**Detecting Solar PV Through Supervised Learning**

Anna Berman, Azu Morales, Viggy Kumaresan, Sicong Zhao, Yifei Wang

As the solar industry grows in size, it’s increasingly important to monitor consumer solar adoption. The first step in understanding consumer energy consumption is to delineate homes with and without solar panels. To address this first step, myself and my colleagues evaluated several modeling approaches to identifying solar PV in high resolution aerial imagery data. A convolutional neural network (CNN) approach yielded high rates of accuracy of solar PV detection from satellite aerial imagery.

**Why Solar PV Detection?**

Solar power currently accounts for 1% of the world’s electricity generation. In fact, estimates of solar energy production predict a potential 65-fold growth by 2050, eventually making solar power one of the largest sources of energy across the globe (Energy Transition Outlook, 2018). Solar photovoltaic (PV) power installed on top of rooftops, or solar PV, is estimated to make up thirty percent of this energy generation. In recent years, solar PV power has already begun playing an increasingly large role in US electricity generation. From 2008 through 2017, there was a 39-fold growth in annual solar generation representing an increase of 75,123 GWh (Environment America, 2018).

As consumer solar PV adoption increases, the consumption habits of solar adopters are of vital interest to a variety of stakeholders (Solar Industry Research Data, 2019). Policymakers, for example, are reliant on accurate measures of both the saturation of solar PV as well as on the energy consumption and production rates associated with these installations. This information is integral the structuring of tax incentives and credits for installation and use of solar PV (Matasci, 2019; Solangi et al. 2011). In another example, solar manufacturers also rely on detailed solar PV market insights. In order to acquire new customers, solar manufacturers depend on this information to target specific customer segments and to create customized marketing and sales strategies (Janes, 2014; Rai et al. 2016).

Traditional data sources such as consumer surveys and market research, are costly and time-consuming to collect, and ultimately can often only give a partial or biased view of the market. Satellite imagery, on the other hand, allows us to see overhead views of households all over the country and does not rely on self-reported data. The utilization of these images to detect consumer solar PV has the potential to result in a more consistent and cost-effective view of solar adoption in the US.

**Previous Research**

Previous studies have explored the use of machine learning for satellite image classification with high rates of accuracy of detection with a low rate of incorrect classification (Mnih et al., 2010; Kubat et al., 1998). Specifically, random forest classification and CNN approaches have been shown to accurately measure size, shape and capacity of solar PV from satellite imagery (Malof et al., 2016; Malof et al., 2017; Golovko et al., 2017).

However, to implement approaches like these, as with any supervised learning approach, a training set of data with labelled classes is required in order to both model the data as well as evaluate performance.Unfortunately, the number of available labelled datasets containing satellite imagery of solar PV is limited. An example of such a dataset is the Distributed Solar Photovoltaic Array Location and Extent Dataset containing geospatial coordinates and border vertices for over 19,000 solar panels across 601 high-resolution images from four cities in California (Bradbury et al., 2016).

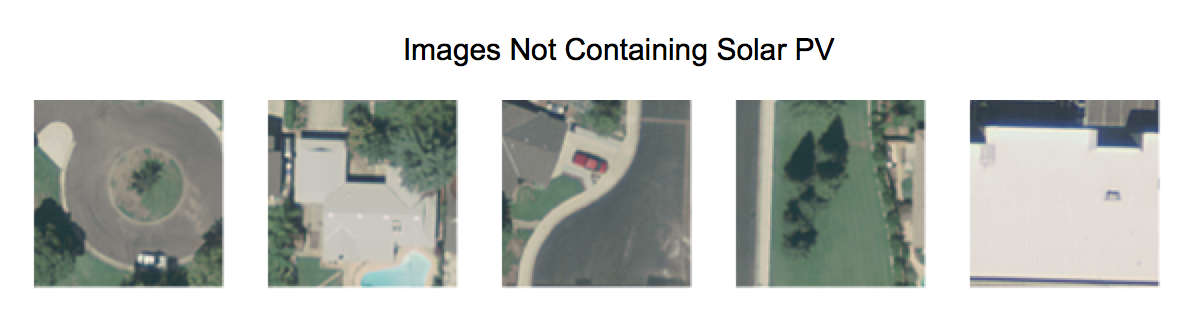
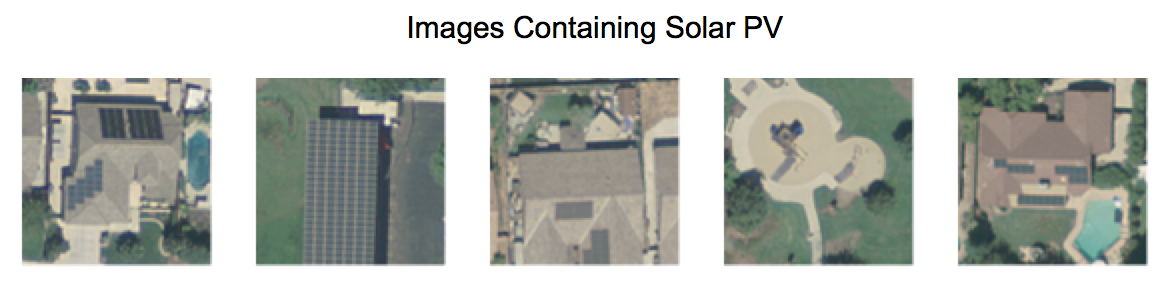
**Data Sources**

Our dataset consists of 1,500 satellite images in TIFF format, each image contains a 101 pixels x 101 pixels size and has three color channels (Red, Green, Blue). In total, each image is represented as 101 x 101 x 3 array, which is a total of 30,603 numerical entries with values ranging from 0 and 255.

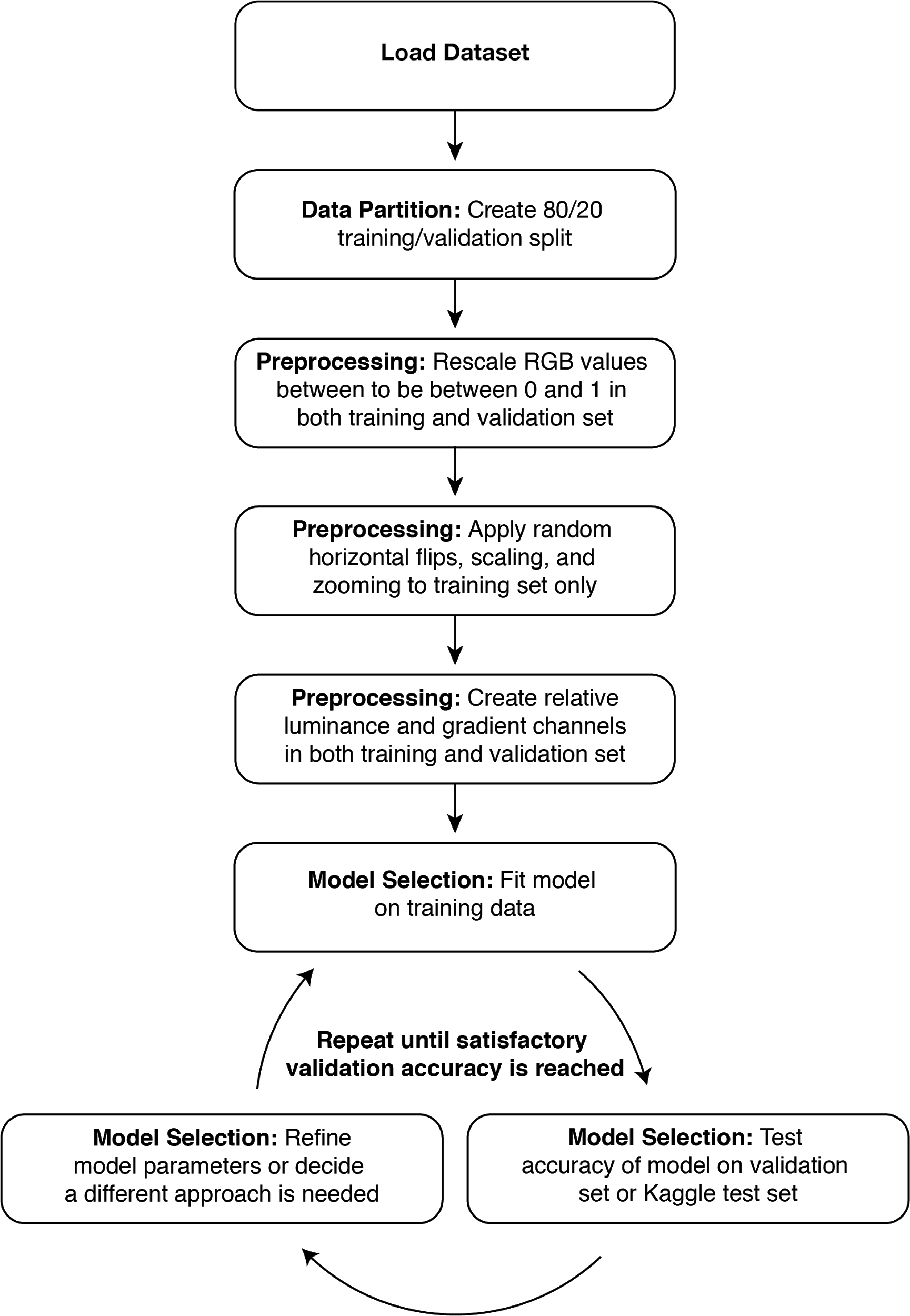
Each image in our dataset has been labeled as one of two classes, either containing a solar PV, or not containing a solar PV (examples of both classes are shown in Figures 1 and 2). The labels for this dataset were created by human observers who visually scanned the imagery and and annotated all of the visible solar PV. There are 555 images that contains a solar PV and 995 images without a solar PV, which means that the classes in our dataset were not entirely equally balanced.

Unfortunately, the images of the solar PV included do not necessarily share standard form. In general, solar PV instances come in a variety of sizes and colors and are not always at the center of the image. However, many solar panels tend to be rectangle with sharp angles and edges. Additionally, the surrounding landscape in images of both classes is also far from standardized. Examples of rooftops, pavement, grass, and residential pools can be seen in both classes. In order for ant model to be successful, this large amount of diversity needs to be addressed. Additionally, a model should predict the a consistent class regardless of orientation of each image. Therefore, rotation and scaling of images should be tolerated in our final model.

In addition to the data described above, an additional 558 number of images and labels were housed on Kaggle.com. These images were used as a test set during our model selection process.

****

**Methods**

****

*Train-Test Split*

Out of the original 1,500 images, we split created a 80-20 training-test split. When partitioning the data, we ensured a ratio of classes proportional to the original dataset in both our training and validation sets.

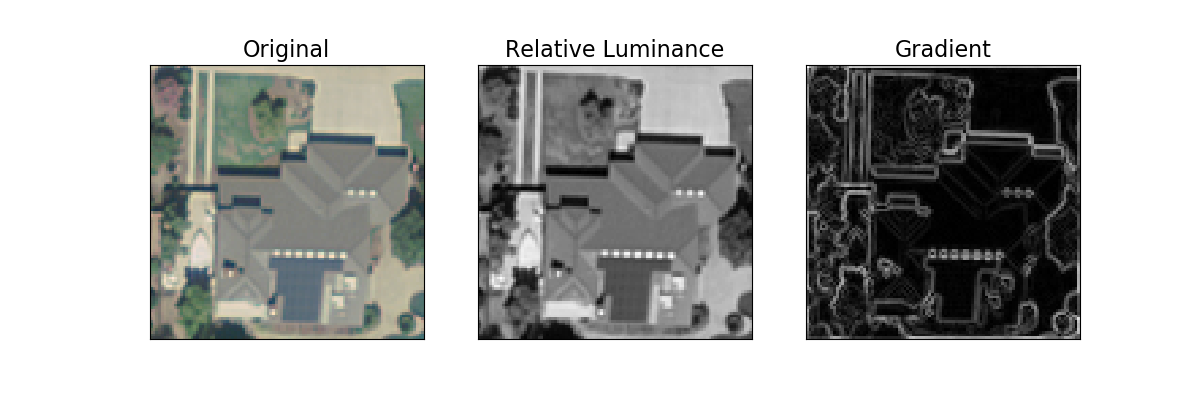
*Data Preprocessing*

When preprocessing our data, we prepared our images in three main steps. In the first step of preprocessing, we rescaled the RGB values of our images to be in the range of 0 and 1 in order to avoid having high values compared to a typical learning rate (Chollet, 2016).

During the second step of preprocessing, we applied a series of random modifications to our data in order to prevent overfitting of our model. Specifically, we applied horizontal flips, image shears, and random zoom. By increasing the diversity of images in our training set we hoped to reduce overfit and ultimately to increase predictive accuracy outside our current dataset (Chollet, 2016; Venkatesh, 2017; Lewinson 2018).



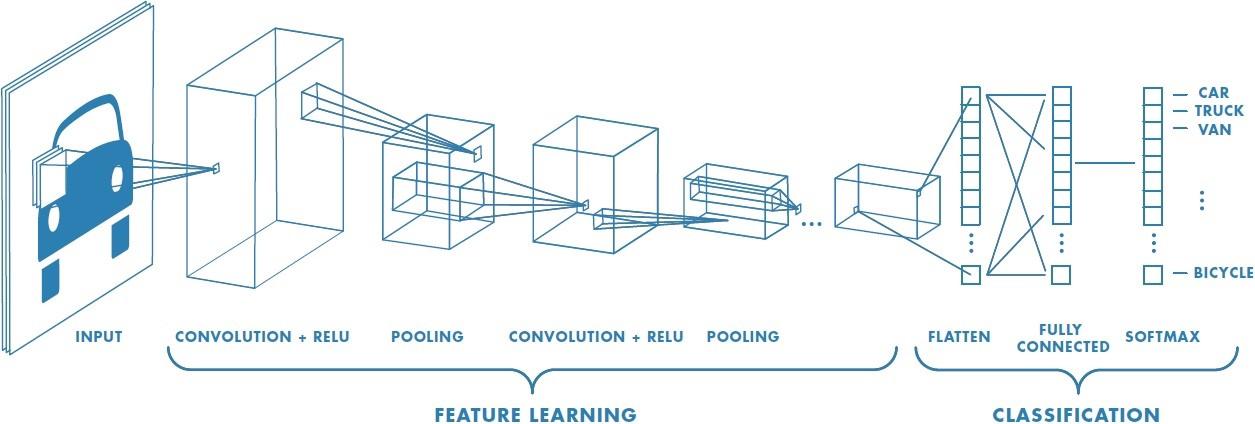
Finally, we added lumiance and gradient features to each of our images in the hopes of creating a clearer delineation between our two image classes (Stokes et al., 1996). We created these features based on properties we know to be true of solar panel, they absorb light and they are angular in shape. By adding features around these two traits, we hoped to increase the accuracy of our predictions.



**Convolutional Neural Networks**

CNN are a type of neural network for processing data that has a grid-like topology. CNNs are widely known in computer vision for being successful for tasks such as classifying images, clustering images, and object recognition (Karpathy, 2018). At they’re core, CNNs use the convolution operation, instead of matrix multiplication in at least one of their layers. As other neural networks, they are conformed by a sequence of layers. The layers of CNN have neurons arranged in three dimensions: width, height and depth. There are different type of CNN architectures, but there are main type of layers (Goodfellow et al., 2016; Skymind, Accessed 2019) (see Figure 8 for a visual representation):

1. Convolutional Layer: This layer applies a convolution operation, the output is passed to the next layer.
2. Pool: This layer performs a down sampling operation, by combining the outputs of neurons at one layer into a single neuron in the next layer.
3. Dropout: This layer performs an analogous feature selection operation as random forest, by randomly turning off some proportion of all nodes.
4. Flatten: This layer flattens the pooled feature map into a column
5. Full-Connection: The layer will compute the class scores, each neuron in this layer will be connected to all the neurons in the previous one.



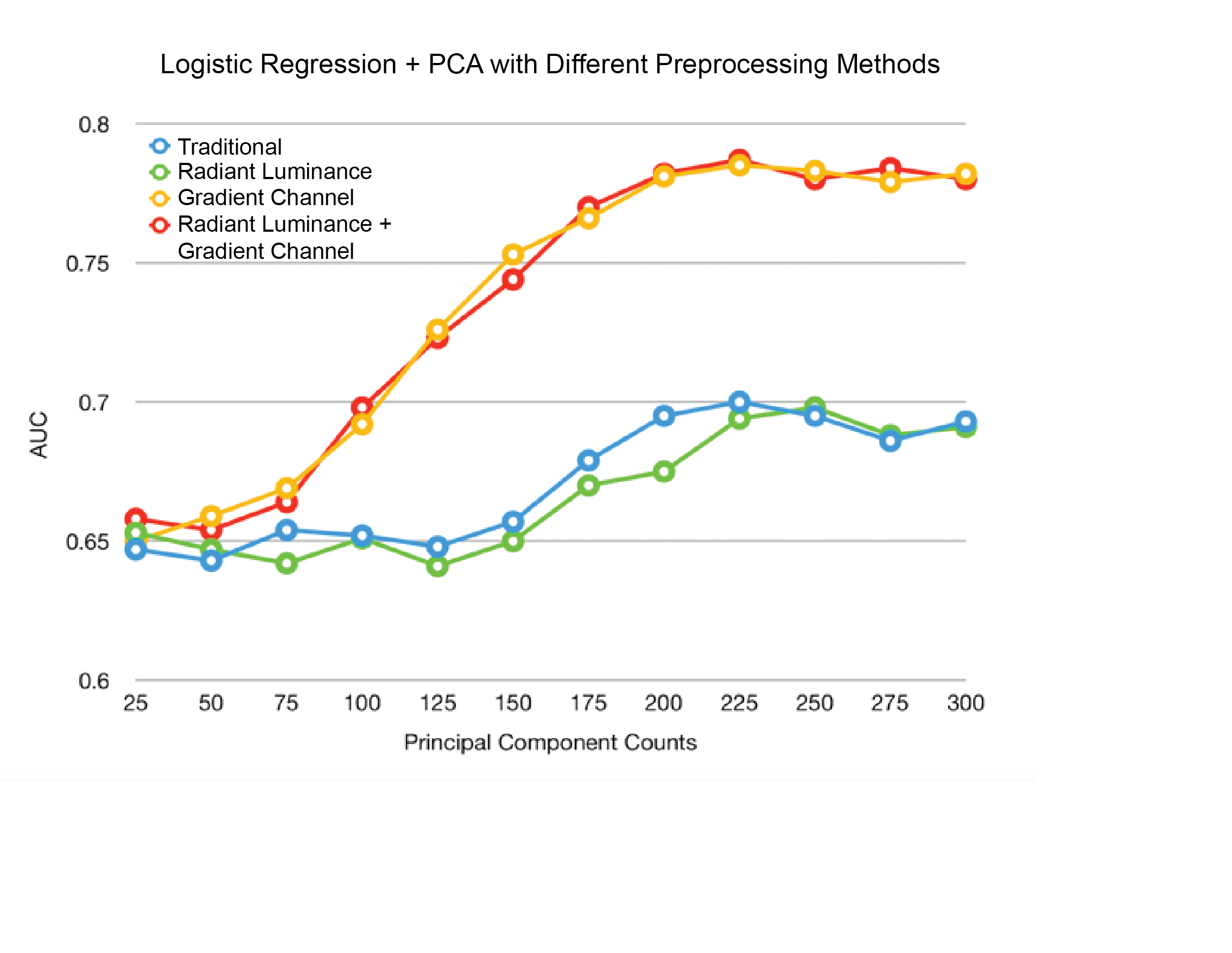
(Prabhu, 2018)**.** The convolutional layer applies a convolution operation, the output is passed to the next layer. The pool layer performs a down sampling operation by combining the outputs of neurons at one layer into a single neuron in the next layer. The flatten reshapes the feature map into a column. The full-connection layer will compute the class scores, each neuron in this layer will be connected to all the neurons in the previous one.

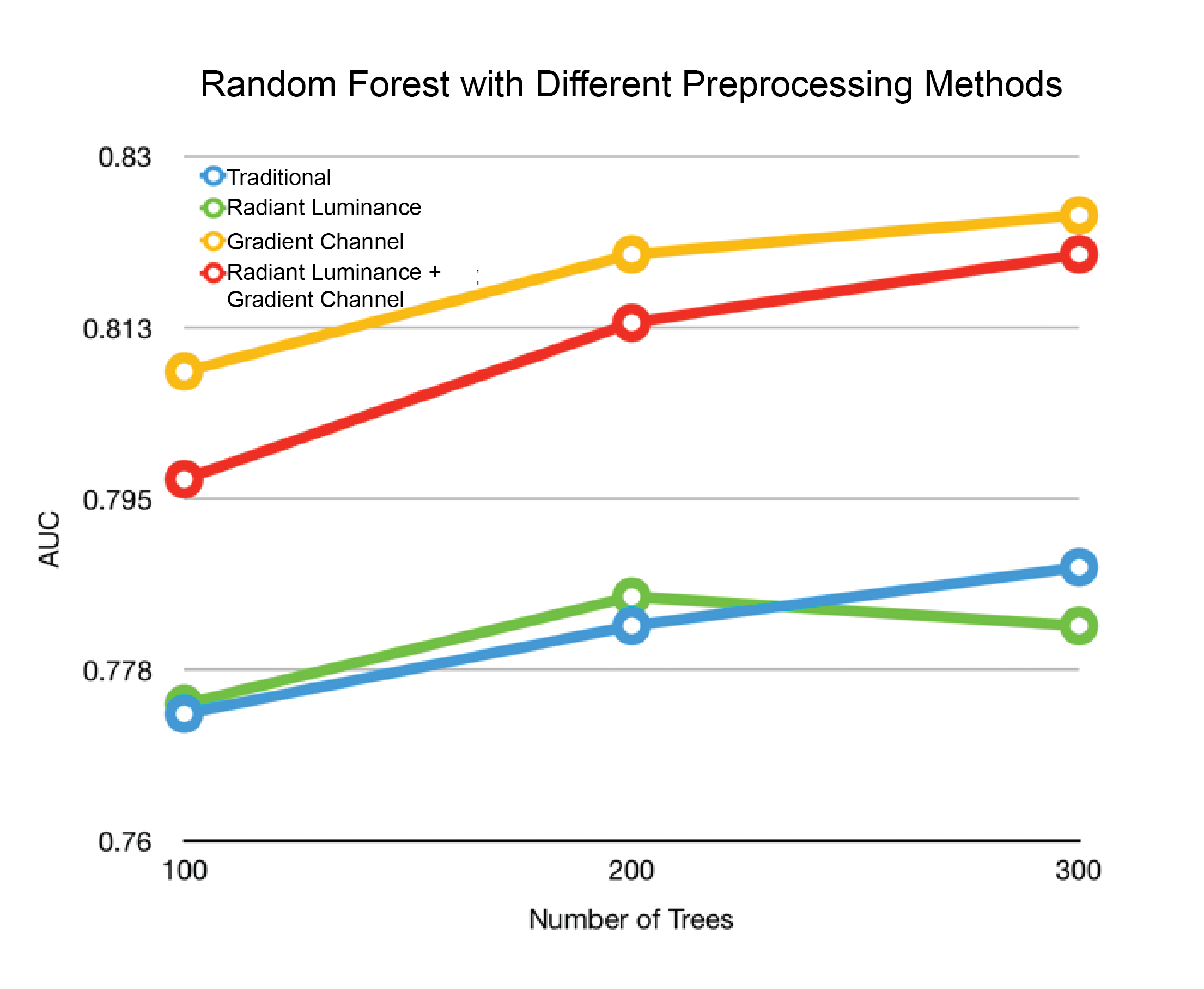
**Results**

Comparing a variety of modeling approached, we found CNN to be the strongest approach. A KNN classifier and a gaussian naive bayes model performed rather similarly to each other with AUCs of 0.639 and 0.643 respectively. A simple logistic regression was able to achieve a surpriseing 0.785 AUC. Random forest modeling performed even stronger with and AUC of 0.824. Finally, our CNN resulted in validation accuracy of 0.957. Although AUC and validation accuracy are not measured on the same scale, the results suggest our CNN model produces meaningful improvement in accuracy.

|  |  |  |
| --- | --- | --- |
| **Model** | **Metric** | **Validation Set Performance** |
| Chance | AUC | 0.500 |
| KNN | AUC | 0.639 |
| Gaussian Naive Bayes | AUC | 0.643 |
| Logistic Regression | AUC | 0.785 |
| Random Forest | AUC | 0.824 |
| CNN | Validation Accuracy | 0.957 |

Interestingly, we also found that the addition of a gradient channel increased the accuracy of many of our models.





**Conclusion**

Our findings align with previous research and suggest that approaches such as traditional logistic regression, k-nearest neighbors, and gaussian naive bayes may not be sufficient for identification with high rates of accuracy. That said, by adding additional features, such as image gradients, performance can be improved even when using simplest models.

Our findings suggest a CNN approach produces the highest accuracy of detection. Instead of handling each pixel as features individually, CNN essentially looks at pixels in context. This type of approach is ideal when considering image recognition analysis because pixels only have meaning in relation to surrounding pixels. Given algorithm behind CNNs in without results, CNN or similar neural network approaches appear to be a feasible and strong choice for similar image detection analyses.

In comparison to previous research on supervised learning and solar PV detection, this analysis was rather small. Our dataset contained only 1,500 images compared to others with almost double the sample size (Malof et al., 2016). By increasing our training data size, we would be able to improve our model and improve its generalizability to other satellite images. Also, additional approaches could yield increased accuracy, such as transferring learning from pre-built models. Future research should also examine the use of supervised learning techniques to not only identify the existence or absence of solar PV, but also estimate the energy consumption associated with each installation. This could be achieved by integrating other disparate data sources, such as geographic data and energy usage patterns.