Current trends on the use of deep learning methods for image analysis in energy applications

Energy and AI 15 (2024) 100330

Contents lists available at ScienceDirect

Energy and AI

journal homepage: www.elsevier.com/locate/egyai

Review

Current trends on the use of deep learning methods for image analysis in energy applications

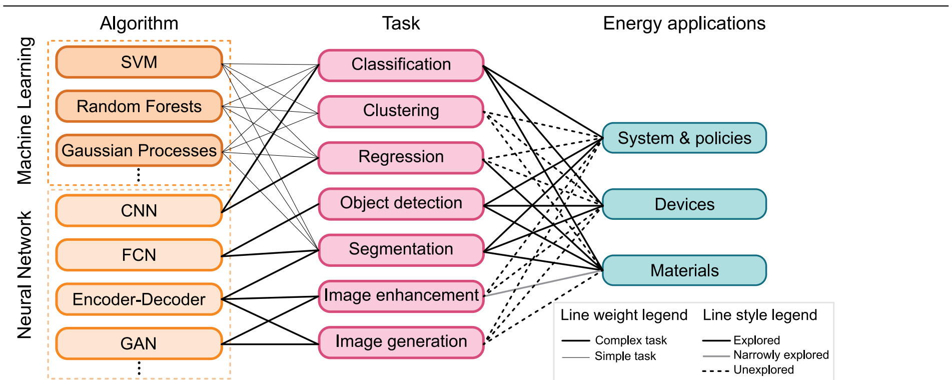
Mattia Casini [1], Paolo De Angelis [1], Eliodoro Chiavazzo, Luca Bergamasco *

Department of Energy, Politecnico di Torino, Corso Duca degli Abruzzi 24, Turin, 10129, Italy

HIGHLIGHTS

- Current application trends of deep learning for image analysis are analyzed.
- Applications related to the energy sector are systematically classified.
- The emerging research trends for energy-related problems are discussed.
- A correlation chart for methods, tasks and energy applications is proposed.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Deep learning
Image analysis
Convolutional neural networks
Energy applications

ABSTRACT

Deep learning methods for image analysis are attracting increasing interest for application in a wide range of different research fields. Here we aim to systematically analyze and discuss the most relevant examples for the energy sector. To this, we perform a comprehensive literature screening on applications of deep learning methods for image analysis, classify the results in application macro-areas, and discuss the emerging trends on the available energy-related cases. The results of the analysis show that, while the exploitation of these methods for energy applications still appears to be at an early stage, the interest during the last years, in terms of number of published works, has considerably grown. To provide a systematic overview on the available energy-related examples, we present a schematic correlation chart mapping algorithms, tasks, and applications. The reported analysis is intended to provide an up-to-date overview on the current application trends and potential developments for energy applications in the next future.

1. Introduction

During the last years, Artificial Intelligence has undergone substantial improvements in several different fields [1]. As of today, Machine Learning (ML) algorithms empower the systems underlying many aspects of our everyday life, such as tailored content recommendation on social networks, traffic prediction in car navigation maps, speech recognition, language translation tools or mail filtering [2]. No less, ML-based techniques have attracted increasing interest in more scientifically-oriented research applications, such as materials science or clinical data analysis [3–6]. One of the most eminent examples of the potential of a ML approach for scientific applications is the recently-released AlphaFold engine for protein structure prediction [7]. Deep

* Corresponding author.
E-mail address: luca.bergamasco@polito.it (L. Bergamasco).
[1] These two authors contributed equally to this work.

https://doi.org/10.1016/j.egyai.2023.100330
Received 28 September 2023; Received in revised form 17 November 2023; Accepted 9 December 2023
Available online 12 December 2023

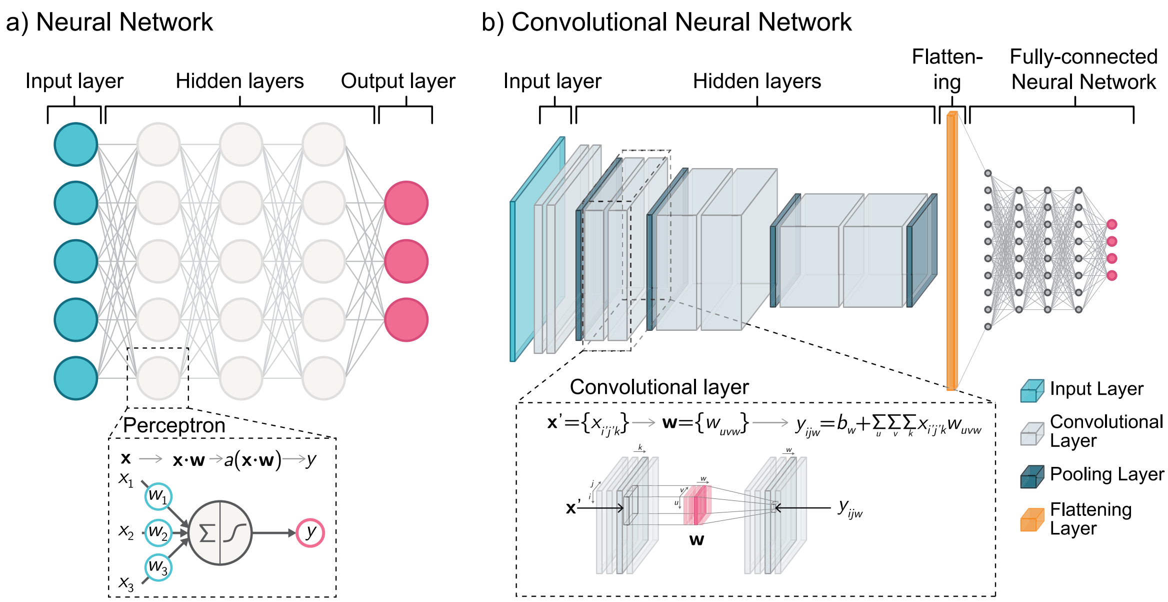


Fig. 1. Schematic representation of prominent DNN architectures and their building blocks: the first illustration (a) shows a feed-forward NN for data processing, and its core component, known as *perceptron*, is highlighted in the close-up box. The second illustration (b) displays a simplified CNN, designed for image processing. The data flows through a sequence of convolutional (close-up box) and pooling layers, extracting important features that are then fed to the fully-connected NN through the flattening layer.

Learning (DL), as a subset of ML models, is considered a particularly promising method, owing to the possibility it provides of analyzing large and intricate data structures [8]. Indeed, DL relies on artificial neural networks with multiple layers, which are intended to mimic the behavior of the human brain. Although still far from this latter goal, DL algorithms can effectively *learn* from data [9].

The advent of Deep Neural Networks (DNN), has opened new and remarkable possibilities to extract information from images, making it faster and easier [10]. Pattern and data recognition from digital images (or videos), known as *computer vision* [11], is typically best accomplished in the deep learning framework via Convolutional Neural Networks (CNNs), also called ConvNet [12]. These latter are specialized neural networks, particularly suitable for pixel analysis as they rely on convolution operations [13]. These networks can be used to perform different tasks on images, such as classification [14], object detection [15] and image generation [16]. These models have applications in important technological areas, such as e.g. facial recognition [17], infrared COVID-19 detection [18] or self-driving cars [19].

Motivated by the increasing range of applications of deep learning methods for image processing, in this work we target an analysis of the recent trends on their utilization for scientifically-oriented research areas, with particular focus on energy-related applications. In this latter field, data-driven analysis based on images can indeed stand as a valuable approach for e.g. screening of geographical areas for renewable potential assessments [20] or monitoring plant installations [21] from satellite images, optimization of energy devices using thermographical images [22], or material design and optimization for energy conversion and storage using microscopic images [23]. These examples all support the opportunities that image analysis provide for energy-related problems at large, especially for those cases where big data processing is required, or when the complexity of the problem hinders classical (physics-based) modeling approaches. Thus, here we aim to screen the available examples on the use of deep learning methods and image analysis for energy problems, to identify whether specific emerging trends in the applications arise. To this, we perform a literature screening on deep learning methods for image analysis on multiple databases, classifying the resulting energy-related applications by research macro-areas and discussing the recent trends and possible developments for these techniques in the near future.

The outline of this document is as follows. We first review the salient aspects of convolutional neural networks and the available main architectures in Section 2. In Section 3, the results of the literature screening are reported, along with a discussion of the emerging trends in energy-related applications. In Section 4, we provide a compact overview on algorithms, tasks and applications. In Section 5, we discuss the current limits and perspective challenges for wider adoption and exploitation of these methods in the energy field. Finally, the conclusions on the proposed analysis are drawn, with an outlook on possible near-future developments for the energy sector.

2. Deep neural networks in computer vision

Computer vision is the branch of computer science dealing with the development of algorithms and techniques to enable computers to extract information from images and videos, based on a partial replication of the functionality of the human visual apparatus [11]. Image analysis algorithms have undergone considerable development during the last decades, thanks also to completely new possibilities provided by Machine Learning techniques [24]. Particularly, the introduction of Deep Learning, has definitely opened to a real boost.

Deep Learning algorithms rely on deep neural networks. These networks, whose architecture design takes inspiration from the human brain, consist of a large set of interconnected neurons, which are organized in multiple layers (hence the name *deep*). Each neuron is in charge of elaborating an output, based on the received input. The layers can be generally summarized to be: an input layer, for input data acquisition and passage; a variable number of *hidden* layers, to perform operations and process information at different levels of detail; an output layer, for the output of the results [25]. A schematic overview of this layout is shown in Fig. 1(a). In order to *learn* from data based on one such layout, the network must be *trained*. The training phase allows to establish the correct *weights*, which are associated with the different connections among neurons and allow the model to adapt to the notions to be learned [26]. The weights are indeed updated during training, to reinforce the useful correlations between neurons and discourage the irrelevant ones. A *bias* is also generally adopted, to offset the output of each neuron, to be fed to an *activation function*. This latter function is used to determine whether a neuron is activated or not, based on the received input and on a threshold value. Several activation functions are available. The most commonly used are the sigmoid and ReLU functions [27]. In general, the greater the number of hidden layers, the deeper the network, and the more complex the information that can be processed. However, an excessive number of layers leads to problems in their management during the learning phase [28]. For this reason, several techniques have been developed to maximize the effectiveness of the network, while maintaining acceptable depths [26].

Thanks to the previous architecture, deep neural networks are inherently well-suited for handling large amounts of data. In general, the network can be fed with any numerical data; however, this data must comply with the layout of the input layer (*i.e.* it must be organized in the form of a row vector of numerical values). In the case of images, one may then think about flattening the three-dimensional data of an image into a row vector to feed the input layer with the numerical pixel values. In the raster format indeed, an image consists of one or multiple matrices of pixel values. For a generic image, the overall dimension is given by $h \times w \times c$, where h is height, w is the width and c is the number of channels (with $c = 3$ for Red, Green and Blue (RGB) color space or $c = 1$ for grayscale or black and white color models). This approach may probably provide a reasonable precision for e.g. the prediction of classes in extremely simple images; however, the ability to analyze more complex images would yield little to no precision. This would be due to the image flattening operation, which causes the lost of most of the information characterizing the image. In images indeed, pixels have (local) dependencies, which are fundamental for information extraction and must be preserved while feeding data to the neural network. Besides this possible loss of information, it is also important to consider that such an operation would imply a considerable usage of (RAM) memory. As an example, a binary classification of a low-resolution $100 \times 100 \times 1$ grayscale image using a neural network with just one hidden layer, would require storage of a total $\approx 1 \times 10^8$ number of variables (*i.e.* $(100 \times 100)^2$ for neural connections between the input and hidden layer, and 100×100 for neural connections between the hidden and the output layer), and $\approx 1 \times 10^4$ bias (*i.e.* $100 \times 100 + 1$ for the neurons in the hidden and the output layers). Considering double precision floating point (64 bits) data, an estimated total of 6.4 Gbit would be required.

Thus, a dedicated class of neural networks have been developed for the purpose: Convolutional Neural Networks (CNNs). These networks are built upon the same deep logic previously discussed (see Fig. 1(b) for a schematic representation), except that here convolution operations are employed on images to preserve their characteristic two-dimensional local features.

2.1. Convolutional neural networks

Digital image representation may rely either on *raster* or *vector* format, depending on the data structure underlying the visual output. In the former case, the image results from a large, discrete set of *pixels*, to which color-representative numerical values are associated; in the latter, the image results from a series of mathematical rules which are used to render curves and other features on a grid of points. Despite the advantage of vector graphics over raster in size scalability with no loss of quality, processing of this format involves particular considerations [29,30] which deviate from the purpose of this work. Thus, here we shall focus on methods for processing of the raster format only. As previously introduced, in the raster format, an image consists of one or multiple matrix pixel values. Image processing typically involves usage of high-definition pictures (e.g. Full HD format is 1920×1080 pixel), which implies a considerable amount of data to be handled. Hence, the primary objective is to reduce this data, while preserving relevant information on local features. In convolutional neural networks, this is accomplished via convolution operations across multiple layers.

With reference to the general CNN layout shown in Fig. 1(b), the input image is fed into a first convolutional layer, where the convolution is performed. This operation implies that a *kernel* (also called *filter*) is applied throughout the original image. Particularly, the kernel (*i.e.* a matrix with predefined size and learnable weights) is made shift with a given *stride* along the spanwise directions of the original image with a zero *padding* (see Fig. 2(a)). The stride and padding are *hyperparameters* of the convolutional layer, and are usually set to one to preserve the input dimension, as shown in Fig. 2(a.i). At each step, an element-wise multiplication between the filter and the original image is performed, and the results are summed (along with the bias) to obtain the convolved output. Each convolutional layer has multiple filters and outputs one *feature map* per filter (see close-up box in Fig. 1(b)), which highlights the specific patterns of the image that activate the filter the most. The weights of the filters are not to be prescribed *a priori*, as they are adjusted during training to enhance feature extraction [26]. In the case of RGB images, the same logic applies on the three color channels, and the result is a single-depth convolved feature matrix. One limit of the so obtained convolved maps is their adherence to the spatial position of the features in the input image, which may lead to incorrect results in case of even small rotation or shifting of the original image. In this view, a more flexible output may be obtained by adjusting the stride of the convolution; however, the use of a dedicated *pooling* layer is more commonly adopted to the purpose (see Fig. 2(b)). The logic of the operation is similar to convolution, except that the size of the pooling filter is smaller than that used for the convolution (typically, 2×2 pixels with 2 pixel stride [26]), and results in a reduction of the size of the feature map. Common criteria for this operation, which is user-specified, are the max and average pooling (where the max or average value of the mapped filter are respectively considered, as shown in Fig. 2(b.i) and Fig. 2(b.ii)). Pooling operates on each feature map separately, to reduce the computational load and memory usage. It also helps to render the model invariant to image distortions, reducing the redundancy of the information. The two layers (convolutional and pooling) form one layer of the convolutional neural network. In deep networks, the usage of multiple layers allows to obtain a multi-level feature extraction, being those closer to the input committed to the extraction of low-level features (such as, e.g., lines), and the "deeper" ones to more high-level features (such as, e.g., shapes).

In this view, one may then think that the development of deeper and deeper networks via integration of additional hidden layers would suffice for tackling progressively more complex problems. This strategy, however, was previously demonstrated to be prone to error propagation [31]. During each training step indeed, the model generates a set of responses on the training data, which are then compared with the correct solutions using a *loss function* [26] (also called cost function). This function, based on different possible user-selected algorithms, quantifies the errors of the model. To reduce these errors during training, the *back-propagation* method can be employed, which allows for a backward analysis of the network, from the output layer to the input layer. Particularly, in this method, all the weights are scrutinized to identify those mostly contributing to the propagation of errors, based on the gradient of the loss function. The targeted weights are then modified using an *optimizer*, in a tentative to reduce the loss function based on the magnitude of the calculated errors. For excessively deep networks, this can lead to two distinct yet similarly-rooted blocking problems: the *vanishing gradient* and the *exploding gradient* problems [26]. As the name suggests, the vanishing gradient problem arises when the (normalized) gradients of the loss function with respect to the model parameters become excessively small as they are back-propagated through the network. Consequently, the initial layers of the network receive minimal weight updates during training, resulting in a stagnation of the learning process. Conversely, the exploding gradient problem occurs when the gradients grow exponentially as they are back-propagated through the network. This leads to substantial weight updates during training, causing instability and frequently impeding the network training process from converging to a meaningful solution. Both issues can severely hinder the training process of deep neural networks; thus, various techniques have been

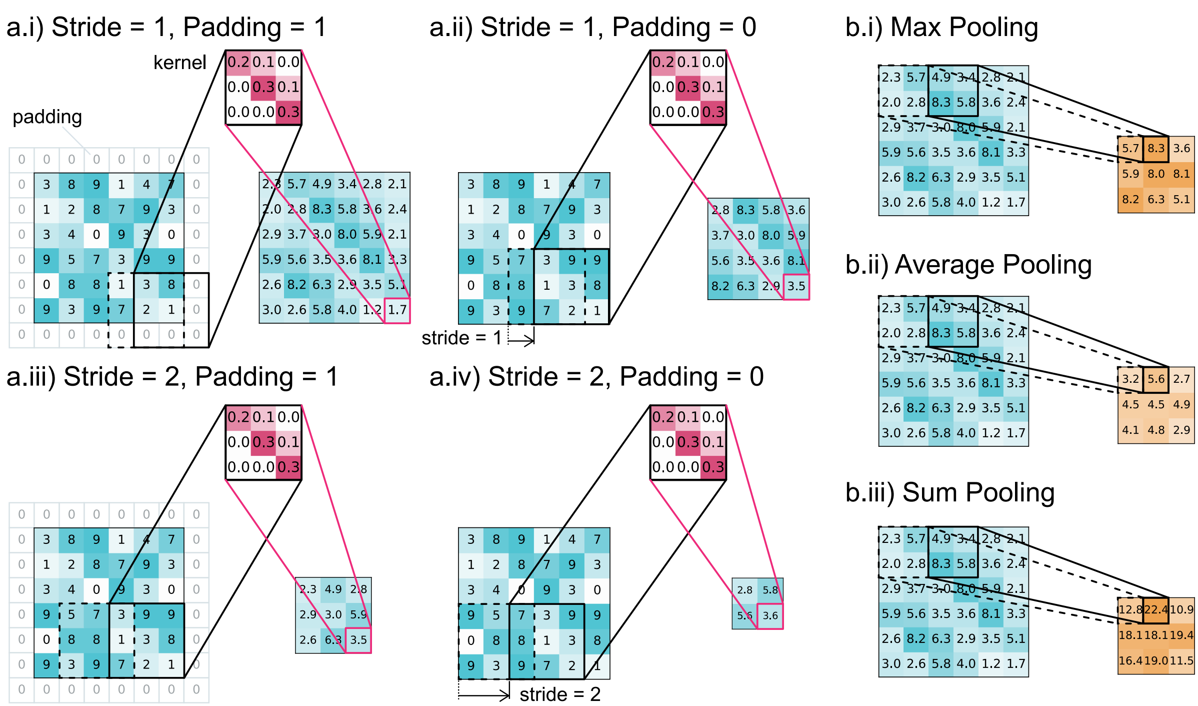


Fig. 2. Effect of the convolutional and pooling layer hyperparameters. In (a) a 6 × 6 input layer (represented by a cyan matrix on the left) is convolved using a 3 × 3 kernel (represented by a magenta matrix on top). By varying the stride and padding hyperparameters, four different feature matrices are produced (shown on the right). Note that the chosen hyperparameters affect the final output dimension, and only certain combinations (e.g unitary stride and padding in (a.i)) keep the input dimension unchanged. In the second example (b), a 6 × 6 feature matrix is subject to max (b.i), average (b.ii), and sum (b.iii) pooling layers to reduce the data and get an output layer (shown as the orange matrix on the right), which is half of the original size (3 × 3).

developed to mitigate these problems, including: weight initialization, changing the activation functions, gradient clipping or architectural modifications [26].

In the context of architectural modifications, particularly, an important advancement in tackling the problem was introduced by the ResNet architectures [32]. The fundamental concept behind ResNets involves incorporating *skip connections* (also called shortcuts), that enable the network to bypass one or more layers facilitating smoother information flow throughout the network. This results in a substantial modification of the convolutional block, which is here called *residual* block. This improvement enhances the learning phase, and ensures a proper back-propagation flow of the gradients of the loss function. The presence of skip connections thus guarantees that, even if the layers within the block experience minimal weight updates, the input information can still propagate to subsequent layers, effectively mitigating the vanishing gradient problem. At the same time, the skip connections restrict the extent of gradient explosion, thanks to bypass of certain layers during the back-propagation process. This finally prevents the gradients from accumulating to excessively high values, and mitigates the instability caused by exploding gradients. This innovative solution has enabled the successful training of significantly deep neural networks, surpassing the 16-layer limit of the VGG16 model to more than 150 layers of the ResNet152 model, resulting in a significantly improved accuracy and performance [31,32].

Finally, in the considered general layout (Fig. 1(b)), *Fully-Connected* Neural Networks (FCNNs) may be adopted as the final layers, to condense the received information into an output vector [33]. This latter output can be used for classification purposes, where each element in the vector indicates the probability that the input image pertains to a specific class, or for regression purposes, where each element in the vector indicates useful features used for regression.

2.2. Fully convolutional networks

Starting from the layout described in the previous section, other architectures may be obtained in case the final, fully connected layer (which does not strictly perform convolution) is replaced with other convolutional layers [26]. In this case, Fully Convolutional Networks (FCNs) are obtained [37]. In these networks, all layers are indeed convolutional, which has two eminent advantages: to provide a more flexible choice on the input data size (which is not any more constrained by the size of the fully connected layer), and a better computational management (due to the reduced number of parameters with respect to those of the fully connected layer). The output of a FCN is then a feature map, which is typically a *down-sampled* (or downsized) version of the input image; this output contains different useful information for, e.g. object detection tasks [38,39], where boxes are typically used to highlight one or more target objects along with an associated classification probability. The same result can be obtained with multiple steps even with a simple CNN; however, the FCN is extremely more efficient in performing this task, as it is able to do it in a single step [37].

The down-sampled output resulting from the FCN part of the network may also be *up-sampled* to a larger-size image (eventually, of the same size of the input image). This operation may be typically achieved via interpolation techniques (e.g. nearest-neighbor, bilinear or bicubic). In this case, the first FCN, also called *encoder*, extracts the feature map that is then fed to another fully convolutional network, where deconvolution operations are adopted and implemented, resulting in the so-called *decoder* of the network (see Fig. 3(a)). Each block of the decoder network consists of a repeating structure of up-sampling layers, followed by a few deconvolution layers and an activaction function. The purpose of this second part of the architecture is to reconstruct the image using the information extracted from the encoder part [40,41]. The final overall layout pertains then to an *encoder–decoder* network. These networks are typically employed for image segmentation tasks, where they receive an image as an input, and provide a *segmented* version of the original image (see, e.g. Fig. 3(b)). Other applications include object detection [42], background removal [43] and noise reduction [44].

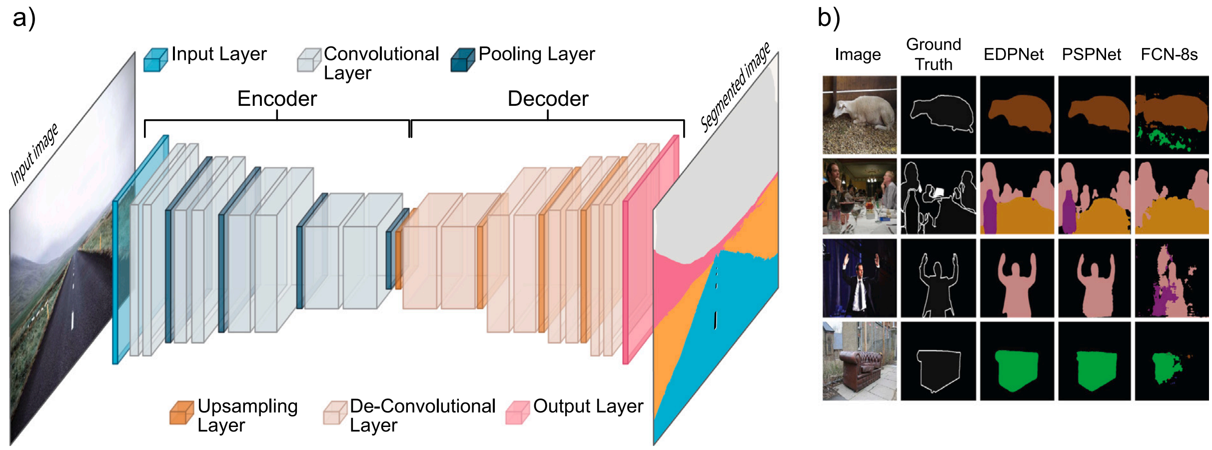


Fig. 3. Example of DNN for image segmentation. In (a) a typical infrastructure of an *encoder–decoder* network is shown. This type of NN involves data passing through a series of convolutional and pooling layers, similarly to the CNN in Fig. 2(a). Afterwards, the extracted features are passed to the decoder, where several upsampling and de-convolutional layers are employed to produce the final segmented image, where each pixel is associated to a certain class (here represented by different colors). In (b), an example of segmentation on the PASCAL VOC 2012 benchmark database [34] is reported. This example compares: Encoding–Decoding Network with Pyramidal representation (EDPNet) by Chen et al. [35], Pyramid Scene Parsing Network (PSPNet) by Zhao et al. [36] and Fully Convolutional Network with an 8-pixel stride (FCN-8s) by Shelhamer et al. [37]. The image is adapted from Ref. [35] under CC BY 4.0 license.

3. Energy applications

3.1. Literature screening

In order to analyze the current trends in the scientific interest and recent applications of Deep Learning methods for image analysis, we first proceed with a screening of the available literature. To this end, we rely on a cross-search on three different databases, namely: Scopus,[2] Web of Science[3] and ScienceDirect,[4] all accessed on date 24/07/2023. These databases were chosen as they provide a comprehensive indexing of scientific documents, and a readily available user interface. All three databases also dispose of several options for searching and analyzing the results. Here we decided to opt for using relevant keywords, to be searched in the abstract, title and keywords in the case of the two former databases, and across the whole article in the case of the latter database. The third database was indeed included for the possibility it provides to search keywords throughout the whole article. We then selected appropriate keywords for our search, that is: "Machine Learning image analysis", "Deep Learning image analysis", "Convolutional Neural Networks image analysis". Different keywords were used along with "image analysis" in a tentative to avoid possible loss of information due to more or less specificity on the tagging of the employed methods within a document.

The results of the research, for the three different databases and keywords, are shown in Fig. 4 (left column), where a grouping by application field has been applied. Note that, this grouping results from the available filtering options for each database and, in case of small discrepancy among the categories, the articles have been manually inserted in the correct collector. We shall first note that the number of articles retrieved from Scopus (Fig. 4(a.i)) and Web of Science (Fig. 4(b.i)) show a similar trend in the applications, with a quite clear bias of the results on "Computer Science", "Engineering" and "Medicine". This polarization may be interpreted based on the large number of articles dedicated to the development of methods ("Computer Science"), a rather generic keyword ("Engineering") and the well-known potential of these methods for the analysis of medical images ("Medicine"). Other notable differences, such as "Mathematics" and "Social Sciences", may be tentatively attributed to either a different coverage or a different application-filtering mechanism of the two databases. As for the adopted keywords, the largest number of results is generally obtained using "Deep Learning" for most of the applications. This may be interpreted assuming it is a more general keyword than "Convolutional Neural Networks" (which are indeed a subset therein), and considering that "Machine Learning" is a quite generic keyword and probably less often used in the context of specific methods (and thus more specific tags).

As expected, the number of articles retrieved from ScienceDirect (Fig. 4(c.i)) is consistently larger. This may be attributed to the different search protocol, where the target keywords can be located anywhere in the article. Note that, this does not necessarily mean that the retrieved source is pertinent, as occasional mention of the selected keywords may occur throughout the body of text (e.g., in the Introduction). A detailed screening of such a large number of full texts may be achieved via text mining techniques [46,47]; however, this goes beyond the purpose of this work. The trends in the applications retrieved from ScienceDirect also shows noticeable differences from the two previous databases, as well as a prominent difference in the results using different keywords. This has led us to conclude that interpretation of the results from a full-text search requires dedicated text mining tools. So, in order to build our database for later analyses, we will adopt the results from the first two databases as a reference, and manually evaluate and add the relevant sources from the third database (as explained next).

Focusing on the number of articles related to "Energy" applications, the first two databases show quite consistent number of results and, based on the adopted filtering criteria, this application seems to still have little representation within the databases. So, for this application, we examine the publication trend during the last years (right column, panels Fig. 4(a.ii), Fig. 4(b.ii) and Fig. 4(c.ii)), using the "Deep Learning" keyword. A well-defined increasing trend emerges from all three databases, demonstrating increasing interest and activity in this application. Thus, in the following, we proceed and analyze in detail the specific applications of the available documents within the energy application field.

Our final database for energy applications is built based on the articles retrieved using the keyword "Deep Learning image analysis" and filtering by "Energy" field from the abstract indexing in Scopus and Web of Science, and from the additional articles from ScienceDirect. For these latter documents, we proceeded by sorting the results of the research by relevance, manually evaluated the pertinence of the sources, and integrated those that were energy related into our database. Starting from the most relevant source, and proceeding in

[2] Scopus database: https://www.scopus.com
[3] Web of Science database: https://www.webofscience.com
[4] ScienceDirect database: https://www.sciencedirect.com

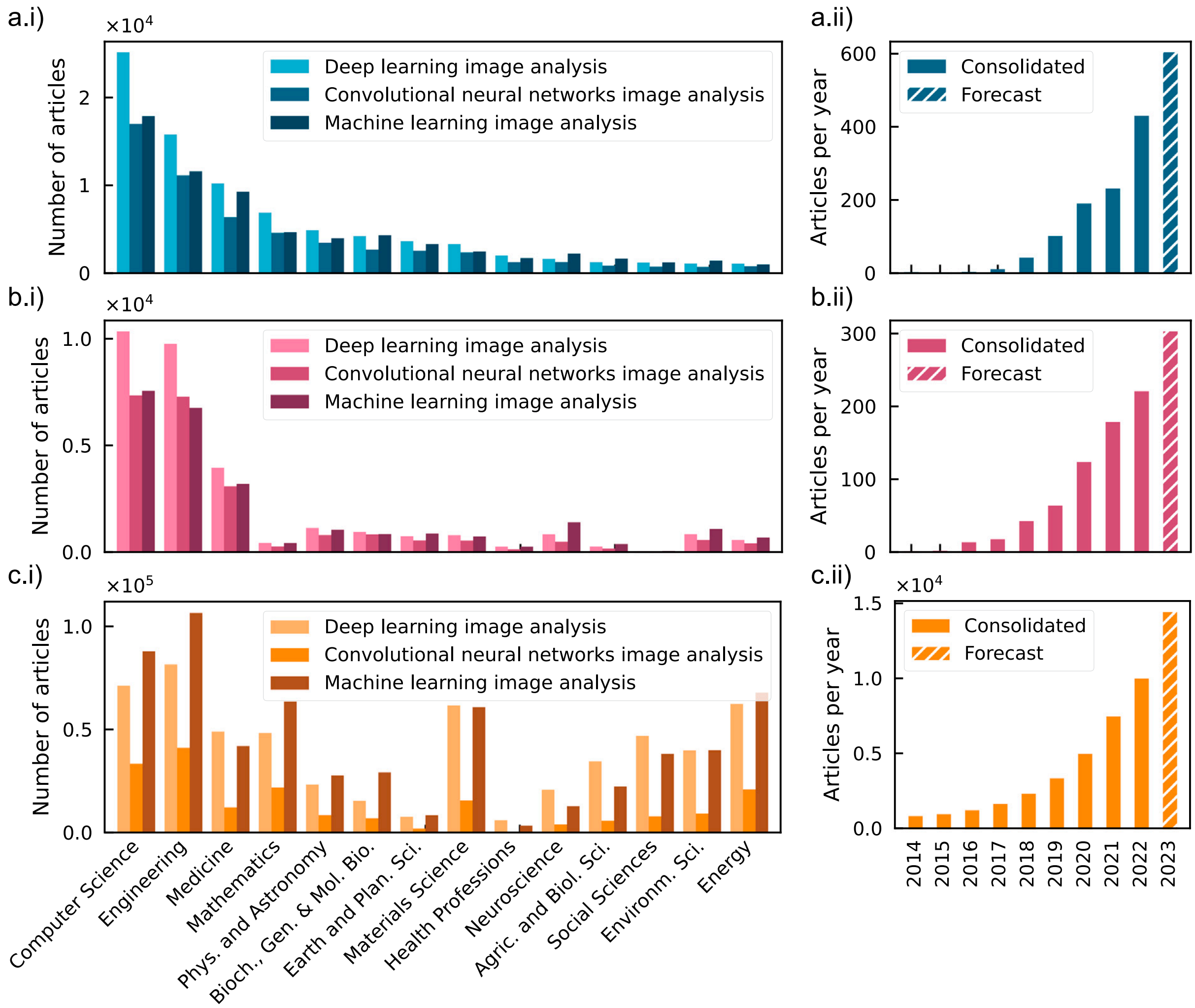


Fig. 4. Results of the analysis on the scientific applications of DNN algorithms for image analysis from different journal article databases: (a) Scopus, (b) Web of Science (WoS), and (c) Science Direct (SciDir). The plots on the left (a.i, b.i, c.i) show the total number of articles obtained by querying each database with the keywords: "Deep Learning Image Analysis" (light color tone), "Convolutional Neural Networks Image Analysis" (medium color tone) and "Machine Learning Image Analysis" (dark color tone). The results of each query were then split according to the scientific area of the article. The abbreviations in the labels are: Phys. and Astronomy: Physics and Astronomy; Bioch., Gen. & Mol. Bio.: Biochemistry, Genetics and Molecular Biology; Earth and Plan. Sci.: Earth and Planetary Sciences; Agric. and Biol. Sci.: Agricultural and Biological Sciences; Environm. Sci.: Environmental Science. The plots on the right show the number of articles published per year, since 2014, for the energy field only, as obtained by the "Deep Learning Image Analysis" keyword. The hatched bar for the year 2023 represents a forecast obtained by interpolating an exponential law, using the number of articles between the years 2020–2022.

descending order, the pertinence was found to be reduced to little after few tens of documents. Using this procedure, 16 pertinent articles (which were not present within the results from the two first databases) were found and included in our database. Based on the adopted screening procedure, our final database for energy applications consists of 152 total articles. Fig. 5 shows a compact visual summary of the adopted protocol for the systematic literature screening, inspired by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [45]. One final consideration on this database is that, articles falling within other categories than "Energy", such as "Materials Science" or "Environmental Science", may still be relevant for our analysis. To overcome possible loss of information, we manually checked the results in these areas and found that the results of interest for our purposes were all already present in the "energy" search.

At an in-depth analysis of the articles classified as energy related in our database, a quite heterogeneous situation of specific applications emerged. Thus, we decided to further categorize the articles based on three different levels of application: *Systems & Policies*, for applications to large power plants or geographic data processing; *Devices*, for energy devices; *Materials*, which includes articles related to materials for energy conversion and storage. An overview of this categorization is shown in Fig. 6, with the actual share on the number of articles in our final database. The specific topical coverage for all the articles in each category is analyzed in the next sections.

3.2. Systems & policies

Consistently with the increasing global demand for energy sustainability and migration from fossil to renewable resources [51], articles falling within this category mostly generally focus on improvement and monitoring of energy conversion and distribution systems in the solar and wind sectors.

The number of solar and wind power plants has increased significantly in recent years, due to the growing interest in renewable energy utilization [52,53]. So, the importance of maintaining high standards of efficiency and reliability has acquired major importance, towards a reduction of costs with an optimized energy production. For this reason, a number of methods have been developed that use neural networks to detect the occurrence of defects or anomalies in photovoltaic modules [54–56] or wind turbines [57–59], and the recognition of the type of defect. For example, in [60,61] images are processed

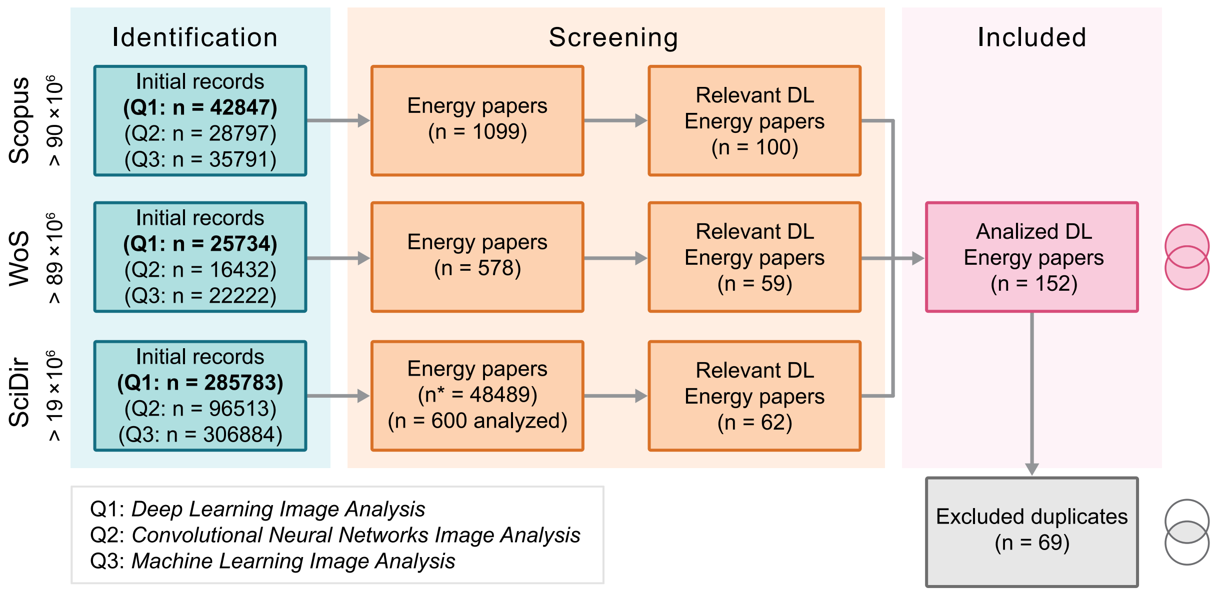


Fig. 5. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram [45] illustrating the evolution of the database throughout our analysis, according to the following stages: *Identification*, *Screening* and *Included*. For each database, namely Scopus, Web of Science (WoS), and Science Direct (SciDir), identification has been obtained according to three queries (Q1, Q2, Q3) using different keywords. The articles retrieved from the "Deep Learning Image Analysis" query underwent then a screening, to refine the selection to only those papers specifically relevant to DL applications in the Energy field. For SciDir, a total of *ca.* n = 600 full texts were selected by pertinence during the screening and analyzed within the total n* results. Finally, in the last step, duplicates were excluded to obtain our final database of included 152 entries.

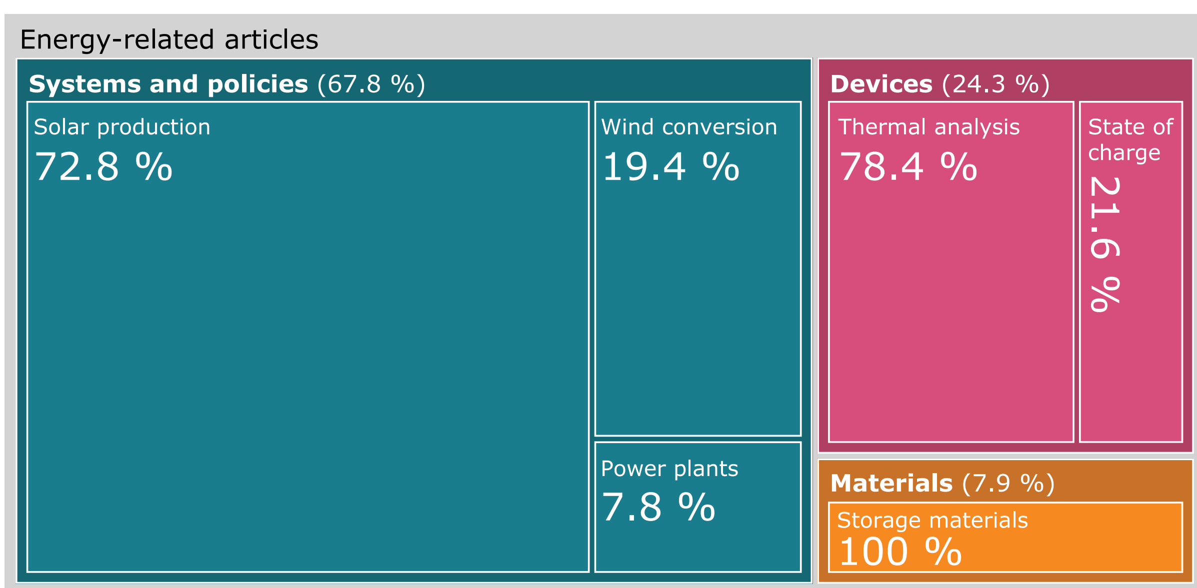


Fig. 6. Tree-map plot of the energy-related articles resulting from the union of the three databases. The energy-related articles were categorized according to the scale of the energy application, i.e. *Systems & Policies*, *Devices*, and *Materials*. The percentage shown in parentheses indicates the percentage of the articles for each of the three categories with respect to the total number of energy-related articles. We further divided each category based on the area of application. For Systems & Policies, the sub-categories were Solar and Wind energy conversion, and Power plants. In the Device category, we found applications related to thermal image analysis and state of charge of batteries. Finally, for the Materials category, all articles fall within energy storage applications. The percentage of each sub-category represents the fraction of articles compared to the total number of articles in its category.

using a convolutional neural network that performs a classification, detecting eventual defects, while in [62] a convolutional encoder decoder model, called U-Net, is used to perform semantic segmentation from electroluminescence images, segmenting any defect in the images. Weather events, that may interfere with the proper operation of power generation, are also included in the analyses, such as partial coverage of solar panels due to snow [63], which uses combined pre-processing techniques, convolutional neural networks and machine learning; other studies analyze points of ice formation on the surface of wind turbine blades [64]. Using these methods, it is thus possible to intervene promptly, ensuring the proper functionality of the plant.

Among the major challenges of these energy sources, other studies address their uncertain and volatile characteristics, which imply the need to ensure the security and stability of energy production, for both solar and wind, by forecasting and organizing the production. Forecasting can be divided into three categories: day-ahead and intra-hours, which are called long and mid term forecasting, and minutes-ahead previsions which are called *nowcasting* [65]. The studies requires no

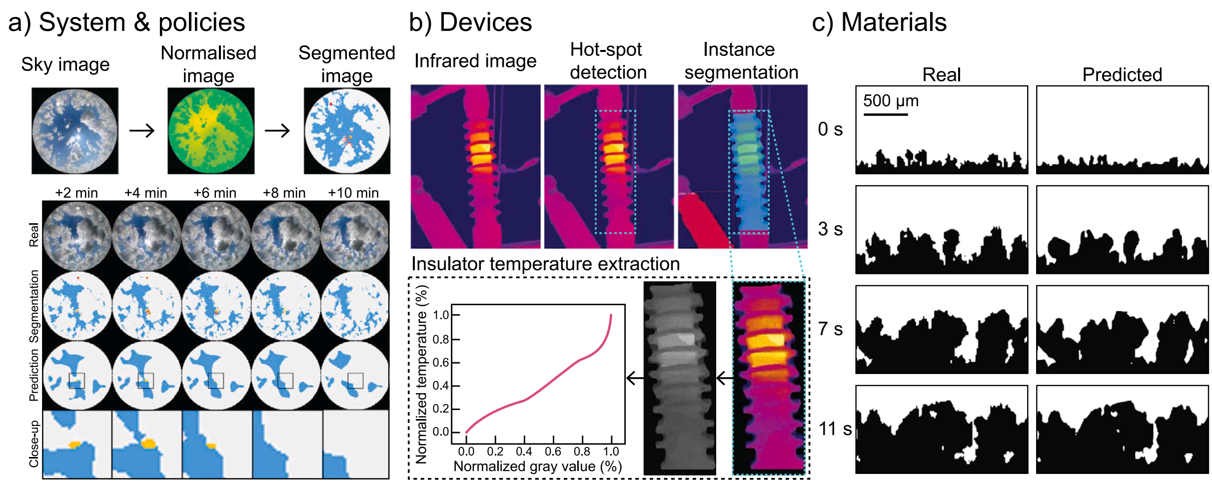


Fig. 7. Examples of how neural networks (NN) are applied to image analysis for each scale of energy application. In the first example (a), the Envisioning CLoud Induced Perturbations in Solar Energy (ECLIPSE) algorithm, developed by Paletta et al. [48], predicts and segments clouds in the sky to determine future irradiance levels and associated uncertainties. This can help to predict the solar energy output and improve energy-grid management. The image is adapted from Ref. [48] under CC BY 4.0 license. The second example (b) shows an infrared detection system developed by Xia et al. [49] used for monitoring and diagnosing faults in electrical devices (in this case, insulators) of the electric power grid. The image is adapted from Ref. [49] under CC BY 4.0 license. The last example (c) is the result of the NN model developed by Kumar et al. [50]. This model predicts the growth of the Solid Electrolyte Interphase (SEI) dendritic electrodeposits on copper electrodes, which is an important degradation mechanism in batteries. Due to its intrinsic multi-scale nature, it is challenging to predict its growth using physics-based models, and here Kumar et al. show how the data-driven models are a valid alternative. The image is adapted from Ref. [50] under CC BY 4.0 license.

numerical data, but only sky images as input for CNN architectures to perform previsions, mostly for the solar case [66–70], but also for wind prediction [71–73]. These data-driven ML and DL methods provide new opportunities to significantly reduce processing time and to continuously refine the output prediction, which supports the emerging interest. Recent research shows that significantly accurate predictions are possible [72,74].

An example of these applications is shown in Fig. 7(a), where the model keeps track of the cloud motion from sky images for short term prediction of irradiance levels and extract information on the local irradiance map [48]. Particularly, here the neural network, which receives an RGB image as an input, is made up of five different modules. The first two are a *Spatial Encoder* and a *Temporal Encoder*, which are a 2D and a 3D convolutional blocks respectively, used to extract the image features and temporal features. The third is a *Future State Prediction* module, which is a convolutional memory-supported module, used to elaborate the features extracted by the previous modules and to provide the short-term predictions. The last two modules are employed to generate the visual output. Particularly, the fourth is a decoder, consisting of convolutional and up-sampling layers, to generate a segmented image in which the sky, clouds and sun are classified. Finally, the last *Irradiance Module*, consists of convolutional layers followed by a fully connected network, and it is used to output the irradiance map. Fig. 7(a) shows five sequential images with a 2 min gap from each other (on the rows, from left to right), and the RGB images, the correct segmentation, the segmentation from the model with a square highlighting the position of the sun, and a close up on the sun to indicate if there is partial or full cloud coverage (on the columns, from top to bottom).

Another relevant application concerns the analysis of the territory, with two different purposes: to evaluate the areas with the highest potential for energy production [20,75], or to count the number of plants already in place. As regards the first purpose, the use of FCN has been suggested to predict annual energy production of a wind plant depending on two input data, the design of the plant and local wind conditions [76]. Other studies analyze the solar potential of rooftops [77]. Here, a FCN performs segmentation of fisheye sky images; then, from the segmented images, features used to calculate the solar potential are extracted. Similarly, satellite images and convolutional encoder–decoder networks have been used to perform data extraction though semantic segmentation for rooftop potentials [78]. Here, the rooftops are classified and differentiated from e.g. streets and other objects; then, a software is employed to transform the segmentation into vector files to compute surfaces and perimeters. Finally, a filter is used to remove unsuitable areas for PV installations (e.g., in this case, those smaller then 5 m^2).

Finally, counting the number of installed plants has particular relevance for the analysis and monitoring of their growth with respect to previous years and for their utilization. Here, relevant works have proposed modifications to a pre-existing CNN, called DeepSolar [79], to perform a segmentation of the images, and obtain the number of elements classified as PV plants [21]. Several other examples within this research field are available [80–84], which have particular relevance in the context of a general assessment on the utilization of renewable sources over large-scale territories (e.g. country-scale level) and may then be helpful for policy makers and energy planning.

3.3. Devices

At the level of energy devices, a quite diversified picture emerges in the specific applications, with some aspects akin to those discussed in the previous section. In general, two main research lines can be outlined: analyses on infrared images for generic thermal monitoring, and analyses on the state of charge of batteries.

Analyses of infrared images include examples on malfunction detection in power and electrical equipment [22], where different CNN architectures are used for fault detection thanks to the infrared image analysis and classification, or on cooling devices [85], where CNNs are used for multi-classification of different fault conditions of a radiator, or for motors [85–87]. In these cases, the analyses help both to flag the problems for prompt action, and to determine the frequency and typology of possible defects. Other efforts focus on the recognition of thermal events [88,89], using FCNNs for classification and object detection on infrared images, detecting the hot spots. Another interesting application relates to combustion systems, where Deep Learning methods can be efficiently exploited for process optimization. Here, several examples of different solutions for the analysis and monitoring of target phenomena are available [90], with the main objective to improve the performance of different devices in operation.

An example application of these methods at device level is shown in Fig. 7(b). Here, a Mask R-CNN (a modified CNN architecture, typically used to perform object detection and image segmentation [91]) was employed to detect target objects in infrared thermographic images (top of the panel) and, following a conversion from RGB to grayscale, to extract temperature values (bottom of the panel) at each point of the grayscale image [49].

Finally, few other works can also be mentioned on other different thermally-related processes, such as the analysis of emissions through image analysis [92], or thermal distributions and dispersion to target and mitigate unintended heat losses; an example of this latter application is reported in [93], where a convolutional encoder–decoder model is trained using numerical (FEM) images to predict the thermal performance of buildings.

As far as energy storage is concerned, a research trend outlines on batteries, and on the monitoring of the state of charge in particular. In this context indeed, image analysis based on Deep Learning methods may stand as a valuable approach to help physics-based modeling of the charging and discharging process in its entire complexity. Examples can be found on the analysis of the state of charge for a better energy management in electric vehicles [94], where multiple Neural Networks, including a CNN for object detection, are used to extract information and monitor the battery power consumption. Other methods can also be used for the prediction of the performance in batteries and fuel cells [95], or for capacity prediction [96]. Fuel cells has been analyzed, for example, using a CNN architecture for predicting different physical quantities resorting to images [97]. Neural networks have also been adopted to recreate synthetic data for the improvement of the battery utilization, using a convolutional encoder–decoder [98]. Finally, an additional relevant application relates to the prediction of the remaining useful life [99], where Fully CNNs and simple CNNs are used to extract physical information from images.

3.4. Materials

We shall now turn the attention on articles specifically related to research on materials for energy purposes. Here, most efforts promptly emerge on development of materials for lithium-ion batteries and fuel cells.

For lithium-ion batteries, image analysis based on Deep Learning methods proves to be particularly effective, as it can help overcoming major modeling challenges related to the high-dimensionality of the problem and to the large number of micro-structure characteristics of the materials. In addition, Deep Learning based approaches may also alleviate the demanding resources required for more conventional trial-and-error methodologies [100]. In this view, different methods based on Deep Learning for image analysis have been developed to encompass different data sources, such as Scanning Electron Microscopy (SEM), Transmission Electron Microscopy (TEM), and Scanning Probe Microscopy (SPM), all discussed theoretically in [23] and more practically in [101]. Another important aspect for batteries relates to the End of Life. Clearly, here the possibility of materials analysis for correct recycling and disposal is key, and Deep Learning can help obtain optimal results by analyzing X-ray images and perform object detection, to identify different elements in images [102].

An example at this level of applications is shown in Fig. 7(c). Here, a method to predict the electro-deposition growth of dendritic copper in electrochemical cells, without prior detailed knowledge of the system, was adopted [50]. A convolutional neural network, supported by a Long Short-Term Memory (LSTM) was used [103]. The convolutions are performed through the image and time, allowing the model to extract the relevant features to predict the electro-deposition growth from a given image. The Figure shows the comparison of the real and predicted evolution (left and right columns, respectively) at different time steps.

In the case of fuel cells, the problem is more related to the repetitiveness of the analysis work, which is also time-consuming and can lead to human errors in the classification [104]. Numerous efforts have been made in this respect, and promising results have been obtained on accuracy and better energy management, using Convolutional encoder–decoder for semantic segmentation [105], and different semantic segmentation algorithms [106–108] used to perform a microstructural analysis. All these works aim to support and accelerate research and development of more and more innovative and performing materials to be used. Finally, an interesting general overview of Machine Learning and Deep Learning methods for image-based material analysis can be found in [109,110], which offer insight into the topic and into possible future directions.

4. Discussion

As a final overview on the presented analysis, we report in Fig. 8 a schematic correlation chart for the available algorithms with respect to different tasks and applications. The algorithms can be divided into two main types: Machine Learning and Neural Networks (these latter being, strictly speaking, a sub-set of the former).

Machine Learning includes algorithms which are typically less demanding in terms of computational resources than neural networks; thus, they can be generally adopted in those cases where the desired level of accuracy can be obtained with a simple dataset (allowing faster project development). Here, typical tasks on image analysis involve regression, classification, segmentation and clustering. For regression, methods include linear, multi-linear or polynomial regression [111] or LASSO regression [112]; whereas, for classification or segmentation include K-Nearest Neighbor [113], Support Vector Machine [114], Naive Bayes [115], Decision Tree [116] and Random Forest [117]. For clustering, instead, typical ML algorithms include K-means clustering [118], DBSCAN [119], Gaussian mixture model [120], Mean-shift [121] or Agglomerative clustering [122].

Neural Networks algorithms include Convolutional Neural Networks (CNN), Fully-convolutional Networks (FCN) and Encoder–Decoder networks, which have all been discussed in this work. Typically, CNN are best suited for classification and regression [66], whereas FCN are mostly adopted for object detection and segmentation [76,77]. Encoder–Decoder may be typically employed for image segmentation, image enhancement, and image generation [41]. Note that, neural networks is a constantly expanding topic, and several other methods may be included into this category, such as the Generative-Adversarial Networks (GANs) [123]. These networks are characterized by two different modules: the *generator*, which is trained to generate data, and the *discriminator*, whose goal is to detect real data from fake data. During training, the discriminator forces the generator to improve in its task in order to generate images that are increasingly similar to the real ones. Once the training is completed, the generator only is used for the required task. Typical applications of these algorithms include image enhancement, which is used to increase the resolution of an image [124], or to restore the missing parts of an image, or for image generation, which is used to generate virtual dataset. Different typologies of GANs are available, among the mostly adopted the following can be mentioned: Vanilla GAN, Conditional GAN (CGAN), Deep Convolutional GAN (DCGAN), CycleGAN, Style GAN or Super Resolution GAN (SRGAN) [123,125].

From our analysis, we observed that several tasks can be related to different specific energy-oriented applications (see Fig. 8). The most explored applications result in classification, object detection and segmentation for different energy applications, and regression mostly for research at the materials level (see also Fig. 7). On the other hand, image enhancement for materials research results to be still little explored; however, potential applications may include 3D geometry reconstruction from 2D microscopic images [108]. Finally, the remaining tasks for the presented applications appear to be still little or not yet explored.

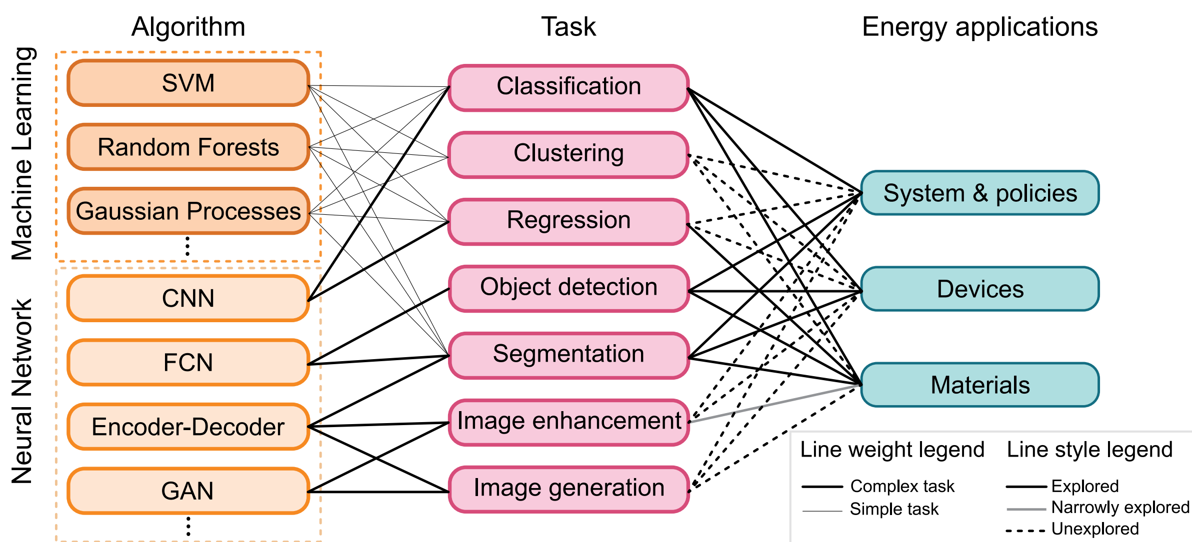


Fig. 8. Schematic correlation chart of available algorithms, tasks and applications, to assist with the decision-making for the choice of different data-driven models for energy applications. In the middle of the diagram, a list of the computer-vision tasks (magenta boxes) presented in this work are shown. Each task can be accomplished using a ML (dark orange boxes) or a NN (light orange) model, listed on the left, depending on the complexity of the problem. Note that some models can perform multiple tasks, which is determined by the training method and hyperparameters used. On the right, the three scales of the energy applications are shown (cyan boxes): *System & Policies*, *Devices* and *Materials*. All data-driven tasks can serve all three scales, but the black solid line indicates where our literature analysis shows significant success in applying one model. The gray line indicates where the community moderately exploits the application, and the dashed black line indicates areas that, based on our analyses, are still little or not yet explored.

Table 1
Overview of key properties, generally observed tasks and related applications in the energy field for Convolutional Neural Networks (CNN), Fully-Convolutional Networks (FCN), Encoder–Decoder networks (E–D) and Generative Adversarial Networks (GAN) with potential future challenges for an increased exploitation in the energy field.

NN	Key properties	General tasks	Energy applications	Energy field	Future challenges
CNN	· Efficient feature extraction · High parameter efficiency · Robustness to image distortions · Computationally intensive	· Classification · Regression	· Fault and anomaly detection · Energy consumption forecasting · Energy production forecasting	· Wind and PV power generation systems · General power plant monitoring · Electro-chemical batteries	· Improve resource handling for large datasets · Reduce data input requirements · Include domain-specific knowledge
FCN	· Pixel-level output · Flexibility on input images · Computationally intensive · Possible contextual limitation	· Object detection · Segmentation	· Energy production forecasting · Recognition of thermal events	· Wind and PV power generation systems · Performance analysis of energy devices	· Improve scaling with input size · Improve contextual inference · Include domain-specific knowledge
E–D	· Good contextual capturing · Versatile architecture for different tasks · Possible complex training · Possible reconstruction loss	· Image segmentation · Image enhancement · Image generation	· Spatial energy production forecasting · Defect identification · Thermal performance prediction · Synthetic data generation · Micro-structural analysis	· Energy-efficient building design · Fuel cells · Electro-chemical batteries	· Improve training times · Improve parallel data processing · Include domain-specific knowledge
GAN	· High-quality image generation · Prone to Unsupervised Learning tasks · Difficult training stability · Resource intensive	· Image enhancement · Resolution increase · Image reconstruction · Image generation	· Novel energy pattern discovery · Generation of synthetic images	· Wind and PV power generation systems · Energy consumption analysis · Power plant monitoring	· Improve resource handling · Improve training stability

In general, and in perspective, suitable applications for DNN based image analysis that still appear to be little or not yet explored, may be expected to include particularly those situations where a physics-based modeling approach can be limited by the complexity of the problem at hand [126]. As an example, liquid foams, although relevant for a number of energy-related problems such as enhanced oil recovery and carbon capture [127–129], are notably challenging from a modeling point of view, due to the wide range of morphological and dynamical features they exhibit at different scales. For this reason, they are indeed typically best modeled as simplified mathematical objects [130]; hence, data-driven approaches based on images may stand as a valuable alternative or complementary tools for e.g. time-evolution prediction of complex systems based on bubble pattern recognition [131] or foam structure analysis and optimization [132] for different energy-related soap film [133,134] and foam-based [135] applications. No less, these methods may be helpful for the analysis of phase-change dynamics and optimization for latent heat storage applications [136,137], particularly, as alternate approaches to e.g. simplified ones [138] for the analysis of the phase-change propagation front based on image data acquisition and processing.

Towards further applications, an overview of the key properties for the different DL methods, along with a summary of the observed energy-oriented applications available is reported in Table 1. The Table also reports on the possible perspective challenges for a wider adoption and further exploitation of these methods for energy applications in the future, discussed in more detail in the next section.

5. Current limits and perspective challenges

Based on our analyses, some specific features that DNNs must hold for their application to energy-related problems emerge with respect to standard image analysis. These specific features can be generally summarized as follows.

Data characteristics and quality. Energy systems can often generate more complex, noisy, or irregular data compared to well-structured and high-quality images typically used in standard DNN applications. For example, these can include sensor data from power grids, satellite images for renewable energy assessments, and thermal images for the analysis of energy efficiency. Therefore, these data may typically require different pre-processing and specific feature extraction techniques.

Time evolution dynamics. The analysis of energy systems often requires considering temporal dynamics, such as for load forecasting, where the data is inherently based on time series and influenced by many factors, such as weather, the usage patterns, and economic activities. Standard DNNs for image analysis are not specifically designed to handle these time-evolution dependencies effectively.

Spatial complexity. In some specific applications, such as smart grid management or mapping of renewable resources, the spatial complexity and geographical context are crucial for proper results. Standard image analysis DNNs are not generally optimized to process and understand this spatial context, as it is required for energy applications.

Scale and Granularity. Energy data can vary significantly in scale and granularity. For example, the analysis of data from a single wind turbine or from a large-scale national power grid, requires different approaches and tuning. Thus, standard image DNNs may not be efficient in handling such different data scales without proper modifications.

Regulatory and safety considerations. Energy systems are very often subject to strict regulatory standards and safety requirements. Therefore, data-driven models to be used in this sector must be highly accurate, reliable, and interpretable. Standard DNNs for image analysis are not generally subject to such specific constraints; thus, applications to the energy field require proper reliability assessment in this sense.

Integration of domain knowledge. The analysis of energy systems often requires the integration of domain-specific knowledge into model development and assessment, such as engineering principles, which is not generally required in standard DNNs for image processing. Therefore, tailored intervention to include this knowledge is generally and often necessary.

Resource constraints. Standard DNNs for image analysis can be resource-intensive, as they are not generally subject to specific constraints in this regard. On the other hand, energy applications may often present resource-limited environments, which limits the direct usage of standard DNNs without proper interventions.

Diverse objectives and metrics. Diverse specific objectives for energy applications, such as optimizing energy efficiency or predicting failures, require different metrics for the evaluation of the results (such as, e.g., accuracy, resilience). These metrics can be very different from those in standard image analysis; thus, specific metrics should be developed on a case-by-case bases along with the related tuned model architectures and training approaches.

The previous required features can then be summarized into the following perspective challenges for the field: developing models that can handle noisy and irregular data, incorporate temporal and spatial dynamics, include domain-specific knowledge into models, and meet the required accuracy to comply with regulatory and safety standards in the sector. In order to address these challenges, adaptation and tuning of state-of-the-art models can be envisioned, involving a multidisciplinary effort that combines data science experience, materials and energy engineering, and regulatory expertise. Regarding the inclusion of domain-specific knowledge into models, in particular, the recently introduced Physics-Informed Neural Networks (PINNs) represent a promising pathway [139,140]. These networks allow indeed inclusion of the information on the governing physical laws into the learning process for a given problem, which is a significant step forward in blending scientific knowledge with neural network models. This feature is particularly relevant for the applications in the energy sector, where the adherence of the results to physical laws and system dynamics is crucial. Thus, the application of PINNs to image analysis opens to remarkable possibilities in the field.

Besides, more and more computationally efficient models are required. Model pruning [141], that is the removal of unnecessary or redundant parameters from a neural network that do not contribute significantly to performance, represents a readily available pathway for faster inference times and lower memory requirements. However, more systematic advancements on computational efficiency may still require improvement of the neural network architectures and optimization of the hardware utilization, levering parallel processing capabilities of the Graphic Processing Units (GPUs) and Tensor Processing Units (TPUs).

6. Conclusions

In this work, the salient aspects of Machine Learning and Deep Learning methods for image analysis have been reviewed and discussed. Based on this methodological ground, a literature screening has been performed on the related application fields, to outline the current research trends and analyze possible emerging opportunities. Particularly, the applications related to the energy sector at large have been analyzed in detail. The analysis has shown that, while usage of Deep Learning methods for images appears to be still at the early stage of exploitation for energy-related applications, an increasing trend clearly emerges in the number of publications per year.

The database of energy-oriented articles, obtained by a cross-search on multiple literature resources, has been examined. The retrieved energy-related applications have been categorized into three different levels, namely: systems and policies, devices and materials. At the systems and policies level, most articles have been found to focus on the improvement and monitoring of energy conversion and distribution systems in the solar and wind sectors. At the device level, a trend appears on application to analyses on infrared images for generic thermal monitoring, and analyses on the state of charge of batteries. Finally, at the materials level, most of the efforts seem to focus on the analysis of microscopic images towards improvement of lithium-ion batteries and fuel cells.

In order to provide a systematic overview of the results obtained, a schematic correlation chart of available algorithms, tasks and applications has been presented and discussed. This chart is intended also as a useful tool to assist with decision-making for the choice of different data-driven models for energy applications. For the sake of clarity, we remark that, considering the large extent and possible different interpretations of "energy-related" applications, this work does not pretend to be completely exhaustive on the coverage of all possible examples. Notwithstanding, the presented literature analysis has been designed to outline the main current and emerging research lines in the field. In this view, we hope that this work will stimulate further activities on application of Deep Learning methods for image analysis in the energy field.

CRediT authorship contribution statement

Mattia Casini: Data curation, Formal analysis, Investigation, Writing – original draft. **Paolo De Angelis:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – review & editing. **Eliodoro Chiavazzo:** Conceptualization, Formal analysis, Methodology, Supervision, Validation, Writing – review & editing. **Luca Bergamasco:** Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

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.

Acknowledgments

The authors acknowledge Paolo Vigo for a preliminary literature search and fruitful discussions during the early stage of this work.

References

[1] Russell SJ. Artificial intelligence a modern approach. Pearson Education, Inc.; 2010.
[2] Johri P, Verma JK, Paul S. Applications of machine learning. Springer; 2020.
[3] Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. Nature 2018;559(7715):547–55.
[4] Trezza G, Bergamasco L, Fasano M, Chiavazzo E. Minimal crystallographic descriptors of sorption properties in hypothetical MOFs and role in sequential learning optimization. npj Comput Mater 2022;8(1):123.
[5] Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 2015;13:8–17.
[6] Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, et al. Machine learning and the physical sciences. Rev Modern Phys 2019;91(4):045002.
[7] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with alphafold. Nature 2021;596(7873):583–9.
[8] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–44.
[9] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press; 2016.
[10] Ravindran S. Five ways deep learning has transformed image analysis. Nature 2022;609.
[11] Szeliski R. Computer vision: Algorithms and applications. Springer Nature; 2022.
[12] Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR; 2019, p. 6105–14.
[13] Khan S, Rahmani H, Shah SAA, Bennamoun M. A guide to convolutional neural networks for computer vision. Synth Lect Comput Vis 2018;8(1):1–207.
[14] Chen L, Li S, Bai Q, Yang J, Jiang S, Miao Y. Review of image classification algorithms based on convolutional neural networks. Remote Sens 2021;13(22):4712.
[15] Dhillon A, Verma GK. Convolutional neural network: a review of models, methodologies and applications to object detection. Prog Artif Intell 2020;9(2):85–112.
[16] Zhu Y, Brettin T, Xia F, Partin A, Shukla M, Yoo H, et al. Converting tabular data into images for deep learning with convolutional neural networks. Sci Rep 2021;11(1):11325.
[17] Lu P, Song B, Xu L. Human face recognition based on convolutional neural network and augmented dataset. Syst Sci Control Eng 2021;9(sup2):29–37.
[18] Brioschi ML, Neto CD, de Toledo M, Neves EB, Vargas JVC, Teixeira MJ. Infrared image method for possible covid-19 detection through febrile and subfebrile people screening. J Therm Biol 2023;112:103444.
[19] Gupta A, Anpalagan A, Guan L, Khwaja AS. Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. Array 2021;10:100057.
[20] Bergamasco L, Asinari P. Scalable methodology for the photovoltaic solar energy potential assessment based on available roof surface area: Further improvements by ortho-image analysis and application to Turin (Italy). Sol Energy 2011;85(11):2741–56.
[21] Mayer K, Wang Z, Arlt M-L, Neumann D, Rajagopal R. Deepsolar for Germany: A deep learning framework for pv system mapping from aerial imagery. In: 2020 International conference on smart energy systems and technologies. IEEE; 2020, p. 1–6.
[22] Jiang A, Yan N, Shen B, Gu C, Zhu H, Huang H. Research on infrared image recognition method of power equipment based on deep learning. In: 2020 IEEE international conference on high voltage engineering and application. IEEE; 2020, p. 1–4.
[23] Ge M, Su F, Zhao Z, Su D. Deep learning analysis on microscopic imaging in materials science. Mater Today Nano 2020;11:100087.
[24] Khan AI, Al-Habsi S. Machine learning in computer vision. Procedia Comput Sci 2020;167:1444–51.
[25] Montavon G, Samek W, Müller K-R. Methods for interpreting and understanding deep neural networks. Digit Signal Process 2018;73:1–15.
[26] Géron A. Hands-on machine learning with scikit-learn, keras, and tensorflow. O'Reilly Media, Inc.; 2022.
[27] Sharma S, Sharma S, Athaiya A. Activation functions in neural networks. Towards Data Sci 2017;6(12):310–6.
[28] Philipp G, Song D, Carbonell JG. The exploding gradient problem demystified-definition, prevalence, impact, origin, tradeoffs, and solutions. 2017, arXiv preprint arXiv:1712.05577.
[29] Carlier A, Danelljan M, Alahi A, Timofte R. Deepsvg: A hierarchical generative network for vector graphics animation. Adv Neural Inf Process Syst 2020;33:16351–61.
[30] Jiang X, Liu L, Shan C, Shen Y, Dong X, Li D. Recognizing vector graphics without rasterization. Adv Neural Inf Process Syst 2021;34:24569–80.
[31] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014, arXiv preprint arXiv:1409.1556.
[32] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, p. 770–8.
[33] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. Pattern Recognit 2018;77:354–77.
[34] Mottaghi R, Chen X, Liu X, Cho N-G, Lee S-W, Fidler S, et al. The role of context for object detection and semantic segmentation in the wild. In: IEEE conference on computer vision and pattern recognition. 2014.
[35] Chen D, Li X, Hu F, Mathiopoulos PT, Di S, Sui M, et al. Edpnet: An encoding–decoding network with pyramidal representation for semantic image segmentation. Sensors 2023;23(6):3205.
[36] Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, p. 2881–90.
[37] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. IEEE Trans Pattern Anal Mach Intell 2017;39(4):640–51. http://dx.doi.org/10.1109/TPAMI.2016.2572683.
[38] Dai J, Li Y, He K, Sun J. R-fcn: Object detection via region-based fully convolutional networks. Adv Neural Inf Process Syst 2016;29.
[39] Tian Z, Shen C, Chen H, He T. Fcos: Fully convolutional one-stage object detection. In: Proceedings of the IEEE/CVF international conference on computer vision. 2019, p. 9627–36.
[40] Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation. In: Proceedings of the IEEE international conference on computer vision. 2015, p. 1520–8.
[41] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder–decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 2017;39(12):2481–95.
[42] Song G, Song K, Yan Y. Edrnet: Encoder–decoder residual network for salient object detection of strip steel surface defects. IEEE Trans Instrum Meas 2020;69(12):9709–19.
[43] Dumitras A, Haskell BG. An encoder–decoder texture replacement method with application to content-based movie coding. IEEE Trans Circuits Syst Video Technol 2004;14(6):825–40.
[44] Konstantinova T, Wiegart L, Rakitin M, DeGennaro AM, Barbour AM. Noise reduction in X-ray photon correlation spectroscopy with convolutional neural networks encoder–decoder models. Sci Rep 2021;11(1):14756.
[45] Moher D, Liberati A, Tetzlaff J, Altman DG, P Group*. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. Ann Intern Med 2009;151(4):264–9.
[46] Tshitoyan V, Dagdelen J, Weston L, Dunn A, Rong Z, Kononova O, et al. Unsupervised word embeddings capture latent knowledge from materials science literature. Nature 2019;571(7763):95–8.
[47] Delen D, Crossland MD. Seeding the survey and analysis of research literature with text mining. Expert Syst Appl 2008;34(3):1707–20.
[48] Paletta Q, Hu A, Arbod G, Lasenby J. Eclipse: Envisioning cloud induced perturbations in solar energy. Appl Energy 2022;326:119924.
[49] Xia C, Ren M, Wang B, Dong M, Xu G, Xie J, et al. Infrared thermography-based diagnostics on power equipment: State-of-the-art. High Voltage 2021;6(3):387–407.
[50] Kumar R, Dhara T, Hu H, Chakraborty M. Visualization-based prediction of dendritic copper growth in electrochemical cells using convolutional long short-term memory. Energy AI 2022;10:100203.
[51] De Angelis P, Tuninetti M, Bergamasco L, Calianno L, Asinari P, Laio F, et al. Data-driven appraisal of renewable energy potentials for sustainable freshwater production in Africa. Renew Sustain Energy Rev 2021;149:111414.
[52] Bologna A, Fasano M, Bergamasco L, Morciano M, Bersani F, Asinari P, et al. Techno-economic analysis of a solar thermal plant for large-scale water pasteurization. Appl Sci 2020;10(14):4771.
[53] Jang D, Kim K, Kim K-H, Kang S. Techno-economic analysis and Monte Carlo simulation for green hydrogen production using offshore wind power plant. Energy Convers Manage 2022;263:115695.
[54] Venkatesh SN, Jeyavadhanam BR, Sizkouhi AM, Esmailifar SM, Aghaei M, Sugumaran V. Automatic detection of visual faults on photovoltaic modules using deep ensemble learning network. Energy Rep 2022;8:14382–95.

[55] Mellit A, Kalogirou S. Assessment of machine learning and ensemble methods for fault diagnosis of photovoltaic systems. Renew Energy 2022;184:1074–90.

[56] Haidari P, Hajiahmad A, Jafari A, Nasiri A. Deep learning-based model for fault classification in solar modules using infrared images. Sustain Energy Technol Assess 2022;52:102110.

[57] Yang X, Zhang Y, Lv W, Wang D. Image recognition of wind turbine blade damage based on a deep learning model with transfer learning and an ensemble learning classifier. Renew Energy 2021;163:386–97.

[58] Xiaoxun Z, Xinyu H, Xiaoxia G, Xing Y, Zixu X, Yu W, et al. Research on crack detection method of wind turbine blade based on a deep learning method. Appl Energy 2022;328:120241.

[59] Attallah O, Ibrahim RA, Zakzouk NE. Cad system for inter-turn fault diagnosis of offshore wind turbines via multi-cnns & feature selection. Renew Energy 2023;203:870–80.

[60] Akram MW, Li G, Jin Y, Chen X, Zhu C, Ahmad A. Automatic detection of photovoltaic module defects in infrared images with isolated and develop-model transfer deep learning. Sol Energy 2020;198:175–86.

[61] Guo J, Liu C, Cao J, Jiang D. Damage identification of wind turbine blades with deep convolutional neural networks. Renew Energy 2021;174:122–33.

[62] Pratt L, Govender D, Klein R. Defect detection and quantification in electro-luminescence images of solar PV modules using u-net semantic segmentation. Renew Energy 2021;178:1211–22.

[63] Zhang X, Araji MT. Snow loss modeling for solar modules using image processing and deep learning. Sustain Energy Grids Netw 2023;34:101036.

[64] Hu Q, Xu X, Leng D, Shu L, Jiang X, Virk M, et al. A method for measuring ice thickness of wind turbine blades based on edge detection. Cold Reg Sci & Technol 2021;192:103398.

[65] Niccolai A, Orooji S, Matteri A, Ogliari E, Leva S. Irradiance nowcasting by means of deep-learning analysis of infrared images. Forecasting 2022;4(1):338–48.

[66] Feng C, Zhang J. Solarnet: A sky image-based deep convolutional neural network for intra-hour solar forecasting. Sol Energy 2020;204:71–8.

[67] Zhang R, Ma H, Saha TK, Zhou X. On sky imaging analysis and deep learning for photovoltaic output nowcasting. In: 2020 IEEE power & energy society general meeting. IEEE; 2020, p. 1–5.

[68] Kong W, Jia Y, Dong ZY, Meng K, Chai S. Hybrid approaches based on deep whole-sky-image learning to photovoltaic generation forecasting. Appl Energy 2020;280:115875.

[69] Zhen Z, Zhang X, Mei S, Chang X, Chai H, Yin R, et al. Ultra-short-term irradiance forecasting model based on ground-based cloud image and deep learning algorithm. IET Renew Power Gener 2022;16(12):2604–16.

[70] Pérez E, Pérez J, Segarra-Tamarit J, Beltran H. A deep learning model for intra-day forecasting of solar irradiance using satellite-based estimations in the vicinity of a PV power plant. Sol Energy 2021;218:652–60.

[71] Liu Y, Qin H, Zhang Z, Pei S, Jiang Z, Feng Z, et al. Probabilistic spatiotemporal wind speed forecasting based on a variational bayesian deep learning model. Appl Energy 2020;260:114259.

[72] Zheng L, Lu W, Zhou Q. Weather image-based short-term dense wind speed forecast with a convlstm-lstm deep learning model. Build Environ 2023;239:110446.

[73] Yu R, Liu Z, Li X, Lu W, Ma D, Yu M, et al. Scene learning: Deep convolutional networks for wind power prediction by embedding turbines into grid space. Appl Energy 2019;238:249–57.

[74] Paletta Q, Arbod G, Lasenby J. Omnivision forecasting: Combining satellite and sky images for improved deterministic and probabilistic intra-hour solar energy predictions. Appl Energy 2023;336:120818.

[75] Bocca A, Bergamasco L, Fasano M, Bottaccioli L, Chiavazzo E, Macii A, et al. Multiple-regression method for fast estimation of solar irradiation and photovoltaic energy potentials over Europe and Africa. Energies 2018;11(12):3477.

[76] Harrison-Atlas D, King RN, Glaws A. Machine learning enables national assessment of wind plant controls with implications for land use. Wind Energy 2022;25(4):618–38.

[77] Lan H, Gou Z, Xie X. A simplified evaluation method of rooftop solar energy potential based on image semantic segmentation of urban streetscapes. Sol Energy 2021;230:912–24.

[78] Sun T, Shan M, Rong X, Yang X. Estimating the spatial distribution of solar photovoltaic power generation potential on different types of rural rooftops using a deep learning network applied to satellite images. Appl Energy 2022;315:119025.

[79] Yu J, Wang Z, Majumdar A, Rajagopal R. Deepsolar: A machine learning framework to efficiently construct a solar deployment database in the United States. Joule 2018;2(12):2605–17.

[80] Hoeser T, Kuenzer C. Synteo: Synthetic dataset generation for earth observation and deep learning–demonstrated for offshore wind farm detection. ISPRS J Photogramm Remote Sens 2022;189:163–84.

[81] Manso-Callejo M-Á, Cira C-I, Garrido RPA, Matesanz FJG. First dataset of wind turbine data created at national level with deep learning techniques from aerial orthophotographs with a spatial resolution of 0.5 m/pixel. IEEE J Sel Top Appl Earth Obs Remote Sens 2021;14:7968–80.

[82] Krapf S, Kemmerzell N, Khawaja Haseeb Uddin S, Hack Vazquez M, Netzler F, Lienkamp M. Towards scalable economic photovoltaic potential analysis using aerial images and deep learning. Energies 2021;14(13):3800.

[83] Yuan J, Yang H-HL, Omitaomu OA, Bhaduri BL. Large-scale solar panel mapping from aerial images using deep convolutional networks. In: 2016 IEEE international conference on big data. IEEE; 2016, p. 2703–8.

[84] Nakhaee A, Paydar A. Deepradiation: An intelligent augmented reality platform for predicting urban energy performance just through 360 panoramic streetscape images utilizing various deep learning models. In: Building simulation, vol. 16. Springer; 2023, p. 499–510.

[85] Nasiri A, Taheri-Garavand A, Omid M, Carlomagno GM. Intelligent fault diagnosis of cooling radiator based on deep learning analysis of infrared thermal images. Appl Therm Eng 2019;163:114410.

[86] Hejazi S, Packianather M, Liu Y. Novel preprocessing of multimodal condition monitoring data for classifying induction motor faults using deep learning methods. In: 2022 IEEE 2nd international symposium on sustainable energy, signal processing and cyber security. IEEE; 2022, p. 1–6.

[87] Sabah AN, Jaffery ZA. Fault detection of induction motor using thermal imaging. In: 2022 IEEE IAS global conference on emerging technologies. IEEE; 2022, p. 84–90.

[88] Grelier E, Mitteau R, Moncada V. Deep learning-based process for the automatic detection, tracking, and classification of thermal events on the in-vessel components of fusion reactors. Fusion Eng Des 2023;192:113636.

[89] Grelier E, Mitteau R, Moncada V. Deep learning and image processing for the automated analysis of thermal events on the first wall and divertor of fusion reactors. Plasma Phys Control Fusion 2022;64(10):104010.

[90] Zhou L, Song Y, Ji W, Wei H. Machine learning for combustion. Energy AI 2022;7:100128.

[91] He K, Gkioxari G, Dollár P, Girshick R. Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision. 2017, p. 2961–9.

[92] Li N, Lu G, Li X, Yan Y. Prediction of nox emissions from a biomass fired combustion process based on flame radical imaging and deep learning techniques. Combust Sci Technol 2016;188(2):233–46.

[93] Kim D-J, Kim S-I, Kim H-S. Thermal simulation trained deep neural networks for fast and accurate prediction of thermal distribution and heat losses of building structures. Appl Therm Eng 2022;202:117908.

[94] Tang X, Chen J, Yang K, Toyoda M, Liu T, Hu X. Visual detection and deep reinforcement learning-based car following and energy management for hybrid electric vehicles. IEEE Trans Transp Electr 2022;8(2):2501–15.

[95] Xing W, Dai S, Shah A, Luo L, Xu Q, Leung P. Emulating spatial and temporal outputs from fuel cell and battery models: A comparison of deep learning and gaussian process models. J Electrochem Energy Convers Storage 2023;20(1):011007.

[96] El-Dalahmeh M, Al-Greer M, El-Dalahmeh M, Short M. Time-frequency image analysis and transfer learning for capacity prediction of lithium-ion batteries. Energies 2020;13(20):5447.

[97] Pang Y, Hao L, Wang Y. Convolutional neural network analysis of radiography images for rapid water quantification in pem fuel cell. Appl Energy 2022;321:119352.

[98] Müller S, Sauter C, Shunmugasundaram R, Wenzler N, De Andrade V, De Carlo F, et al. Deep learning-based segmentation of lithium-ion battery microstructures enhanced by artificially generated electrodes. Nature Commun 2021;12(1):6205.

[99] Couture J, Lin X. Image-and health indicator-based transfer learning hybridization for battery rul prediction. Eng Appl Artif Intell 2022;114:105120.

[100] Faraji Niri M, Mafeni Mase J, Marco J. Performance evaluation of convolutional auto encoders for the reconstruction of li-ion battery electrode microstructure. Energies 2022;15(12):4489.

[101] Zhang G-X, Song Y, Zhao W, An H, Wang J. Machine learning-facilitated multiscale imaging for energy materials. Cell Rep Phys Sci 2022;101008.

[102] Sterkens W, Diaz-Romero D, Goedemé T, Dewulf W, Peeters JR. Detection and recognition of batteries on X-ray images of waste electrical and electronic equipment using deep learning. Resour Conserv Recy 2021;168:105246.

[103] Sherstinsky A. Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. Physica D 2020;404:132306.

[104] Hwang H, Choi SM, Oh J, Bae S-M, Lee J-H, Ahn J-P, et al. Integrated application of semantic segmentation-assisted deep learning to quantitative multi-phased microstructural analysis in composite materials: Case study of cathode composite materials of solid oxide fuel cells. J Power Sources 2020;471:228458.

[105] Colliard-Granero A, Batool M, Jankovic J, Jitsev J, Eikerling MH, Malek K, et al. Deep learning for the automation of particle analysis in catalyst layers for polymer electrolyte fuel cells. Nanoscale 2022;14(1):10–8.

[106] Hwang H, Ahn J, Lee H, Oh J, Kim J, Ahn J-P, et al. Deep learning-assisted microstructural analysis of ni/ysz anode composites for solid oxide fuel cells. Mater Charact 2021;172:110906.

[107] Yang Y, Li N, Wang B, Li N, Gao K, Liang Y, et al. Microstructure evolution of lithium-ion battery electrodes at different states of charge: Deep learning-based segmentation. Electrochem Commun 2022;136:107224.

[108] Nagda V, Kulachenko A, Lindström SB. Image-based 3d characterization and reconstruction of heterogeneous battery electrode microstructure. Comput Mater Sci 2023;223:112139.
[109] Choudhary K, DeCost B, Chen C, Jain A, Tavazza F, Cohn R, et al. Recent advances and applications of deep learning methods in materials science. npj Comput Mater 2022;8(1):59.
[110] Jacobs R. Deep learning object detection in materials science: Current state and future directions. Comput Mater Sci 2022;211:111527.
[111] Maulud D, Abdulazeez AM. A review on linear regression comprehensive in machine learning. J Appl Sci Technol Trends 2020;1(4):140–7.
[112] Muthukrishnan R, Rohini R. Lasso: A feature selection technique in predictive modeling for machine learning. In: 2016 IEEE international conference on advances in computer applications. IEEE; 2016, p. 18–20.
[113] Peterson LE. K-nearest neighbor. Scholarpedia 2009;4(2):1883.
[114] Noble WS. What is a support vector machine? Nature Biotechnol 2006;24(12):1565–7.
[115] Murphy KP, et al. Naive bayes classifiers. Univ Br Columbia 2006;18(60):1–8.
[116] Quinlan JR. Induction of decision trees. Mach Learn 1986;1:81–106.
[117] Biau G, Scornet E. A random forest guided tour. Test 2016;25:197–227.
[118] Steinley D. K-means clustering: a half-century synthesis. Br J Math Stat Psychol 2006;59(1):1–34.
[119] Khan K, Rehman SU, Aziz K, Fong S, Sarasvady S. Dbscan: Past, present and future. In: The fifth international conference on the applications of digital information and web technologies. IEEE; 2014, p. 232–8.
[120] Reynolds DA, et al. Gaussian mixture models. Encycl Biom 2009;741(659–663).
[121] Cheng Y. Mean shift, mode seeking, and clustering. IEEE Trans Pattern Anal Mach Intell 1995;17(8):790–9.
[122] Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? J Classification 2014;31:274–95.
[123] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Commun ACM 2020;63(11):139–44.
[124] Yuan Y, Liu S, Zhang J, Zhang Y, Dong C, Lin L. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2018, p. 701–10.
[125] Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, Bharath AA. Generative adversarial networks: An overview. IEEE Signal Process Mag 2018;35(1):53–65.
[126] Chen B, Huang K, Raghupathi S, Chandratreya I, Du Q, Lipson H. Automated discovery of fundamental variables hidden in experimental data. Nat Comput Sci 2022;2(7):433–42.
[127] Majeed T, Kamal MS, Zhou X, Solling T. A review on foam stabilizers for enhanced oil recovery. Energy Fuels 2021;35(7):5594–612.
[128] Fasano M, Morciano M, Bergamasco L, Chiavazzo E, Zampato M, Carminati S, et al. Deep-sea reverse osmosis desalination for energy efficient low salinity enhanced oil recovery. Appl Energy 2021;304:117661.
[129] Føyen T, Brattekås B, Fernø M, Barrabino A, Holt T. Increased co2 storage capacity using co2-foam. Int J Greenh Gas Control 2020;96:103016.
[130] Brakke KA. The surface evolver. Exp Math 1992;1(2):141–65.
[131] Poletaev I, Tokarev MP, Pervunin KS. Bubble patterns recognition using neural networks: Application to the analysis of a two-phase bubbly jet. Int J Multiph Flow 2020;126:103194.
[132] Knüpfer L, Heitkam S. A machine learning approach to determine bubble sizes in foam at a transparent wall. Meas Sci Technol 2022;33(6):067001.
[133] Afanasyev Y, Andrews GT, Deacon C. Measuring soap bubble thickness with color matching. Amer J Phys 2011;79(10):1079–82.
[134] Falciani G, Bergamasco L, Bonke SA, Sen I, Chiavazzo E. A novel concept of photosynthetic soft membranes: a numerical study. Discov Nano 2023;18(1):9.
[135] Kennedy MJ, Conroy MW, Dougherty JA, Otto N, Williams BA, Ananth R, et al. Bubble coarsening dynamics in fluorinated and non-fluorinated firefighting foams. Colloids Surf A 2015;470:268–79.
[136] Choi SH, Ko HS, Sohn DK. Bubble-driven flow enhancement of heat discharge of latent heat thermal energy storage. Energy 2022;244:123168.
[137] Ribezzo A, Falciani G, Bergamasco L, Fasano M, Chiavazzo E. An overview on the use of additives and preparation procedure in phase change materials for thermal energy storage with a focus on long term applications. J Energy Storage 2022;53:105140.
[138] Ribezzo A, Bergamasco L, Morciano M, Fasano M, Mongibello L, Chiavazzo E. Experimental analysis of carbon-based phase change materials composites for a fast numerical design of cold energy storage systems. Appl Therm Eng 2023;120907.
[139] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 2019;378:686–707.
[140] Banerjee C, Nguyen K, Fookes C, Karniadakis G. Physics-informed computer vision: A review and perspectives. 2023, arXiv preprint arXiv:2305.18035.
[141] Yeom S-K, Seegerer P, Lapuschkin S, Binder A, Wiedemann S, Müller K-R, et al. Pruning by explaining: A novel criterion for deep neural network pruning. Pattern Recognit 2021;115:107899.