Detecting Solar Panels in Aerial Imagery

Keerthana Kesavan
Department of Data Science
University of Colorado Boulder
keke2547@colorado.edu

Mohal Khandelwal
Department of Data Science
University of Colorado Boulder
mokh8410@colorado.edu

Puja Kumari
Department of Data Science
University of Colorado Boulder
puln8181@colorado.edu

ABSTRACT

In the US and other countries, interest in solar energy and other green energy sources is growing. The installation of solar panels has grown recently, and they are a significant source of renewable energy. Aerial photography offers an affordable way to survey huge areas for the presence of solar panels. The use of evidence-based decision-making in these areas would be supported by automating the process of data collection on solar panel installations. The objective of this project is to identify solar panels using aerial photography data using various machine learning classifiers. The proposed method involves pre-processing the aerial photos, feature extraction from those images, and then categorization using machine learning techniques. The outcomes of tests performed on a dataset of aerial photos demonstrate that the suggested method is capable of accurately detecting solar panels.

KEYWORDS

solar energy, solar panels, aerial photography, feature extraction, classifier model, geospatial information, real-time object detection, knn, neural network

1 Introduction

Numerous factors, including lower costs, solar incentive programs, investment tax credits, and high efficacy, attracted millions of Americans to environmentally friendly solar energy. In the past ten years, rooftop solar photovoltaic (PV) installations have significantly expanded in the United States. The U.S. Energy Information Administration (EIA) reports that the number of solar panel installations in the country has grown significantly in recent years. Over 47 gigawatts of solar photovoltaic (PV) systems were installed in the United States as of 2020. (GW). Compared to the installed capacity of 26 GW in 2016, this is a huge increase. With a roughly three-fold increase in installed capacity between 2016 and 2020, the residential sector has had the fastest rise in terms of the number of solar panel installations. The commercial and industrial sectors have also experienced tremendous expansion, with installed capacity more than doubling during the same time frame.

Declining costs and supportive governmental regulations are what are driving the increase in solar panel installations in the United States. The cost of implementing solar energy has decreased because of federal and state incentives including tax credits and subsidies. The rise in installations is also a response to the need for clean, renewable energy and growing environmental concerns about the effects of conventional energy sources.

Comprehensive geographic data on PV installations is an essential tool for efficiently managing, promoting, and planning the development of solar energy. Having geospatial information about photovoltaic (PV) installations is important for several reasons:

1. **Resource management:** It is possible to plan for the deployment of solar panels in places with the greatest potential by using geospatial data about PV installations to estimate the potential for solar energy generation in various locations. Additionally, this data can be used to track the development and dispersion of solar panels through time.
2. **Grid integration:** Accurate geographical data is helpful in integrating PV installations into the power grid. This data can be used to plan for grid upgrades or adjustments that may be required to handle the rising penetration of solar energy and to assess the impact of solar energy generation on the system, including the distribution of load and the possibility for congestion.
3. **Compliance with policies and regulations:** Geospatial data can be used to track adherence to policies and regulations linked to PV installations, including zoning and environmental laws. The effectiveness of government incentives and programs aimed at boosting the use of solar energy can also be assessed using this information.
4. **Research and development:** The assessment of the effects of shadowing or environmental conditions on the performance of PV installations is one example of how geospatial information can enhance research and development in the field of solar energy.

Surveys and utility interconnection filings are used to gather up-to-date information on the characteristics of solar PV installations. As a result, it is either limited, lacking in geospatial information, or frequently unavailable. However, we can try to address the problems in a relatively quick and scalable manner with the use of computer vision techniques. By gathering this data on PV installations, decision-makers will be able to make informed choices and use it to study solar development in other nations.

Fig 1.1 Solar panels on a rooftop

2 Related Work

An efficient method for checking for solar panels across wide areas is aerial imaging. However, it can take a while and be inaccurate to manually locate solar panels in aerial images. The recent acceleration of computing power and the expanding availability of digital images have helped to develop the computer vision discipline. One of the most difficult issues in computer vision is object detection, which is a hotly researched topic. In addition to traditional object detection techniques, some previous work has also explored the use of multispectral imagery and remote sensing data to enhance the accuracy of solar panel detection in aerial images. The combination of these methods has been shown to provide a more comprehensive picture of the solar PV market, improving the accuracy and reliability of the information obtained through object detection in aerial imagery. The results of these studies have demonstrated the potential of aerial imagery and remote sensing data in accurately detecting and mapping solar panels, providing valuable information for monitoring and planning the growth of the solar industry.

2.1 Previous Work in Digital Imagery Object Detection:

1. Numerous research on the detection of automobiles, roads, buildings, fruits, and other objects in digital pictures have been conducted. For instance, the following research was done on the identification of automobiles in UAV photos. A new technique for car detection developed by Moranduzzo and Melgani (2014) yielded promising results by combining an SVM classifier with two sets of HOG features (for the horizontal and vertical directions).
2. Redmon et al's, You Only Look Once is a real-time object detection technique that identifies items in a picture using a single neural network. In order to forecast the existence and placement of items within each cell, YOLO divides the image into a grid of cells.
3. RetinaNet is a one-stage object detection technique developed by Lin et al. in 2017 to address the issue of class imbalance in object detection. RetinaNet accurately detects objects of various sizes and scales using a pyramid of characteristics and a focus loss function.

There are also many other computer vision and machine learning-based approaches that have been used for object detection, including template matching, feature-based methods, and scale-invariant feature transform (SIFT).

2.2 Previous Work on PV Array Detection in Aerial Imagery:

1. Several studies have been done on the identification of solar panels in aerial imagery. The U.S. Geological Survey provided the imagery (100 photos), which was then hand labeled. Studies that used the SVM classifier yielded a 94% detection rate, proving the viability and promise of this strategy.
2. In 2016, Kim et al. proposed a machine learning-based approach to detect PV arrays in aerial imagery. The authors used a support vector machine (SVM) classifier trained on manually extracted features to distinguish between PV arrays and other objects in the images. The proposed method achieved a detection accuracy of 87.5%.
3. A transfer learning-based method to identify PV arrays in aerial imagery was put forth by Dong et al. in 2019. To increase the detection accuracy, the scientists employed a pre-trained CNN, such as VGG-16, as a feature extractor and fine-tuned it on a dataset of PV arrays. The detection precision of the suggested approach was 96.2%.
4. Li et al. suggested an attention-based technique to find PV arrays in aerial imagery in 2020. To automatically extract pertinent information and increase the detection accuracy, the authors employed a dual-attention technique.

3 Proposed Work

An efficient method for checking for solar panels across wide areas is aerial imaging. However, it can take a while and be inaccurate to manually locate solar panels in aerial images. The recent acceleration of computing power and the expanding availability of digital images have helped to develop the computer vision discipline. One of the most difficult issues in computer vision is object detection, which is a hotly researched topic.

3.1 Data

There are 2058.tif files total among the data that was gathered for this project. A picture with a resolution of 101X101 pixels is represented by each file. The pictures come from cropped sub-images of overhead satellite photography data taken in Fresno, California, in 2013, a city in the United States with a lot of solar panel installations. Each image has been ortho-rectified and has a resolution of 0.3 meters per pixel. 1500 of the 2058 photos used in the study have labels indicating whether solar panels are present (Y=1) or not (Y=0). To assess the model's performance on unobserved data, the remaining 558 photos were left unlabeled. 33.7% of the photographs contain solar panels, which demonstrates

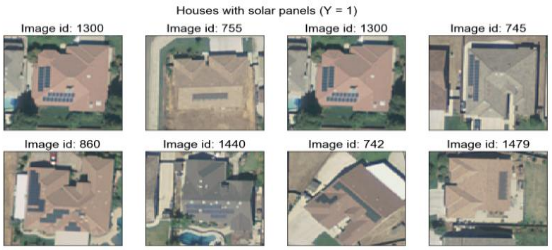

Fig 3.1.1 Positive examples of solar panels

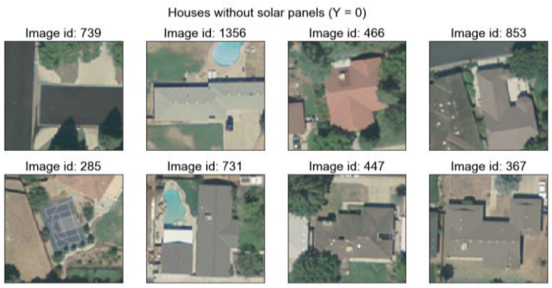

Fig 3.1.2 Negative examples of solar panels

a class divide. The images are in color, which means that each one has three color channels (red, green, and blue), creating a 101 × 101 x 3 pixel array. Figures 3.1.1 and 3.1.2 represent aerial satellite views of homes, some of which have rooftop solar panels and others of which do not.

3.2 Initial Data Analysis

3.2.1 Finding contours and Edge Detection: Identifying contours in solar panel photos is a crucial step in determining whether an image contains solar panels. The shapes and bounds of the solar panels can be recognized using contours, which are the edges of objects in an image. In this perspective, it's crucial to understand why we found contours:

1. Finding the contours enables us to distinguish the panels from the background, which is important because solar panels can appear visually identical to their surroundings. This is crucial for precisely locating and describing the panels' shape.
2. Calculating the panels' dimensions and form. The contours offer details on the size and shape of the solar panels, which can be used to determine whether the photos contain panels. For instance, it is likely that solar panels are present if the outlines are broad and rectangular.
3. Finding the panel orientation: The contours might also reveal information about the panel orientation. Understanding how the panels are positioned in the image can help you better understand how this affects their effectiveness.
4. Images that lack solar panels can be removed by applying a filter that uses outlines to identify their presence. This can help when studying huge datasets because it can save time and processing resources.

3.2.2 Statistical Properties: Statistical features are quite important in this aerial photography project. These characteristics make the underlying distributions and patterns in the picture data visible,

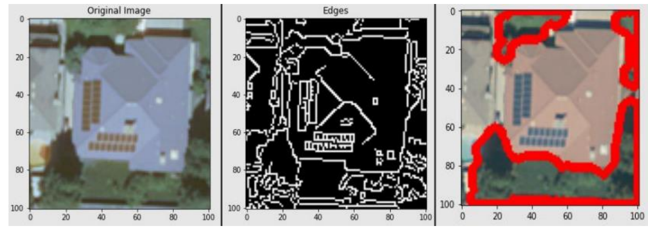

Fig 3.2.1 Finding contours and edge detection.

which makes it easier to choose the best techniques and models for classifying images. The asymmetry of the pixel value distribution is shown by kurtosis and skewness. Understanding these statistical aspects can help in making decisions about the best preprocessing procedures, feature extraction approaches, and hyperparameter optimization for classification models. Therefore, when working with image data in a solar panel detection project that makes use of aerial photography, considering statistical properties is essential for precise and reliable classification outcomes. The correlation on the RGB channels of the image data is crucial. The presence and placement of solar panels in an image can be determined using the RGB channels, which offer information about the intensity of red, green, and blue colors in each pixel. The existence of solar panels in the image can be determined by comparing the RGB channels in order to find spatial patterns and correlations between the pixel values. It is possible to improve the contrast between the solar panels and their surroundings, making them simpler to discern and notice, by figuring out the relationships between the RGB channels. Additionally, correlation on the RGB channels can help to lessen noise and boost the detection algorithm's overall accuracy.

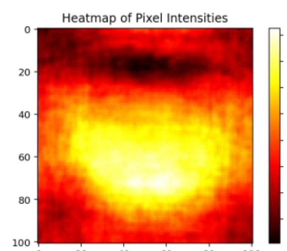

Fig 3.2.2 Heatmap of Pixel Intensities

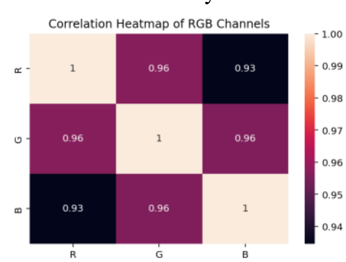

Fig 3.2.3 Correlation of RGB channels

Fig 3.2.3 is a correlation heatmap of RGB channels of the images in the dataset's mean and standard deviation. The average brightness of the images in each color channel is shown by the mean value. The standard deviation shows the degree of variance in the pixel values for each color channel across the dataset. How should you interpret the values you got? Mean: [0.50514135 0.51940146 0.47810523] Approximately 0.51, 0.52, and 0.48 are the mean values for the red, green, and blue channels, respectively. This shows that the dataset's photos tend to be slightly more green than red or blue overall. Std: [0.17305651 0.13798065 0.11417361] The red, green, and blue channels' respective standard deviation values are roughly 0.17, 0.14, and 0.11 accordingly. This implies that while there is some variation in the pixel values for each color channel across the sample, the variation is not very significant. These numbers can be helpful for analyzing the dataset's features and choosing the best image preparation methods. The brightness or color balance of the photos, for instance, can be altered using the mean values to match a target. The values of the standard deviation can be used to find outliers in the dataset or to normalize the pixel values.

3.2.3 Data pre-processing: We have rescaled the images. Our original images have RGB coefficients between, but such values would be too high for our models to process (when compared to given a typical learning rate), so we target values between 0 and 1 instead by scaling with a 1/255 factor. We performed pre-processing by identifying two key properties were identified: the ability of solar panels to absorb light and their angular shape. By adding features that capture these properties, the accuracy and effectiveness of the image classification algorithms was improved. Each image had gradient elements added in order to accurately depict the angular design of solar panels. These characteristics signify variations in color saturation and brightness throughout the image, which can be used to locate the corners and edges of solar panels. The algorithms can more effectively discriminate between solar panels and other items in the image that do not have the same angular shape by recognizing these traits. We added brightness cues to each image to capture the characteristic of solar panels absorbing light. These characteristics assist distinguish solar panels from other things that may have similar angular shapes but may not absorb as much light. They represent the amount of light absorbed by various sections of the image. The algorithms can identify solar panels more correctly and decrease false positive detections by adding these features to the training dataset.

Fig 3.2.4 Luminance and Gradient Features

4Methods

HOG Features: HOG is an abbreviation for "Histogram of Oriented Gradients." Typical applications for these features include object detection and recognition. HOG features' main objective is to extract local gradient data from a picture, which can be used to identify the edges and forms of objects in the image. An image is initially divided into small cells, and the gradient magnitude and direction for each pixel within each cell are determined, in order to calculate HOG features. While the gradient direction indicates the edge's orientation, the gradient magnitude indicates how strong the edge is. The gradient directions are then oriented-based binned into a histogram, where the histogram bins reflect various edge orientations. The local gradient information for the entire image is captured by these histograms, which are averaged over several cells. Machine learning algorithms can then use the generated HOG features as input to carry out object detection and recognition tasks. To further improve the accuracy and durability of the algorithms, HOG features are frequently combined with other feature descriptors and methods, such as support vector machines (SVM) or convolutional neural networks (CNN).

Test-Train Split: The split between training and test was 80-20. A ratio of classes proportional to the original dataset should be used when splitting the data in both the training and validation sets.

The detection will be carried out by both traditional machine learning and contemporary deep learning methods. We'll employ the following techniques:

1. HOG feature-trained **Logistic Regression**. Following preprocessing, a baseline model—consisting of a fundamental logistic regression model—will be trained using HOG features. The baseline model provides a statistic that we can use to compare the performance of the simpler models listed below, giving us a decent place to start when trying to solve the solar panel detection problem. By calculating the receiver operating characteristic (ROC) curve's area under the curve (AUC), we assessed the performance of our models. Due to the imbalance in the target class, AUC provides a reliable indicator of how well a machine learning model is performing. This is because AUC is high when both the true positive and true negative rates are high.
2. HOG feature-trained **K nearest neighbors**. The distribution of gradient orientation in a picture is described by a particular class of feature extraction known as HOG features. Important details about the texture and geometry of objects in a picture can be recorded using HOG features. This technique uses the KNN algorithm to identify solar panels in satellite photos by training it on HOG features. In order to forecast the label of a brand-new, unseen image, the algorithm locates the K nearest neighbors in the feature space. In the case of solar panel detection, the algorithm would decide as to whether a solar panel is present in an image by using its HOG properties.
   Prior to training the algorithm, we had to first select a set unique hyperparameters for the KNN model. The range of neighbors (1 to 30) that were considered while calculating the separation of a test data point and the training data point was tested with. We also explored whether the impact of a neighbor should be weighted according to the proximity between the test point and a neighbor, using various distance metrics (Manhattan or Euclidean). Finally, we evaluated the results using 5-fold cross-validation while tuning the hyperparameters by randomly selecting samples from the hyperparameters space.
3. In the realm of computer vision, **Convolutional Neural Networks** (CNNs), a subset of Artificial Neural Networks, are frequently used. CNNs maintain the spatial organization of an image, unlike conventional artificial neural networks, and approach the image from a human perspective. An input layer, seven convolutional

layers, and an output layer made up the CNN structure. ReLu activation and batch normalization are performed after each convolution operation since they have both been shown to be efficient in handling the vanishing gradient problem and accelerating the convergence of the cost function. When training CNN, we gave the minority class 1 more weight, which significantly increased the AUC compared to when we gave both classes the same amount of weight. With 16 filters and a kernel size of 3x3, our initial CNN had two layers (input layer, hidden layer, and output layer). The filter size was chosen in accordance with widely used image classifiers, which employs a 3x3 filter. Other filter sizes, such 5x5 and 7x7, were also tested, but 3x3 appeared to perform the duty of picture classification the best. We first chose the filter size and activation function, and then we simultaneously raised the number of layers and filters until we noticed overfitting. For our image classification challenge, we discovered that a CNN with seven convolutional layers with global max pooling at the end performs better than one with a fully connected layer.

5 Evaluation

The performance of the logistic regression model on HOG (Histogram of Oriented Gradients) features can be evaluated using the area under the curve (AUC) in a receiver operating characteristic (ROC) curve. The receiver operating characteristic (ROC) curve provides a graphical representation of the performance of a binary classifier system, here for the logistic regression model. The curve is plotted by evaluating the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold values.

True Positive Rate (TPR) = True Positives / (True Positives + False Negatives)

False Positive Rate (FPR) = False Positives / (False Positives + True Negatives)

ROC curve:

The ROC (Receiver Operating Characteristic) curve is a graphical representation of a binary classification system's performance. It is often used to analyze and compare the performance of various binary classification models or algorithms. At various threshold values, the ROC curve plots the True Positive Rate (TPR) versus the False Positive Rate (FPR).

The True Positive Rate (TPR) is the proportion of accurately categorized positive instances to total positive instances. In other words, it is the proportion of positive examples that the classifier properly identifies. The False Positive Rate (FPR), on the other hand, is the ratio of mistakenly classified negative occurrences to the total number of real negative instances.

The ROC curve visualizes and compares the trade-offs between TPR and FPR at various threshold levels. A threshold value is a value that the classifier uses to decide whether a projected probability is positive or negative. We can modify the balance of TPR and FPR by varying the threshold value. A lower threshold, for example, may categorize more cases as positive, resulting in a higher TPR but also a higher FPR.

The ROC curve's area under the curve (AUC) is a single metric that represents the classifier's overall performance. The results of the examination vary from 0.5 (random guess) to 1 (excellent classifier). A higher AUC value implies that the classifier is effective at distinguishing between positive and negative cases. An AUC of 0.5 shows that the classifier is no better than chance.

The area under the curve (AUC) of the ROC curve is a single metric that represents the classifier's overall performance. The evaluation results vary from 0.5 (random guess) to 1 (excellent classifier). A higher AUC value shows that the classifier can efficiently distinguish between positive and negative events. An AUC of 0.5 shows that the classifier is no better than random guessing.

The ROC curve is a helpful tool for evaluating and comparing binary classifier performance. It can help to pick the appropriate threshold value for a given classification problem by providing insights into the trade-offs between sensitivity (TPR) and specificity (1-FPR). It should be emphasized, however, that the ROC curve is susceptible to class imbalance and does not provide information about the classifier's overall accuracy or mistake rate. When evaluating binary classifiers, other metrics such as precision, recall, and F1 score should be considered in addition to the ROC curve.

Next for the evaluation of the KNN(K-Nearest Neighbours) and CNN (Convolution Neural Network) models, 5 - fold cross validation and 3 - fold cross validation can be used respectively.

5-fold cross-validation:

5-fold cross-validation is a typical technique in machine learning for evaluating model performance. The 5-fold cross-validation method divides the dataset into five roughly equal folds, with each fold serving as a test set for the model and the remaining four folds serving as the training set. The model is trained using the training set and assessed using the test set five times, with each fold serving as the test set once.

We can acquire a more trustworthy estimate of the model's performance by employing 5-fold cross-validation, which accounts for the variability in the training and testing data. In contrast, utilizing a single train/test split can produce misleading findings if the split is not representative of the total dataset. Another advantage of employing 5-fold cross-validation is that we may make better use of our data. We may use all the data for both training and testing by rotating through the five folds as test sets, which is very beneficial when working with smaller datasets.

In addition to model evaluation, 5-fold cross-validation can be utilized for hyperparameter adjustment. Hyperparameters, which are parameters that are set prior to train the model, such as the learning rate or regularization strength, can have a major impact on the model's performance. We can identify the set of hyperparameters that offers the highest performance on the data by using 5-fold cross-validation to evaluate alternative hyperparameter values.

Overall, 5-fold cross-validation is an effective strategy for evaluating and optimizing machine learning models. It enables us to make better use of our data, obtain a more trustworthy assessment of the model's performance, and fine-tune the hyperparameters for maximum performance.

The performance of a model in machine learning is evaluated using validation approaches, one of which is known as "k-fold cross-validation." This is a statistical method for estimating machine learning model performance by dividing the dataset into k subsets, or folds, where k is an integer. The data is divided into k subsets in this procedure, with one subset serving as the validation set and the remaining k-1 subsets serving as training.

3-fold cross-validation:

The 3-fold cross-validation approach is one of several variants of the k-fold cross-validation technique, where k = 3. The dataset is partitioned into three equal portions in this method, and the model is trained and assessed using three separate subsets of the dataset. Two subsets are utilized for training the model in each iteration, and the remaining one is used for validation. This process is performed three times, with the model's average performance across all three folds serving as the final evaluation statistic.

The key advantage of utilizing 3-fold cross-validation over single train-test splits is that it delivers a more accurate estimate of the model's performance. This is due to the model being assessed using three separate subsets of the dataset, and the performance measures being averaged across all three rounds, which reduces the variability in the results.

Another benefit of employing 3-fold cross-validation is that it decreases the possibility of overfitting. When a model is overly complex, it learns the noise in the training data rather than the underlying pattern. We can train the model on multiple subsets of the data and check that it generalizes well to new data by using cross-validation.

The key advantage of utilizing 3-fold cross-validation over single train-test splits is that it delivers a more accurate estimate of the model's performance. This is due to the model being assessed using three separate subsets of the dataset, and the performance measures being averaged across all three rounds, which reduces the variability in the results.

Another benefit of employing 3-fold cross-validation is that it decreases the possibility of overfitting. When a model is overly complex, it learns the noise in the training data rather than the underlying pattern. We can train the model on multiple subsets of the data and check that it generalizes well to new data by using cross-validation.

F1 score, Precision, Recall:

A popular metric for assessing the effectiveness of classification model performance is the F1 score. It evaluates how well a model predicts the appropriate class labels for a given dataset. Precision and recall are two additional measures that are factored into the F1 score. The capacity of the model to accurately forecast the positive class is measured by precision, and its proficiency in properly identifying all positive cases is measured by recall.

The precision is computed by dividing the overall number of accurate positive predictions by the total number of genuine positives. When a model predicts a positive instance, it is more likely to be accurate since it produces fewer erroneous positive predictions when it has a high accuracy score. Precision is crucial in applications like fraud detection and medical diagnosis where false positives can have serious repercussions.

On the other hand, recall is computed by dividing the total number of positive instances in the dataset by the proportion of true positives. When a model has a high recall score, it suggests that it correctly recognizes many of the positive cases in the dataset, which means that it is likely to recognize positive instances when they occur. In applications like illness diagnosis or security screening, where false negatives might have serious repercussions, recall is especially crucial.

A harmonic mean of memory and precision makes up the F1 score. As 2 * (precision * recall) / (precision + recall), it is calculated. When both precision and recall are crucial, the F1 score is an effective tool for assessing models. When the dataset is unbalanced between positive and negative cases, it is especially helpful. In these situations, the F1 score, which accounts for both precision and recall, can provide a more realistic assessment of the model's overall performance.

6 Results

Baseline Model. While the logistic regression model trained on HOG features achieved an AUC score of 0.718, indicating the model can accurately distinguish positive from negative samples, it appears to perform less well than other models discussed here. The blue curve in the ROC curve plot Figure 6.1 for this model shows a decent balance between true-positive and false-positive rates, however the PR curve plot exposes an imbalance between recall and precision rates, with lower recall rates for higher precision rates, as shown by the blue lines on its PR plot. AUC scores and PR curves must be compared in order to assess a model's performance in relation to other models. AUC scores indicate how accurately a model ranks positive and negative data, and PR curves show how the accuracy and recall tradeoffs change as threshold values are increased or decreased. In order to determine which model works best and pinpoint possible areas for development, compare the AUC scores and PR curves of different models that are offered inside a certain task and dataset. This comparison should help establish which model has proven most suitable.

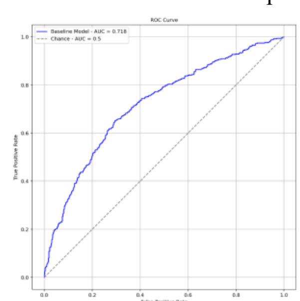

Fig 6.1 ROC plot

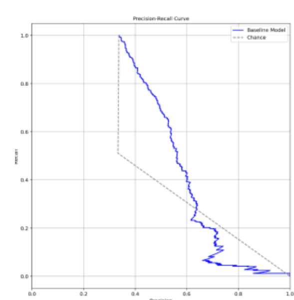

Fig 6.2 Precision-Recall plot

KNN Model. The ideal K-neighbor value for the HOG-KNN pipeline, according to our iterative procedure described in the techniques section, was five. L1 had the best performance in terms of distance measures. In terms of picking an efficient weighting technique, we discovered that distance-based weighing gave us the most satisfying results in our tests, yielding an AUC score of 0.812. This superior performance is represented in Figure 6.4 by its green curve. The model significantly outperformed the baseline model by achieving greater TPR and TNR rates than the baseline model at each threshold. A PR curve is also shown in Figure 6.5, which demonstrates that the KNN model (green curve) offered a greater precision to recall ratio than the baseline model (blue curve). The misclassifications are still present despite its relative success, as shown in Figure 6.4. As you can see in the graph below, this model performs somewhat better when the rooftop and solar panel form an increasing color contrast and there isn't a corresponding contrast in the image. Predictions that turn out to be true less frequently. If there are no solar panels visible but our model predicts that there should be, we can look for signs of solar panels by looking for similar contrast in the image (even if there are none), which suggests their presence (for example). This model might be able to anticipate, for instance, that a house will have a shade like that of solar panels. Examples of false-positive predictions are scenarios in which our model predicts the absence of solar panels in a photo when they are present. On the other hand, false-negative prediction examples occur when actual solar panels do exist, but our model incorrectly predicts they do not. This is apparent if an object containing solar panels appears in an image but is excluded from being displayed correctly in search results. A model may misclassify itself if the contrast between the color of a solar panel and that of the rooftop is significantly higher or lower than expected.

Fig 6.3 KNN output

CNN Model. After threefold cross-validation, our final CNN model with ideal hyperparameters outperformed even KNN by a significant margin, achieving an AUC score of 0.992. Figure 5's (red curve) illustration demonstrates that its robustness rests in its capacity to accurately distinguish between positive and negative samples. However, we discover some misclassifications in Figure 4 when we examine the false positives and false negatives of our model. It was difficult for the algorithm to distinguish between positive and negative samples because most false-negative cases had solar PVs that either covered a small section of a picture or had colors that blended in with neighboring features like roof colors or shadows. On the other hand, this model's misinterpretation of shadows and dark hues as solar panels resulted in many false-positive occurrences. For instance, this algorithm correctly recognized solar panels in photos with dark-gray basketball courts (image ID 285) and rooftops (image ID 313).

Our model performed best when there was a high contrast between the solar panel and the roof, as shown by an analysis of true positives. This finding suggests that the performance of our model could be further improved by including features that account for color contrast and spatial relationships between these components of a model roof-solar power system. Consequently, even though the final CNN model with the best hyperparameters had an exceptional AUC score, there is still room for advancement.

6.1 Model Accuracy plots

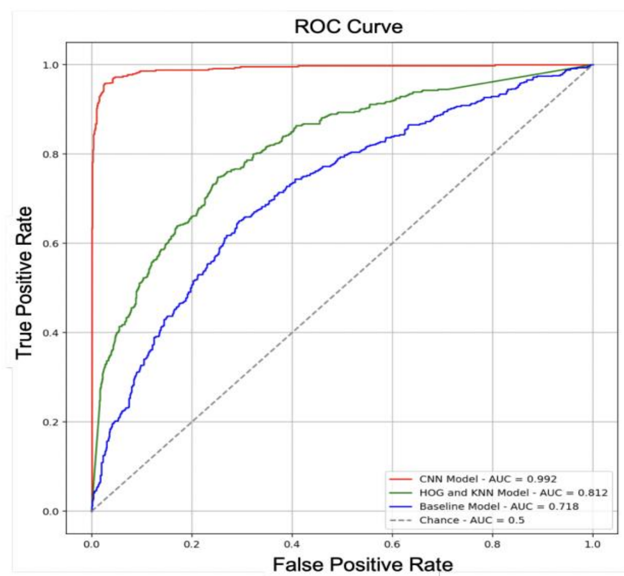

Fig 6.4 ROC Curve for all three models

Model	Metric	Validation Set Performance
Logistic Regression	AUC	0.718
KNN	AUC	0.812
CNN	AUC	0.992

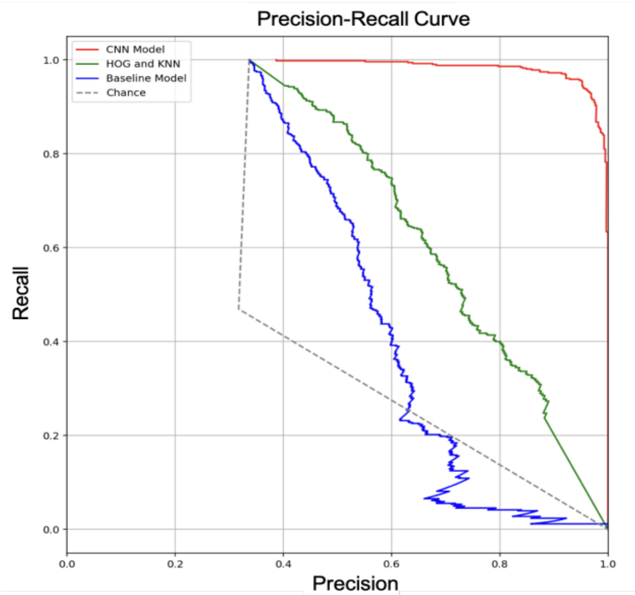

Fig 6.5 PR plot for all three models

7 Conclusion

As renewable energy sources become more and more popular all over the world, the ability to identify solar panels in aerial photography is becoming more and more crucial. Being able to precisely locate and track solar panels can aid in maintenance and monitoring efforts as solar panels can be a significant investment for both businesses and individuals. Additionally, the detection of solar panels in aerial imagery can assist governments and organizations in monitoring the progress of the transition to a more sustainable energy future. Solar panels are a crucial part of this transition.

Solar panels can be found in aerial imagery using a variety of methods and tools, such as object detection algorithms, machine learning, and remote sensing. These techniques can deliver precise and fast information about the placement and condition of solar panels and can be very effective when used in combination. The precision and effectiveness of solar panel detection are likely to keep improving with technological advancements, bringing more advantages to people, businesses, and society at large.

In general, being able to spot solar panels in aerial images is a useful tool for encouraging sustainable energy practices and making it possible to effectively monitor and maintain solar installations. The significance of this technology will only rise as the usage of renewable energy sources expands, and continual research and development in this field will be necessary to guarantee the ongoing success of renewable energy efforts.

8 Future Work

Significant progress has been made in the field of correctly identifying solar panels using aerial imagery for the project utilizing aerial photography to detect solar panels. The model could still perform better in a few specific instances, especially when the solar panels are small or have colors that mix in with the surroundings. One method is to raise the input image resolution or employ sophisticated object detection methods like object suggestions or region-based approaches in situations when the solar panels are small. The model can gather more specific information about the solar panels and help identify them from other items in the image by raising the resolution of the input images. One method is to add more features to the training dataset that capture the distinct color and texture attributes of solar panels in order to improve the model's performance in scenarios where the solar panels' colors merge with the surroundings. To help the model distinguish between solar panels and other items with similar colors and textures, the dataset could be supplemented with texture-based characteristics like local binary patterns (LBP) or color-based features like hue, saturation, and value (HSV).

Occlusion and clutter can make object detection less accurate, which is one of the difficulties in using aerial photography to detect solar panels. Future work may investigate several strategies to overcome these issues and enhance the project's performance. Using sophisticated object detection techniques, such as region-based techniques that concentrate on certain areas of the image where solar panels are likely to occur, is one option to tackle occlusion and clutter in the photographs. This method can lessen the amount of clutter in the image and enhance the recognition of solar panels that are partially obscured. Using deep learning methods like convolutional neural networks (CNN), which can automatically learn to detect characteristics that are robust to occlusion and clutter, is an alternative strategy. Incorporating more prior knowledge and domain knowledge into the model is another option. For instance, if the position and orientation of the solar panels are known in the image, this knowledge can be utilized to direct the object detection algorithm to concentrate on these areas and increase the detection's accuracy. Additionally, the model is better able to distinguish solar panels from other objects in the image by considering information about the anticipated size, shape, and orientation of solar panels. Expertise and knowledge in the installation and maintenance of solar panels could be used for this.

Ensemble models are a prominent strategy in the field of machine learning to enhance model performance. This method might be used for the aerial photography-based solar panel detection project. The accuracy and resilience of the solar panel identification model may be improved in the future work in this field by investigating different ensemble strategies as bagging, boosting, or stacking. In order to use ensemble models in this project, many object detection models with various architectures would need to be trained, and their predictions would need to be combined to give a result. Examples include using gradient and luminance characteristics in one model while HOG features in another. The solar panel detection algorithm's overall accuracy and dependability could be increased using this technique. In addition, strategies like bagging, boosting, or stacking could be used to improve the model's performance even further.

9 References

[1] Zhang, Y., Zhang, D., & Fan, X. (2017). A deep learning-based approach for photovoltaic panel detection in high-resolution aerial imagery. Renewable Energy, 107, 500-509.
[2] Dong, H., Liu, Z., & Chen, Y. (2019). A transfer learning-based approach for photovoltaic panel detection in aerial imagery. Remote Sensing, 11(17), 1997.
[3] Li, S., Li, C., & Zhang, X. (2020). Dual-attention mechanism for photovoltaic panel detection in aerial imagery. Remote Sensing, 12(18), 2857.
[4] Redmon, J., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 779-788).
[5] Goldberger, J., Roweis, S., Hinton, G., & Salakhutdinov, R. (n.d.). Neighbourhood Components Analysis. 8. Camilo, J., Wang, R., Collins, L. M., Bradbury, K., & Malof, J. M. (2018). Application of a semantic segmentation convolutional neural network for accurate automatic detection and mapping of solar photovoltaic arrays in aerial imagery. ArXiv:1801.04018 [Cs]. http://arxiv.org/abs/1801.04018
[6] Hu, J., Shen, L., Albanie, S., Sun, G., & Wu, E. (2019). Squeeze-and-Excitation Networks.
[7] Malof, J. M., Bradbury, K., Collins, L. M., & Newell, R. G. (2016). Automatic Detection of Solar Photovoltaic Arrays in High Resolution Aerial Imagery. Applied Energy, 183, 229–240. https://doi.org/10.1016/ j.apenergy.2016.08.191
[8] Zou, Z., Shi, Z., Guo, Y., & Ye, J. (2019). Object Detection in 20 Years: A Survey. ArXiv:1905.05055 [Cs]. http://arxiv.org/abs/1905.05055
[9] W. Yin, S. Lingxin, L. Maohuan, S. Qianlai and L. Xiaosong, "PV-YOLO: Lightweight YOLO for Photovoltaic Panel Fault Detection," in IEEE Access, vol. 11, pp. 10966-10976, 2023, doi: 10.1109/ACCESS.2023.3240894.
[10] V. Golovko, A. Kroshchanka, S. Bezobrazov, A. Sachenko, M. Komar and O. Novosad, "Development of Solar Panels Detector," 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 761-764, doi: 10.1109/INFOCOMMST.2018.8632132.