

Received January 3, 2022, accepted January 25, 2022, date of publication January 31, 2022, date of current version February 9, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3148138

Visual Vocabulary Based Photovoltaic Health Monitoring System Using Infrared Thermography

WAQAS AHMED¹, MUHAMMAD UMAIR ALI²,
SHAIK JAVEED HUSSAIN³, (Senior Member, IEEE),
AMAD ZAFAR⁴, AND SULAIMAN AL HASANI³

¹Secret Minds, Entrepreneurial Organization, Islamabad 44000, Pakistan

²Department of Unmanned Vehicle Engineering, Sejong University, Seoul 05006, South Korea

³Department of Electrical and Electronics, Global College of Engineering and Technology, Muscat 112, Oman

⁴Department of Electrical Engineering, Ibadat International University, Islamabad 54590, Pakistan

Corresponding authors: Shaik Javeed Hussain (s.javeedhussain@gcet.edu.om) and Amad Zafar (amad.zafar@iiui.edu.pk)

The authors would like to thank the Global College of Engineering and Technology, Muscat, for funding the research under an internal research funding grant.

ABSTRACT Photovoltaic (PV) systems have gained global acceptance in terms of green, replenishable energy resources to meet energy demand with no emissions. However, PV systems are susceptible to operational and environmental stresses. Moreover, PV panels monitoring is necessary to keep their performance and efficiency intact due to their lack of supervisory control. Therefore, this study monitors PV panels based on health into three sub-classes: healthy, hotspot, and faulty through infrared thermography. First, Thermographs key points are selected using an 8×8 uniform pixel grid, and speed-up robust features (SURF) are extracted from grid intersection points. Afterward, due to its simplicity, the k-mean clustering algorithm creates single-level clusters based on actual observations similarities and similar observations closeness within-cluster and dissimilarity to other clusters observations are used to transform features into visual words. Finally, shallow classifiers are utilized because of low training time and high prediction speed. After extensive testing and compressive analysis, the proposed approach was found economical, fast, and showed high testing accuracy of 97% through a multi-class shallow classifier (support vector machine) with low computational complexity and less storage size. Thus, this approach can monitor megawatt PV systems with high accuracy and keep performance and emissions mitigation potential high while lowering payback time.

INDEX TERMS Visual vocabulary, photovoltaic (PV) system healthy, hotspot.

I. INTRODUCTION

Energy is an essential entity in the modern era. While its major production technologies involve dreadful gas emissions such as CO_2 , NO_x , water vapors, etc., also known as greenhouse gases (GHG), they are responsible for global warming and climate change. Therefore, green renewable energy resources such as solar photovoltaic (PV) systems, wind energy systems, etc., have gained wider acceptance across the globe as the next energy security resources owing to energy needs and climatic concerns. Moreover, fossil fuels are limited and depleting fast. On the other hand, green renewable energy resources, such as PV systems, have a global potential of 1500-50,000 EJ per year [1]–[4]. Since then, the energy sector has become the leading contributor to global climate change, with 60% of global

GHG emissions [5]. Hence, research has focused on green renewable energy resources in multiple aspects.

However, despite the tremendous green energy potential offered by the PV system, its output varies with multiple parameters such as solar radiation, orientation, geographical location, temperature, etc. [4], [6]. In addition, the PV system, due to its no supervisory operation (unlike fossil fuel plants) [7], its performance is susceptible to operational and environmental stresses such as arcs, line-to-line, short-circuiting, open circuiting, glass breakage, interconnection faults, bird dropping, dust accommodation, shadowing, etc., which introduce defects in the PV system such as hotspots, faults, and mismatches, etc. [1], [3], [4], [6], [8]–[10]. More detail on PV system faults is provided at [10]. Therefore, the literature categorizes faults at three levels; cell-level (mechanical cracks, corrosion, UV degradation, thermal degradation, etc.), module-level (open-circuit, short-circuit, etc.), and string level (mismatches, open circuit, short

The associate editor coordinating the review of this manuscript and approving it for publication was Diego Oliva¹.

circuit, partial shadowing, etc.) [1]. For example, fig 1. shows defective PV panels due to environmental and operational stresses. Moreover, when photon transmission to PV cells is blocked due to environmental stresses such as shadowing, dust, bird droppings, snow, etc. (surface area varies in each case), due to series operation [2], PV cells under stress, such as bird dropping, operate in the reverse region and offer more resistance to healthy cells current. As a result, these reverse regions heat up (Joule's effect [11]), and a localized heated point appears, known as a hotspot [2], [4], as shown in Fig 2.

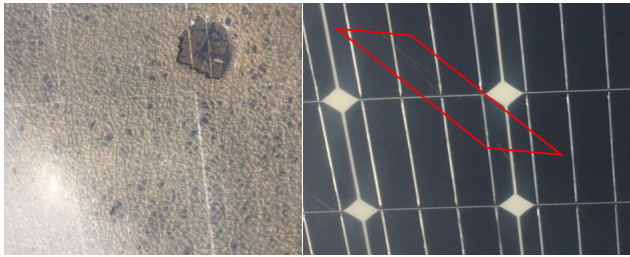


FIGURE 1. Protection glass breakdown and PV panels cracks.

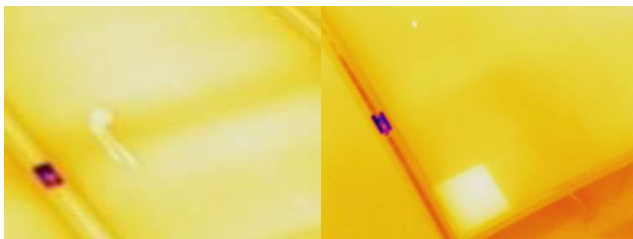


FIGURE 2. Hotspots (bird dropping and cell) [4].

Some of the PV panels' issues/defects are identifiable with the naked eye, such as protection glass breakdown, cracks, corrosion, etc. [4]. Moreover, for a large-scale PV system, automation is necessary. Literature has utilized multiple approaches such as thermography, electrical signals, electroluminescence, photoluminescence, etc., using different machine learning and deep neural network-based classification approaches to classify and differentiate PV systems based on health, defects, stresses, etc. [2], [4], [9], [12]–[14]. Electroluminescence images are high resolution but can damage PV cells due to the current injection. While the electrical signal-based approach necessitates sensors, increasing cost and complexity, the VI characteristics curve indicates an anomaly; locating the source or fault in the string is difficult. In contrast, thermographs show abnormal behavior of PV panels suffering from defects and are non-invasive thermal approaches [15], [16]. In addition, the thermographic approach is reliable, fast, economical, and has vast outdoor applications [4], [16]. Thermographs of PV panels are shown in Fig 2. More detail on the PV system's fault diagnosis approaches is available in [10].

A. PV SYSTEM FAULT DIAGNOSIS- LITERATURE REVIEW

Literature has extensively analyzed the impact of PV system defects/faults on energy production. For instance, 1 megawatt PV system with 20% defective panels reduce approx. 306,081.5 kWh energy potential per annum if defective panels are not timely identified and treated. And to produce this loss energy using fuel mix in Pakistan approx. 151.12 tons of CO₂ equivalent are produced and released into the atmosphere as per a theoretical study [3]. Therefore, literature has proposed multiple approaches (including machine learning, deep learning etc.) for PV system health monitoring and fault diagnosis.

Authors in [1] identified PV module mismatch faults through thermographs, which can be incorporated with a maximum power point tracker to adjust to a new operating point. Li *et al.* [17] proposed a six PV module defects diagnostic model through convolution neural network-based deep learning using unmanned aerial vehicles on a large-scale PV farm with high accuracy. Ali *et al.* [18] used PV panel thermographs and color image descriptors as features to train machine learning algorithms (k-nearest neighbor) into three health-based classes.

While Ali *et al.* [2] extracted thermographs texture features, mean of the histogram of oriented gradient (HOG), local binary pattern (LBP), and red-green-blue (rgb) color features to classify PV panels based on health using shallow classifiers. Niazi *et al.* [12] also extracted thermograph texture features and applied principal component analysis (PCA) on HOG features to reduce dimensionality and classified PV panels based on health using shallow classifiers (machine learning approach). Finally, Ahmed *et al.* [4] used deep convolutional neural networks (isolated, transfer-learned trained isolated, and transfer learned pre-trained) and classified PV panels based on health and defects.

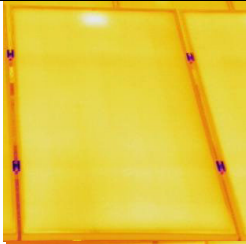

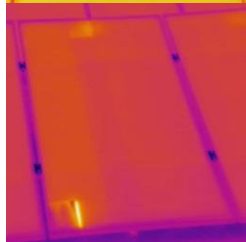
Deng *et al.* [15] proposed a subtraction network with a voting strategy to differentiate abnormal panels from normal panels having abnormal panels' features. Turhal *et al.* [13] proposed a common vector approach for fault detection. The authors of [6], proposed a computer vision approach to detect soil and dust on PV panel surfaces based on gray-level co-occurrence matrix features extraction. Zyout *et al.* [8] proposed a deep convolutional neural network-based approach to characterize the surfaces of PV panels and defect identification.

While Ebner *et al.* [14] used thermographs, electroluminescence, and photoluminescence (PL) approaches to find the production failures/defects in PV cells and modules, the authors [7] proposed a simple linear iterative clustering, a super-pixel technique for PV system hot spot detection using thermal imaging. Jaffery *et al.*, in [16], proposed PV panel early fault diagnosis through thermograph-based detection and fault classification through fuzzy logic. Artificial intelligence-based PV system health monitoring and fault diagnosis through thermographs are provided in Table 1. More details on image features are provided at [19].

TABLE 1. PV system monitoring for fault diagnosis.

Ref	Approach	Features	Classifier	PV States	Accuracy	Limitations
[4]	Isolated and pre-trained convolutional neural networks	-	Convolutional neural network	Three health-based conditions (healthy, hotspot, and faulty).	96% (Isolated) 100% (Pre-trained)	High execution time, GPU requirement, and have architectural complexity which increases further with PV Conditions
[20]	Manual features extraction (machine learning)	Texture (contrast, correlation, energy, entropy, and homogeneity)	Naïve Bayes	Two health-based conditions (Defective and Non-Defective panels)	98.4%	<p>Approach accuracy is limited to PV health-based conditions</p> <p>Texture features (extraction using the Global Local Occurrence Matrix) have a high dimensionality of the matrix. However, they are very sensitive to the size of the texture samples that are processed, and they perform better in simple situations where the textures are visually easily distinguishable but consume much memory.</p> <p>Furthermore, the GLCM features are ineffective for historical document photos with much noise.</p> <p>While LBP features are sensitive to image rotation, they produce rather long histograms, resulting in decreased distinctiveness and the need for large storage; the amount of textural information extracted is limited if the spatial support is small; it loses local textural information because it only considers signs of differences between neighboring pixels; and it is highly sensitive to noise and blurring.</p> <p>Furthermore, HOG is extremely sensitive to picture rotation, and the number of its cells is a compromise between storage and usable characteristics.</p>
[21]	Manual features extraction	Texture	Artificial Neural network classifier	Two (normal and potential faulty PV modules)	91.7%	
[2]	Manual features extraction	HOG, RGB, LBP, texture (3 RGB, 12 contrast, 12 correlation, 3 energy, 1 HOG, and 10 LBP)	Support Vector Machine	Three health-based conditions (healthy, non-faulty hotspot, and faulty)	92%	
[12]	Manual features extraction	Texture and HOG (52 texture (contrast, correlation, energy, and homogeneity) and 130 HOG features (PCA applied to reduce dimensionality))	Naïve Bayes	Three health-based conditions (Defective, Non-defective with hotspots, and non-defective without hotspots)	94.1%	
[18]	Manual features extraction	Red-Green Scale-Invariant Feature Transform (rgSIFT) descriptor	K- Nearest Neighbor	Three health-based conditions (normal, hotspot, and defective)	98.66 (Training)	It requires an optimal non-overlapped image division size of 71x71 pixels. A single deflection to left or right reduces the system accuracy. While this window is suitable for K-NN, and rest of the shallow classifiers suffer from low performances.
[4]	Transfer learning on isolated neural networks and pre-trained networks	-	CNN	Five faults (bird drop, single, patchwork, horizontally aligned string, and block)	97.62% (Isolated-learned) 100% (pre-trained)	Time, GPU requirement, architectural complexity increase with PV states
[22]	CNN	-	CNN	11 faults	92.5%	
[23]	CNN (open-source frameworks, i.e., Tensor flow)	-	CNN	3 health-based conditions (correctly functioning modules and defective modules)	99%	It took approx. 30 min computational time with low mid-range CPU, needs GPU for fastness, storage requirements, architectural complexity increases with PV states

TABLE 2. Details of PV panels [2].

Condition	Thermograph	Temperature	Possible Causes	Possible Solution	Dataset Size
Healthy		50-60° C	No hotspot, no damage	-	32.38%
Hotspot		65-80° C	Bird dropping, Shadow etc.	Cleaning (Timely monitoring)	31.43%
Faulty		Above 80° C	Damage	Replacement (Timely monitoring)	36.19%

Therefore, considering the limitations of literature such as manual features extraction (texture, HOG, RGB and LBP etc.), the introduction of the new dataset with abnormalities (rotation, visuals, memory etc.), deep learning approaches requirements of GPU, storage, architectural complexity, and mainly excess time consumption, a fast and least limited automatic PV monitoring approach is needed to classify the thermographic images of PV panels with high accuracy.

Concisely, overgrowing environmental concerns about energy production through conventional technologies with limited and fast depleting resources, the world has focused on alternative energy resources such as PV systems due to their cleanliness, tremendous potential, and availability for the next millions of years. However, due to operational and environmental stresses, the PV system suffers from low performance, reducing GHG mitigation potential and payback time. Moreover, considering the limitations of the machine and deep learning approaches, this study employs infrared thermography to categorize PV panels into three sub-classes based on health: healthy (working fine), hotspot (temporarily arose due to bird drop, etc.), and inoperable (faulty). Timely identification will result in required cleaning of PV panel suffering from the hotspot and revert its status as healthy and faulty panels can be replaced on time. Infrared thermograph features are extracted via speeded-up robust features (SURF) (captures the rotation information as well, unlike HOG, LBP etc.). Furthermore, the features vector size is

reduced by rejecting redundant features to decrease memory requirement, which increases the training and testing speed of the classifier.

Moreover, all thermograph features are categorized based on health using the k-mean clustering algorithm because of its dependence on actual observations and simplicity to achieve high accuracy and fast response.

Moreover, it is important that this approach can be extended to other-dimension issues such as using medical images/radiographs (such as Alzheimer's, tuberculosis, etc.) to perform least dependent-features based classification into no, mild, strong (demented), etc. conditions with high accuracy and fast response, since feature-based classification in the medical field is already being utilized to detect diseases with higher accuracy and reduce individual analysis, hence saving time and effort. Hence, feature-based disease diagnosis is provided at [24].

The rest of the study is organized as follows: Section II focuses on the approach, Section III on the feature extraction mechanism, Section IV on the results, Section V on the discussion, and finally Section VI on the conclusion and future prospects.

II. MATERIAL AND METHODOLOGY

Infrared thermography was obtained from a 42.24kW PV (240Watt PV panels, 22 panels in a string and eight strings, $240 \times 22 \times 8$) located in Pakistan's city Lahore, using

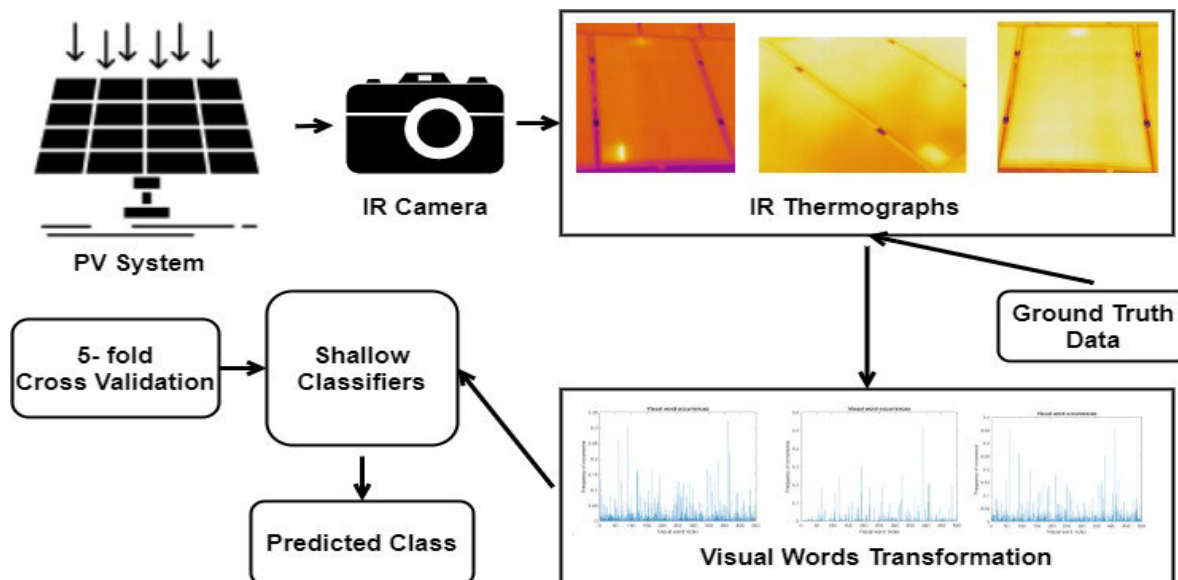


FIGURE 3. Proposed approach.

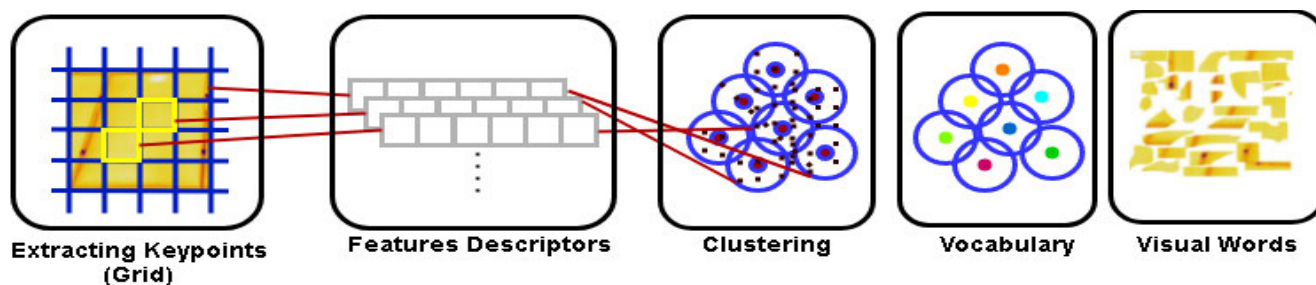


FIGURE 4. Thermographs to visual words [25].

FLIR VUE-Pro 640, 8-bit thermal image bit depth and 640×512 /pixel spatial resolution. Thermographs were captured with 32-40 Celsius temperature, wind speed of 6.9m/s, and irradiance level of 700W/m². Detail of the experimental setup is provided at [12]. Total 315 infrared thermographs are used in this study which is segregated into three classes based on health: healthy, hotspot, and faulty PV panel thermographs. Table 2. presents the in-depth detail of PV panels based on health, while a simplified proposed approach is presented in Fig 3.

The dataset was segregated into training and testing datasets with an 80:20 ratio randomly and with an equal proportion of each class for proper training and testing. Moreover, shallow classifiers were used due to their fast training and validation response [2], [4]. In addition, 5- fold cross-validation approach was utilized to avoid over-fitting.

III. FEATURES EXTRACTION

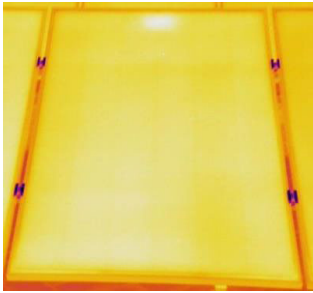
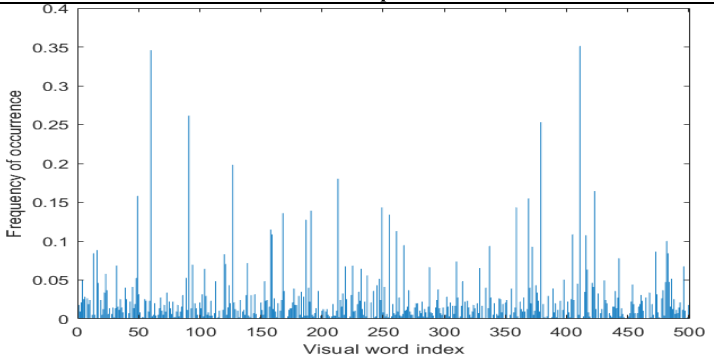
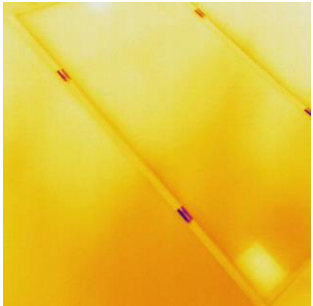
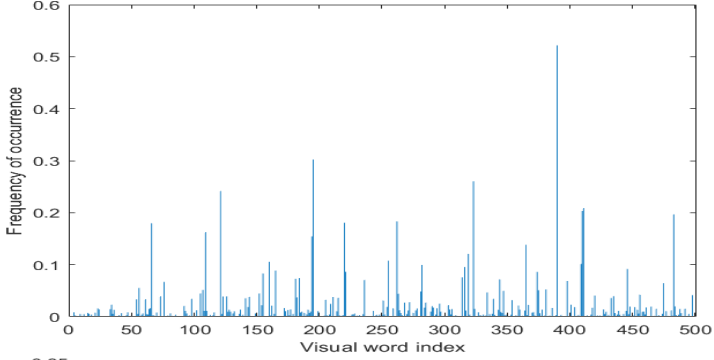
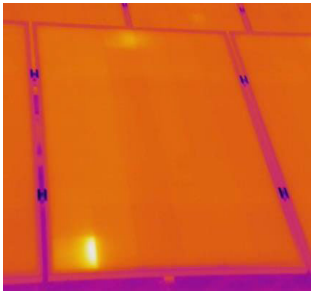
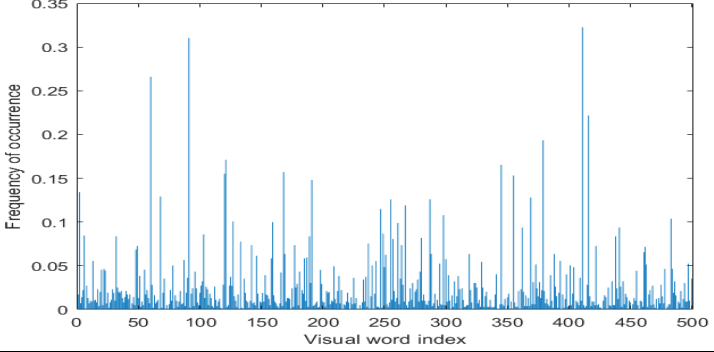
In literature, thermograph features are extracted to train and test (shallow and deep) classifiers using a manual approach or deep neural networks [2], [4], [12], which have their

corresponding advantages and disadvantages [19]. In this study, a visual vocabulary is created from the thermographs by extracting features descriptors from each health-based class of thermographs; finally, a bag of features/ feature histograms are formed. More detail is provided at [26].

A uniform 8×8 size pixel grid is defined as point selection to extract the thermograph features, where grid junctions define locations for feature extraction. Variation in grid size changes the size of the feature. Moreover, a four-element [32, 6, 96, 128] vector is used in which each element relates to the size of a square block from which the upright speeded-up robust features descriptors are extracted. It has the advantage that it can also capture the rotation information if required (limitation of LBP and HOG etc. features). More detail on SURF is provided at [27]. Using this information, the features from the training dataset are extracted.

Afterward, the most redundant parameters/features (20%) are reduced from each class of dataset to reduce the vector size to 80% strongest points, lessen memory requirement and increase the execution time (All classes' features are balanced based on class with least strong features to improve clustering).

TABLE 3. Visual words/ feature histogram representation of PV IR thermography.

PV Class	Thermograph	Visual Words Representation
Healthy		
Hotspot		
Faulty		

Finally, the K-means clustering algorithm is used for features segmentation because of its simplicity and dependence on actual observations rather than the dissimilarity between every pair of observations in the data, creating a single level of clusters. In addition, it has the advantage that observations within each cluster are as close to each other as possible and as far from objects in other clusters as possible. Using k-mean clustering, 500 visual word/feature histogram is created.

More detail on K- mean clustering is provided at [28], [29]. Finally, transformation is provided in Fig 4; further explanation about transformation can be found in [25].

IV. RESULTS

For visual vocabulary, 3167829 training features were extracted, and 80% of features from all classes considering the least strong features (782262) of a health-based class, feature vectors were formed with a total of 2346786 features.

Afterward, K- means clustering was used to create a 500-word visual vocabulary, and clustering took approx. 2 minutes. A reduced representation of thermographs with visual words occurrences in the form of the histogram for healthy, hotspot, and faulty are provided in Table 3. The frequency of occurrence in visual words for PV-system health-based classification is different; healthy, hotspot, and faulty occurrence frequency illustrates the difference detected by the proposed approach for their classifications.

Results were processed using MATLAB 2021a using core i7, 7th generation, 16 GB RAM, NVIDIA GeForce GTX 1060, 1.5TB SSD, and 64-bit operating system specifications. The performance of various shallow classifiers such as a tree, linear discriminant, Naïve Bayes, SVM, KNN trained using the visual words approach are provided in Table 4. True positive rate (TPR), positive predictive value (PPV), false discovery rate (FDR), and false-negative rate (FNR) are used as an accuracy measurement metric.

TABLE 4. Results of classification of PV panels based on health.

PV Class		Classifier																			
		Tree				Linear Discriminant				Naïve Bayes				SVM				KNN			
		TPR	FNR	PPV	FDR	TPR	FNR	PPV	FDR	TPR	FNR	PPV	FDR	TPR	FNR	PPV	FDR	TPR	FNR	PPV	FDR
Faulty		84.6	15.4	88.5	11.5	98.9	1.1	98.9	1.1	95.6	4.4	88.8	11.2	98.9	1.1	100	0	98.9	1.1	98.9	1.1
Healthy		90.2	9.8	85.1	14.9	91.5	8.5	96.2	3.8	93.9	6.1	97.5	2.5	97.6	2.4	96.4	3.6	95.1	4.9	96.3	3.7
Hotspot		77.2	22.8	78.2	21.8	96.2	3.8	91.6	8.4	84.8	15.2	89.3	10.7	96.2	3.8	96.2	3.8	96.2	3.8	95	5
Training Accuracy (%)		84.1				95.6				91.7				97.6				96.8			
Training Time (sec)		3.5				3.77				8.86				5.63				8.92			
Prediction Speed (observations/sec)		910				740				720				590				290			
Validation		5- fold cross validation																			
Features		Speeded-up robust features (SURF)																			
Features Selection Point		8x8 uniform pixel grid																			
Clustering Algorithm		K-mean																			
Clustering Time		7.63 seconds/ iteration (approx...)																			
Convergence (clustering)		19 iterations																			

TABLE 5. Testing of the proposed approach on new dataset.

Symbol	Faulty	Healthy	Hotspot
Faulty	1	0	0
Healthy	0	0.95	0.05
Hotspot	0.05	0	0.95
Accuracy	97%		

After extensive training and testing, the multi-class SVM shows the highest training accuracy, as illustrated in Table 4. To further validate the testing accuracy for the proposed approach, a new images dataset is fed to the trained SVM model to check the performance against online testing; the online testing results are provided in Table 5.

V. DISCUSSION

PV panels’ importance in terms of green replenishable energy resource and effective tool against GHG emissions mitigation in global climate change perspective has increased the global penetration of PV systems in the main grid (fossil fuel dominant). However, stresses such as environmental (bird dropping, shading, snow, dust etc.), operational (DC side, MPPT etc.), aging, and other phases (installation, transportation, manufacturing etc.) introduce defects and failures causing pre-mature aging and early retirement of PV panels. This concisely impacts the PV system performance,

TABLE 6. Thermograph based literature-accuracy.

Reference	Accuracy
Ahmed et.al., [4]	96%
Umair et.al., [2]	92%
Niazi et.al., [12]	94.1%
Umair et.al., [18]	98.66%

energy potential, GHG mitigation potential, and payback time. Therefore, catering to this issue, monitoring PV panels is essential. Moreover, stresses on PV panels are not identifiable through the naked eye. Therefore, an economical, fast, accurate approach is preferable, hence literature has focused on different approaches.

In literature for PV panels, classification based on health includes different machine and deep learning-based approaches, but these approaches also have associated limitations. Literature classifying PV panels based on health into three categories: healthy, hotspot, and faulty, with their accuracy in Table 6. In literature, Ahmed *et al.* [4] used an isolated deep convolutional neural network for health-based classification and attained 96% accuracy but took several minutes of models training and validation.

In contrast, Niazi *et al.* [12] and Umair *et al.* [2] used texture, HOG, LBP features-based machine learning approaches and achieved 94.1% and 92% accuracy, respectively, but these features fail in low visual images rotation of images, inappropriate cell size etc. Another machine learning approach presented by Umair *et al.* [18] used rgSIFT descriptor and achieved 98.66% accuracy. The authors defined the limitation that it requires optimal non-overlapped image division-size 71×71 pixels. A single deflection to left or right reduces the system accuracy.

The proposed approach transformed images into visual words/features thermographs. Extracted features of images through SURF by defining an 8×8 uniform grid step and transformation using a k-mean clustering algorithm, the details about the feature extraction is provided in Section 3. Changing the grid step is a trade-off between the features descriptor and clustering time. At the same time, 80% strongest features increase the classification accuracy by reducing unwanted dimensionality and lowering transformation time; further reduction in features impacts the accuracy of the classifier. The proposed approach resulted in 97.6% training accuracy and 97% testing accuracy on the new dataset for online testing of the SVM multi-class trained model (see Tables 4 and 5). Furthermore, the proposed approach is fast, with adequate memory storage, no high operating system requirements, highly accurate and efficient in detecting PV panels based on health into its subclasses, i.e., healthy, hotspot, and faulty. Therefore, the proposed approach can be easily employed to monitor high-power PV plants.

VI. CONCLUSION

The PV system is susceptible to multiple environmental and operational stresses with tremendous potential, outdoor operation, green nature, and without supervisory control (unlike fossil fuel-based systems). Therefore, timely monitoring of PV system health avoids capital energy output loss while maximizing GHG emissions mitigation potential. This study transformed PV panel infrared thermography into visual words through the SURF feature detector and k-mean clustering algorithm. The proposed scheme classified PV panels into three health-based sub-classes: healthy, hotspot, and faulty class. After extensive training and validation, it was found that the extracted featured trained SVM model has the highest training accuracy of 97.6 % to classify the PV panel into their respective class based upon their health condition. Furthermore, the proposed models showed similar accuracy (97%) against the new dataset, which was not used for the model's training. Therefore, the proposed approach can be utilized for the field monitoring of large PV plants.

ACKNOWLEDGMENT

(Waqas Ahmed and Muhammad Umair Ali are co-first authors.)

REFERENCES

- [1] Y. Hu, W. Cao, J. Ma, S. J. Finney, and D. Li, "Identifying PV module mismatch faults by a thermography-based temperature distribution analysis," *IEEE Trans. Device Mater. Rel.*, vol. 14, no. 4, pp. 951–960, Dec. 2014.
- [2] M. U. Ali, H. F. Khan, M. Masud, K. D. Kallu, and A. Zafar, "A machine learning framework to identify the hotspot in photovoltaic module using infrared thermography," *Sol. Energy*, vol. 208, pp. 643–651, Sep. 2020.
- [3] W. Ahmed, J. A. Sheikh, S. H. Farjana, and M. A. P. Mahmud, "Defects impact on PV system GHG mitigation potential and climate change," *Sustainability*, vol. 13, no. 14, p. 7793, Jul. 2021.
- [4] W. Ahmed, A. Hanif, K. D. Kallu, A. Z. Kouzani, M. U. Ali, and A. Zafar, "Photovoltaic panels classification using isolated and transfer learned deep neural models using infrared thermographic images," *Sensors*, vol. 21, no. 16, p. 5668, Aug. 2021.
- [5] (Sep. 30, 2021). *Goal 7 Affordable and Clean Energy*. [Online]. Available: <https://www.jointsdgfund.org/sustainable-development-goals/goal-7-affordable-and-clean-energy>
- [6] K. A. Abuqaoud and A. Ferrah, "A novel technique for detecting and monitoring dust and soil on solar photovoltaic panel," in *Proc. Adv. Sci. Eng. Technol. Int. Conf. (ASET)*, Feb. 2020, pp. 1–6.
- [7] M. Alsafasfeh, I. Abdel-Qader, and B. Bazuin, "Fault detection in photovoltaic system using SLIC and thermal images," in *Proc. 8th Int. Conf. Inf. Technol. (ICIT)*, May 2017, pp. 672–676.
- [8] I. Zyouat and A. Oatawneh, "Detection of PV solar panel surface defects using transfer learning of the deep convolutional neural networks," in *Proc. Adv. Sci. Eng. Technol. Int. Conf. (ASET)*, Feb. 2020, pp. 1–4.
- [9] R. Al-Mashhadani, G. Alkaws, Y. Baashar, A. A. Ahmed, N. F. Hani, and W. Hashim, "Deep learning methods for solar fault detection and classification: A review," *Inf. Sci. Lett.*, vol. 10, no. 2, p. 13, 2021.
- [10] A. Haque, K. V. S. Bharath, M. A. Khan, I. Khan, and Z. A. Jaffery, "Fault diagnosis of photovoltaic modules," *Energy Sci. Eng.*, vol. 7, no. 3, pp. 622–644, Mar. 2019.
- [11] (Sep. 30, 2021). *The Joule Heating Effect*. [Online]. Available: <https://www.comsol.com/multiphysics/the-joule-heating-effect>
- [12] K. A. K. Niazi, W. Akhtar, H. A. Khan, Y. Yang, and S. Athar, "Hotspot diagnosis for solar photovoltaic modules using a naive Bayes classifier," *Sol. Energy*, vol. 190, pp. 34–43, Sep. 2019.
- [13] U. Turhal and Y. Onal, "Photovoltaic fault detection and classification using common vector approach based on $I-V$ curve," *Acta Phys. Polonica A*, vol. 137, no. 3, pp. 421–429, Mar. 2020.
- [14] R. Ebner, B. Kubicek, and G. Ujvari, "Non-destructive techniques for quality control of PV modules: Infrared thermography, electro- and photoluminescence imaging," in *Proc. 39th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Nov. 2013, pp. 8104–8109.
- [15] J. Deng, T. Minematsu, A. Shimada, and R. Taniguchi, "Identify solar panel defects by using differences between solar panels," in *Proc. 15th Int. Conf. Quality Control Artif. Vis.*, vol. 11794. Bellingham, WA, USA: SPIE, Jul. 2021, Art. no. 1179415.
- [16] Z. A. Jaffery, A. K. Dubey, Irshad, and A. Haque, "Scheme for predictive fault diagnosis in photo-voltaic modules using thermal imaging," *Infr. Phys. Technol.*, vol. 83, pp. 182–187, Jun. 2017.
- [17] X. Li, Q. Yang, Z. Lou, and W. Yan, "Deep learning based module defect analysis for large-scale photovoltaic farms," *IEEE Trans. Energy Convers.*, vol. 34, no. 1, pp. 520–529, Mar. 2018.
- [18] M. U. Ali, S. Saleem, H. Masood, K. D. Kallu, M. Masud, M. J. Alvi, and A. Zafar, "Early hotspot detection in photovoltaic modules using color image descriptors: An infrared thermography study," *Int. J. Energy Res.*, vol. 46, no. 2, pp. 774–785, Feb. 2022.
- [19] A. Humeau-Heurtier, "Texture feature extraction methods: A survey," *IEEE Access*, vol. 7, pp. 8975–9000, 2019.
- [20] K. Niazi, W. Akhtar, H. A. Khan, S. Sohaib, and A. K. Nasir, "Binary classification of defective solar PV modules using thermography," in *Proc. IEEE 7th World Conf. Photovoltaic Energy Convers. (WCPEC), A Joint Conf. 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC*, Jun. 2018, pp. 0753–0757.
- [21] V. S. B. Kurukuru, A. Haque, M. A. Khan, and A. K. Tripathy, "Fault classification for photovoltaic modules using thermography and machine learning techniques," in *Proc. Int. Conf. Comput. Inf. Sci. (ICCSIS)*, Apr. 2019, pp. 1–6.
- [22] R. H. F. Alves, G. A. D. D. Júnior, E. G. Marra, and R. P. Lemos, "Automatic fault classification in photovoltaic modules using convolutional neural networks," *Renew. Energy*, vol. 179, pp. 502–516, Dec. 2021.

[23] D. Manno, G. Cipriani, G. Ciulla, V. Di Dio, S. Guarino, and V. L. Brano, "Deep learning strategies for automatic fault diagnosis in photovoltaic systems by thermographic images," *Energy Convers. Manage.*, vol. 241, Aug. 2021, Art. no. 114315.

[24] U. K. Lopes and J. F. Valiati, "Pre-trained convolutional neural networks as feature extractors for tuberculosis detection," *Comput. Biol. Med.*, vol. 89, pp. 135–143, Oct. 2017.

[25] (Sep. 12, 2021). *Image Classification With Bag of Visual Words*. [Online]. Available: <https://www.mathworks.com/help/vision/ug/image-classification-with-bag-of-visual-words.html>

[26] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *Proc. Workshop Stat. Learn. Comput. Vis. (ECCV)*, May 2004, vol. 1, no. 1–22, pp. 1–2.

[27] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.

[28] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A K-means clustering algorithm," *J. Roy. Stat. Soc., C, Appl. Statist.*, vol. 28, no. 1, pp. 100–108, 1979.

[29] (Dec. 18, 2021). *k-Means Clustering*. [Online]. Available: https://www.mathworks.com/help/stats/k-means-clustering.html#bq_679x-19



WAQAS AHMED received the degree in electrical engineering with a specialization in the power and energy domains. He received multiple awards throughout his academic and professional life. He has worked on green renewable energy production methodologies and climate change. For three years, he was in-charge of the Power Systems Laboratory, University of Wah. He is currently the Chief Researcher of the Entrepreneurial Organization, Secret Minds. His research interests include

machine and deep learning, green renewable energy resources, and energy sustainability to tackle climate change.



MUHAMMAD UMAIR ALI received the B.S. degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2009, the M.S. degree in electrical power engineering from the University of Wah, Pakistan, in 2015, and the Ph.D. degree from Pusan National University, South Korea, in 2020, under the supervision of Prof. Hee-Je Kim.

After completing his Ph.D. degree, he joined The University of Lahore, Pakistan, as an Assistant Professor. Since 2021, he has been working as an Assistant Professor with Sejong University, South Korea. He is a reviewer and an editorial member of various SCIE journals. His current research interests include modeling, state estimation, and prediction of various dynamical systems.



SHAIK JAVED HUSSAIN (Senior Member, IEEE) received the B.Tech., M.Tech., and Ph.D. degrees in electronics and communication engineering from Jawaharlal Nehru Technological University Hyderabad. He has worked on brain-computer interfacing funded project by CSRI with the MITS Research Laboratory, Department of Science and Technology. He is working on a project robotics in health care systems funded by TRC, Oman. He is currently working as the Head of the Electrical and Electronics Department, Global College of Engineering and Technology, Muscat, Oman. He is a reviewer and a technical committee member of several IEEE journals, conferences, and other reputed journals. His research interests include machine vision, robotics, human-computer interface, medical imaging, applied machine learning, and renewable energy.



AMAD ZAFAR received the Ph.D. degree in intelligent control and automation from Pusan National University, Busan, South Korea, in 2019. He is currently working as an Associate Professor with the Department of Electrical Engineering, Ibadat International University, Islamabad, Pakistan. He has more than 12 years of experience in research and academia in the field of electrical and brain engineering. He has authored more than 45 scientific peer-reviewed journals, conferences, and book chapters. He taught numerous courses in the field of electrical and brain. His research interests include the modeling, estimation, prediction, machine learning, optimization, and brain-computer interface.



SULAIMAN AL HASANI received the B.Eng. degree (Hons.) in electrical and computer systems engineering and the Ph.D. degree in electrical and computer systems engineering with the main focus on signal processing and magnetic resonance imaging (MRI) application from Monash University, Melbourne, Australia. He worked on projects with the Monash Centre for Synchrotron Science, Monash Biomedical Imaging, and the University of Glasgow. He worked with the Department of Electrical and Computer Systems Engineering, Monash University, and then with the Global College of Engineering and Technology, Muscat, Oman, as a Lecturer. He is currently working with the Global College of Engineering and Technology as the Deputy Dean (A&R). His research interests include signal processing, image processing, compressive sensing, deep learning, medical imaging, and biomedical signal acquisition and reconstruction.

...