

Full length article

Deep learning-based automated defect classification in Electroluminescence images of solar panels



Hazem Munawer Al-Otum

EE Department, Faculty of Engineering, Jordan University of Science & Technology, Jordan

ARTICLE INFO

Keywords:

Electroluminescence imaging
Convolutional neural network
Deep learning
Defects detection and classification

ABSTRACT

The rise of deep learning approaches has boosted the modern computing capabilities and paved the way to novel outstanding applications. Recently, deep neural networks (DNN) have caught an increasing attraction in the field of PV cells detection and fault classification. Convolutional neural networks (CNN) were proved as DNNs that provide more reliable results in diagnosing these faults. On the other hand, Electroluminescence imaging (EL) analysis has been recognized as a powerful tool for addressing PV models malfunctions. In this work, an effective fault detection and classification approach is developed using multi-scale CNN-based models using two scenarios: a) a transfer learning-based approach using two selected DNNs (ResNet18 and ShuffleNet), and b) an independent light-depth CNN (denoted as CNN-ILD). CNN-ILD is a light CNN based on the insertion of parallel convolutional branches with low-to-high kernel sizes. This construction allowed CNN-ILD to capture a wide variety of features ranging from elementary to specific textures and patterns. The public ELPV dataset has been adopted after proper preprocessing and class categorization. The experimental results have shown promising classification results (88.41%-98.05%) of PV cell defects in EL images. CNN-ILD has a privilege in terms of the computational power/time and stability over the other two proposed pretrained DNNs.

1. Introduction

Recently, the tremendous development in solar photovoltaic (PV) systems has broadly revealed a huge increase in solar power plants. The huge demand on solar systems is vastly growing and becoming widespread in domestic as well as commercial applications [1]. As reported by the International Energy Agency (IEA), the total capacity of the power that depends on renewable resources is expected to grow up to 50% between 2019 and 2024. The PV-based systems are expected to occupy about 60% of the additional capacity [2], with an increase in the global PV-based capacity from around 593.9 GW (in 2019) to around 1500 GW (in 2030) [3]. Such a tremendous growth, urged the need for creating and developing appropriate technologies for PV faults detection and classification. A crucial issue is to assess the PV module performance, using a detailed inspection, even before its installation at the operating site, which allows to take proper actions to avoid wastage in labor and time [4,5].

In PV fabrication market, there are many types of available solar panels, with the crystalline silicon-based cells (c-Si) being very popular. The c-Si cells are fabricated in two types: single/mono (mc-Si) and polycrystalline (pc-Si) cells. These cells have a good, tested energy

conversion efficiency that approaches 20%-25% for mc-Si and pc-Si cells, respectively. These cells cover up to 97% of the PV market due to its well-known fabrication techniques and low-power cost. In such panels, various defects and faults may appear including a) malfunctions, like micro cracks, finger-interruptions, and dislocated areas; b) assembly defects, like soldering defects, disconnections, and lamination flaws; and c) faults due to transportation and field operation, like cracks and breaks. Many of these defects could dramatically decrease the module output power, which urges the need for tools for proper defect detection and classification [6-8].

The commonly utilized approaches for evaluating the PV operational conditions include I-V curve tests, thermal/infrared imaging (IR), and Electroluminescence (EL) imaging [9-16]. In the I-V curve approach, the operating condition of the whole PV module can be evaluated. However, I-V curve tests cannot properly detect and locate minor defects, such as cracks and finger-interruptions [9,11]. In IR imaging approaches, the PV panel is captured using a thermal camera to record the variation in temperature between defect-free, and defected regions on the panel surface [12,13]. Recently, EL imaging has been adopted as a powerful tool to evaluate the quality of PV models. EL imaging is a highly effective fault detection technique that can be implemented to detect diverse

E-mail address: hazem-ot@just.edu.jo.

types of PV defects such as cracks, finger interruption, broken parts, and short circuits [14]. Typically, EL imaging can operate in laboratory conditions within a dark environment to ensure the separation of emitted and surrounding radiation. It can also be in outdoor under low light conditions. EL equipment is easy to integrate into production lines and easy to operate [14–16]. Using EL-testing, hidden defects in the structure of the PV cells can be detected non-destructively. This is a wealth of data about the area uniformity of the PV module surface.

Recently, machine learning has become a vastly increasing tool that can be used in analyzing PV cells. Many research works were devoted to EL imaging. Good examples include the utilization of the independent component analysis (ICA) [17], anisotropic diffusion with support vector machine (SVM) [18], random forests (RF) [19], etc. Ever since around 2010, deep neural networks (DNN) have pushed fault detection and classification to another peak. Specifically, the employment of the convolutional neural networks (CNN) has caught a great attention in this field. The development of AlexNet [20], VGG [21], ResNet [22], GoogleNet [23], MobileNet [24] and other DNNs had triggered new achievements of using CNNs in EL image analysis applications.

In [25], Pierdicca et al. utilized thermal infrared imaging, attached to a drone, to capture about 880 defective and 2500 normally operating PV solar images (from 50-m constant height). The authors developed a transfer learning-based approach, using VGG-16 [21], that led to a moderate detection accuracy of about 69%. In [26], Bartler et al. produced a dataset containing 98,280 cell-images that were captured using EL image cropping with distortion and perspective correction (as a postprocessing stage). The effect of oversampling and cell augmentation was assessed to overcome the imbalanced source dataset. Again, VGG-16 was utilized, in a transfer-learning fashion, to classify the input samples into two classes (normal and defected). An accuracy of around 92% was achieved. In [27], Balzategui et al. implemented an approach for analyzing high-resolution EL images (pc-Si cells) using a CNN-based model. The approach classified the cells into two types, either flawed or flawless cells. To reduce the effect of inadequate data size, a sliding window was adopted. Results showed an accuracy of about 85%. In [28], Karimi et al. introduced an unsupervised approach to classify EL images into degraded and non-degraded categories. They applied the introduced supervised and unsupervised learning with only pc-Si cells being investigated. Otamendi et al. [29], proposed an end-to-end deep learning pipeline that detects, locates and segments cell-level anomalies from entire photovoltaic modules via EL images. Evidently, the prescribed approaches were elaborated on a binary-classification basis, i.e., the input images were classified into either functional or defective cells/modules, with no further defect classification being investigated.

Other approaches were developed to classify the defects of EL images into different categories. In [19] Karimi et al. proposed to classify EL images using SVM, RF and CNN models. Here, three classes were adopted, namely: good, corroded and cracked. Similar classification was investigated by Korovin et al., and was applied to heterojunction solar cells [30]. In [31], Tang et al. tackled the EL-based classification issue of mc-Si PV cells by dividing the dataset into four classes (250 images per class). Data augmentation was used to increase the classification accuracy. A CNN model was developed with an achieved average accuracy of about 92%. When compared to VGG models, the time for testing classification was only about 120 ms, whereas the VGG-based testing consumed about 3323 ms. To avoid overfitting, Tang et al. [6], utilized data augmentation using a generative adversarial network (GAN). They produced a dataset for EL images with around 1800 images. Here, they adopted a four-class scenario, namely: microcrack, breakage, finger-interruption and flawless with 450 EL images per class. The comparison of the proposed CNN model with the existing models (ResNet-50, MobileNet, Inception V3, and VGG-16) has shown an average testing accuracy of about 83%. In [32], Lin et al. tackled the problem of imbalanced EL dataset by applying a cropping stage to yield single solar cells which enabled to reduce the computational cost as well as to improve the detection of small-to medium-sized defects. Here, two light-

CNN networks were proposed with achieved a recall and a precision of about 98.8% and 99.4%, respectively. In [33], Chen et al. developed an automatic and open-source pipeline to analyze EL images, with an ability to recognize various EL defects including intra-cell, cracks, oxygen-induced defects, and soldering disconnections. Here, the approach was based on utilizing different machine learning models like YOLO, ResNet18, ResNet-150, ResNet-152, and RF models. Simulations showed that an average F1-score of about 0.83 was achieved on real world EL datasets.

In the prescribed approaches, the utilized datasets were problem specific. Evidently, it is not adequate to objectively judge approaches when utilizing different datasets. In fact, the only available public EL image dataset is the dataset named 'ELPV' and was published in [34]. In [13], Deitsch et al. utilized the ELPV dataset in developing a CNN-based networks for EL image analysis. Here, EL images were classified into four categories: non-defected, possibly normal, possibly defected, and defected. An average accuracy of about 88.4% was achieved that outperformed another introduced model based on the SVM approach. In [33], Akram et al. adopted the ELPV dataset to propose a light CNN model for EL image fault detection and classification with an accuracy of about 93%. In [8], Demirci et al. proposed Deep Feature-Based (DFB) method, for detection and classification of defects in PV modules, using the EL images. Here, the image features were extracted by utilizing various combinations of a variety of DNNs. Then, classical machine learning methods were deployed to combine the obtained features. Next, feature selection is obtained using the minimum Redundancy Maximum Relevance (mRMR) algorithm which led to reduce the dimensions of the feature vectors while increasing the classification performance with a reduced number of features. The classification was elaborated using 2-class and 4-class datasets with the highest scores being achieved using the combination of DFB and SVM. Classification scores of 94.5% and 90.6% were obtained, respectively. However, the scheme required heavy complexity due to the utilization of four different DNNs (ResNet-50, VGG-16, VGG-19, and DarkNet-19) for feature extraction along with other machine learning approaches for feature selection.

It is worthy to note that, in all the prescribed research works, the classification was conducted based on utilizing 2–4 classes with combining the mc-Si and pc-Si cells, having similar features, in the same class. Up to our knowledge, the only work that utilized further split of these cells was recently published in [36], however, the proposed approach, in [36], utilized the PVEL-AD dataset [37]. Nevertheless, the achieved average accuracy was about 93.6%.

Analyzing the available literature, it is worthy to note that:

- 1- Various works were elaborated on different EL image datasets which makes their achievements problem specific.
- 2- Most of the surveyed works compared the performance of classification and object detection using the validation set which is improper. In fact, the validation set is a part of the training set and does not reflect the accurate performance when applying the approach to samples from outside the training/validation dataset.
- 3- Despite of the achieved good success, however, there is still a potential possibility for further improvement by applying: a) better preparation of the dataset, to be more balanced, for better detection and classification performance, b) further exploitation of classes of the PV faults in EL images, and c) further improvement of classification accuracies by either using transfer learning on existing models, or by developing light DNNs.

Consequently, the main contribution of this work can be briefly summarized in proposing an effective approach for detection and classification of faults, in EL images of the PV cells, based on deploying multi-scale CNN models. The approach was elaborated in two scenarios: a) a transfer learning-based approach using two selected DNNs (ResNet18 and ShuffleNet), and b) an independent light-depth CNN (CNN-ILD) that is based on inserting parallel convolutional branches

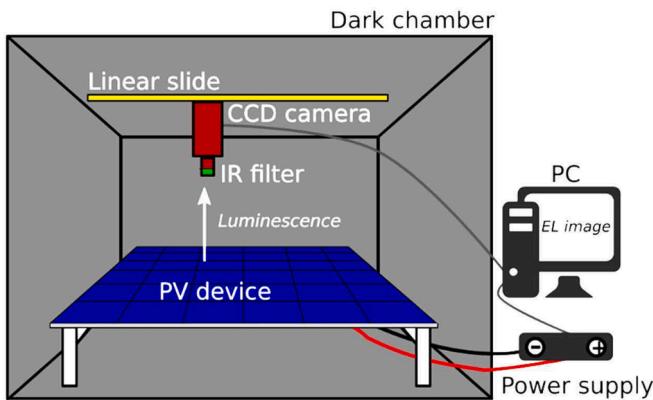


Fig. 1. EL imaging setup.

with low-to-high kernel sizes to enable better performance in capturing the multiscale spectral features of defects. Moreover, the work has adopted the public ELPV dataset with proper preprocessing and class categorization. The experimental results have shown high promising classification scores with a privilege, granted to the proposed light weight CNN-ILD, in terms of the computational power and stability.

In detail, the technical novelty and contribution can be outlined as follows:

- 1- Proposing a modified effective fault detection and classification approach using CNN-based models. In fact, such methods were already deployed by previous research works and lacks scientific contribution to the domain. In this context, an effective fault detection and classification approach is developed using multi-scale CNN-based models in a transfer learning-based fashion. Two selected DNNs were tested, the former is the ResNet18 that was selected due to its remarkable performance in image classification contest with an ability to construct stable and very deep networks. The second is the ShuffleNet which was selected due to its computation-efficient features, even with devices having limited computing power, beside its promising performance in classification tasks. Here, the deployed ResNet18 and ShuffleNet have medium- and high-depth constructions with 71 and 172 layers, respectively.

- 2- Proposing a learning-from-scratch approach with a light-depth CNN-based model based on inserting parallel convolutional construction to capture multiscale image spectral features.
- 3- Applying a proper class categorization by splitting mc-Si and pc-Si cells, having similar features, into separate classes to properly achieve better classification performance.
- 4- Performing a testing stage, in the proposed approaches, against the testing dataset. The testing dataset includes samples that were not being used during the training/validation process.

The rest of this article is structured in the following manner: in [Section 2](#), preliminaries on EL imaging, PV modules, and their defects are presented briefly. In [Section 3](#), the main methodology of the general framework, the utilized dataset, the proposed CNN-models are explained in detail. In [Section 4](#), experimental results, study findings, and comparisons are worked out. Finally, conclusions are capitalized in [Section 5](#).

2. Preliminaries on EL imaging dataset and PV module defects

2.1. EL imaging for solar cells and modules

EL imaging is a state-of-art imaging technique employed to test PV cells and modules, that was originated by Fuyuki et al. [\[14\]](#). The Electroluminescence phenomenon can be obtained as a result of the radiative recombination of electrons and holes, especially in a semi conductor material. The excited electrons release their energy as photons. The electrons and holes may be separated either by doping the material to form a p-n junction (LED), prior to recombination, or through excitation by impact of high-energy electrons accelerated by a strong electric field.

In EL imaging, when the cell is forward biased, it emits energy in the 1100-nm range [\[38\]](#). The emitted energy can be captured by specialized CCD-cameras. The entire operation is performed in dark conditions to ensure high capturing performance as shown in [Fig. 1](#). Examples of the captured EL images are shown in [Fig. 2](#) (taken from the ELPV dataset [\[34\]](#)).

2.2. PV cell defects

The structural quality of the solar cells and modules can be assessed using EL imaging tests. Here, different types of defects can be found,

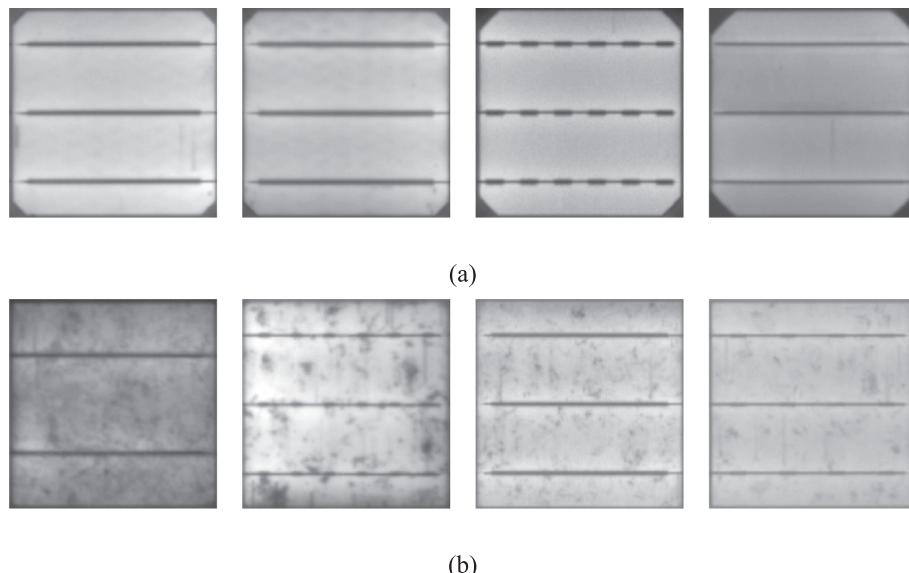


Fig. 2. Examples on EL images taken from the ELPV dataset [\[34\]](#): a) mono c-Si cells (mc-Si), and b) poly c-Si cells (pc-Si).

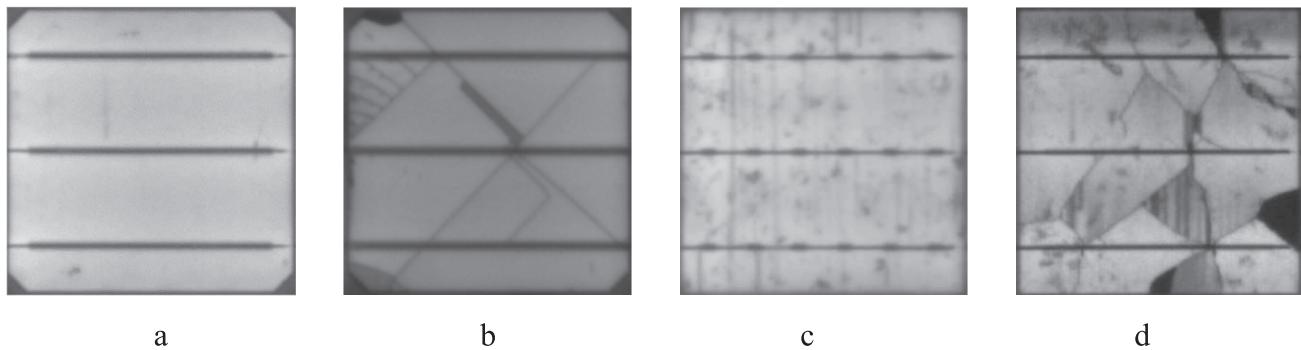


Fig. 3. Illustrative examples on PV cells' defects with mainly: a) finger-interruptions in mc-Si, b) cracks in mc-Si, c) finger-interruptions in pc-Si, b) cracks in pc-Si cells [34].

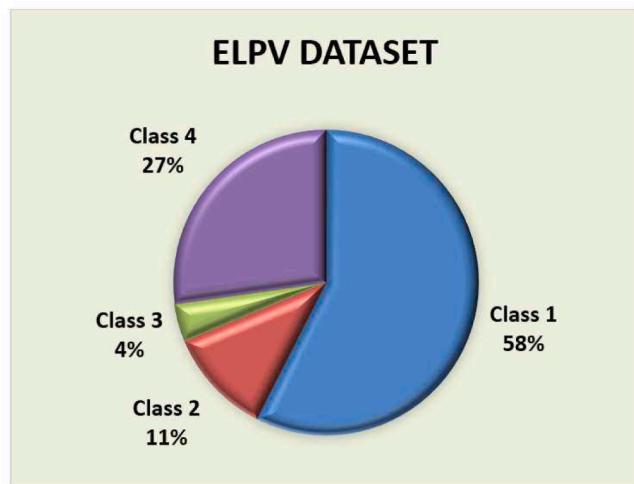


Fig. 4. The distribution of EL images in the ELPV Dataset.

including microcracks, cell cracks, finger-interruptions, disconnected cells, soldering defects, PID defects, diode failure, etc. Fig. 3 demonstrates illustrative examples on PV cells that are mainly defected with finger-interruptions and cracks in both mc-Si and pc-Si cells (taken from the ELPV dataset [34]). Apparently, some of these defects may not dramatically affect the module's output power, others may lead to the appearance of hot spots and a loss of power production or even failures. This imposes the need for urgent procedures to perform automatic inspection of defects using EL imaging technology [39].

3. Methodology

In this section, the proposed methodology adopted to classify defects in PV modules, based on the EL imaging concept, is discussed in detail. Apparently, the methodology is carried out in three phases: a) pre-processing of EL images and dataset preparation, b) training the DNNs using the produced EL image dataset, and c) testing the EL images using the proposed trained DNNs.

3.1. EL imaging dataset and data augmentation

3.1.1. The ELPV dataset

As prescribed, the commonly available, in literature, public EL image dataset is the dataset named as 'ELPV'. This dataset was created by Buerhp et al. in [34]. This distinctive dataset contains 2624 EL image samples of 300x300 pixels with 8-bpp grayscale images of functional and defective solar cells. These images were extracted from 44 different solar modules with varying degree of degradations. In ELPV, all

Table 1
ELPV augmentation operations.

#	Operation	Details
1	image translation	by (1,1), (2,2), (-1,-1) and (-2,-2)
2	image flipping	horizontal and vertical flips
3	image rotation	by 90°, 180°, and 270°
4	image resizing	resizing by 1.01 and 1.02 with cropping to the original size
5	image adjusting	by adjusting image intensity values to (0.1–0.9) and (0.15–0.85) of its original values
6	other manipulations	Histogram equalization, random cropping, slight noise addition, etc.

extracted images were normalized with respect to size and perspective. Moreover, possible camera-induced distortion was eliminated prior to solar cell extraction. In ELPV, the cell images were taken from either mc-Si or pc-Si PV modules. According to the malfunction level, the dataset images were labeled and classified into four classes, namely: a) non-defected (0% defect rate), b) possibly normal (33% defect rate), c) possibly defected (66% defect rate), and d) defected (100% defect rate). The distribution of the images into these classes is shown in Fig. 4, where it is clear that the samples' distribution is unbalanced. Such a distribution may lead to the overfitting issue during the training stage.

3.1.2. Improved ELPV dataset

As reported in literature, there is a lack of large datasets with a balanced number of images per class. This is considered as one of the major restrictions for achieving high performance in automatic

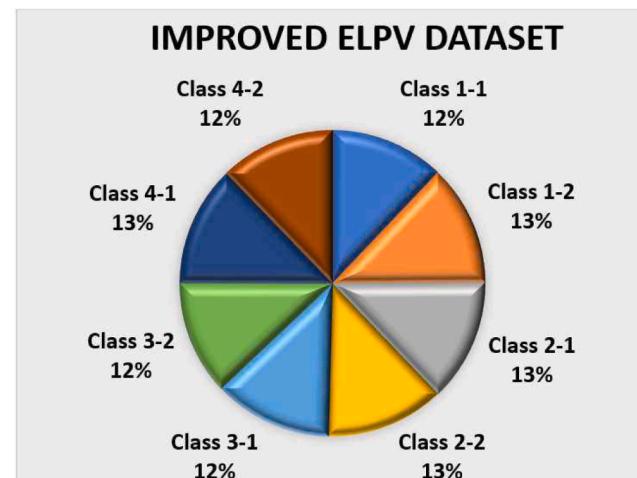


Fig. 5. The distribution of EL images in the 'improved' ELPV Dataset.

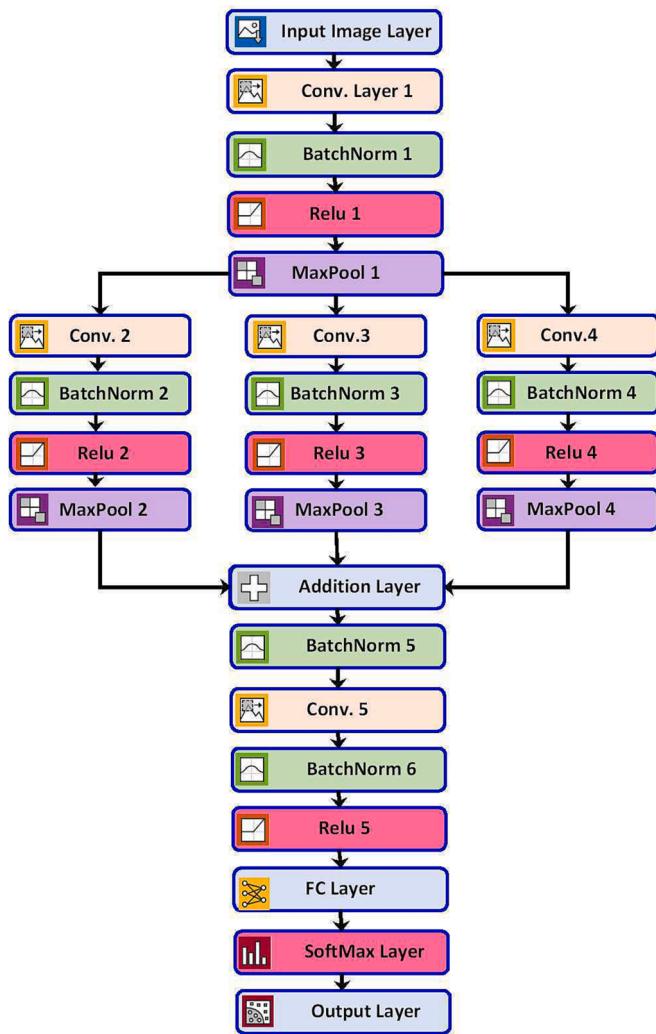


Fig. 6. The block diagram of the proposed CNN-ILD.

classification of PV anomalies, especially for DNN training. To achieve almost balanced classes, it is recommended to perform data augmentation (DA). The DA is a technique that utilizes computational methods for increasing the training set while preserving its original label. DA can be achieved either by using simple geometric transformations (flipping, rotation, cropping, etc.), or by utilizing more advanced methods such as Generative Adversarial Networks (GANs) [6]. In this work, for the sake of simplicity, DA has been performed using various image processing manipulations. The utilized DA operations are depicted in Table 1.

The employed DA led to expand the ELPV dataset with almost balanced number of EL samples per class. Moreover, each of the ELPV classes were further decomposed into two sub-categories: one for mc-Si, and the other for pc-Si cells. In fact, the concatenation of mc-Si and pc-Si cells in one class may lead to misclassification of the samples and, hence, may reduce the classification accuracy as will be seen in next sections. The adopted procedure for DA and class-splitting led to produce an ‘improved’ ELPV dataset having the distribution of EL images as shown in Fig. 5. Clearly, all 8 classes have almost balanced number of samples (with around 12–13% of images per class) with a total number of samples of around 8400 EL images.

3.2. CNN-based automatic defect classification

CNN models have been proven as powerful tools in image detection and classification applications. CNNs are DNNs constructed for

processing structured data arrays, such as images, which are widely implemented in computer and visual applications. For DNNs, it is common to use transfer learning with a pretrained DNN instead of training it from scratch. Transfer learning is commonly utilized, where pre-trained models are used as the starting step for another task. Here, one can use DNNs with comparatively little data, as in most real-world problems, in a transfer learning-based fashion which considerably reduces the time and computational costs of the training stage.

In this work, to bridge the gap in research and to override possible limitations, two scenarios were deployed for efficient defect classification of EL images using CNN-based models. The former is based on using a light-depth CNN, while the second is based on a transfer learning approach. In the former scenario, the depth of the developed CNN is comparably low, while in the second, two pretrained DNNs were selected, one with medium number of layers, while the other is with a high depth. The details of these scenarios are described in the subsequent sections.

3.2.1. Defect classification using a light-depth CNN model

The proposed independent light-depth CNN model (CNN-ILD) is designed as a multi-scale CNN and considered light-depth since it only contains 25 layers. Fig. 6. depicts the block diagram of the proposed CNN-ILD model. As depicted in Fig. 6., the proposed CNN-ILD model accepts input images of the size 246x246 pixels. The CNN-ILD is constructed of the following building layers:

Table 2
CNN-ILD hyperparameters.

Layer Type	Parameters	Activations	Learnables
Input Image	246 × 246 × 3 with “zero center” normalization	246 × 246 × 3	—
Conv. Layer 1	[3,3], Num of Filters = 8, Stride=[1,1].	246 × 246 × 8	weights = 3 × 3 × 3 × 8 Bias = 1 × 1 × 8 offset = 1 × 1 × 8 scale = 1 × 1 × 8
BatchNorm 1	8 channels	246 × 246 × 8	—
Relu 1	—	246 × 246 × 8	—
Maxpool 1	[2,2], Stride=[2,2].	123 × 123 × 8	—
Conv. 2	[5,5], Num of Filters = 16, Stride=[1,1].	123 × 123 × 16	weights = 5 × 5 × 8 × 16 Bias = 1 × 1 × 16
Conv. 3	[7,7], Num of Filters = 16, Stride=[1,1].	123 × 123 × 16	weights = 7 × 7 × 8 × 16 Bias = 1 × 1 × 16
Conv. 4	[3,3], Num of Filters = 16, Stride=[1,1].	123 × 123 × 16	weights = 3 × 3 × 8 × 16 Bias = 1 × 1 × 16
BatchNorm 2–4	16 channels	123 × 123 × 16	offset = 1 × 1 × 16 scale = 1 × 1 × 16
Relu 2–4	—	123 × 123 × 16	—
Maxpool 2–4	[2,2], Stride=[2,2].	61 × 61 × 16	—
Addition Layer	—	61 × 61 × 16	—
BatchNorm 5	16 channels	61 × 61 × 16	offset = 1 × 1 × 16 scale = 1 × 1 × 16
Conv. 5	[3,3], Num of Filters = 32, Stride=[1,1].	61 × 61 × 32	weights = 3 × 3 × 16 × 32 Bias = 1 × 1 × 32
BatchNorm 6	32 channels	61 × 61 × 32	offset = 1 × 1 × 32 scale = 1 × 1 × 32
Relu 5	—	61 × 61 × 32	—
FC Layer	output classes = 8 / 4	1 × 1 × 8	weights = 8 × 119072 / 4 × 119072 bias = 8 × 1 / 4 × 1
SoftMax	—	1 × 1 × 8	—
Output Layer	Classification layer with crossentropyex	—	—

- A. Convolution Layers (Conv1-Conv5):** represent 2D convolution layers that are mainly used to extract and learn features from input images. The first convolution layer (Conv1) is used to extract elementary features like shape and boundary. Conv2-4 layers encode simple textures constructed from compositions of colors and edges, while Conv.5 extracts textures and specific patterns from images. In CNN-ILD, a crucial point is the insertion of the parallel construction of Conv2-Conv4 layers. Here, low-, medium- and high-level-sized filters (3×3 , 5×5 and 7×7 , respectively) are utilized to cope multiscale spectral features. Consequently, various features are captured by the parallel combination of these filters.
- B. Activation Layers (Relu1-Relu5 and SoftMax):** As in most of the CNN networks, activation layers are incorporated as the non-linear detection layers. The rectified linear unit (Relu) is the activation layer utilized at the input as well as the intermediate layers due to its superior performance, accelerated convergence speed, and simple learning strategy when compared to other activation functions. The SoftMax activation function is commonly utilized for multivariate classification tasks and can be adopted as the readout layer to normalize the probability distribution of the output predicted classes. Therefore, for the last layer in CNN-ILD, SoftMax activation is utilized.
- C. Batch Normalization Layers (BatchNorm1-6):** The employment of the batch normalization layers improves the CNN-ILD training and allows the network to learn faster while reducing the overfitting situations. It operates by normalizing the input to each layer of the network (with zero mean and unit standard deviation) which allows the training to be unified.
- D. Maxpooling Layers (MaxPool 1-4):** by applying maxpooling, the dimension of the inputs, coming from activations, is reduced based on the statistics of neighboring cells. This leads to extract more learned features and improves the network performance in feature representation formulation.
- E. The Fully Connected Layer (FC):** represents the actual component that performs the decision step on the network. The FC layer is constructed of a multi-layer perceptron, where each neuron is connected to all the neurons in preceding layers. FC can learn weights employed in categorizing specific object classes.

For better comprehension, the CNN-ILD specifications are shown in Table 2. Here, the layers parameters (if any), activations and learnables are depicted in detail.

3.2.2. Defect classification using transfer learning-based CNN models

Transfer learning is an efficient procedure that leads to a significant reduction in the training time and computational load as well. As part of transfer learning trade-off, network testing performance could be relatively moderate. In this work, two pretrained DNNs were selected for ELPV dataset: the ResNet18 with medium depth of 18 (71 layers), and the ShuffleNet with high depth of 48 (172 layers). More details on these DNNs are given below:

- 1) RESNETS: are considered as one of the most powerful DNNs that achieved remarkable performance in image classification contest. ResNets are CNN-based networks with an architecture that enables to construct very deep networks, with a very large number of layers, by overcoming the ‘vanishing gradient’ issue. RESNETs have many architecture versions, with varying number of layers, while keeping the same concept including ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-110, ResNet-152, ResNet-164, etc. ResNets utilize the skip connection concept that allows to handle the activation from one layer and suddenly feed it into another layer even much deeper in the network. ResNets are built of the so called “the residual block” and learn residual functions with reference to the layer inputs, instead of learning unreference functions. Instead of hoping each few stacked layers directly, residual nets let these layers fit a residual

Table 3
Adopted parameters for the training stage.

Parameter	Proposed CNN-ILD	ResNet18	ShuffleNet
Number of Classes	4/8	4/8	4/8
Used Solver	SGDM	SGDM	SGDM
Inti. Learn. R.	0.0001	0.0001	0.0001
Valid. Frequency	5	5	5
Max Epochs	50	20	20
Mini Batch	16	16	16
Number of Images / Training	5880	5580	5880
L2Norm	0.0001	0.0001	0.0001
Grad. Th. Meth.	L2Norm	L2Norm	L2Norm
Layers/Depth	25	71/18	172/50

mapping with the **residual blocks** being stacked on top of each other to form the network. For instance, the ResNet-18 has 18 layers built using these blocks. Here, residual learning is performed on multiple layers. Let $\mathcal{H}(x)$ denote the required underlying mapping, then the stacked nonlinear layers follow the residual mapping $\mathcal{F}(x) = \mathcal{H}(x) - x$. Thus, the original mapping is recast into $\mathcal{H}(x) = \mathcal{F}(x) + x$. Such a configuration enables to train very deep networks and there is empirical evidence that this type of networks is easier to be optimized and can attain high accuracy from significantly increased depth. In this work, the ResNet-18 has been selected in a transfer learning scenario for EL image classification. ResNet-18 is a CNN-based network with a medium number of layers (71) and depth of 18. It was trained on more than a million images from the ImageNet database [38] and can classify images into 1000 object categories (keyboard, mouse, pencil, and many animals).

- 2) The ShuffleNet: is a computation-efficient CNN-based network that was developed for mobile devices with very limited computing power. The main concept of the ShuffleNet is to apply pointwise group convolution and channel shuffle operation. ShuffleNet is constructed of repeated building blocks that employs the concept of efficient group convolutions to maintain an excellent trade-off between performance capabilities and computational power. The idea of group convolution is to separate the input tensor (of g subtensors) each one 1/g distinct channels of the initial tensor. Then a pointwise (1x1) convolution is applied to each subtensor and finally the subtensors are concatenated together. It also utilizes the concept of channel shuffling. The shuffling operation can be obtained by randomizing the order in which the data is provided to the DNN during training. This enables the model to learn the general data patterns, rather than memorizing its data specific order. ShuffleNet is constructed of repeated building blocks and starts with a stage of convolution-max pooling that is followed by a number of stages, each of which contains a number of shuffle blocks, and finally closes with a global fully-connected block. Despite of the large number of layers (172) and depth (50), the ShuffleNet architecture enables to reduce the computational cost while maintaining a high accuracy.

It is worthy to note that the main reasons for selecting the ResNet18 and ShuffleNet, for defect classification, using transfer learning are owned to:

- a) the training feasibility on a single-GPU computer,
- b) the excellent trade-off between representation capability and computational cost,
- c) both ResNet18 and ShuffleNet were not employed, up to our knowledge, in EL image classification using the ELPV dataset.

4. Experimental results and discussion

4.1. Initialization

In this section, experimental evaluations are elaborated to assess the

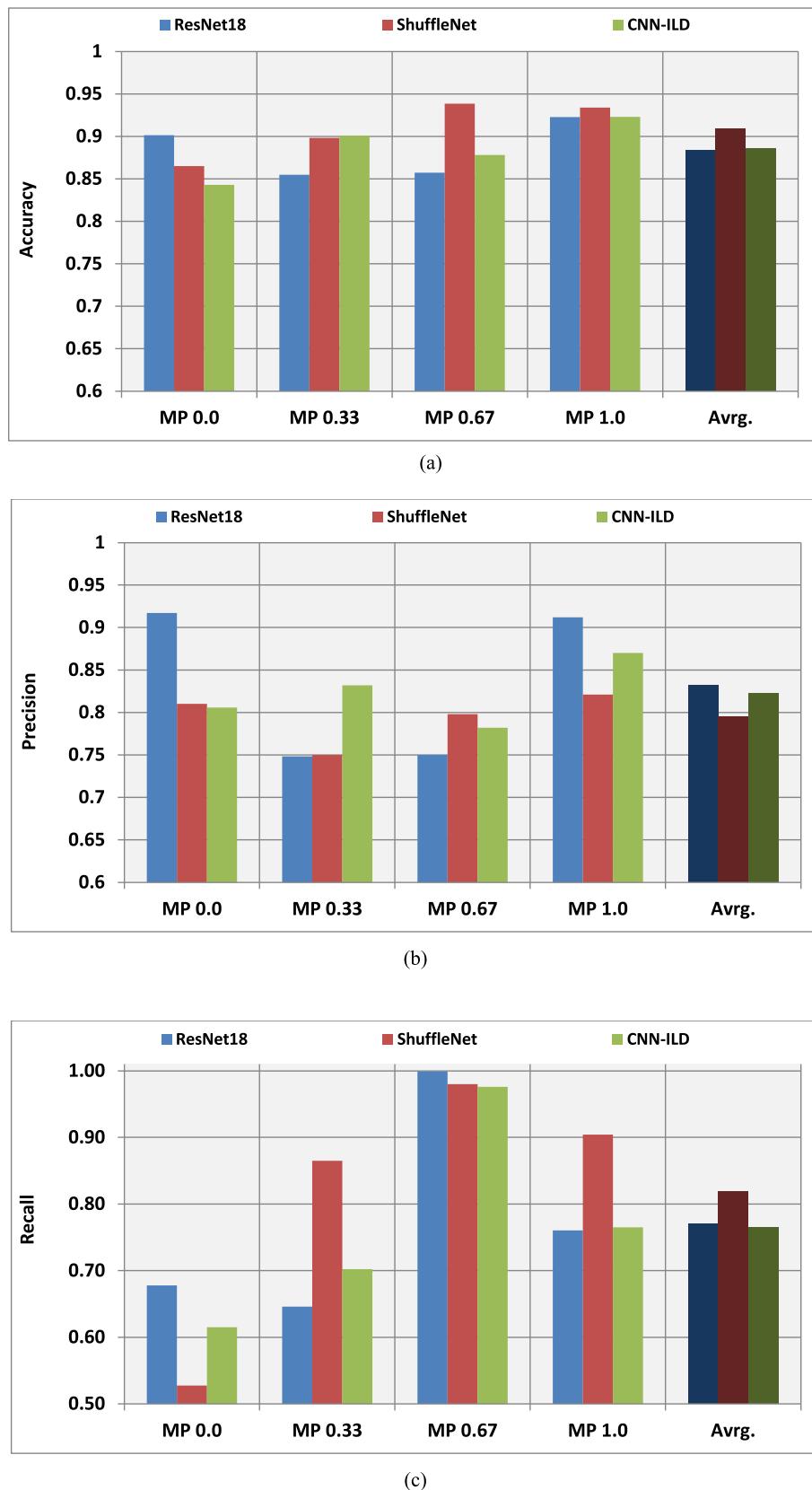
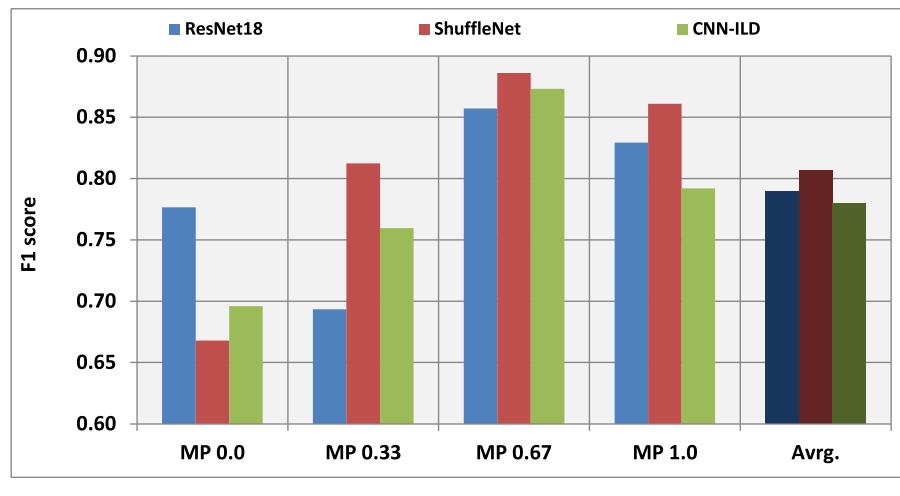


Fig. 7. The 4-class experimental performance (on the validation set) in terms of the: a) Accuracy, b) Precision, c) Recall, and d) F1-score.



(d)

Fig. 7. (continued).

statistical validity of the prescribed scenarios. All experiments were carried out using MATLAB® R2020a, supported with Deep Learning Toolbox, on a windows-10 environment. The training was performed using a computer with Intel(R) Core(TM) i7-10510U CPU @ 2.30 GHz., 16 GB RAM memory with NVIDIA GeForce MX250 graphics unit. The adopted parameters of the training stage are shown in Table 3. Here, the improved ELPV dataset was randomly split into 70%, 15%, and 15% for training, validation, and testing, respectively. All DNNs were trained using a Stochastic Gradient Descent (SGD) optimizer with a learning rate of $n = 10^{-4}$. The number of epochs was chosen to be 50 for the proposed CNN-ILD and 20 for the pretrained ResNet18 and ShuffleNet networks.

4.2. Evaluation metrics

To quantitatively asses the performance of the proposed EL-image classification using the prescribed DNNs, various evaluation metrics were adopted, namely: a) the accuracy (Acc), b) the precision (P), c) the F1-score (F1), and d) the recall (R). The accuracy is a measure of the model performance including all classes. The precision is defined as the truly positive results divided by the predicted positive results, i.e., it reflects the fraction of the predicted defects that were identified correctly, while the recall is defined as the ratio of true positive predictions to actual positives, i.e., it represents the fraction of actual defects that were correctly labeled by the model. The F1-score represents the harmonic mean of the precision and the recall. These metrics are defined according to:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = 2 \frac{PxR}{P + R} \quad (4)$$

where, TP , TN , FP , and FN are the true positive, true negative, false positive and false negative predictions. The obtained cell classification results are discussed in the following sections.

		Precision			
		MP 0.0	MP 0.33	MP 0.67	P 1.0
TRUE LABEL	MP 0.0	285 17.8%	38 2.4%	22 1.4%	24 1.5%
	MP 0.33	51 3.2%	287 17.9%	61 3.8%	42 2.6%
	MP 0.67	46 2.9%	53 3.3%	282 17.6%	55 3.4%
	P 1.0	18 1.1%	22 1.4%	35 2.2%	279 17.4%
		Accuracy			
		Recall	71.3%	71.8%	70.5%
					69.8%
					80.81%

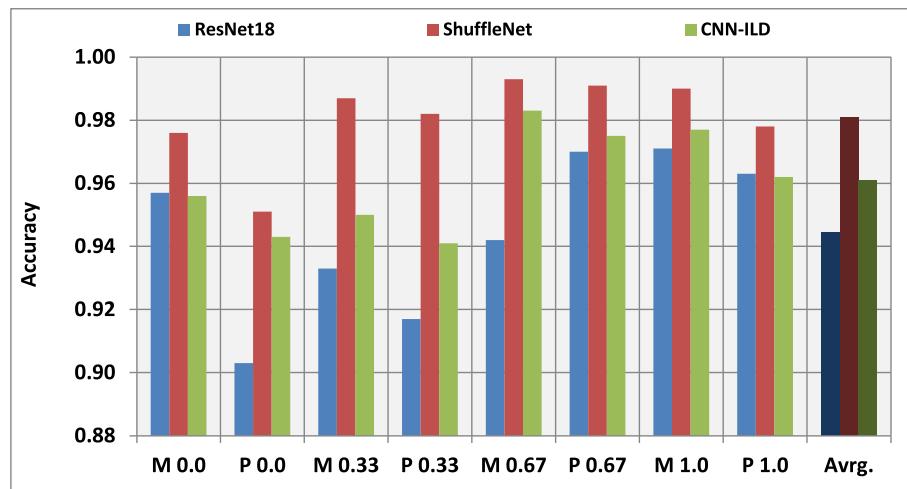
Fig. 8. The 4-class confusion-chart of the proposed NCC-ILD for the testing dataset.

4.3. EL image cell classification results

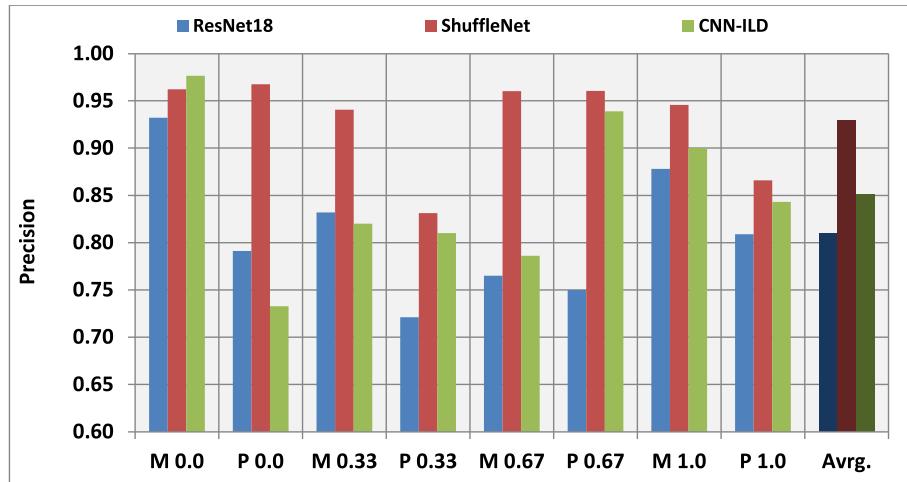
In this section, two approaches were elaborated, the former is carried out using the 4-class scenario with combining the mc-Si and pc-Si cells in the same class, while in the second, the mc-Si and pc-Si cells are split into separate classes which produces an 8-class classification task.

4.3.1. The 4-Class classification results

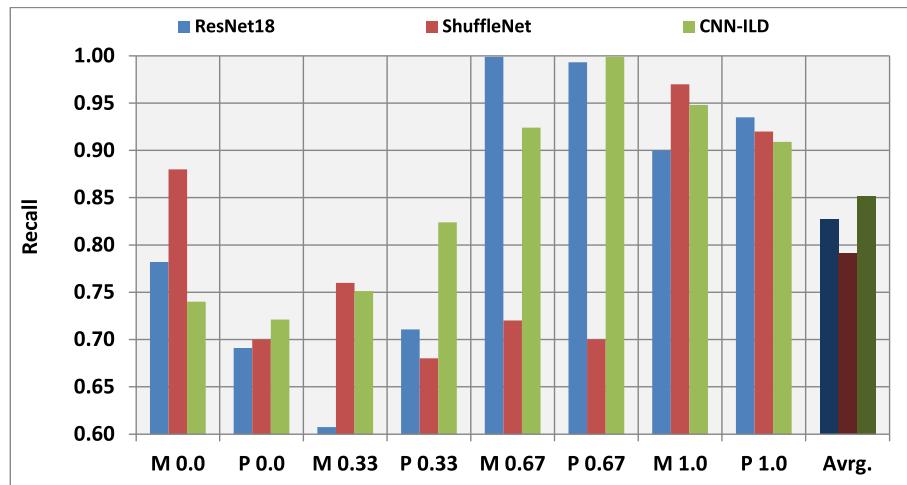
The performance of the EL image defect classification was elaborated using 4 classes according to the cell defective portion (0.0, 0.33, 0.67 and 1.0 rates). Here, the mc-Si and pc-Si cells having the same classification were combined in one class as supposed by the ELPV dataset. Henceforth, these classes are denoted as MP 0.0, MP 0.33, MP 0.67, and MP 1.0, respectively. In fact, this scenario was adopted by previous works (like [6,8,13]). The experimental results, for the validation set, are illustrated in Fig. 7.a-d, in respect to the accuracy, precision, recall and F1-score. The average value of the accuracy, precision, recall and F1-score are also depicted in Fig. 7a-d, respectively. Furthermore, the confusion-chart of the proposed NCC-ILD is demonstrated in Fig. 8. for the testing dataset. As prescribed, the testing dataset contains new



(a)

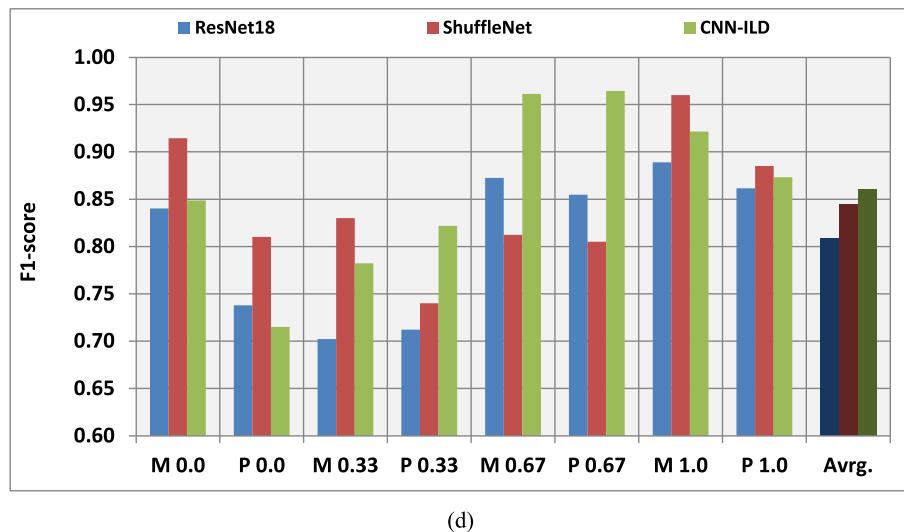


(b)



(c)

Fig. 9. The 8-class experimental performance (on the validation set) in terms of the: a) Accuracy, b) Precision, c) Recall, and d) F1-score.



(d)

Fig. 9. (continued).

		Precision								
		Accuracy								
TRUE LABEL	PREDICTED LABEL	M 0.0	172 10.8%	8 0.5%	2 0.1%	3 0.2%	0 0.0%	2 0.1%	0 0.0%	0 0.0%
		P 0.0	12 0.8%	182 11.4%	0 0.0%	0 0.0%	5 0.3%	2 0.1%	3 0.2%	6 0.4%
		M 0.33	0 0.0%	0 0.0%	173 10.8%	13 0.8%	2 0.1%	3 0.2%	7 0.4%	5 0.3%
		P 0.33	6 0.4%	5 0.3%	15 0.9%	170 10.6%	0 0.0%	3 0.2%	0 0.0%	0 0.0%
		M 0.67	0 0.0%	0 0.0%	5 0.3%	10 0.6%	169 10.6%	9 0.6%	6 0.4%	10 0.6%
		P 0.67	6 0.4%	5 0.3%	3 0.2%	4 0.3%	15 0.9%	159 9.9%	0 0.0%	3 0.2%
		M 1.0	4 0.3%	0 0.0%	2 0.1%	0 0.0%	7 0.4%	13 0.8%	172 10.8%	18 1.1%
		P 1.0	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	9 0.6%	12 0.8%	158 9.9%
		Recall	86.0%	91.0%	86.5%	85.0%	84.5%	79.5%	86.0%	79.0%

Fig. 10. The 8-class confusion-chart of the proposed NCC-ILD for the testing dataset.

samples that were not used during the training process and represents about 15% of the total number of samples.

4.3.2. The 8-Class classification results

The performance of the EL image defect classification was also carried out using 8 classes by splitting the mc-Si and pc-Si cells (with similar features) into separate classes. According to the cell defective portion, these classes were denoted as: M 0.0, P 0.0, M 0.33, and P 0.33, M 0.67, P 0.67, M 1.0, and P 1.0, where M and P denote the mc-Si and pc-Si cells, respectively. The experimental results, for the validation dataset, are illustrated in Fig. 9. Here, the average value of the accuracy, precision,

recall and F1-score are also depicted in Fig. 9.a-d, respectively. Furthermore, the confusion-chart of the proposed NCC-ILD is depicted in Fig. 10 for the testing dataset.

For better visualization, illustrative examples on the evolution curves of the accuracy and loss functions during training of the proposed DNNs are shown in Fig. 11. Clearly, no overfitting is observed in the tested DNNs.

4.4. Discussions and comparisons

In this section, the elaborated results are discussed and analyzed. A

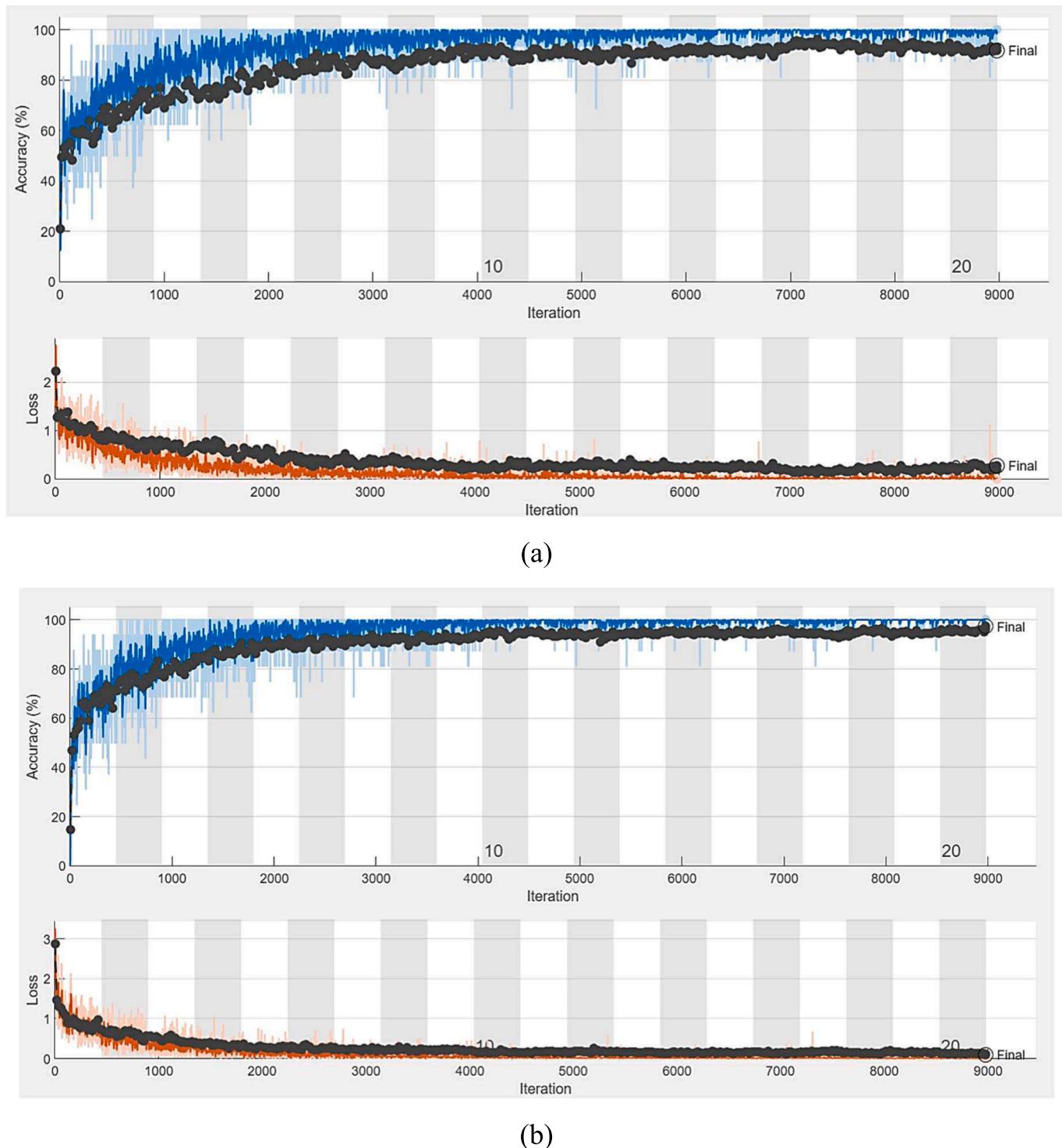


Fig. 11. Illustrative examples on the evolution curves of the training process using: a) ResNet18, b) ShuffleNet, and c) CNN-IDL networks.

comparative analysis is also provided with other existing, recently published, techniques.

4.4.1. Results discussions

In this section, it would be helpful to provide a more in-depth discussion of the obtained results. Referring to Figs. 7–11, results discussion will be presented, in subsequent subsections, based on the achieved scores for accuracy, precision, recall, and F1-score.

A. Accuracy

As shown in Fig. 7.a, and Fig. 9.a, high average validation accuracies were accomplished using the proposed DNNs with an average in the range of (88.41–90.89%) and (94.31–98.05%) in case of the 4-class and 8-class scenarios, respectively. For better visualization, Fig. 12.a demonstrates a comparison of the achieved validation accuracies using the prescribed DNNs, while the achieved accuracy improvements (in case of 8-class over the 4-class scenarios) is depicted in Fig. 12.b. Here, the

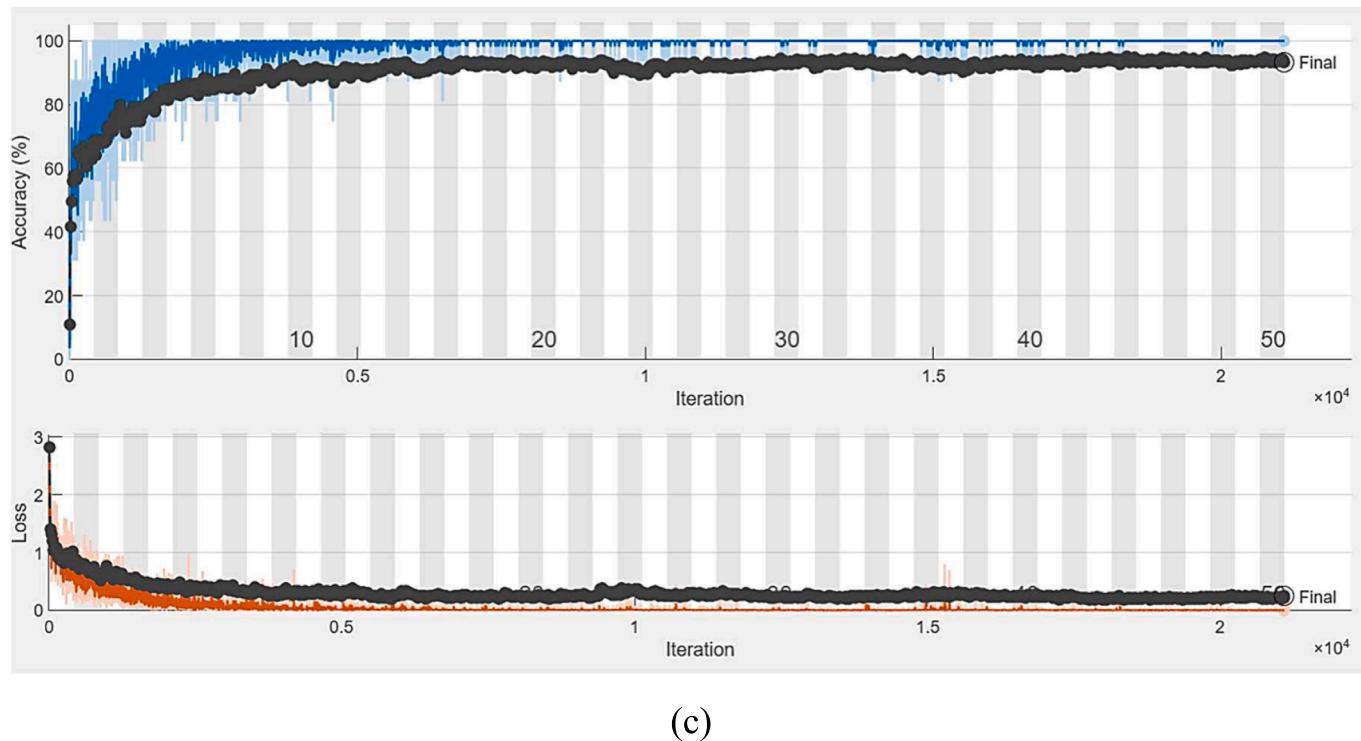


Fig. 11. (continued).

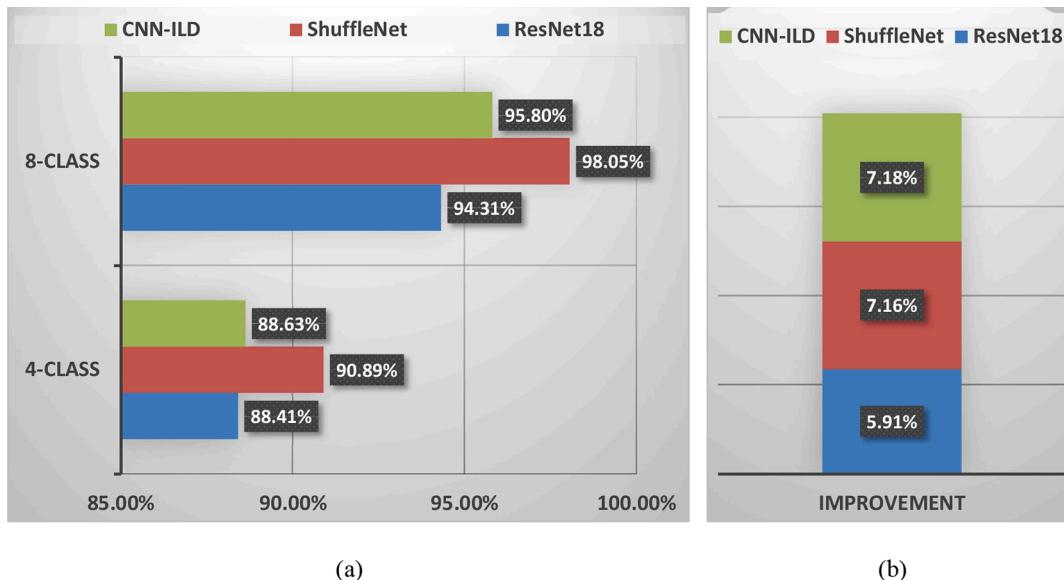


Fig. 12. Comparison of the achieved validation accuracy using the prescribed DNNs: a) average accuracies, and b) obtained improvement of using the 8-class over the 4-scenario.

proposed CNN-IID attained a high average accuracy of about 95.8%, while ShuffleNet came in the first place with an average accuracy of 98.05%. Moreover, an improvement, in the range of (5.91–7.18%), has been achieved when using the 8-class over the 4-class scenario.

The obtained average testing accuracies are depicted in Fig. 13.a, with illustrating the attained accuracy improvements (in Fig. 13.b). Clearly, the proposed DNNs exhibited moderate accuracies (around 70%) in case of the 4-class scenario, while exhibiting very good testing accuracies (around 84%) in case of the 8-class approach. Here, an improvement of about 14% has been achieved with the 8-class over the 4-class scenario, which proves the correctness of splitting the classes into

8 (instead of 4). Moreover, all the proposed DNNs exhibited very close performance when applied to the testing dataset. This stable behavior is a key feature of the prescribed DNNs.

B. Precision, Recall and F1-score

The obtained average precision, recall and F1-score metrics were depicted in Figs. 7-11. For better visualization, these metrics along with corresponding achieved improvement are illustrated in Fig. 14.a-c, respectively. Clearly, the proposed DNNs exhibited good-to-high precision, recall and F1-score values. Here, these metrics have been

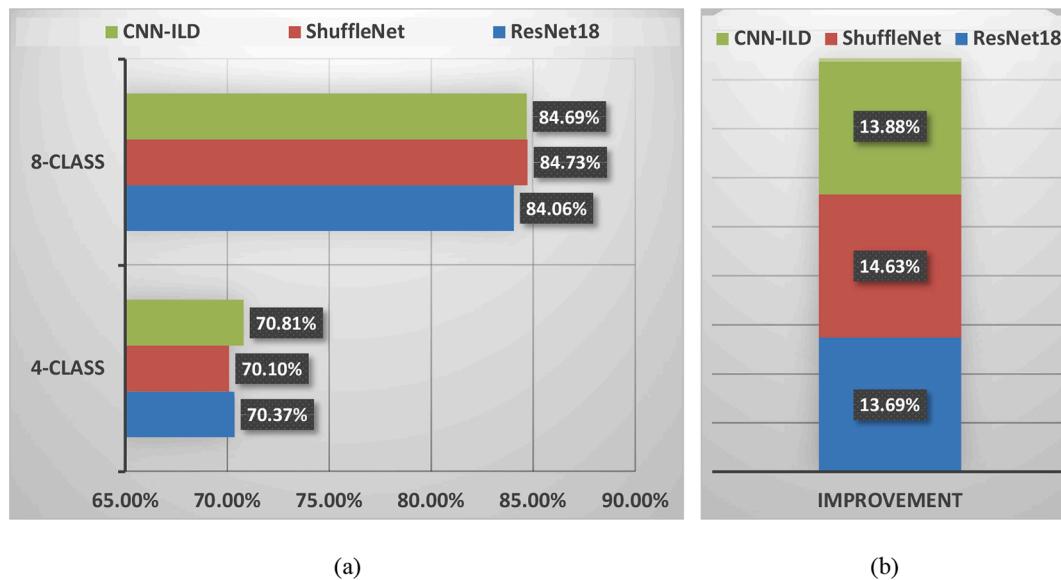


Fig. 13. Comparison of the achieved testing accuracy using the prescribed DNNs: a) average and b) obtained improvement of using the 8-class over the 4-scenario.

improved, when adopting the 8-class over the 4-class scenario (regardless of individual improvement for the ShuffleNet at precision). Moreover, the proposed CNN-ILD exhibited steady performance which, again, proves the suitability of adopting CNN-ILD, in an 8-class fashion, for faults classification of EL images.

Another key issue can be noted when observing the confusion-charts in Fig. 8 and Fig. 10. In Fig. 8, it is very clear that many classification errors can be observed in the neighboring classes (especially MP0.33 and MP0.67). This due to the fact that these classes have very close tiny features. For example, in case the MP0.33 class 61 samples were incorrectly classified as MP0.67, and 51 samples were classified as MP0.00. When turning to the 8-class scenario, it is very clear that a small number of samples were misclassified into other classes. Furthermore, referring to Fig. 11, the issue of overfitting is not observed in all the three proposed DNNs. This was expected due to performing DA before training. Again, this is considered as another key feature of the proposed DNNs.

4.4.2. The training computational cost

To evaluate the computation complexity of a DNN, many measures can be used, including network latency, error rate, GPU and ARM speeds [41]. In this regard, a very common metric is the number of float-point operations (FLOPs) that is defined as the number of multiply-adds. Nevertheless, FLOPs is considered as an indirect metric and doesn't usually reflect the actual DNN speed since various DNNs with similar FLOPs may have different speeds.

Another metric for DNN computational complexity is the total network latency that represents the time it takes for a DNN to process input data and produce output. In fact, latency is one of the more crucial aspects of deploying a DNN into a production environment.

Another practical measure for computational complexity that is usually implemented is the normalized elapsed time. This measure is useful to make proper time comparison between various networks. In this work, to evaluate the computational complexity of the developed DNNs, two metrics were adopted, namely:

1. The total latency L_T (in cycles) as well as the number of frames that the DNN can handle (F_{DNN}) (per second). The resulting L_T and F_{DNN} for the prescribed DNNs (for both the 4- and 8-class scenarios) are illustrated in Table 4. Clearly, the proposed CNN-ILD exhibited the best performance, in terms of L_T and F_{DNN} for both the 4-class and 8-

class scenarios. Here, the lower L_T and the higher the F_{DNN} , the better is the DNN time efficiency.

2. The normalized consumed elapsed time for DNN training. This normalized training time is depicted for the prescribed DNNs in Fig. 15. Here, the three DNNs were trained using MATLAB environment on a computer with a single GPU (see specifications on sec.4.1) for both the 4-class and 8-class scenarios. Evidently, the proposed CNN-ILD is a light-weight network that is expected to consume less computational power and time than ResNet18 and ShuffleNet. As clear from Fig. 15, a reduction in time consumption was achieved with all of the three models, in case of utilizing the 8-class scenario, when compared to the 4-class scenario. In case of the 4-class scenario, the proposed NCC-IDL outperformed the ResNet18 and ShuffleNet by about 9.2% and 20.8%, respectively. In case of the 8-class scenario, similar results were achieved with about 10.4% 17.9% when compared to ResNet18 and ShuffleNet, respectively. This achievement is counted as a positive point in favor of the proposed CNN-ILD approach.

4.4.3. Comparative analysis

In this section, the obtained experimental results of utilizing the three tested DNNs, for classification of EL images, are compared with other existing state-of-art works. To objectively judge various DNN approaches, comparisons are carried out with research works that utilized the ELPV public dataset, including [5,6,8,12,13,29,30].

A. "Detection only" Comparative Analysis

To start with, the proposed CNN-ILD is compared with other existing state-of-art works in case of binary classification. Here, the employed DNN is only able to detect whether the cell is operative (no defect) or defective without any further classification of the existing defect type. Comparative results (with works in [5,8,12,29,30]) are given in Table 5, where it is obvious that CNN-ILD exhibited superior performance in terms of the accuracy, precision, recall and F1-score. This proves the validity of the proposed CNN-ILD in fault detection tasks of EL images.

B. Classification Comparative Analysis

The obtained experimental results of utilizing the three tested DNNs for classification of EL images are compared with other existing state-of-art works, including [6,8,13]. Comparisons are provided in Table 6.

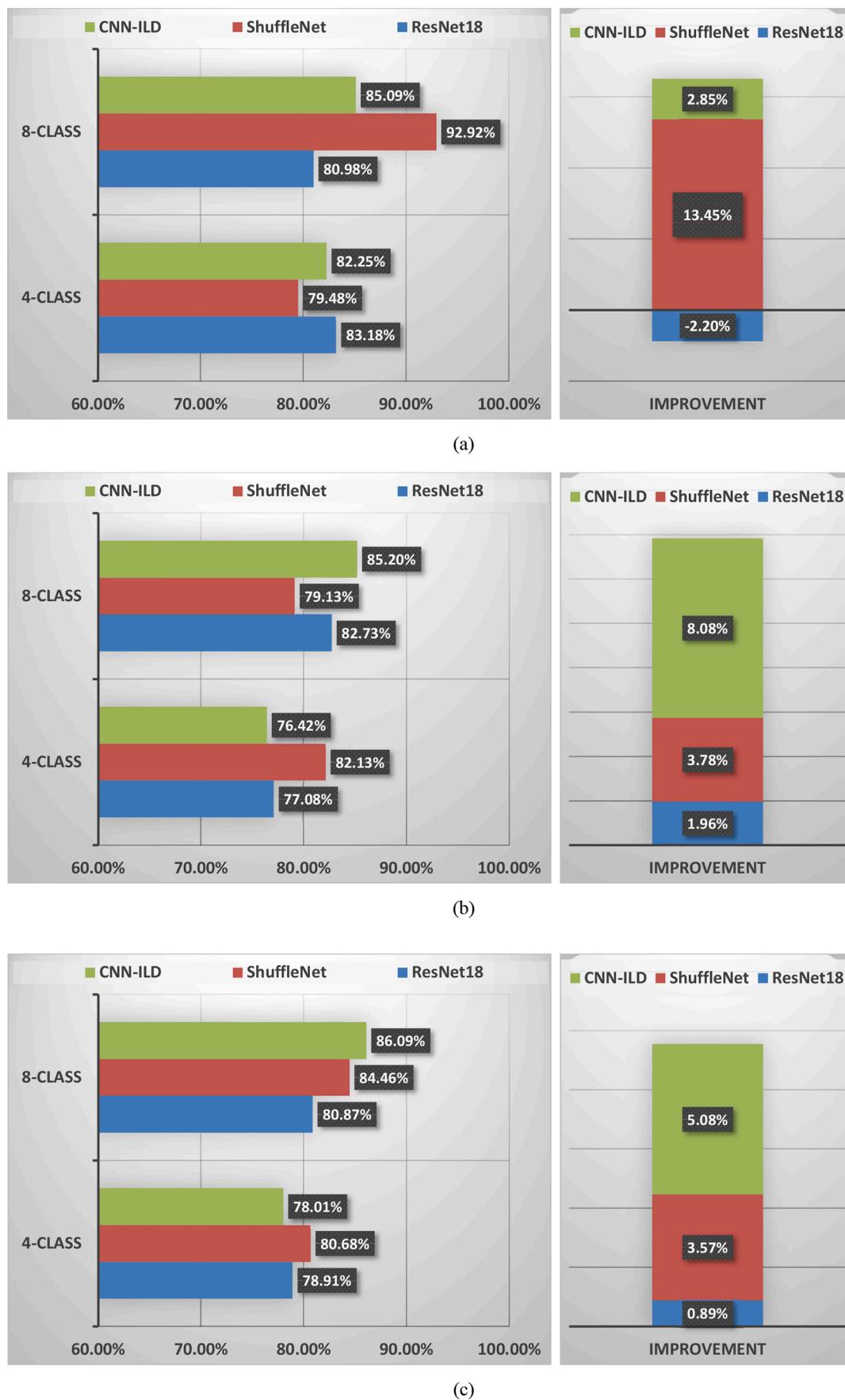


Fig. 14. Comparison of the achieved performance metrics using the prescribed DNNs in terms of the average: a) precision, b) recall, and c) F1-score.

Table 4

The obtained total latency of the prescribed DNNs.

Trained DNN	4-class scenario		8-class scenario	
	L_T (cycles)	F_{DNN} (frames/sec)	L_T (cycles)	F_{DNN} (frames/sec)
ResNet18	22 834 943	9.26	21 086 928	9.97
ShuffleNet	26 186 656	7.91	22 662 718	8.97
CNN-ILD	21 099 487	10.07	18 705 603	11.04

Here, 4-class and 8-class scenarios are included. Clearly:

- Most works were based on 2- to 4-class scenarios. No prior works were elaborated using 8-class DNNs. As prescribed, the only work that considered the 8-class scenario is [36] and was based on using another dataset (called PVEL-AD [37]) rather than the well-known ELPV dataset [34].

- When comparing selected DNNs, in case of 4-class scenario, the three DNNs have comparable results close to results in [13] and overcomes results obtained in [6,8].
- Ultimately, the results of the 8-class scenario outstand all other results as shown in Table 6.

4.4.4. Illustrative examples

To demonstrate the applicability of the proposed CNN-ILD, especially when adopting the 8-class scenario, illustrative classification examples are shown in Table 7. The predicted labels of the misclassified cases are shaded with red background color. In fact, in these examples the classification errors are higher when utilizing the 4-class scenario. Moreover, the error achieved with the 8-class scenario (see #4) has misclassified the sample from being classified as P 0.33 (true label) to be classified as P 0.67 (predicted label). This overestimates the error level in the investigated sample. This situation appeared in many similar cases (8-class scenario). On the contrary, as clear form Table 7, the situation is

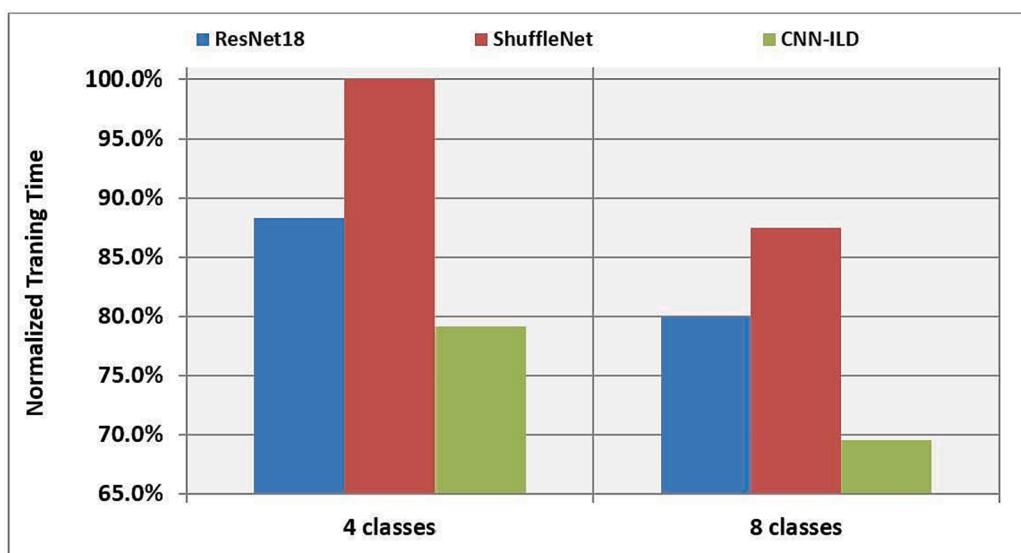


Fig. 15. Normalized elapsed training time for the investigated CNN-models.

Table 5

Comparison of DNNs detection performance.

#	Network Name	Number of Classes	Classification Method	Accuracy	Precision	Recall	F1-score
1	Tang et al. [5]	2	InceptionV3 and V4	0.9300	–	–	–
2	Demirci et al., [8]	2	Light CNN	0.8933	0.9044	0.9542	0.9286
3	Akram et. al, [12]	2	Light CNN	0.928	0.9250	0.9200	0.9230
4	Otamendi et al., [29]	2	CNN	0.8400	–	–	–
5	Korovin [30] (with combined B&C classes)	2	Light CNN	0.8520	–	–	0.8880
6	CNN-ILD	2	Light CNN	0.9984	0.9982	0.9982	0.9982

Table 6

Comparison of DNNs classification performance.

#	Network Name	Number of Classes	Classification Method	Accuracy	Precision	Recall	F1-score
1	Deitsch et al., [13]	4	SVM	0.824	–	–	0.825
2	Deitsch et al., [13]	4	CNN	0.884	–	–	0.889
3	Demirci et al., [8]	4	Light CNN	0.826	0.807	0.916	0.858
4	Tang et al., [6]	4	CNN	0.830	–	–	–
5	CNN-ILD	4	Light CNN	0.886	0.812	0.764	0.781
6	ShuffleNet	4	CNN	0.909	0.793	0.821	0.796
7	ResNet18	4	CNN	0.884	0.832	0.771	0.789
8	CNN-ILD	8	Light CNN	0.958	0.85	0.853	0.849
9	ShuffleNet	8	CNN	0.981	0.927	0.793	0.850
13	ResNet18	8	CNN	0.944	0.820	0.827	0.812

Table 7

Illustrative examples on EL image Classifications.

#	EL image	Cell Type	True Label		Predicted Label	
			8-classes	4-classes	8-classes	4-classes
1		<i>mc-Si</i>	M 0.0	MP 0.0	M 0.0	MP 0.0
2		<i>pc-Si</i>	P 0.0	MP 0.0	P 0.0	MP 0.33
3		<i>mc-Si</i>	M 0.33	MP 0.33	M 0.33	MP 0.0
4		<i>pc-Si</i>	P 0.33	MP 0.33	P 0.67	MP 0.67
5		<i>mc-Si</i>	M 0.67	MP 0.67	M 0.67	MP 0.0
6		<i>pc-Si</i>	P 0.67	MP 0.67	P 0.67	MP 1.00
7		<i>mc-Si</i>	M 1.0	MP 1.0	M 1.0	MP 1.00
8		<i>pc-Si</i>	P 1.0	MP 1.0	P 1.0	MP 1.00

unlikely different when coming to the 4-classes scenario, where errors have random behavior (see # 2, # 4 and # 5). In fact, a detailed analysis of many of the tested samples has shown similar situations, i.e., overestimating the error level (in case of the 8-class scenario) and exhibiting random behavior of misclassification (in case of the 4-class scenario). This proves the correctness adopting the 8-class approach over the 4-class scenario.

5. Conclusions

Recently, CNN-based approaches caught an increasing attraction in

the field of PV cells detection and fault classification. On the other hand, EL imaging analysis has been recognized as a powerful tool for addressing malfunctions in PV models. In this work, an effective fault detection and classification approach is developed using multi-scale CNN-based models using two scenarios: a) a transfer learning-based approach using the pretrained ResNet18 and ShuffleNet, and b) an independent light-depth CNN (denoted as CNN-ILD). The employed DNNs have various number of layers, namely the developed CNN-ILD (with 25 layers), while the pretrained ResNet18 and ShuffleNet have 71 and 172 layers, respectively. The main feature of the developed CNN-ILD is based on the insertion of the parallel convolutional branches with low-to-high

kernel sizes. This gave an ability to capture a wide range of features that range from elementary (shape and boundary) to specific textures patterns. On the other hand, the adopted improved ELPV dataset, with almost-balanced number of samples per class, along with the proper class categorization have led to high classification performance. The experimental results have shown that:

- For fault detection tasks, with no further classification, the proposed CNN-ILD exhibited superior performance in terms of the accuracy, precision, recall and F1-score (up to 99.8%).
- The proposed DNNs exhibited high fault classification performance in EL images. Here, an average accuracy (88.41–90.89%) and (94.31–98.05%) was achieved, in case of the 4-class and 8-class scenarios, respectively. Moreover, the proposed DNNs exhibited good-to-high precision, recall and F1-score records. Here, these metrics have been improved, when adopting the 8-class scenario over the 4-class scenario.
- the proposed CNN-ILD exhibited the best performance, in terms of the latency and the number of handled number frames for both the 4-class and 8-class scenarios. Here, the number of handled frames (in CNN-ILD) was around 11.04 frames/s. Moreover, when considering the normalized elapsed time, a clear privilege was granted to the CNN-ILD with a reduction in time consumption of about 12% and 19% in case of the 4-class and 8-class scenarios (compared to ShuffleNet).

Regardless of these achievements, it is worthy to note that:

- The computational cost of CNN-ILD can be reduced by adopting proper layers in its construction.
- The tested DNNs were verified using the ELPV dataset, however, other datasets (like PVEL-AD [37]) can be adopted along with using other time-efficient DNN models (like SqueezeNet and MobileNet-v2).
- it would be useful to extend the results of this work to other issues like thermographic fault diagnosis in commutators and induction motors.

Despite of the good success, achieved using the proposed DNNs, however, there is still potential opportunities for further improvement in designing more suitable time-efficient models for practical industrial applications. These issues can be tackled in a forthcoming paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- [1] E.H.S. Oviedo, L. Travé-Massuyès, A. Subias, C. Alonso, M. Pavlov, Feature extraction and health status prediction in PV systems, *Adv. Eng. Inf.* 53 (2022), 101696.
- [2] IEA, 2019. Renewables 2019. Paris <https://www.iea.org/reports/renewables-2019>.
- [3] International Energy Agency, 2018. Renewables 2018, Market analysis and forecast from 2018 to 2023. Available at: <https://www.iea.org/renewables2018/>.
- [4] A.B. Saka, L.O. Oyedele, L.A. Akanbi, S.A. Ganiyu, D.W.M. Chan, S.A. Bello, Conversational artificial intelligence in the AEC industry: A review of present status, challenges and opportunities, *Adv. Eng. Inf.* 55 (2023) 101869.
- [5] W. Tang, Q. Yang, X. Hu, W. Yan, Convolution neural network based polycrystalline silicon photovoltaic cell linear defect diagnosis using electroluminescence images, *Expert Syst. Appl.* 202 (2022) 117087.
- [6] W. Tang, Q. Yang, K. Xiong, W. Yan, Deep learning based automatic defect identification of photovoltaic module using electroluminescence images, *Sol. Energy* 201 (November 2019) (2020) 453–460, <https://doi.org/10.1016/j.solener.2020.03.049>.
- [7] R. Pimpalkar, A. Sahu, R.B. Patil, A. Roy, A comprehensive review on failure modes and effect analysis of solar photovoltaic system, *Mater. Today.: Proc.* 77 (2023) 687–691.
- [8] M.Y. Demirci, N. Beşli, A. Gümuşçü, Efficient deep feature extraction and classification for identifying defective photovoltaic module cells in Electroluminescence images, *Expert Syst. Appl.* 175 (2021) 114810.
- [9] Denio H. Aerial solar thermography and condition monitoring of photovoltaic systems, IEEE Proceeding of the 38th photovoltaic Spec. Conference (2012), pp. 000613–000618.
- [10] H. Hui, C. Gao, Z. Yue, et al., Polycrystalline silicon wafer defect segmentation based on deep convolutional neural networks, *Pattern Recogn. Lett.* 130 (2018).
- [11] C. Schuss, K. Remes, K. Leppanen, J. Saarela, T. Fabritius, B. Eichberger, T. Rahkonen, Detecting defects in photovoltaic panels with the help of synchronized thermography, *IEEE Trans. Instrum. Meas.* 67 (5) (2018) 1178–1186.
- [12] M.W. Akram, G. Li, Y.i. Jin, X. Chen, C. Zhu, X. Zhao, A. Khaliq, M. Faheem, A. Ahmad, CNN based automatic detection of photovoltaic cell defects in electroluminescence images, *Energy* 189 (2019) 116319.
- [13] S. Deitsch, V. Christlein, S. Berger, C. Buerhop-Lutz, A. Maier, F. Gallwitz, C. Riess, Automatic classification of defective photovoltaic module cells in electroluminescence images, *Sol. Energy* 185 (February) (2019) 455–468, <https://doi.org/10.1016/j.solener.2019.02.067>.
- [14] Fuyuki, T., Kondo, H., Kaji, Y., Yamazaki, T., Takahashi, Y., & Uraoka, Y. (2005). One shot mapping of minority carrier diffusion length in polycrystalline silicon solar cells using electroluminescence, 1343–1345. 10.1109/pvsc.2005.1488390.
- [15] Kasemann, M., Kwapił, W., Walter, B., Giesecke, J., Michl, B., The, M., ... & Glunz, S. W. (2008, September). Progress in silicon solar cell characterization with infrared imaging methods. In *Proceedings of the 23rd European Photovoltaic Solar Energy Conference, Valencia, Spain* (pp. 965–973).
- [16] I. Berardone, G.J. Lopez, M. Paggi, Analysis of electroluminescence and infrared thermal images of monocrystalline silicon photovoltaic modules after 20 years of outdoor use in a solar vehicle, *Sol. Energy* 173 (2018) 478–486.
- [17] D.-M. Tsai, S.-C. Wu, W.-Y. Chiu, Defect detection in solar modules using ICA basis images, *IEEE Trans. Ind. Inf.* 9 (1) (2012) 122–131.
- [18] S.A. Anwar, M.Z. Abdullah, Micro-crack detection of multi crystalline solar cells featuring an improved anisotropic diffusion filter and image segmentation technique, *EURASIP J. Image Video Process.* 2014 (1) (2014) 1–17.
- [19] A.M. Karimi, J.S. Fada, M.A. Hossain, S. Yang, T.J. Peshek, J.L. Braid, R.H. French, Automated pipeline for photovoltaic module electroluminescence image processing and degradation feature classification, *IEEE J. Photovolt.* 9 (5) (2019) 1324–1335.
- [20] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* 25.
- [21] Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [22] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2016–December. IEEE Computer Society, pp. 770–778, [Online]. Available: <http://image-net.org/challenges/LSVRC/2015/>.
- [23] Szegedy, C., Liu, W., Jia, Y., et al., 2015. Going deeper with convolutions. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9.
- [24] Howard, A.G., Zhu, M., Chen, B., et al., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications, *arXiv preprint arXiv:1704.04861*.
- [25] R. Pierdicca, E.S. Malinverni, F. Piccinini, M. Paolanti, A. Felicetti, P. Zingaretti, Deep Convolutional Neural Network for Automatic Detection of Damaged Photovoltaic Cells, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-2 (2018) 893–900.
- [26] A. Bartler, L. Mauch, B. Yang, M. Reuter, L. Stoicescu, Automated Detection of Solar Cell Defects with Deep Learning, 2018 26th European Signal Processing Conference, 2018: 2049–2053.
- [27] J. Balzategui, L. Eciolaza, N. Arana-Arexolaleiba, J. Altube, J.-P. Aguerre, I. Legarda-Ereno, A. Apraiz, Semi-automatic quality inspection of solar cell based on Convolutional Neural Networks, 24th IEEE Int. Conference on Emerging Technologies and Factory Automation, 10–13 Sept. 2019, Zaragoza, Spain.
- [28] Karimi, A. M., Fada, J. S., Liu, J., Braid, J. L., Koyuturk, M., & French, R. H. (2018). Feature Extraction, Supervised and Unsupervised Machine Learning Classification of PV Cell Electroluminescence Images. In 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC) (pp. 0418–0424). IEEE. 10.1109/PVSC.2018.8547739.
- [29] U. Otamendi, I. Martínez, M. Quartulli, I.G. Olaizola, E. Viles, W. Cambarau, Segmentation of cell-level anomalies in electroluminescence images of photovoltaic modules, *Sol. Energy* 220 (2021) 914–926.
- [30] A. Korovin, A. Vasilev, F. Egorov, D. Saykin, E. Terukov, I. Shakhray, L. Zhukov, S. Budennyy, Anomaly detection in electroluminescence images of heterojunction solar cells, *Sol. Energy* 259 (2023) 130–136.
- [31] W. Tang, Q. Yang, W. Yan, Deep learning-based model for Defect Detection of Mono-Crystalline-Si Solar PV Module Cells in Electroluminescence Images Using Data Augmentation, 2019 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), 1–4 Dec. 2019, Macao.

- [32] K.-M. Lin, H.-H. Lin, Y.-T. Lin, Development of a CNN-based hierarchical inspection system for detecting defects on electroluminescence images of single-crystal silicon photovoltaic modules. *Materials Today, Communications* 31 (2022) 103796.
- [33] X. Chen, T. Karin, A. Jain, Automated defect identification in electroluminescence images of solar modules, *Sol. Energy* 242 (2022) 20–29.
- [34] Buerhop-Lutz, C., Deitsch, S., Maier, A., Gallwitz, F., Berger, S., Doll, B., ... & Brabec, C. J. (2018, September). A benchmark for visual identification of defective solar cells in electroluminescence imagery. In *35th European PV Solar Energy Conference and Exhibition* (Vol. 12871289).
- [36] H. Acikgoz, D. Korkmaz, U. Budak, Photovoltaic Cell Defect Classification based on Integration of Residual-Inception Network and Spatial Pyramid Pooling in Electroluminescence Images, *Expert Syst. Appl.* 229 (2023) 120546.
- [37] B. Su, Z. Zhou, H. Chen, PVEL-AD: A Large-Scale Open-World Dataset for Photovoltaic Cell Anomaly Detection, *IEEE Trans. Ind. Inf.* 19 (1) (Jan. 2023) 404–413, <https://doi.org/10.1109/TII.2022.3162846>.
- [38] K.G. Bedrich, Quantitative electroluminescence measurements of PV devices, Loughborough University, 2017. Doctoral dissertation.
- [39] M. Dhimish, V. Holmes, B. Mehrdadi, et al., The impact of cracks on photovoltaic power performance, *J. Sci.: Adv. Mater. Devices* 2 (2017) 199–209.
- [41] R. Desislavov, F. Martínez-Plumed, J. Hernández-Orallo, Trends in AI inference energy consumption: Beyond the performance-vs-parameter laws of deep learning, *Sustainable Comput. Inf. Syst.* 38 (2023), 100857.