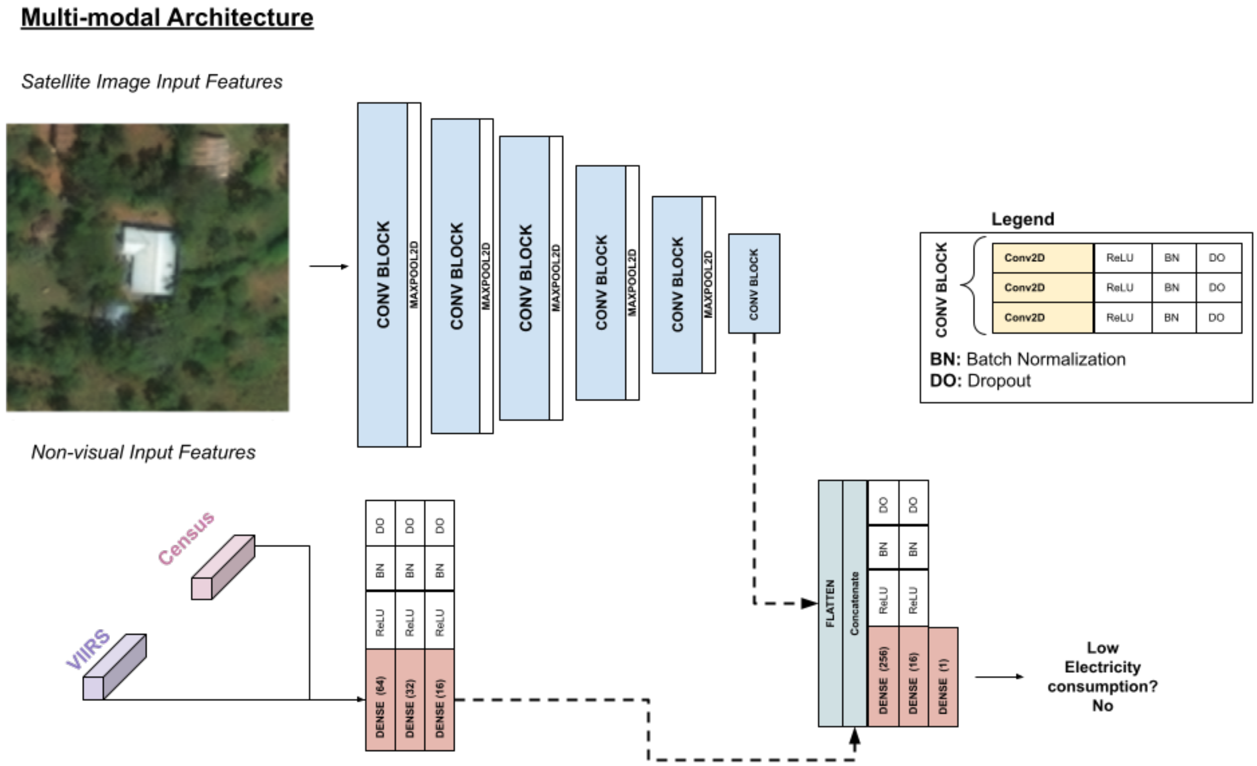


Figure 3. Multi-modal architecture combining the CNN image-based encoder with an MLP to predict consumption levels using visual images and non-visual public data sources.

predict consumption of yet to be connected customers can be challenging. For energy practitioner looking to apply our approach, we show that learning about buildings from using a segmentation task, provides useful weight tuning needed for appropriate prediction of consumption tiers.

### A.3. Multi-modal architecture: Encoder and MLP

Figure 3 shows the multimodal architecture used to combine satellite images with public data sources.