An efficient and portable solar cell defect detection system

Mohamed A. El-Rashidy

Received: 25 March 2021 / Accepted: 22 May 2022 / Published online: 18 June 2022
© The Author(s) 2022, corrected publication 2022

Abstract

The photovoltaic (PV) system industry is continuously developing around the world due to the high energy demand, even though the primary current energy source is fossil fuels, which are a limited source and other sources are very expensive. Solar cell defects are a major reason for PV system efficiency degradation, which causes disturbance or interruption of the generated electric current. In this study, a novel system for discovering solar cell defects is proposed, which is compatible with portable and low computational power devices. It is based on K-means, MobileNetV2 and linear discriminant algorithms to cluster solar cell images and develop a detection model for each constructed cluster. It can extract the distinct features between defective and nondefective solar cell clusters and overcome the confusion between different cell shapes. The proposed system was assessed using a benchmark dataset of electroluminescence images. The results demonstrate that it achieved the highest accuracy at a substantial rate compared to recent studies.

Keywords Classification · Clustering · MobileNetV2 approach · Photovoltaic system · Electro luminescence images

1 Introduction

Energy in various forms and types, such as coal, oil and natural gas, is a crucial element of economic development that cannot be overlooked. Energy has been a primary driver of the world economy since the beginning of industrial, agricultural and service activities. The rapid growth in economic and demographic terms, particularly after WWII, contributed to high energy consumption rates. This continuous increase in oil, natural gas and coal consumption has resulted in the consumption of about half of its conventional energy reserves. Thus, most of the world’s countries have recently resorted to renewable energy, particularly solar energy, to compensate for the energy shortage in the electric grid.

Many countries have begun using solar energy as an alternative source to conventional energy to reduce reliance on fuel and deliver electricity to isolated places such as deserts and mountains, where it is difficult to deliver electricity from the electrical grid due to their isolation and the high cost of connecting it to the network. Generally, the solar cell industry has gained a reputation worldwide, and the cost of constructing a solar cell system has decreased because of the use of current technologies. At the end of 2015, this industry has considerably grown in competition rather than the conventional energy sources, having extended to more than 30 countries [1]. In 2017, the total capacity of solar electricity generated was 178 GW. It is sufficient to provide 1% of the world’s electricity needs. Currently, the photovoltaic (PV) system has a participation rate of 200 GW [2]. The International Energy Agency forecasts that by 2050, PV systems are expected to be the primary source of electricity generation worldwide [3].

Solar energy is converted into electrical energy using arrays of solar cells. These arrays are the primary components in constructing the PV systems that contribute to energy production. An aluminum frame and a glass cover surround the solar cell arrays. These frames are called solar cell panels. These panels protect the PV systems from environmental factors such as wind, dust, rain and snow. However, since the crystal structure of the manufactured panels is fragile, various factors can cause cracks and damage some cells in the arrays during the manufacturing and installation of panels. Faulty welding during manufacturing results in damage to solar cells. Dropping solar cell panels during installation is also a factor that cracks
some cells. In addition to the environmental factors during the work of solar power plants, such as snowfall, tree branches, heat stress and winds. PV system efficiency degradation is mostly due to cell damage and cracks in solar panels. Therefore, it causes a disturbance or an interruption of the generated electric current. It can cause damage to the electrical appliances reliant on the PV system’s grid. Thus, monitoring the produced energy’s efficiency and disturbance by solar panels is crucial, as it aids in identifying damaged cells and replacing them to ensure the efficiency of PV systems.

Most of the defects in a solar cell that reduces the PV system’s efficiency are invisible. Conversely, the obvious clefts in the cells may not reduce cell efficiency. Thus, visual identification of the cracked and damaged cells is challenging. Infrared imaging was employed to assess the efficiency of the solar panels. Since the damaged cell cannot convert solar energy into electricity, the temperature of the cell rises. Thus, to identify the damaged cells, infrared cameras are used to image the emitted rays from the panels. Infrared cameras have a low nonrelative resolution, which can prevent small fault detection such as micro-fissures. Furthermore, a hot area in an IR image is not always a defect [4]. Additionally, high and low resistance series are complicated to differentiate in infrared images [5].

Recently, electroluminescence (EL) imaging has become one of the most extensively used approaches [6–17] for detecting defects in solar cells as an alternative approach to infrared camera imaging and visual identification. EL imaging can image solar panels with high accuracy, which also includes imaging the infrared radiation emitted by photoelectric cells, allowing it to discover microcracks. Visual identification using EL images is a major problem. It necessitates considerable experience and studies to examine solar cell images. Additionally, it is unsuitable for large solar power plants with thousands of solar panels. In this context and with the increased interest in relying on PV energy as a viable alternative to conventional energy and the production of large quantities of solar cells, automatic detection of solar cell defects and solar station efficiency has become an imperative.

Various research applications to automatically detect solar cell defects have been conducted, but there have been few investigations on EL imaging. Furthermore, these earlier recent studies [6–17] that relied on EL imaging were unable to detect solar cracks with high accuracy, due to the different solar cell shapes and low density or backlighting of EL images. They also relied on complicated approaches and devices with great computational power to discover the cell cracks.

In this study, an automatic recognition approach for detecting solar defects using EL images is proposed. The proposed approach’s contribution aims at two goals. The first goal is to determine the distinguishing features between defective and nondefective solar cells for each group of homologous cells and to overcome the confusion between the various cell shapes, resulting in improved detection accuracy. The second goal is to reduce the detection model’s size and complexity cost, making it compatible with portable devices or any device with low computing power.

The proposed approach consists of two phases. The first phase is based on K-means, MobileNetV2 and linear discriminant algorithms. This phase’s objective was to perform data training. It finds the clusters of homologous solar cells, constructs a detection model that can identify the defective solar cell with the highest possible accuracy for each cluster of homologous cells and builds a classifier model of the solar cell image of the constructed clusters, which helps overcome the confusion between the different cell shapes and increases the detection accuracy. The second phase uses the constructed linear discriminant classifier to classify the cell image to the cluster label that has the closest set of solar images. Then, it uses the assigned cluster’s constructed detection model to identify the defective solar cell.

To verify the proposed approach, several experiments were conducted, with a benchmark dataset of EL images consisting of 2,624 solar cells extracted from panel images [18]. The proposed approach achieved the highest accuracy compared to various recent research frameworks [6–8].

The rest of this study summarizes the related work, describes the proposed method, presents and discusses the experimental results. Finally, the last section concludes the study.

2 Related works

In the recent past, several real time automated approaches have been developed to discover the cracks in solar panels. These approaches were based on machine learning functionalities, which have the ability to analyze the images and automatically recognize the patterns.

In [6], a novel framework was proposed, and it used Deep Siamese based on CNN model. The framework was dependent on the symmetric structure of two input images to detect solar cracks. EL solar images were grouped into two clusters. Then, the Deep Siamese model was trained by the clustered images to compare the symmetrical structure and to detect the solar cracks. In [7], new CNN architecture was proposed for selecting features and detecting solar cracks. It aimed to fit the proposed architecture by training data and achieve the highest accuracy. Therefore, it added
data augmentation process to overcome the overfitting and data scarcity problems.

In [8], two different methods were proposed for automatic detection of the photoelectric cell cracks. It aimed to search about the best method that achieves the highest accuracy and low computing cost. Each of the two methods requires different hardware specifications to be used in the classification process. The proposed method that was based on Support Vector Machine (SVM) required a low computing cost. The most demanding hardware method was that relied on CNN, which required a GPU for efficient execution. The proposed method that was based on SVM achieved high accuracy compared to the proposed method that was based on CNN. In [9], new CNN-based deep learning approach was developed to analyze and discover the abnormal regions from the image, it used UAV for imaging the large-scale solar panels in the farms. It was based on high computational power devices to analyze the aerial solar cell images.

In [10], a modified Kirsch operator was presented to discover the solar cracks, it aimed to reduce space and time complexity cost. It proposed new customized CNN architecture to select the effective features for each cracked solar image. Then, multiple classification support vector machine technique was used to group the extracted features of the cracked solar cells and identify the abnormal cells. In [11], deep two-dimensional CNN method was proposed to extract the most effective features from 2-D scalograms of the solar panel images, it based on the extracted features to effectively analyze and discover solar cracks in severe conditions.

In [12], long short-term memory approach was presented for extracting the effective features of abnormal cells, a softmax regression classifier was utilized to detect the solar cracks using the extracted features. In [13], a high-precision deep neural network model named PVP Net was proposed to discover solar cracks from solar panel images. It used stochastic gradient descent technique to optimize the calculating process of the proposed PVP Net parameters. In [14], unsupervised approach was presented to identify the unknown solar defects, it used the deep embedded restricted cluster method for grouping the solar cell images. It modified the clustering method using the point-wise convolution layers to reduce the computing cost and the number of learning parameters. In [15], a hybrid approach was proposed to detect the solar defects. It was intended to improve the classification accuracy, which relied on ANN, LSTM and CNN to accommodate different training data.

In [16], deep belief network was presented to identify the solar cracks. It used the genetic algorithm as the optimizer for calculating the training parameters. The genetic algorithm was based on the Boltzmann machine reconstruction error as a fitness function. The genetic algorithm enhanced the training speed and deep belief network accuracy. In [17], deep CNN based on the transfer learning was utilized, it used VGG16 architecture to discover the defects in the cells of PV system.

In these researches, some studies were aimed to increase the accuracy based on hybrid complex techniques, which required high computational power devices to detect the solar cracks in the real time. Otherwise, the other studies were focused on reducing the computing cost of the proposed approaches, but this goal had a negative impact on the detection accuracy.

3 Proposed system

3.1 Methodology

Identification of defective solar cells entirely depends on the features of the cell images. These features must be distinct and capable of distinguishing between defective and nondefective solar cells. This improves the accuracy of solar cell detection without confusing the identification of defective and nondefective solar cells. There are different characteristics between the solar cell images’ structure shapes, whether it is defective or nondefective cells, as shown in Fig. 1. These differences are accompanied by a similarity in most of the cell features, which causes confusion in the classification process and reduces the accuracy of the defective cells’ identification. The proposed approach’s first objective is to determine the distinguishing features between defective and nondefective solar cells for each group of homologous cells and to overcome the confusion between the different cell shapes, resulting in improved detection accuracy.

There has been a recent global trend to spread solar power stations, which can generate thousands of megawatts over vast areas outside cities in the desert, mountains and agricultural lands. Solar panels have proliferated to generate a few kilowatts of electricity within cities on the roofs of factories, barns, gas stations and homes due to the attractive prices of small PV equipment. It is difficult to use computers with high computational power to identify a defective solar cell in so many places, especially in small applications. Thus, the proposed approach’s second objective is to reduce the detection model’s size and complexity cost, making it suitable for portable devices, or devices with low computational power.

As shown in Fig. 2, the proposed approach consists of two phases. The first phase is based on K-means, MobileNetV2 and linear discriminant algorithms. This phase’s objective was to conduct data training. It constructs the clusters of homologous structure cells, a detection model
that can identify the defective solar cell with the highest possible accuracy for each cluster of homologous cells and a classifier model of the cell image of the constructed clusters, which aids in overcoming the confusion between the different cell shapes and increases the detection accuracy.

The second phase will be used to discover the defective solar cell, it classifies the cell image to the cluster label that
has the closest set of images and then uses the detection model of this cluster to identify it. After preprocessing the images of the solar panels, the defective and nondefective solar cell images will be stored in dataset \( D = \{ i_{m_1}, \ldots, i_{m_n} \} \), where \( n \) is the total number of cell images. Each cell image \( i_{m_i} \) has a class label “defective” or “nondefective,” and it consists of \( X_i \times Y_i \times Z_i \), where \( X \), \( Y \), and \( Z \) are the number of length pixels, width pixels, and colors, respectively.

As shown in Fig. 2, the training phase consists of three steps. The first step is to train the layers of MobileNetV2 by \( D \) dataset using Model 1’s architecture, as shown in Fig. 3, where \( \text{Conv2D} \) is the 2D convolution layer (Fig. 4), \( \text{avgpool} \) is the average pooling layer (Fig. 5), \( \text{FC} \) is the fully connected layer, and \( \text{BMV2} \) is the block of MobileNetV2, as shown in Fig. 6. Each \( \text{BMV2} \) can be repeated \( n \) times with stride \( s \) and an expansion factor \( t \). The training phase aims to update all training parameter values “2,225,153 parameters” of Model 1 layers to fit the training data by minimizing the difference value between the actual class label of \( i_{m_i} \) and the predicted label during the training process.

As shown in Fig. 1, the characteristic difference between the solar cell shapes may cause confusion in identifying defective and nondefective solar cells when Model 1 tries to fit the training dataset. Thus, after training Model 1, the training parameter values may not be able to distinguish between the defective and nondefective solar cell images with the highest possible accuracy.

The second step is proposed to overcome the confusion in detecting solar cracks resulting from the different characteristics of solar cell shapes. It is based on a \( K \)-means algorithm to cluster the training solar cell images into \( K \) clusters, in which each cluster has a set of cells with a high intra-similarity shape and low inter-similarity shape from the other clusters.

Model 2 in the second step is proposed to reduce the difference value between the actual class label and the predicted label for each cluster cell and obtain the highest possible detection accuracy. To achieve this goal, Model 2 was constructed by adding a new layer with extra training parameters “128 kernels” to Model 1, as shown in Fig. 4. This process aims to update the additional training parameter values “164,097 parameters” of Model 2’s new layer, which freezes the other layers’ parameter values in the training process. Model 2 will be trained \( K \) times to construct \( K \) classifiers, one classifier for each cluster. To fit the training cell images for cluster \( I \), the additional training parameter values of \( C_i \) classifier will be updated. Each cluster contains a set of cell images with high similarity of the structure shape, ensuring that the different structure shapes of solar cells are not confused. The updated process of the additional training parameters will be focused on distinct defective and nondefective solar cells of a specific cluster without considering the different cell structures of other clusters.

The third step constructs a linear discriminant classifier, which classifies the input cell image to the closest cluster of cell structure features. The linear discriminant classification algorithm will be trained using the cell images of the \( K \) clusters, with each image associated with its cluster label. The constructed linear discriminant classifier can classify the input images into \( K \) cluster labels after training. The detection phase classifies the input cell image into the closest cluster and discovers the defective cell. After constructing linear discriminant and \( K \) Model 2 classifiers, the detection phase employs the linear discriminant classifier to classify the input image into \( L \) cluster labels, where \( L \) is between 1 and \( K \) clusters. Then, it uses Model 2 classifier no. \( C_L \) to identify the defective cell.

### 3.2 MobileNetV2

MobileNetV2 was proposed to reduce deep learning’s complexity cost and approach size. It is suitable for portable devices or devices with low computational power to support clustering, classification, feature extraction and object recognition. The execution of deep learning on
personal portable devices will be developed in the user practice. It allows accessing the detection system anywhere and anytime, and also adds benefits for energy consumption, security and privacy. MobileNetV2 is based on the MobileNetV1 model’s architecture. The idea behind MobileNetV1 is to split the convolution layer into two separate layers, as shown in Fig. 5. The first layer, called convolution depth-wise, applies a weighted filter using one convolutional filter for each input channel. The second layer, called a convolution point-wise, applies linear combinations of the input channels using a \((1 \times 1)\) convolutional filter.

The traditional convolutional layer’s computational cost is calculated using Eq. 1, where it has a layer input \(L_i\) \((X_i \times Y_i \times Z_i)\), and applies kernels \(k \times D_i \times D_j\) to produce an output layer \(L_j\) \((X_j \times Y_j \times D_j)\) \([19]\).

Computational cost of traditional convolutional layer

\[
= X_i \times Y_i \times D_i \times D_j \times k \times k \tag{1}
\]

Compared to a standard convolutional layer, the MobileNetV1 block reduces the computation cost by almost a \(k^2\) factor. The computational cost of the MobileNetV1 block was calculated using Eq. 2, which is the sum of the convolution depth-wise and point-wise layers \([19]\).

Computational cost of MobileNetV1 block

\[
= X_i \times Y_i \times D_i \times (D_j + k^2) \tag{2}
\]

MobileNetV2 is based on the MobileNetV1 concept, as it is an efficient building block. MobileNetV2 uses kernel size \(k = 3\). Thus, the computational cost is reduced by almost nine times smaller than the traditional convolution layer \([20]\). However, the MobileNetV2 model presents two new features to the block structure of MobileNetV1. The
first feature is linear bottlenecks between the block layers, and the second feature is the shortcut connections between the bottlenecks. Figure 6 shows the block structure of the MobileNetV2 model.

The MobileNetV2 model’s first feature preserves the information of the block input, and experimental evidence indicated that the use of linear layers is crucial since it prevents nonlinearity from destroying much information [20]. The MobileNetV2 model’s second feature optimizes a gradient for propagation across many layers, which leads to improving the model accuracy. It adds shortcut connections between the bottlenecks similar to the residual connections. The MobileNetV2 block output after using the shortcut connections is calculated using Eq. 3.

\[
\text{MobileNetV2 block output } (a_{L+3}^+) = g(Z_{L+3}^+ + a_L^+) \tag{3}
\]

\[
Z_{L+3}^+ = W_{L+3}^+ a_{L+2}^+ + b_{L+3}^+ \tag{4}
\]

where \(a_L^+\) represents the MobileNetV2 block input, \(a_{L+3}^+\) denotes the MobileNetV2 block output, \(g\) activation function is Relu, \(a_{L+2}^+\) is the activation output of the second layer, \(W_{L+3}^+\) represents the weight matrix of the third layer and \(b_{L+3}^+\) represents the bias vector of the output layer of the MobileNetV2 block.

### 3.3 K-means

K-means is an effective clustering technique for data with numerical features. It aims to group the dataset into clusters with lower inter-distance and higher intra-distance. This goal can be attained by reducing the sum of the square error in Eq. 5 [21]. It consists of four steps. The first step randomly selects an arbitrary K of the solar cell images, each of these images initially represents the cluster’s center called the seed object. Based on the distance between the cell image and the seed object, the second step assigns each solar cell image in the dataset to the closest seed object. Next, it calculates the centers of the K clusters to update the seed objects. The third step tests the seed objects’ positions, if the updated seed objects do not move, the clustering process will be stopped. Otherwise, the process will proceed to the second step.

\[
\text{SSE} = \sum_{i=1}^{K} \sum_{im \in \text{Cluster}_i} \text{dist}(im, \text{Mean}_i) \tag{5}
\]

where SSE is the sum of the squared error for all solar images in the dataset, \(im\) is the solar cell image and \(\text{Mean}_i\) is the seed object of Cluster_i.

### 3.4 Linear Discriminant Analysis

Linear discriminant analysis (LDA) is the most popular classification algorithm for data with more than two clusters. The LDA algorithm is based on reducing the training dataset to a lower-dimensional dataset. It consists of three steps to perform the dimension reduction. The first step is to compute the separation distance between the different clusters, called the between-cluster matrix \(BCM\), which is calculated using Eq. 6 [22].

\[
BCM = W^T S_B W \tag{6}
\]

\[
S_B = \frac{1}{n} \sum_{k=1}^{K} n_k S_{Bk} \tag{7}
\]

\[
S_{Bk} = (m_k - m)(m_k - m)^T \tag{8}
\]

\[
m_k = \frac{1}{n_k} \sum_{im \in k} im \tag{9}
\]

\[
m = \frac{1}{n} \sum_{i=1}^{n} im \tag{10}
\]

where \(W\) represents the transformation matrix of LDA, \(m_k\) represents the mean of the \(k\)th cluster, \(m\) represents the mean of all clusters, \(S_{Bk}\) represents the between-cluster variance of the \(k\)th cluster and \(S_B\) represents the total between-cluster variance.

The second step is to calculate the distance between the solar cell images of each cluster and its mean, called the within-cluster matrix \(WICM\), which is calculated using Eq. 11 [22].

\[
WICM = W^T S_W W \tag{11}
\]

\[
S_W = \frac{1}{n} \sum_{k=1}^{K} S_{Wk} \tag{12}
\]

\[
S_{Wk} = \sum_{im \in k} (im - m_k)(im - m_k)^T \tag{13}
\]

where \(S_W\) represents the total within-cluster variance, \(im\) represents the \(i\)th image cell in the \(k\)th cluster, \(m_k\) represents the mean of the \(k\)th cluster and \(S_{Wk}\) represents the within-cluster variance of the \(k\)th cluster.

The third step is to identify the lower-dimensional dataset, which minimizes the within-cluster variance and maximizes the between-cluster variance as in Eq. 14.

\[
\arg \max \frac{BCM}{WICM} = \arg \max \frac{W^T S_B W}{W^T S_W W} \tag{14}
\]

The solution to this problem in Eq. 14 can be achieved by calculating the diagonal eigenvalue vector \(E\) as in Eq. 15, which transforms the dataset features into lower-dimensional data to obtain the maximum ratio of between-cluster variance \(S_B\) against within-cluster variance \(S_W\) [22].

\[
S_W W = E S_B W \tag{15}
\]
4 Experimental results and discussion

Many experiments were developed to test and assess the proposed approach using the fivefold cross-validation testing technique. These experiments were applied to a benchmark dataset of EL images [18]. It consists of 2,426 solar cell images and is used to detect solar defects automatically. The dataset images contain both defective and nondefective solar cells with varying degrees of degradation. The images were analyzed by experts who rated the solar images by the likelihood of a defect within each cell. They were extracted from 44 different solar panels, which were either monocrystalline or polycrystalline.

Recall, precision, accuracy and f-score metrics were used to evaluate the conducted experiments [7]. As shown in Eq. 16, the recall metric was used to calculate the percent of correctly classified images of the defective solar cells. As shown in Eq. 17, the precision metric is the ratio between the correctly classified images of the defective solar cells to all images that are classified as defective cells. As shown in Eq. 18, the accuracy metric was used to calculate the percent of the overall classified images, whether it is defective “positive instances” or nondefective cells “negative instances.” The F1-score metric in Eq. 19 was also used to calculate the percent of the overall classified images, but it is an efficient metric for unbalanced datasets [7].

\[
\text{Recall} = \frac{TP}{TP+FN} \quad (16)
\]

\[
\text{Precision} = \frac{TP}{TP+FP} \quad (17)
\]

\[
\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (18)
\]

\[
F1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (19)
\]

where \( TN, TP, FN \) and \( FP \) are the true negative, true positive, false negative and false positive solar cells, respectively.

The proposed approach consists of two phases. The first phase is based on K-means, MobileNetV2 and linear discriminant algorithms. As shown in Fig. 2, it consists of three steps. The first step fits the training data and minimizes the difference value between the actual class labels and predicted labels as closely as possible. To develop and evaluate the first step, two experiments were implemented.

The first experiment conducted a thorough investigation to determine the best classifier that can fit the training data with the highest accuracy. It used 10 machine learning classification algorithms support vector machine (SVM), random forest, naïve Bayes, K-nearest neighbor, LDA and pre-trained MobileNetV1, Inception, MobileNetV2, ResNet and Xception on the ImageNet dataset. The pre-trained models were frozen with the parameter values of all layers except for the output layer. Only the output layer parameter values “1,281 parameters” were updated during the training. MobileNetV2 achieved the highest accuracy compared with the other classifiers, as shown in Table 1. Although, it has size and complexity costs lower than the other pre-trained models.

The second experiment employed the pre-trained MobileNetV2 on the ImageNet dataset, which achieved the highest accuracy in the first experiment, but it started with the Model 1 architecture layers’ parameter values as initial values. All the parameter values “2,225,153 parameters” were updated during the training in this experiment. As shown in Figs. 7 and 8, Model 1 architecture that updates the parameter values of all layers could fit the training data and achieve the highest accuracy compared to the model that only updates the output layer in the first experiment.

When the classification algorithm tries fitting the training dataset, the characteristic difference between the solar cell shapes, as shown in Fig. 1, may confuse the identification of defective and nondefective solar cells. This will lead to a decrease in the classification accuracy of the Model 1 classifier in the second experiment. Thus, the second step was used to overcome the confusion of the identification of defective and nondefective solar cells resulting from the different characteristics between the solar cell shapes. The third experiment used the K-means algorithm to cluster the training solar cell images into \( K \) clusters. This experiment finds the best \( K \) value “number of clusters,” which can group the set of cells into clusters with high intra-similarity cell shapes and low inter-similarity cell shapes from the other clusters. In this experiment, the dataset was clustered using different \( K \) values ranging from 1 to 12. As illustrated in Fig. 9, SSE in Eq. 5 droops in the location of \( K = 8 \), indicating the appropriate number of clusters.

The fourth experiment was developed to evaluate the effectiveness of the proposed system’s clustering step. This experiment was conducted using the trained Model 1 on the \( D \) dataset. It was trained eight times to construct eight classifiers, one for each constructed cluster. The output layer parameter values were updated only during the training to fit the training cell images of the constructed cluster, and the other parameters were frozen. As shown in Figs. 10 and 11, the detection accuracy was improved based on the Model 1 classifiers of the clusters, indicating the importance of grouping the solar cells into homogeneous clusters.

The proposed approach minimizes the difference value between the actual class label and the predicted label for each cluster cell and to achieve the highest possible accuracy of the detection. Thus, Model 2 is proposed,
which uses the transfer learning approach, adds a new layer with extra training parameters “164,097 parameters” to the trained Model 1, as shown in Fig. 4.

The fifth experiment was conducted using Model 2’s proposed architecture for each constructed cluster. It was trained eight times to construct eight classifiers, one classifier for each constructed cluster. It was assisted by the trained Model 1 layers on the training dataset $D$. The parameter values of all layers in the pre-trained Model 2 were frozen, except for the additional layer and the output layer. Only the parameter values of the additional and output layers were updated during the training. As shown in Tables 2, 3, 4 and 5, the proposed architecture of Model 2 achieved the highest accuracy compared to the

| Algorithm      | Training (%) | Testing (%) |
|----------------|--------------|-------------|
|                | Recall | Precision | Accuracy | F1 score | Recall | Precision | Accuracy | F1 score |
| SVM            | 78.18  | 78.39     | 78.18     | 77.76    | 74.12  | 74.23     | 74.12    | 73.57    |
| Random forest  | 100.00 | 100.00    | 100.00    | 100.00   | 75.15  | 75.08     | 75.15    | 74.96    |
| Naive bayes    | 67.76  | 67.63     | 67.77     | 66.66    | 67.68  | 67.60     | 67.68    | 66.50    |
| K-nearest      | 84.24  | 84.56     | 84.24     | 84.00    | 72.44  | 72.52     | 72.45    | 71.76    |
| LDA            | 93.71  | 93.70     | 93.71     | 93.70    | 71.07  | 70.96     | 71.07    | 70.90    |
| Mobile netV1   | 76.58  | 77.97     | 76.58     | 75.53    | 74.23  | 75.42     | 74.23    | 73.03    |
| Inception      | 77.05  | 77.46     | 77.05     | 76.46    | 72.86  | 72.99     | 72.87    | 72.16    |
| Mobile netV2   | 77.31  | 77.55     | 77.31     | 76.85    | 75.99  | 76.18     | 75.99    | 75.47    |
| Resnet         | 79.09  | 79.53     | 79.10     | 78.61    | 75.31  | 75.69     | 75.30    | 74.63    |
| Exception      | 74.62  | 75.51     | 74.62     | 73.59    | 73.09  | 73.88     | 73.09    | 71.98    |

Fig. 7 Training results (in %) of updating and freezing Model 1 layers to detect the defective cells

Fig. 8 Testing results (in %) of updating and freezing Model 1 layers to detect the defective cells
pre-trained Model 1. These results show the ability and effectiveness of the proposed additional layer parameter values to fit the training cell images for each constructed cluster.

The sixth experiment was performed to use the third step of the training phase. The LDA classification algorithm was trained by the dataset images assisted by the eight cluster labels. After training, the constructed linear discriminant classifier was evaluated to classify the input images into the constructed cluster labels. As shown in Table 6, the LDA classifier achieved the highest accuracy without overfitting compared with SVM, random forest,
naive Bayes and K-nearest classifiers, which indicates the ability of the constructed LDA classifier to assign the input cell image to the closest cluster of cell structure features.

The seventh experiment was conducted to evaluate the proposed approach’s detection phase after constructing the eight classifiers of the proposed Model 2 and LDA classifier in the training phase. It was used to identify defective cells. It used an LDA classifier to classify the input images to the eight cluster labels. Then, it used the Model 2 classifier depending on the classified cluster label to identify the defective cell. The proposed approach achieved the highest accuracy compared to the recent research studied Deep Siamese CNN [1], SVM trained on KAZE/VGG

| Algorithm   | TN  | FP  | FN  | TP  | Recall (%) | Precision (%) | Accuracy (%) | F1 Score (%) |
|-------------|-----|-----|-----|-----|------------|---------------|--------------|--------------|
| C1 classifier | 1330 | 22  | 4   | 564 | 99.29      | 96.24         | 98.64        | 97.74        |
| C2 classifier | 392  | 8   | 20  | 780 | 97.50      | 98.98         | 97.66        | 98.23        |
| C3 classifier | 1324 | 58  | 10  | 1880| 99.47      | 97.00         | 97.92        | 98.22        |
| C4 classifier | 628  | 0   | 2   | 1050| 99.80      | 100.00        | 99.88        | 99.90        |
| C5 classifier | 208  | 4   | 0   | 28  | 100.00     | 87.50         | 98.33        | 93.33        |
| C6 classifier | 308  | 2   | 6   | 644 | 99.07      | 99.69         | 99.16        | 99.38        |
| C7 classifier | 180  | 0   | 7   | 749 | 99.07      | 100.00        | 99.25        | 99.53        |
| C8 classifier | 50   | 7   | 2   | 229 | 99.13      | 97.03         | 96.87        | 98.07        |
| Overall results of model 1 classifiers (K-fold = 5) | 4420 | 101 | 51  | 5924| 99.14      | 98.32         | 98.55        | 98.73        |

| Algorithm   | TN  | FP  | FN  | TP  | Recall (%) | Precision (%) | Accuracy (%) | F1 score (%) |
|-------------|-----|-----|-----|-----|------------|---------------|--------------|--------------|
| C1 classifier | 330  | 8   | 1   | 141 | 99.29      | 94.63         | 98.12        | 96.90        |
| C2 classifier | 98   | 2   | 5   | 195 | 97.50      | 98.98         | 97.66        | 98.23        |
| C3 classifier | 329  | 29  | 10  | 450 | 97.82      | 93.94         | 95.23        | 95.84        |
| C4 classifier | 157  | 1   | 2   | 260 | 99.23      | 99.61         | 99.28        | 99.42        |
| C5 classifier | 39   | 2   | 0   | 19  | 100.00     | 90.47         | 96.66        | 95.00        |
| C6 classifier | 78   | 2   | 0   | 160 | 100.00     | 98.76         | 99.16        | 99.37        |
| C7 classifier | 45   | 0   | 0   | 189 | 100.00     | 100.00        | 100.00       | 100.00       |
| C8 Classifier | 26   | 2   | 1   | 43  | 97.72      | 95.55         | 95.83        | 96.62        |
| Overall results of model 1 classifiers (K-fold = 5) | 1102 | 46  | 19  | 1457| 98.71      | 96.93         | 97.52        | 97.81        |

| Algorithm   | TN  | FP  | FN  | TP  | Recall (%) | Precision (%) | Accuracy (%) | F1 score (%) |
|-------------|-----|-----|-----|-----|------------|---------------|--------------|--------------|
| C1 classifier | 1336 | 0   | 0   | 584 | 100.00     | 100.00        | 100.00       | 100.00       |
| C2 classifier | 412  | 0   | 0   | 788 | 100.00     | 100.00        | 100.00       | 100.00       |
| C3 classifier | 1328 | 0   | 0   | 1944| 100.00     | 100.00        | 100.00       | 100.00       |
| C4 classifier | 603  | 17  | 25  | 1035| 97.64      | 98.38         | 97.50        | 98.01        |
| C5 classifier | 212  | 0   | 0   | 28  | 100.00     | 100.00        | 100.00       | 100.00       |
| C6 classifier | 316  | 0   | 0   | 644 | 100.00     | 100.00        | 100.00       | 100.00       |
| C7 classifier | 174  | 2   | 6   | 754 | 99.21      | 99.74         | 99.15        | 99.47        |
| C8 classifier | 46   | 0   | 6   | 236 | 97.52      | 99.69         | 99.47        | 99.54        |
| Overall results of model 2 classifiers (K-fold = 5) | 4427 | 19  | 37  | 6013| 99.39      | 99.69         | 99.47        | 99.54        |
features [3], proposed CNN [3] and VGG structure-based networks [2], as shown in Table 7.

The proposed approach grouped the training cell images into eight clusters, each cluster contains a set of cell images with high similarity to the structure shape. It constructed eight classifiers of the proposed Model 2, one classifier for each constructed cluster assisted by the trained Model 1 layers. The proposed Model 2 architecture’s additional training parameter focused on distinguishing between the defective and nondefective solar cells for each constructed cluster and ignored the different structure cells of other clusters. This led to not being confused between the different structure shapes of solar cells and achieving the highest accuracy at a significant rate compared with the recent approaches, as shown in Table 7.

### 5 Conclusion

In this study, a proposed system for detecting solar cell defects is presented, which is based on an efficient model that works on low computational power devices. To determine the distinguishing features between defective and nondefective solar cells for each group of homologous cells, and to identify the defective cells without confusion between the different cell shapes, it depended on K-means, MobileNetV2, and linear discriminant algorithms. A comprehensive study conducted to evaluate the proposed system used a benchmark dataset of EL images to cluster the solar cells into eight homologous clusters, construct eight classifiers based on deep learning approaches; one classifier for each constructed cluster, and train the LDA algorithm to assign each cell image to the closest constructed cluster. Compared with SVM, random forest, naïve Bayes, K-nearest, Inception, ResNet, Exception,
MobileNetV1, Deep Siamese CNN, SVM trained on KAZE/VGG features, new proposed CNN architecture and VGG structure-based network algorithms, the proposed system achieved the highest accuracy by a significant rate.

**Funding** Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

**Declarations**

**Conflict of interest** The author has no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

**References**

1. Report of the Deutsche Bank Markets Research (DBMR), https://www.db.com/en/docs/solar_report_full_length. (Accessed in March 2021)
2. Report of the International Energy Agency (IEA) 2020–2021 overview, https://www.iea.org/reports/renewable-energy-market-update/2020-and-2021-forecast-overview. (Accessed in March 2021)
3. Report of the International Energy Agency (IEA) forecasting, https://www.iea.org/news/how-solar-energy-could-be-the-largest-source-of-electricity—by-mid-century. (Accessed in March 2021)
4. Ebner R, Zamin S and Ujjari G (2010), “Defect analysis in different photo voltaic modules using electro luminescence (EL) and infrared (IR) –Thermography”, In: Proceedings of 25th European Photovoltaic Solar Energy Conference and Exhibition (25th EUPVSEC)
5. Jahn U, Herz M, Kontges M, Parlevliet V, Paggi V and Tsanakas I (2021), Report Review on infrared and electroluminescence imaging for PV field application reports IREA-PVPS. T13–10: 2018. International Energy Agency. ISBN:978–3–906042–10–8. (Accessed in March 2021)
6. Acharya A, Sahu P and Jena S (2021) Deep neural network based approach for detection of defective solar cell. Mater Today: Proc 39(5):2009–2014
7. Akram M, Li G, Jin Y, Chen X, Zhu Ch, Zhao X, Khaliq A, Faheem M, Ahmad A (2019) CNN based automatic detection of photovoltaic cell defects in electroluminescence images. Energy 189:116319
8. Deitsch S, Christlein V, Berger S, Buerhop-Lutz C, Maier A, Gallwitz F, Riess Ch (2019) Automatic classification of defective photovoltaic module cells in electroluminescence images. Sol Energy 185:455–468
9. Li X, Yang Q, Lou Z, Yan W (2019) Deep learning based module defect analysis for large-scale photovoltaic farms. IEEE Trans Energy Convers 34(1):520–529
10. Li X, Li W, Yang Q, Yan W, Zomaya AY (2020) An unmanned inspection system for multiple defects detection in photovoltaic plants. IEEE J Photovolt 10(2):568–576
11. Aziz F, UHaq A, Ahmad S, Mahmoud Y, Jalal M, Ali U (2020) A novel convolutional neural network-based approach for fault classification in photovoltaic arrays. IEEE Access 8:41889–41904
12. Appiah A, Zhang X, Ayawli B, Kyeremeh F (2019) Long short-term memory networks based automatic feature extraction for photovoltaic array fault diagnosis. IEEE Access 7:30089–30101
13. Huang C, Kuo P (2019) Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting. IEEE Access 7:74822–74834
14. Li X, Li W, Yang Q, Yan W, Zomaya AY (2020) Edge-computing-enabled unmanned module defect detection and diagnosis system for large-scale photovoltaic plants. IEEE Internet Things J 7(10):9651–9663
15. F. Wang et al (2019)Deep Learning Based Irradiance Mapping Model for Solar PV Power Forecasting Using Sky Image, In: Proceedings of 2019 IEEE Industry Applications Society Annual Meeting, Baltimore, MD, USA
16. Tao C, Wang X, Gao F, Wang M (2020) Fault diagnosis of photovoltaic array based on deep belief network optimized by genetic algorithm. Chin J Electr Eng 6(3):106–114
17. Pierdicca R, Malinverni E, Piccinini F, Paolanti M, Felicetti A, Zingaretti P (2018) Deep convolutional neural network for automatic detection of damaged photovoltaic cells”. ISPRS Int Arch Photogramm Remote Sens Spat Inform Sci. https://doi.org/10.5194/isprs-archives-XLII-2-893-2018
18. Buerhop-Lutz C, Deitsch S, Maier A, Gallwitz F, Berger V, Doll B, Hauch J, Camus C and Brabec C J (2018), A Benchmark for Visual Identification of Defective Solar Cells in Electroluminescence Imagery, In: Proceedings of 35th European Photovoltaic Solar Energy Conference and Exhibition, 2018.
19. Sandler M, Howard A, Zhu M, Zhmoginov A and Chen L (2018) MobileNetV2: Inverted Residuals and Linear Bottlenecks, In: Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA
20. Howard A, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto T and Hartwig Adam, (2017) Mobilenets: Efficient convolutional neural networks for mobile vision applications”, CoRR.abs/1704:04861
21. Na S, Xumin L and Yong G “Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm”, In: Proceedings of 2010 Third International Symposium on Intelligent Information Technology and Security Informatics, Jian, China, 2010
22. Alaa Th, Tarek G, Abdelhameed I, Ella H (2017) Linear discriminant analysis: A detailed tutorial. AI Commun 30(2):169–190

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.