Fault Detection of Solar PV system using SVM and Thermal Image Processing

Karuppiah Natarajan*, B. Praveen Kumar **†, Vankadara Sampath Kumar***

*Department of Electrical and Electronics Engineering, Vardhaman College of Engineering, Hyderabad (501 218), Telangana, India.

**Department of Electrical and Electronics Engineering, Bharat Institute of Engineering and Technology, Telangana, India

***Department of Electrical and Electronics Engineering, Bharat Institute of Engineering and Technology, Telangana, India

*atarajankaruppiah@gmail.com, praveenbala038@gmail.com, sampath.vankadara62@gmail.com

† Praveen Kumar B, Bharat Institute of Engineering and Technology, Hyderabad, Telangana, India, Tel: +91 96296 32445, praveenbala038@gmail.com

Received: 10.04.2020 Accepted: 07.05.2020

Abstract- Installation of photovoltaic plants across the globe increases, in the recent years, due to the energy demand across the world. Solar energy is free of cost, inexhaustible and a non-polluted source to the environment. The efficiency of any power plant depends on its fault free operation. Due to the occurrence of fault in PV system, the reliability and power output is reduced. In PV plants, the internal and external faults normally result in an increase in temperature which is easily sensed by different methods. In this paper, an algorithm based on thermal image processing, is proposed to extract the features of the PV cells in operation. These extracted features are then compared with the features of the healthy PV module using Support Vector Machine. SVM is a classifier tool which classifies whether the PV modules are defective or non-defective. An experimental set up is created and the performance of the algorithm is verified by testing it with faulty data sets which are obtained by creating different types of faults intentionally. The algorithm successfully identifies the defective PV module and its performance is validated experimentally.

Keywords PV faults, image processing, thermal images, texture information, maximum power.

1. Introduction

Nowadays the rapid growth in renewable energy has reduced the usage of fossil fuels. Since PV technology is an eco-friendly resource its usage has increased many folds. This reduces the greenhouse gas emissions to a great extent [1-4]. Due to the climatic changes in the recent years, the PV panels experience stresses and thereby the performance is much affected. Normally the power output of a PV module depends on the irradiation (W/m²) and the temperature (°C) [5-6]. The environmental changes have a direct impact on the amount of irradiance and temperature which results in the deterioration of PV module’s efficiency [7-8]. In addition to this, partial shading, humidity, high temperature, bird droppings, dust accumulation etc., affects the efficiency of the PV system. Also, hotspots caused by the internal faults results in power loss of the PV system [9-13].

Researchers of today work to maximize the performance of the PV system thereby ensuring its reliable operation [14-16]. Depending on the characteristic curves of PV system, its state of operation is classified. A PV module gives its power output based on the irradiance and temperature level. The power output also depends on the maximum current, voltage and power of the module. Based on the manufacturer’s data and the irradiance, temperature level the normal power output of the PV module is estimated [17-18]. This power output is termed as the normal operating state of the PV module. If the power output varies drastically from its expected normal output then the PV module is said to be faulty. Normally PV plant is an array of PV modules in which number of PV cells are connected in series and parallel depending on its desired power output. If the power output varies from the expected level, one or more PV cells could be faulty, depending on the severity of fault. Hence a fault monitoring and diagnostic system is to be developed for
the better operation of the PV plant [19-20]. The different types of possible faults on the PV system is listed in Table 1.

Table 1. Types of Faults

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Fault Location</th>
<th>Types of Fault</th>
</tr>
</thead>
</table>
| 1     | On PV module   | 1. Encapsulation failure  
|       |                | 2. Back sheet adhesion loss  
|       |                | 3. Cells cracking  
|       |                | 4. Broken interconnection  
|       |                | 5. Shading & soiling  
|       |                | 6. Hotspots  
|       |                | 7. Module corrosion  
|       |                | 8. Potential induced degradation  
|       |                | 9. Light induced power degradation  |
| 2     | Inverter Failure | 1. Manufacturing / Designing failure  
|       |                | 2. Control Problems  
|       |                | 3. Electrical component failure  |
| 3     | Other Failure Modes | 1. Unbalance of the system  
|       |                | 2. Junction box failure  
|       |                | 3. By-pass diode failure  
|       |                | 4. Mismatch fault  
|       |                | 5. Ground Fault  
|       |                | 6. Line-line fault  
|       |                | 7. Arc Fault  |

Apart from the monitoring system it is necessary to have fault detection system to avoid the failure [21-22]. In general, major fault has been developed from several minor faults. Thus, the major fault is prevented by identification at the minor level [23-24]. In earlier literatures, the electrical data was used as a measure for identifying the fault [25-26]. Later the similar problems were resolved using soft computing techniques by considering it as an optimization problem. ANN based fault detection technique measures the MPP voltage and current from the MPPT controller and classifies the operating state as, normal state, degraded state, shaded state and the short circuit fault state. But this method was not successful in preventing major faults due to its limitations [27-28]. Another soft computing-based method is decision tree-based algorithm. In this method irradiation, current and voltage are the reference values and are used to train the algorithm and it flows like branches and leaves like the tree. Based on the number of trees and branches the type and level of fault was determined [29]. Mutually comparing algorithm, minimum covariance determinant estimator-based algorithm, analysis of variance-based algorithm are some other soft computing techniques used for determining fault in PV system [30]. Recent literature regarding partial shading, MPPT algorithms, Lab VIEW and reconfiguration techniques for solar panel applications and fault detection were analyzed [31-36].

Performance Ratio of the PV system is the ratio between the instantaneous values to the nominal value. Many researchers have developed fault monitoring and diagnostic system which do not take account of the instantaneous value of power output. Hence the severity of fault could not be estimated. A new kind of fault detection method has been discussed in this paper. This approach is based on the thermal image processing with a machine learning tool. The thermal camera is used for taking samples from the PV system and thereby extracting features of it. Using the machine learning tool like SVM, the features are fed into it and by comparing it with the reference samples it is classified into two types as defective and non-defective.

2. Creation of Faults

Different types of faults that are intentionally created are cracks, hot spots due to shading and soiling. The nature of different types of faults, possibility of occurrence of faults and its effects are discussed. These faults are intentionally created on a 3*3 PV array as shown in Figure 1 and its behaviour is checked to identify the healthy and faulty PV cell.

![Fig. 1. Connection of PV cells in 3*3 array](image)

2.1. Crack

Cracks on the PV panels are formed during transportation, severe climatic disturbances such as strong
winds, hail storms and improper handling of the panels during installation etc. The formation of cracks results in the decreased PV power output. Micro cracks are mainly due to the thermal stresses on the solar panel. During the production, these micro cracks occur on the surface of the panel during its lamination. Also, when materials like copper and silicon expands it creates micro cracks resulting in a higher resistance of the PV cell. Figure 2 shows the intentional cracks created on PV cells numbering 12, 23 and 31 in a 3*3 PV array due to thermal stress on the module.

![Cracks created on cells 12, 23 and 31](image)

Fig. 2. Intentional cracks made on 3*3 PV array

2.2. Soiling

Soiling refers to the accumulation of soil, dust particles, etc. on the solar module. This soil accumulation restricts the amount of solar irradiance to pass into the module. This leads to the reduction of power output of the solar module. The factors affecting soiling and the power loss are climatic conditions, tilt angle of modules and type of liquid used for cleaning.

**Climatic conditions:** The dry condition with continuous flowing of wind would cause accumulation of soil and dust particles on the solar module and this process is called as cementing. This cemented dust particles are mostly irremovable which causes a fixed reduction of power output from the module and further damaging the module permanently.

**Tilt angle of modules:** The optimum tilt angle is the latitude of a particular location. Due to space constraints and to reduce the shadowing effect, sometimes the tilt angle is kept low from the actual. In such cases more dust accumulation occurs on the solar panel and decreases the power output.

**Type of liquid used for cleaning:** This factor has a direct impact on the efficiency of solar module. The chemical composition of the liquid used for cleaning contributes to this cause. Generally small droplets would be left over on the panel after the cleaning process. When this left-over cleaning liquid is evaporated, the chemical substance left over accumulates on the surface of the panel and restricts the amount of sunlight falling on it. Figure 3 shows the intentional soiling fault created on PV cells numbering 11 and 31 in a 3*3 PV array.

![Accumulation of soil on cells 11 and 31](image)

Fig. 3. Intentional accumulation of soil on 3*3 PV array

2.3. Hotspots due to shading

Hot spots appear when a solar cell within a module generates less current than the string current of the module. Hot spots are due to shading of trees and internal module failures. This results in temperature difference which affects the output current. Figure 4 shows the intentional hot spots created on PV cells numbering 13, 21, 32 and 33 in a 3*3 PV array. Full shading is shown in PV cells numbering 13 and 21. Similarly partial shading is shown in PV cells numbering 32 and 33.

![Full shading (on cell 13, 21) and partial shading (on 32, 33)](image)

Fig. 4. Intentional Hot spot on 3*3 PV array

2.4. Combined Faults

Faults due to soiling, hot spots, internal failures and cracks are combined on the PV array and its behaviour is identified using thermal imaging and the faulty and non-faulty samples are classified using SVM. Figure 5 shows the combined faults created on PV cells numbering 13, 23, 32 and 31 in a 3*3 PV array. Crack is shown in PV cells numbering 13, 23 and 32. Similarly soiling is shown in PV cells numbering 31. Figure 6 shows the combined faults created on PV cells numbering 13, 21 and 33 in a 3*3 PV array. Shading is shown in PV cells numbering 13 and 21. Similarly soiling is shown in PV cells numbering 33.
3. Methodology

The fault detection in the PV array is done using a thermal imaging camera and Support Vector Machine. The flowchart of the proposed algorithm is shown in Figure 7. The image of a healthy PV panel is captured using a thermal imaging camera. Then its features are extracted using MATLAB image processing code. The extracted features are then trained and considered as standard data set for Support Vector Machine. Now intentionally different faults are introduced on the surface of PV panel. Then the combined faulty PV panel’s image is captured using a thermal imaging camera. Its features are again extracted using MATLAB code and trained and then compared with the standard data set using SVM. If the data lies between the hyperplane in SVM then the PV panel is said to be in healthy state. Otherwise it is considered as faulty.

Contrast is the information about the local variation in the GLCM expressed as per equation (1)

\[ \text{Contrast} = \sum_{i,j=0}^{n-1} (i-j)^2 m_{ij} \]  

(1)

Correlation is the measure of the joint probability occurrence of the specific pixel pairs as per equation (2)

\[ \text{Correlation} = \sum_{i,j=0}^{n-1} (i-\mu)(j-\mu)m_{ij}/(\sigma^2) \]  

(2)

Energy is the texture information which provides the sum of squared elements in the GLCM. It also known as the uniformity or the angular moment as shown in equation (3)

\[ \text{Energy} = (\sum_{i,j=0}^{n-1} m_{ij}^2)^{1/2} \]  

(3)

Homogeneity is the measure of the closeness of the distributed elements in the GLCM to the GLCM diagonal as in equation (4)

\[ \text{Homogeneity} = \sum_{i,j=0}^{n-1} m_{ij} / \sqrt{1 + |i-j|^2} \]  

(4)

Where i, j are the rows and columns of the matrix

- \( m_{ij} \) is the pixel location
- \( n-1 \) is the total no of rows and columns
- \( \mu \) is the mean
- \( \sigma \) is the variance

The texture information of the healthy panels at standard test condition (STC) is shown in Table 2.

<table>
<thead>
<tr>
<th>Texture</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast</td>
<td>0.1282</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.7223</td>
</tr>
<tr>
<td>Energy</td>
<td>0.5407</td>
</tr>
<tr>
<td>Homogeneity</td>
<td>0.9399</td>
</tr>
<tr>
<td>Entropy</td>
<td>5.3294</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>11.8869</td>
</tr>
<tr>
<td>Mean</td>
<td>88.8742</td>
</tr>
</tbody>
</table>
For classifying the PV cell using the derived texture information, SVM based classifier is used. SVM is one of the standard and simplest machine learning algorithms used for the classification problems as compared to other soft computing techniques like neural networks.

3.1 Data Raw Input

The texture information of the thermal image is treated as the raw input to the SVM classifier. Based on this textural data, the SVM classifies the PV cell into two classifications as defective and non-defective. The steps followed in classifying the samples as healthy and faulty using SVM is shown in Figure 8.

![Diagram of SVM process](image)

**Fig. 8.** Steps involved in SVM

3.2 Pre-processing

The SVM classifier will not recognize the raw inputs obtained from the feature extraction. Hence it is necessary to convert all the raw data into the numerical format which the SVM classifier can recognize.

3.3 Normalization

Data normalization is the important step to be carried out to execute the classification using SVM. The use of normalization process is to use the small sized attributes in the classification without neglecting them. If X is the pre-processed data, then the data normalization can be performed using the normalization function in the equation (5).

$$X' = a + \frac{(X-X_{\text{min}})(b-a)}{X_{\text{max}}-X_{\text{min}}}$$ (5)

Where, $X'$ is the Normalized data value, $X_{\text{max}}$ and $X_{\text{min}}$ are the maximum and minimum data value in the domain of X, a is the maximum value in the output range, b is the minimum value in the output range.

3.4 Kernel selection

The SVM tries to obtain a separating hyper plane with the maximal margin for the classification. It chose a kernel function for obtaining the hyper plane. The kernel function selection is shown in the equation (6).

$$K(x_i,x_j) = \varphi(x_i)^T\varphi(x_j)$$ (6)

SVM uses four kinds of kernels based on the problem definition. In this problem, the classification is a linear problem. For this the SVM uses the linear kernel function for classifying the defected and non-defected PV cell.

3.5 Modeling Parameters

For the Linear kernel, C is the required parameter need to select before the kernel function. C parameter determines the influence of the misclassification on the objective function. It tells the optimization to avoid the misclassifying of each training data. The accuracy of the SVM classifier directly depends on this parameter. LIBSVM (Library for Support Vector Machine) is used for optimizing the C parameter by the cross validation.

For the large values of C, the optimization chose a small margin hyper plane and gives better accuracy, For the small value of C, the optimizer chooses a large margin, and it misclassifies many points. The sequences used here is $C = [10^{-5}, 10^{-4}, \ldots, 104, 105]$. This classifier is trained on the basis of pair that gives the better cross-validation accuracy.

3.6 Training the Classifier

After the selection of kernel and the model parameter, the SVM classifier is trained on the training data set. The textural information of the healthy PV panel is selected as the trained data set. It is trained to classify the defected PV cell.

3.7 Testing a New Data

The SVM classifier is capable of classifying the test data, when it was trained once. The textural information of the unknown PV cell is trained with the SVM classifier to characterize it as defective or non-defective.

The following steps are involved in the fault classification algorithm,

**Step-1:** Obtaining trained data sets from the healthy PV panel.

**Step-2:** Capturing thermal image of PV cell using FLIR T420bx.
Step-3: Window analysis using MATLAB tool for obtaining attribute

Step-4: Finding contrast, Correlation, energy, homogeneity, entropy, mean and standard deviation for the thermal image.

Step-5: Formulating the test data set.

Step-6: SVM classifier for obtaining fault level

Step-7: Non-defective cell satisfies the following condition

Hyper plane 1 ≤ testing data set ≥ Hyperplane 2

Step-8: Else it is a defective cell.

4. Results and Discussion

Before implementing the methodology to a higher end PV plant, the same is implemented with a lower order PV array. It is a 3*3 array of solar cell with series-parallel connections. Faults are intentionally created at various cells and thermal images of different cells are taken using thermal camera (FLIR E8). The features of the images such as standard deviation, mean, median etc. are examined using MATLAB code. Features are trained and tested in SVM and classified. Figure 9 shows a normal image of lower order PV system with 3*3 array of cells. Figure 10 shows the thermal images of 3*3 PV array with cracks alone, Figure 11 shows the thermal images of 3*3 PV array with soil accumulation alone, Figure 12 shows the thermal images of 3*3 PV array with full and partial shading and Figure 13 shows the thermal images of 3*3 PV array with combined faults.
Table 3. Training data of Cell 11 and Cell 21

<table>
<thead>
<tr>
<th>Texture</th>
<th>Cell 11</th>
<th></th>
<th></th>
<th></th>
<th>Cell 21</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quadrant 1</td>
<td>Quadrant 2</td>
<td>Quadrant 3</td>
<td>Quadrant 4</td>
<td>Quadrant 1</td>
<td>Quadrant 2</td>
<td>Quadrant 3</td>
<td>Quadrant 4</td>
</tr>
<tr>
<td>Mean</td>
<td>29.56</td>
<td>58.92</td>
<td>39.28</td>
<td>36.16</td>
<td>143.72</td>
<td>206.72</td>
<td>219.08</td>
<td>215.44</td>
</tr>
<tr>
<td>Entropy</td>
<td>5.56</td>
<td>5.92</td>
<td>5.28</td>
<td>5.16</td>
<td>143.72</td>
<td>206.72</td>
<td>219.08</td>
<td>215.44</td>
</tr>
</tbody>
</table>

A 5 kWp PV plant is installed in 9° 42' 20.5956" N, 78° 5' 38.544" E, research location. This plant is being used for this research work and is shown in Figure 14. The 5 kWp PV plant is constructed with twenty numbers of monocrystalline PV panels of 250Wp each. FLIR T420bx model thermos graphi c camera is being used in this work. Its resolution is 320*240 pixels. Each thermal image has 320*240 data in the matrix. Figure 15 shows the thermal image of the complete layout of PV system.

![Fig. 14. Normal image of the 5kWp PV system](image1)

![Fig. 15. Thermal image of the 5kWp PV system](image2)

To start or initialize the fault detection system, two images of a single PV cell is captured using a thermal image camera taken under the same environmental condition. The first thermal image is captured from the PV panel which generates rated power under specified environmental condition. This is used as the reference value or the training data set for the SVM classifier. The second thermal image captured is the entire PV system with the manual defects which are intentionally made. This image is used for generating the testing data set. The thermal and Gray scale images of healthy and defective PV cell is shown in Figure 17, 18, 19 and 20 respectively.

![Fig. 16. PV Array Block diagram with fault location](image3)

Figure 16 shows the layout of PV array with intentionally created faults.
The PV system is divided into twenty windows by using the MATLAB coding for obtaining the attributes of each panel. The attributes of the healthy panel is obtained by using the Gray Level Co-occurrence Matrix tool (GLCM). GLCM extracts the texture information of the image. The following required data are obtained from the textural information and is used for the fault classification.

SVM classifier is used to classifying all training vectors into two classifications. The training data set has derived from the textural information of healthy panel. The condition for validating the training data set has prepared with the bounding limits. The feature from the validation has selected to describe the best hyper parameter. The feature information from the individual window has trained with the hyper parameter of the SVM. This data are accumulating on the classifier region. The hyper parameters are the factor which forms a boundary between the two classifications. The data set that lies between the hyper parameter is classified under the non-defective classification and the remaining data sets are classified under the defective classification. The table shows the feature details of each window and the classification.

The defective classification is the output of this work. Two thermal images give the fault classification and it eliminates the need of electrical, environmental data analysis. It reduces the time period for fault detection. The faults in the PV system have intentionally created for the output verification. Soil, dust, leaves are spread over the some PV cells, some other have minor and major cracks, some PV cells are degraded due to the high temperature and one PV cell has diode failure. These are the well-known faults whose electrical characteristics has also verified for confirming the faulty state. The PV cells are named as P11, P12, P13, P14 and up to P45. Totally there are twenty panels constructed as 4×5 PV array. Table 4 shows the textural information and SVM classifier output.

<table>
<thead>
<tr>
<th>Panel No</th>
<th>Texture Information</th>
<th>Classifier Output</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contrast</td>
<td>Correlation</td>
</tr>
<tr>
<td>P11</td>
<td>0.1244</td>
<td>0.7222</td>
</tr>
<tr>
<td>P12</td>
<td>0.1735</td>
<td>0.8885</td>
</tr>
<tr>
<td>P13</td>
<td>0.2221</td>
<td>0.8876</td>
</tr>
</tbody>
</table>

Fig. 17. Thermal image of a Non-Defective PV Cell

Fig. 18. Gray scale image of a Non-Defective PV Cell

Fig. 19. Thermal image of a Defective PV Cell

Fig. 20. Gray scale image of a Defective PV Cell
PV system with very less computation time. This fault analysis and the results obtained are 97% accurate with the defects are intentionally created in the PV system for thermal output near to the expected values. In this work, the cells in the non-defective classification gives the abnormal thermal output. It is inferred that, the cells which are in the considered electrical parameters like current and the PV cell into two classifications. It is validated without considering the abnormal thermal output near to the expected values. In this work, the defects are intentionally created in the PV system for the analysis and the results obtained are 97% accurate with the comparison of test and training results. This fault classification technique is used in real time for the large PV system with very less computation time.

5. Conclusions

In this paper, a simple image processing-based machine learning algorithm is performed for classifying the PV cell into two classifications. It is validated without considering the electrical parameters like current and power. It is inferred that, the cells which are in the defective classification gives the abnormal thermal output. The cells in the non-defective classification give the thermal output near to the expected values. In this work, the defects are intentionally created in the PV system for the analysis and the results obtained are 97% accurate with the comparison of test and training results. This fault classification technique is used in real time for the large PV system with very less computation time.

References


[31] B. Meenakshi Sundaram, B. V. Manikandan, B. Praveen Kumar, “Combination of Novel Converter Topology and Improved MPPT Algorithm for Harnessing Maximum Power from Grid Connected


