Machine and deep Learning-Powered analysis of photovoltaic properties in 4-terminal FASnI$_3$/CIGS tandem solar cells

A. Maoucha [a], T. Berghout [b], F. Djeffal [a,*]

[a] LEA, University of Batna2, Batna 05000, Algeria
[b] University of Batna 2, Batna 05000, Algeria

ARTICLE INFO

ABSTRACT

This work presents a comprehensive numerical and machine learning-based analysis of lead-free four-terminal (4 T) FASnI$_3$/CIGS tandem thin-film solar cells. Using SCAPS-1D, we evaluated the photovoltaic performance of the top and bottom sub-cells under various material and structural configurations. The FASnI$_3$ top cell and CIGS bottom cell were optimized individually, achieving power conversion efficiencies (PCE) of 18.80 % and 15.47 %, respectively. Machine learning (ML) and deep learning (DL) approaches were employed to identify key performance-influencing parameters. Feature importance analysis revealed that the buffer layer donor density significantly impacts the Jsc and fill factor of the bottom sub-cell, while the top sub-cell's performance is predominantly governed by the electron transport layer and perovskite properties. Despite higher complexity in the top cell's behavior, attributed to environmental variability, the ML/DL framework effectively pinpointed the most critical design factors. These findings contribute to the accelerated development of high-efficiency and sustainable lead-free tandem solar cells.

1. Introduction

The increasing global demand for renewable energy has driven intensive research into next-generation photovoltaic technologies aimed at achieving superior efficiencies and reduced fabrication costs [1–3]. Silicon-based solar cells continue to dominate the photovoltaic industry due to their exceptional efficiency and long-term stability. However, these cells present inherent limitations, including energy-intensive manufacturing processes and structural rigidity, which constrain their applicability in flexible photovoltaic systems [4]. To overcome these challenges, researchers have proposed alternative photovoltaic structures such as thin-film solar cells (TFSCs). By using direct bandgap semiconductors such as Perovskite, CIGS, GaAs, and CdTe, these innovations present a promising, cost-effective alternative to conventional silicon-based solar cells. Leveraging these advancements, tandem solar cells have emerged as a next-generation solution, effectively overcoming the efficiency limitations imposed by the Shockley-Queisser limit on single-junction cells [5]. By depositing thin-film materials with appropriate bandgaps, tandem cells optimize the solar spectrum absorption, minimize thermal losses, and significantly improve power conversion efficiency [6,7]. In addition to enhancing efficiency, tandem solar cell

architectures offer advantages such as lightweight, flexible designs, improved land-use efficiency, and the potential for versatile material combinations, including perovskites with silicon or CIGS [8]. These performances make tandem solar cells as promising candidates for high-efficiency and scalable photovoltaic technologies [9,10]. The four-terminal (4 T) tandem solar cell architecture has recently gained attention as a promising advancement in photovoltaic technology, offering significant advantages over conventional designs. Unlike traditional two-terminal (2 T) tandem cells, the 4 T structure allows independent optimization of the top and bottom sub-cells, resulting in greater flexibility in material combinations and device configurations. This architecture mitigates current-matching constraints, enabling more efficient utilization of the solar spectrum and enhancing overall power conversion efficiency. With these capabilities, 4 T tandem cells have emerged as a next-generation solution for achieving high efficiency and scalability in advanced photovoltaic systems [9,10]. It is chosen because it allows independent optimization of the top and bottom sub-cells without current-matching constraints, leading to greater flexibility in material selection and improved overall efficiency. This architecture also simplifies simulation and analysis, making it ideal for studying novel lead-free tandem combinations like FASnI$_3$/CIGS. Recently,

* Corresponding author.
E-mail address: faycal.djeffal@univ-batna2.dz (F. Djeffal).

Perovskite/CIGS tandem structures, which combine materials from the perovskite and chalcopyrite families, have gained considerable attention due to their potential to achieve high power conversion efficiency (PCE) while ensuring material abundance and low toxicity [11,12]. Developing high-performance tandem solar cells, however, involves numerous challenges. The efficiency of these devices is affected by a complex interplay of design and elaboration factors, including the properties of different layers, such as absorber materials, the carrier transport layer selections, and mechanisms of interfacial recombination [13]. Conventional experimental methods for optimizing these parameters are time-consuming, resource-intensive, and constrained by the expansive parameter space. In this regard, numerical simulations, such as those using Technology Computer-Aided Design (TCAD) tools, have been involved in addressing these challenges by enabling accurate modeling of current–voltage (I–V) characteristics under different conditions [14,15]. Nevertheless, these simulations alone are insufficient to fully elucidate the complex relationships between design parameters and device performance. In recent years, Machine Learning (ML) techniques have demonstrated transformative potential in accelerating the analysis and optimization of complex systems. By uncovering hidden patterns and identifying critical parameters, ML enables researchers to efficiently navigate the multi-dimensional design space of tandem solar cells [16]. When combined with accurate numerical simulations, ML offers a robust and powerful tool for predicting device performance and optimizing the design of high-efficiency solar cells [16,17]. In this context, the present study introduces an innovative methodology that integrates numerical simulations with machine and deep learning techniques to investigate and analyze the critical design parameters of 4 T lead free FASnI3/CIGS tandem solar cells. Accurate numerical simulations, incorporating key physical mechanisms and dissimilar parameter variations across different layers of top (FASnI$_3$) and bottom (CIGS) sub-cells, such as absorber layer thickness, absorber acceptor density, transport layer thickness, transport layer donor and acceptor densities, transport layer materials and optical window properties, are employed to generate a robust dataset representing the I–V characteristics of the tandem device. Subsequently, various ML algorithms are utilized to evaluate the influence of design parameters on key Figures of Merit (FoMs), including power conversion efficiency (PCE), fill factor (FF), open-circuit voltage (Voc), and short-circuit current density (Jsc) for each sub-cell. Furthermore, the proposed methodology leverages deep learning classification and feature importance evaluation to identify and rank the most impactful design parameters and sub-cells. This comprehensive analysis not only highlights the role of high-efficiency absorber and carrier transport layer materials but also provides insights into the underlying mechanisms driving energy conversion performance. By bridging the gap between accurate numerical modeling and Machine-Deep Learning analysis, the proposed approach contributes to the development of efficient tandem TFSCs.

Despite significant progress in tandem solar cell research, most existing studies have focused on either experimental fabrication or numerical modeling of two-terminal architectures, with limited efforts devoted to 4 T lead-free tandem configurations. In particular, there is a noticeable lack of studies that combine physics-based simulations with ML and DL methods to analyze and optimize 4 T tandem devices based on eco-friendly materials such as FASnI$_3$ and CIGS. While some ML-based investigations exist for single-junction cells or monolithic tandems, comprehensive ML/DL analyses that capture the distinct contributions and parameter sensitivities of independently optimized sub-cells in a 4 T structure remain unexplored. This study fills that gap by integrating SCAPS-1D numerical simulations with advanced ML/DL techniques to investigate and optimize the performance of lead-free FASnI$_3$/ CIGS 4 T tandem solar cells.

In this study, SCAPS-1D was employed to simulate a wide range of material and structural configurations for both the top and bottom sub-cells. The simulation outputs, including parameters such as open-circuit voltage (Voc), short-circuit current density (Jsc), fill factor (FF), and power conversion efficiency (PCE), were compiled into structured datasets. These datasets were then used as input for machine learning and deep learning models. This integration enables a comprehensive exploration of complex parameter interactions and allows for the identification of features that most significantly influence device performance. Although Long Short-Term Memory (LSTM) networks are typically applied to sequential or time-series data, in this work LSTM was chosen due to its ability to learn complex, non-linear dependencies among input variables. Its use is particularly beneficial for handling large datasets with high dimensionality and overlapping class features, which are common in simulation-based photovoltaic modeling.

2. Device designing and methodology

2.1. Tandem SC structure and materials

Tandem solar cells are capable of reaching significantly higher efficiencies than their single-junction counterparts. In this study, we propose a 4 T tandem solar cell design using an FASnI3 perovskite as the top cell and CIGS as the bottom cell to capture the solar spectrum more effectively. The lead-free perovskite top cell focuses on converting short-wavelength sunlight into photocurrent, while the longer-wavelength photons pass through to the CIGS bottom cell, ensuring efficient use of the entire spectrum. The 3D structure shown in Fig. 1 highlights the design of the proposed device, featuring a lead-free perovskite top sub-cell with carefully selected layers to ensure efficient charge transport and light absorption. For the electron transport layer (ETL), materials like Zn(S,O), ZnMgO, and C60 were explored to evaluate their impact on device performance. Meanwhile, the hole transport layer (HTL) employs materials such as PEDOT: PSS and Cu$_2$O to facilitate effective hole transfer. FASnI$_3$ is used as eco-friendly absorber layer. The latter plays a crucial role in capturing sunlight and generating electron-hole pairs. The top electrode, made of transparent conductive materials (AZO, ITO,..). As illustrated in this figure, incident sunlight excites the perovskite absorber layer, inducing photon absorption and the generation of electron-hole pairs. The electron transport layer (ETL) and hole transport layer (HTL) effectively facilitate the selective extraction and transport of charge carriers, thereby driving the generation of photocurrent. This innovative design not only optimizes the performance of the perovskite solar cell but also ensures a lead-free composition, making it an eco-friendly alternative. The design of the proposed bottom cell features a thin-film structure built on soda lime glass (SLG), a popular choice for substrates due to its affordability and the beneficial presence of sodium and potassium. This multi-layered structure is designed to optimize performance and ensure efficient operation. A molybdenum back contact was employed as a conductive layer, facilitating efficient hole transport beneath the absorber. Furthermore, the CIGS absorber layer is crucial for photon absorption, effectively generating electron-hole pairs upon sunlight exposure, which is essential for the photovoltaic process. To create the heterojunction structure and improve device efficiency, a buffer layer composed of an n-type semiconductor (such as CdS, WS2, WS3, ZnMgO, ZnS, or ZnSe) was utilized. The structure is completed with a front contact layer of aluminum-doped ZnO (AZO), a transparent, conductive n-type material that enables light transmission to the CIGS absorber while facilitating efficient electron transport. This structure, as illustrated in Fig. 1, is carefully engineered to maximize light absorption and charge carrier dynamics, contributing to the overall efficiency of the 4 T tandem TFSC (See Fig. 1(a)). Besides, it also gives insights about the spectral of light filtering (See Fig. 1(b)).

2.2. Numerical modeling

The device configurations were simulated using SCAPS-1D, a powerful tool developed by Prof. Marc Burgelman and his team at Ghent University [18,19]. SCAPS-1D is a versatile software designed for modeling a wide range of solar cell structures, providing precise
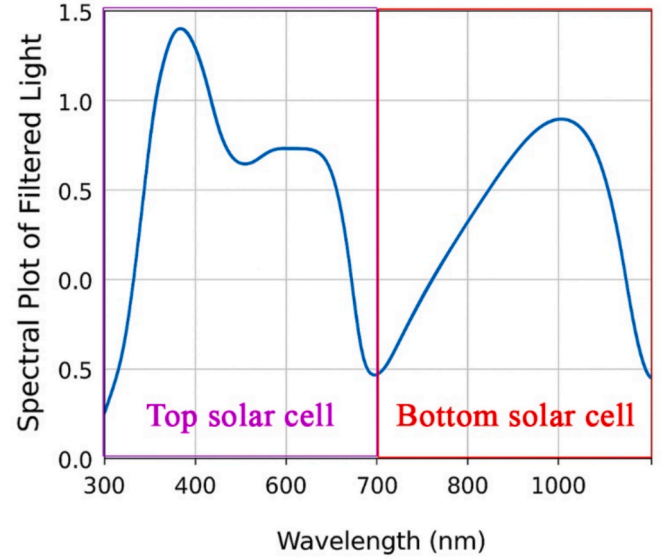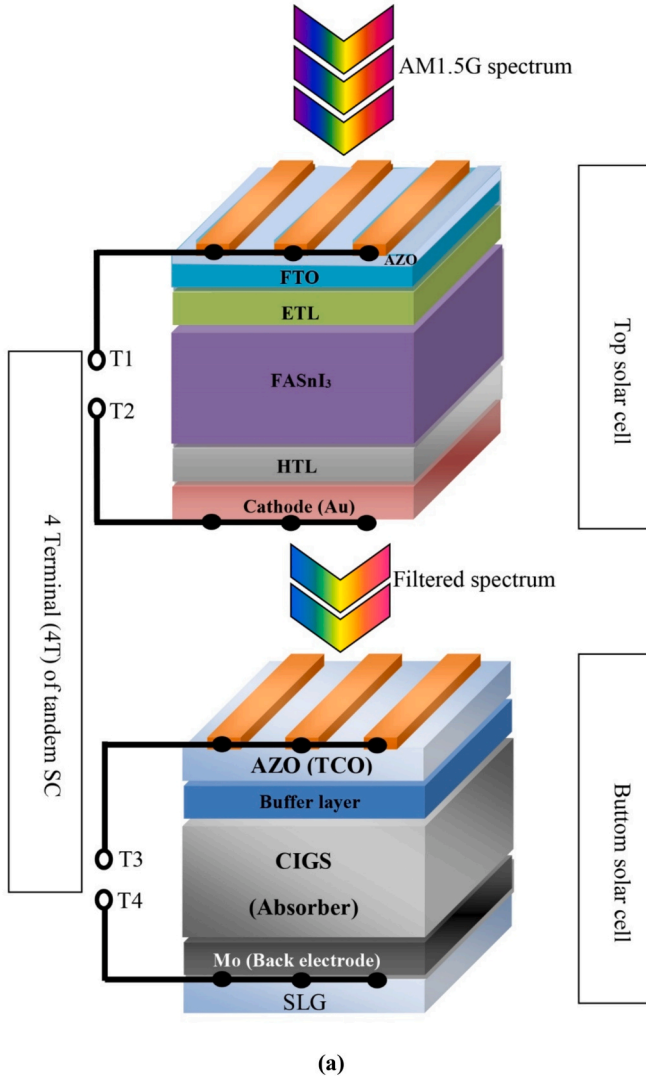
Fig. 1. (a) 3D structure of 4 T Perovskite/CIGS tandem solar cell; (b) Spectral plot of filtered light for top and bottom sub-cells.

predictions of their performance under various conditions. It allows researchers to evaluate factors such as bulk and interface defects, optical absorption properties, carrier densities, and bandgap variations. The software calculates key performance metrics, including open-circuit voltage, photogenerated current density, fill factor, and power conversion efficiency, by solving the governing numerical equations that describe the device's physical behavior. These include Poisson's equation, the charge carrier continuity equations, and the drift–diffusion equation [14,20]:

$$\frac{dp_n}{dt} = G_p - \frac{p_n - p_{n0}}{\tau_p} - p_n u_p \frac{dE}{dx} - u_p E \frac{dp_n}{d_x} + D_p \frac{d^2 p_n}{dx^2} \tag{1}$$

$$\frac{dn_p}{dt} = G_n - \frac{n_p - n_{p0}}{\tau_n} - n_p u_n \frac{dE}{dx} - u_n E \frac{dn_p}{d_x} + D_n \frac{d^2 n_p}{dx^2} \tag{2}$$

$$\frac{d}{dx}\left(\varepsilon(x)\frac{d\phi}{dx}\right) = q[p(x) - n(x) + N_{d+}(x) - N_{a-}(x) + p_t(x) - n_t(x)] \tag{3}$$

where ε represents the dielectric permittivity, q is the electron charge, and G denotes the generation rate. The diffusion coefficient is symbolized by D, while ϕ and E stand for the electrostatic potential and electric field, respectively. The free carriers, holes $p(x)$ and electrons $n(x)$, along with their trapped counterpart's $p_t(x)$ and $n_t(x)$, describe the carrier dynamics. Ionized doping concentrations are represented by N_{d+} for

donors and N_a for acceptors, while x corresponds to the device thickness.

This study uses a 4-terminal (4 T) tandem structure, where the top and bottom sub-cells operate independently, unlike monolithic (2 T) configurations that require current matching. Therefore, electrical integration between sub-cells is not applicable here. The FTO layer's optical absorption was included in the filtered spectrum (Equation (4), and its electrical properties were defined in SCAPS-1D. This has been clarified in the revised manuscript. The critical material parameters used for simulating FASnI3 top sub-cell and CIGS bottom sub-cell in SCAPS-1D are summarized in Table 1. These values were selected based on relevant literature and validated sources to ensure accurate numerical modeling. Where applicable, SCAPS default values were used, and all parameters were chosen to reflect realistic device behavior and fabrication conditions. It is noteworthy, that SCAPS-1D simulates each sub-cell independently and may not fully capture the complex interactions in tandem cells. However, this approach is commonly used as an effective approximation to analyze the performance of tandem devices. We have clarified this limitation in the revised manuscript.

Simulations were conducted at an operating temperature of 300 K, with illumination applied through the front contact under the Air Mass 1.5 global solar spectrum, providing a light intensity of one sun (1000 W/m²). To evaluate the accuracy of the simulation, the current–voltage (I-V) characteristics generated using SCAPS-1D were compared to experimental data [21,22], as shown in Fig. 2. The results demonstrate a

Table 1
Key material parameters used in SCAPS-1D simulations for the FASnI₃ top sub-cell and CIGS bottom sub-cell.

Parameter	Top Sub-Cell (FASnI$_3$)	Bottom Sub-Cell (CIGS)	Unit	Reference
Bandgap (Eg)	1.4	1.15	eV	[21,22]
Electron affinity (χ)	4.1	4.5	eV	[21,22]
Relative permittivity (er)	25	13.6	–	[23,24]
Effective density of states, CB (Nc)	2.2×10^{18}	2.2×10^{18}	cm^{-3}	[23]
Effective density of states, VB (Nv)	1.8×10^{19}	1.8×10^{19}	cm^{-3}	[23]
Electron mobility (μ_e)	10	100	cm^2/V·s	[24,25]
Hole mobility (μ_h)	5	25	cm^2/V·s	[24,25]
Absorber thickness	500–800	1000–1600	nm	This work
Acceptor doping density (Na)	1×10^{17}	$1 \times 10^{16} - 1 \times 10^{18}$	cm^{-3}	[23,26]
Defect density (N_t)	1×10^{15}	1×10^{15}	cm^{-3}	[26]
Recombination model	SRH	SRH	–	SCAPS default
Thermal velocity	1×10^7	1×10^7	cm/s	SCAPS default
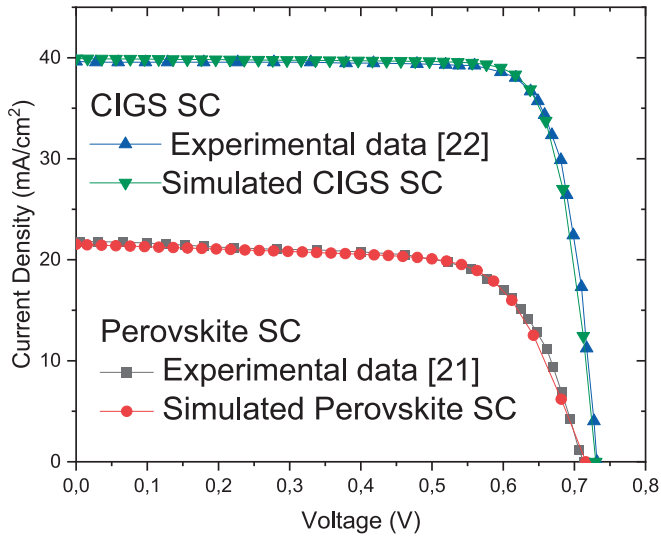
Fig. 2. I-V characteristics of experimental data and simulation results associated with both the conventional Perovskite and CIGS solar cells.

strong alignment between the simulated and experimental curves, highlighting the reliability of the simulation model. The input parameters for these simulations were based on experimental values reported in Refs. [21,22], including layer thicknesses, doping concentrations, and material selections, to ensure a meaningful and representative comparison. These findings underscore SCAPS-1D's ability to precisely capture the electrical and optical properties of solar cells, with the simulated I-V curves aligning closely with experimental observations. The excellent overlap between the two datasets, as illustrated in Fig. 2, reinforces the robustness and accuracy of the modeling approach.

To accurately simulate the proposed tandem structure, the Perovskite/CIGS double-junction solar cell can be represented as two single-junction solar cells with distinct bandgap energies working in tandem. The top cell is made of lead-free perovskite, while the bottom cell uses a CIGS absorber material. When illuminated with the standard AM1.5 G spectrum, the top cell captures and converts the shorter-wavelength light, while the remaining unabsorbed light is transmitted to the bottom cell. This transmitted light mimics the optical behavior of the top layer. To model this radiation splitting, an optical transmission filter is

introduced, with its behavior described by the following equation [23,24]:

$$S(\lambda) = S_0(\lambda) \cdot e^{(-\alpha AZO.dAZO)} \cdot e^{(-\alpha FTO.dFTO)} \cdot e^{(-\alpha ETL.dETL)} \cdot e^{(\alpha perov.dperov)} \cdot e^{(-\alpha HTL.dHTL)}$$

(4)

where α, $S_0(\lambda)$, d represent the absorption coefficient of each material, the global AM1.5G spectrum, and the thickness of different layers respectively.

This filtered spectrum was numerically calculated in MATLAB using Equation (4), based on the optical properties and thicknesses of each top sub-cell layer. The resulting attenuated spectrum was then exported and directly used as an external spectral input in SCAPS-1D for simulating the bottom sub-cell under realistic illumination conditions. It is worth noting that the gold (Au) contact in the top sub-cell is implemented in an interdigitated or grid configuration, allowing a significant portion of light to pass through. As a result, the reflective effects from this contact are minimal and have a negligible impact on the overall light absorption in the top cell. Therefore, reflection from the Au layer was not included in the attenuation model of Equation (4).

Fig. 3 shows the J–V curves simulated for the top cell under the AM1.5G spectrum and for the bottom cell under both the full and filtered spectra to reflect the light reaching each sub-cell.

The 4 T tandem PCE was calculated as the sum of the individually optimized PCEs of the FASnI₃ (top) and CIGS (bottom) sub-cells, each simulated under standard AM1.5G illumination. To account for optical losses, the transmitted spectrum from the top cell was used as the input for the bottom cell. Electrical losses were neglected due to the electrically independent nature of the 4 T configuration.

The overall efficiency of 4 T Perovskite/CIGStandem solar cells is largely influenced by key design parameters. To gain deeper insights, we employ a machine learning-driven approach, leveraging a comprehensive dataset specifically curated for this analysis. This method enables a systematic exploration of how design variations impact device performance.

3. Joint machine learning and deep learning for High-efficient 4 T Perovskite/CIGS tandem solar cell investigation

This section is dedicated to conducting the proposed machine learning and deep learning analyses in this work. The objective of the machine learning analysis is to determine the feature importance in relation to the recorded figures of merit (FoMs) for each dataset. The
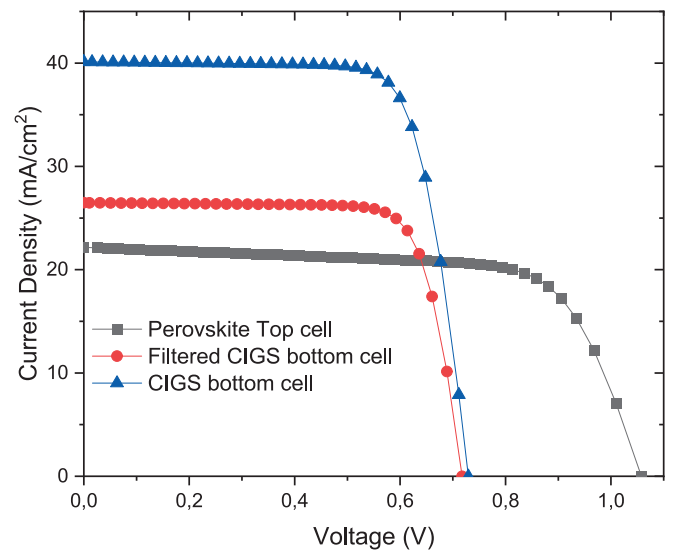
Fig. 3. I-V characteristics of the Perovskite Top cell, CIGS bottom cell and filtered CIGS bottom cell.

second objective involves using deep learning for the classification of observations based on efficiency values into three main categories: good, worst, and extremely worst. The primary aim of both analyses is to enhance the understanding of feature significance and improve the classification accuracy of efficiency levels. The flow diagram Fig. 4 provides clearer insights into the methodology followed in this section by introducing accurate connections between different studies and sub-studies, ultimately leading to important conclusions. This flow diagram will be discussed in the next subsection for further clarification.

3.1. Dataset description

In this work, the generated dataset for perovskite solar cells includes two subsets: The bottom sub-cell and the top sub-cell datasets. The bottom sub-cell subset consists of 252 samples, each with 10 features. These features encompass both categorical and numeric variables, such as material types (CdS, WS2, ZnMgO, ZnS, ZnSe) for the buffer layer and numeric variables like shallow acceptor density, layer thicknesses, and solar cell performance indicators (Voc, Jsc, FF, Efficiency). The dataset reflects a wide range of values, with some features showing considerable variation, such as CIGS shallow acceptor density (ranging from [1.0000 x 1014, 1.4000 x 1018] and performance metrics like efficiency (from 1.34 % to 15.48 %). The top sub-cell subset comprises 418 samples with 13 features, which include material categories (Cu$_2$O, PEDOT: PSS for HTL; C60, Zn (O, S), ZnMgO for ETL) and numeric variables like perovskite shallow acceptor density, layer thicknesses, and solar cell performance (Voc, Jsc, FF, Efficiency). Similar to the bottom sub-cell, this dataset shows variability in feature values, with the perovskite thickness ranging from 100 to 1000 nm and Voc varying from 0.5898 V to 1.1493 V. The data from both subsets are utilized for predictive modeling and analysis in the study of solar cell performance, with an emphasis on optimizing the efficiency of perovskite-based devices. For further clarity and reference, Tables 2 and 3 provide a detailed overview of the dataset statistics, highlighting the main feature ranges and categories for both the bottom sub-cell and top sub-cell subsets. These tables clearly showcase the distribution of variables, including both categorical and numeric features, along with their respective ranges and statistical measures. From both tables Table 2 and Table 3 The top and bottom sub-cells of a Perovskite/CIGS tandem solar cell contribute differently to the overall efficiency, especially regarding the key performance indicators of Voc, Jsc, FF, and efficiency. For the bottom sub-cell, the average Voc is 0.6949 V, Jsc is 24.97 mA/cm2, FF is 73.47 %, and efficiency is 12.76

Table 2
Summary table for bottom sub-cell data (252 samples x 10 features).

Variable	Type	Categories	Mean	Std. Deviation	(Min, Max)
Buffer	Categorical	CdS, WS2, ZnMgO, ZnS, ZnSe	–	–	–
CIGS shallow acceptor density (1/cm^3)	Numeric	–	1.16 x 10^{18}	5.09 x 10^{17}	[1.0000 x 10^{14}, 1.4000 x 10^{18}]
p-CIGS thickness (nm)	Numeric	–	1511.3	313.93	[500, 1650]
ZnO thickness (nm)	Numeric	–	73.06	17.18	[10, 80]
Buffer donner density (1/cm^3)	Numeric	–	1.99x 10^{18}	2.83 x 10^{18}	[1.0000 x 10^{14}, 8.0000 x 10^{18}]
Buffer thickness (nm)	Numeric	–	46.19	9.55	[10, 50]
Recorded FoMs:					
Voc (V)	Numeric	–	0.69	0.1052	[0.1910, 1.0253]
Jsc (mA/cm^2)	Numeric	–	24.97	4.83	[1.9228, 28.4220]
FF (%)	Numeric	–	73.47	14.95	[16.0865, 81.9555]
Efficiency (%)	Numeric	–	12.76	3.96	[1.3390, 15.4798]

%. These values indicate that the bottom sub-cell has a moderate performance in terms of voltage and current generation, with some variation in efficiency due to the material properties, such as CIGS thickness and buffer layer characteristics. In contrast, the top sub-cell exhibits a higher average Voc of 0.9029 V, which is indicative of better voltage generation due to the Perovskite material's favorable properties. The Jsc is slightly lower at 22.17 mA/cm2, but the fill factor (FF) is also important, with a mean value of 68.38 %, which affects the overall current extraction efficiency. The efficiency of the top sub-cell is 13.77 %, reflecting the higher voltage generation but slightly lower current
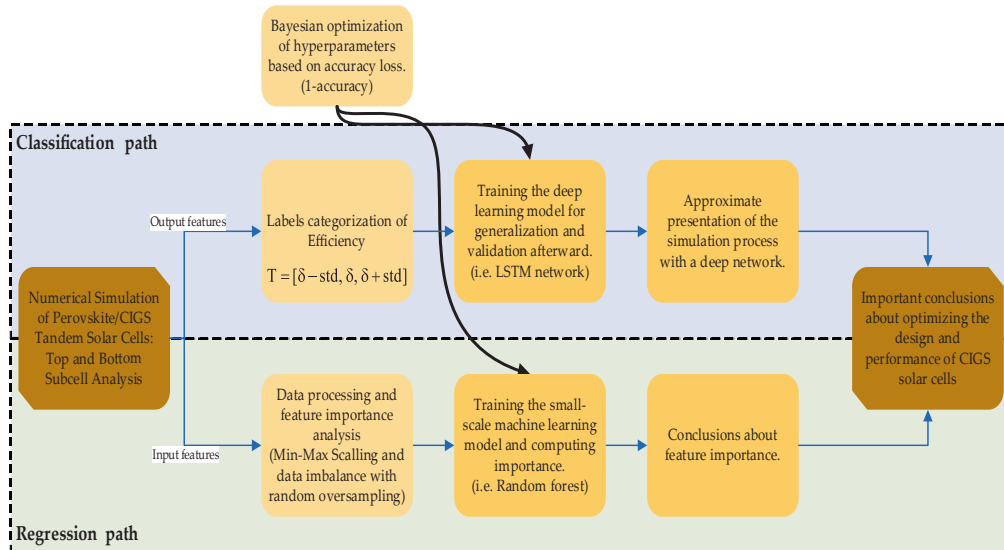
Fig. 4. Flow diagram of joint small-scale machine learning and deep learning for the investigation of high-efficiency 4 T Perovskite/CIGS tandem solar cells under Bayesian optimization.

Table 3
Summary table for top sub-cell data (418 samples x 13 features).

Variable	Type	Categories	Mean	Std. Deviation	(Min, Max)
HTL	Categorical	Cu$_2$O, PEDOT: PSS	–	–	–
ETL	Categorical	C60, Zn (O, S), ZnMgO	–	–	–
Perovskite shallow acceptor density (1/cm^3)	Numeric	–	8.65x 10^{16}	1.53x 10^{17}	[1.0000e + 14, 1.4400 x 10^{18}]
Perovskite thickness (nm)	Numeric	–	293.06	151.53	[100, 1000]
FTO thickness (nm)	Numeric	–	33.61	14.02	[10, 100]
ETL donner density (1/cm^3)	Numeric	–	3.17x 10^{18}	3.24 x 10^{18}	[1.0000 x 10^{14}, 8.0000 x 10^{18}]
ETL thickness (nm)	Numeric	–	50.72	11.06	[10, 100]
HTL accepter density (1/cm^3)	Numeric	–	1.01x 10^{18}	9.09x 10^{17}	[1.0000 x 10^{14}, 8.0000 x 10^{18}]
HTL thickness (nm)	Numeric	–	96.39	14.02	[30, 120]
Recorded FoMs:					
Voc (V)	Numeric	–	0.90	0.1795	[0.5898, 1.1493]
Jsc (mA/cm^2)	Numeric	–	22.17	0.0347	[21.9065, 22.2203]
FF (%)	Numeric	–	68.38	3.68	[47.6065, 74.1546]
Efficiency (%)	Numeric	–	13.77	3.14	[7.7049, 18.8096]

output compared to the bottom sub-cell. Both sub-cells have complementary roles in determining the overall efficiency of the tandem solar cell. While the top sub-cell generates higher voltage, the bottom sub-cell contributes more significantly to current generation, balancing out the overall performance. The combined effect of these parameters from both sub-cells ultimately influences the global efficiency of the tandem solar cell. Optimizing the Voc, Jsc, and FF in both sub-cells is crucial to maximizing the overall efficiency.

3.2. Data processing

Data processing in this work focuses on reducing data complexity in the initial stages and making the dataset coherent for further analysis, particularly for both machine learning and deep learning tasks [25]. To achieve this, a key question is: what steps are necessary to prepare the data effectively? One of the main challenges from a machine learning perspective is the scale of both categorical and numerical variables. The first step is to normalize the data and ensure that categorical variables align with numerical ones for future analysis. Categorical variables are converted into natural numbers (e.g., 1, 2, 3, …) and then normalized using min–max normalization alongside other numerical variables [26]. The second step involves defining an additional categorical variable to account for variations in efficiency values. As previously mentioned, this new categorical variable classifies the data into three categories, good, worst, and extremely worst, based on efficiency variations. Formula (5) defines the thresholds T for categorization, where σ represents the mean and std represents the standard deviation of efficiency values. This

approach ensures that the dataset is appropriately structured for classification tasks and ready for subsequent analysis.

$$T = [\sigma - std, \sigma, \sigma + std] \tag{5}$$

After the processes of normalization, transforming categorical variables, and thresholding, we have obtained the following Fig. 5. Fig. 5 presents a detailed comparison of the t-Stochastic Neighborhood Embedding (t-SNE) visualizations and class proportions for the bottom and top sub-cell datasets [27]. Fig. 5(a, b) shows the t-SNE visualizations for the bottom and top sub-cell datasets, with data points color-coded by efficiency categories. The overlapping classes in most cases make the separation complex, highlighting the challenge of distinguishing the classes in the two-dimensional t-SNE space. Moving to Fig. 5(c), the class proportions for the bottom sub-cell dataset are displayed in a bar chart. The counts for the efficiency categories are as follows: Extremely Worst (28), Good (187), and Worst (37), showing a more imbalanced distribution with a dominant "Good" class. Finally, Fig. 5(d) depicts the class proportions for the top sub-cell dataset, with the following counts: Extremely Worst (142), Good (260), and Worst (14). The "Good" class again dominates, but the "Extremely Worst" category is notably more frequent than in the bottom dataset, while "Worst" remains the least represented. Overall, Fig. 5 provides a comprehensive view of both the t-SNE visualizations and the class distribution, offering insights into the structure and class separation within the two datasets.

Fig. 6 demonstrates the effect of random oversampling applied to overcome the class imbalance observed in Fig. 5. While Fig. 5 likely shows a skewed distribution of classes, with some categories like "Extremely Worst" being underrepresented compared to others, random oversampling, which carried by replicating existing simples in in minority classes, helps balance the dataset by increasing the number of samples in the minority classes. In Fig. 6, both the bottom and top sub-cell datasets are balanced, with each class ("Extremely Worst," "Good," and "Worst") having an equal number of samples, 187 for the Bottom Sub-cell Dataset and 260 for the top sub-cell dataset. This balance ensures that all classes have a fair opportunity to contribute to the training process, preventing bias toward the majority class. However, this approach also introduces increased data complexity. By duplicating or generating synthetic minority class samples, the feature space becomes more crowded, leading to greater overlap between classes. While the class proportions are now equal, the increased overlap and scatter in the feature space make it more difficult for the model to distinguish between classes, thus adding complexity to the data. Therefore, while random oversampling addresses class imbalance and provides more equitable representation of each class, it also increases the challenge of defining clear decision boundaries due to the heightened complexity of the data.

3.3. Feature importance analysis

Figs. 7 and 8 present the feature importance analysis of the input parameters influencing the photovoltaic performance metrics for the bottom and top sub-cells. This analysis was conducted using a Random Forest Regressor optimized with Bayesian Optimization, ensuring accurate estimation of feature contributions and robust model performance [28]. It is worth noting that Bayesian optimization was performed using MATLAB's default settings, which include 30 maximum objective evaluations and the "expected improvement per second" acquisition function. Given the exploratory nature of this study, the default search space was deemed sufficient to tune the LSTM hyperparameters without extensive manual intervention. The analysis of the bottom sub-cell reveals that specific features have a dominant influence on its photovoltaic performance. For Voc, the critical factors are the CIGS shallow acceptor density, with a high importance score of 55.87, and the Buffer, contributing 25.48. Other factors, such as p-CIGS thickness and ZnO thickness, show minimal or even negative importance, indicating their limited or adverse impact on Voc. In the case of Jsc, the Buffer emerges as the most significant contributor, with an
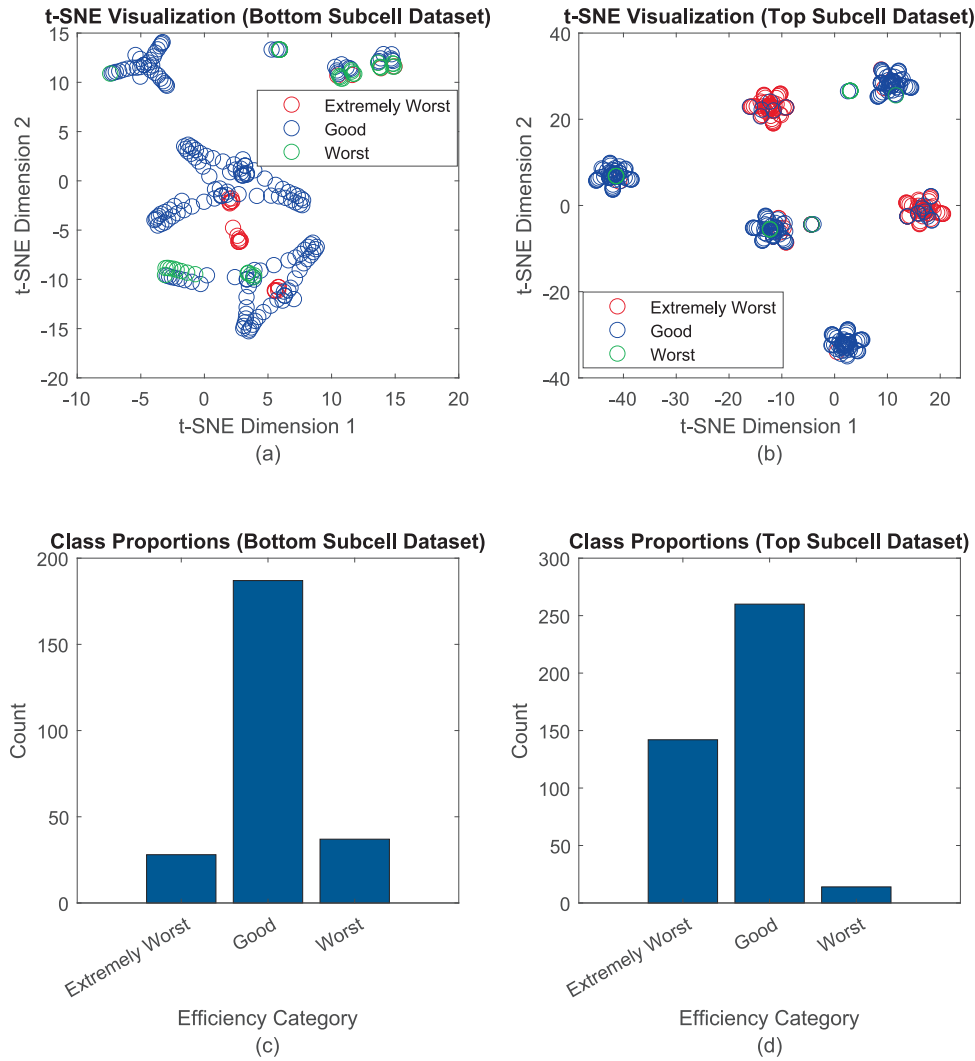
Fig. 5. T-sne visualizations and class proportions of efficiency categories for sub-cell datasets: (a) t-sne visualization for bottom sub-cell dataset; (b) t-sne visualization for top sub-cell dataset; (c) class proportions for bottom sub-cell dataset; (d) class proportions for top sub-cell dataset.

importance score of 79.01, followed by the Buffer donor density, which scores 56.55. Notably, ZnO thickness (−13.92) and CIGS shallow acceptor density (−14.08) have negative importance, suggesting that they adversely affect the short-circuit current density. For FF, both the Buffer (46.24) and Buffer donor density (44.36) are pivotal. Other features, such as p-CIGS thickness and ZnO thickness, contribute marginally or negatively. When examining Efficiency, the Buffer (44.64) and Buffer donor density (35.40) remain the most significant contributors. The CIGS shallow acceptor density provides a moderate positive influence (16.85), while the impact of other features is minor.

For the top sub-cell, the analysis highlights the critical role of the ETL and HTL layers. In terms of Voc, the ETL dominates with a significant importance score of 88.07, while the HTL contributes 5.09. Features such as Perovskite shallow acceptor density and ETL donor density have a smaller impact, and FTO thickness (−0.56) negatively influences Voc. Regarding Jsc, the Perovskite thickness (81.47) is the most important factor, followed by FTO thickness (0.51), indicating their substantial influence. Features such as HTL (4.57) and ETL (2.71) have moderate contributions, while others, like HTL accepter density, have minimal or negative importance. The FF is most strongly affected by the ETL donor density (48.20) and the ETL (36.33). Other features, such as Perovskite shallow acceptor density and HTL, play a secondary role. For Efficiency, the ETL (67.71) and ETL donor density (19.14) stand out as the most influential features. Minor positive contributions come from Perovskite

shallow acceptor density (6.06) and FTO thickness (1.12), while HTL accepter density (−0.23) has a slight negative impact.

The feature importance analysis reveals that for the bottom sub-cell, the Buffer layer and its donor density play a pivotal role, particularly in enhancing Jsc and FF. On the other hand, the top sub-cell's performance is heavily reliant on the ETL and its donor density, which significantly influence both efficiency and FF. Negative feature importance, such as ZnO thickness for Jsc in the bottom sub-cell and FTO thickness for Voc in the top sub-cell, suggests areas where material properties or design could be optimized for better performance. This analysis provides a clear roadmap for targeted improvements in tandem solar cell design.

Physically, the strong influence of buffer layer donor density on Jsc can be attributed to its role in establishing the electric field at the heterojunction interface. A higher donor density enhances the built-in potential, improving charge separation and facilitating the extraction of photogenerated carriers, thereby increasing Jsc. Similarly, the importance of the buffer layer material type reflects its impact on band alignment and recombination rates, which directly affect carrier transport efficiency. Also, it is important to note that negative feature importance values indicate that these features detract from model accuracy when included, typically due to weak or inverse correlations with the target variable or introducing irrelevant variance. In this context, such features may obscure more predictive signals and are effectively considered as having adverse or noisy influence on the model outcome.
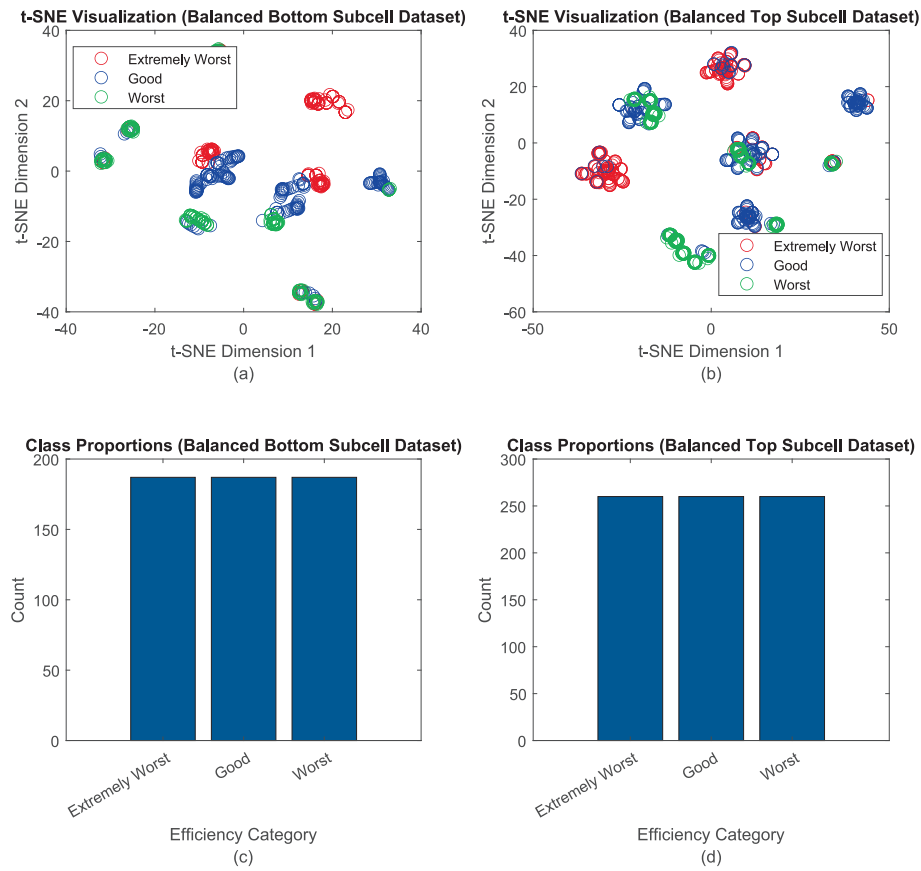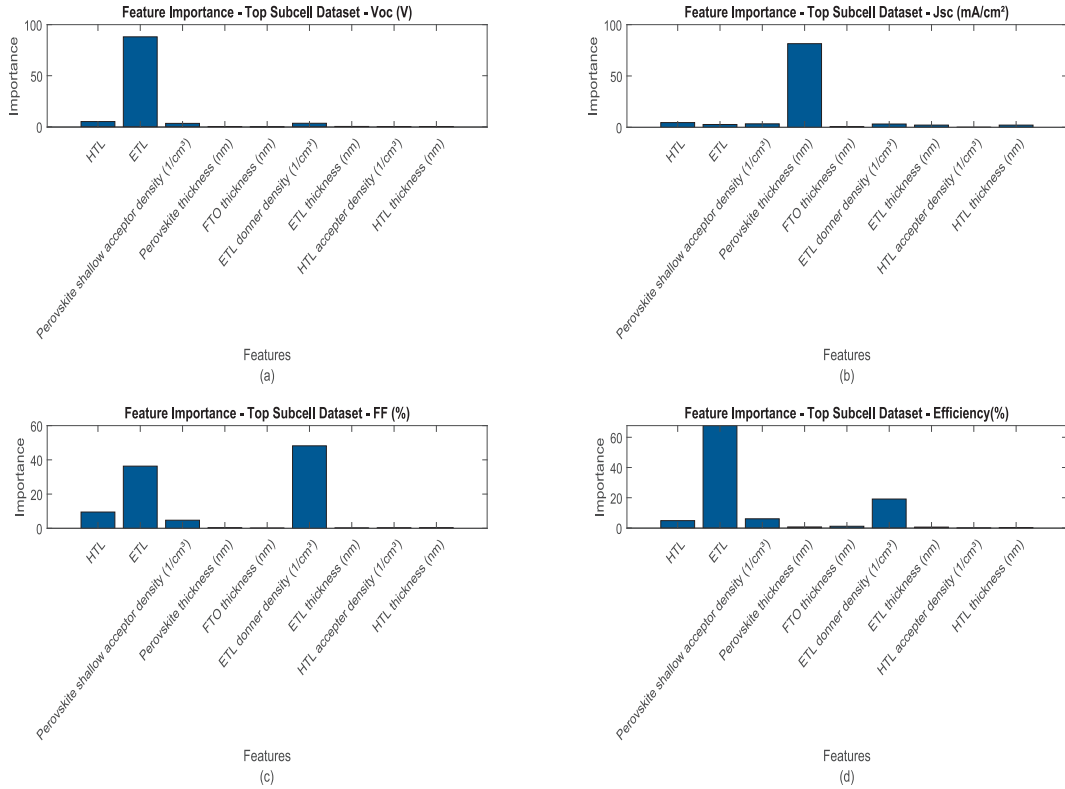
Fig. 6. Effect of random oversampling on class balance and data complexity: (a-b) Increased data complexity and overlap after random oversampling. (c-d) (a) Class proportions in the bottom sub-cell dataset before and after oversampling.
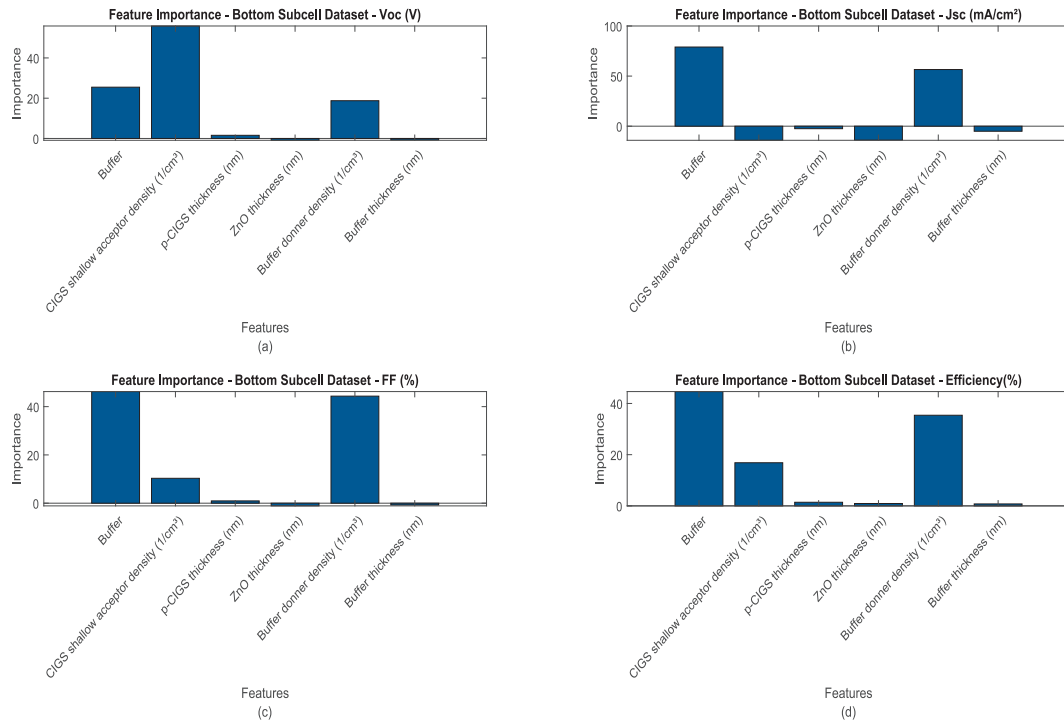
Fig. 7. Feature Importance for Bottom Sub-cell (CIGS-based) Performance Metrics: (a) Voc (V); (b) Jsc (mA/cm^2); (c) FF (%); (d) Efficiency (%).

Fig. 8. Feature Importance for Top Sub-cell (Perovskite-based) Performance Metrics: (a) Voc (V); (b) Jsc (mA/cm^2); (c) FF (%); (d) Efficiency (%).

3.4. Deep learning analysis

For the investigation of high-efficiency 4 T Perovskite/CIGS tandem solar cells, a combination of two sub-cells (top Perovskite and bottom CIGS) is examined to improve energy conversion efficiency. To optimize the performance of these cells, we leverage advanced machine learning algorithms, specifically a Long Short-Term Memory (LSTM) neural network, which excels in modeling temporal dependencies in complex data [28]. In this context, we employ Bayesian optimization to fine-tune hyperparameters of the LSTM model for better classification of material properties and their correlation with efficiency metrics. The optimization process involves defining an objective function that evaluates the LSTM model's performance based on several hyperparameters, such as the number of hidden units, the maximum number of epochs, mini-batch size, learning rates, regularization terms, and dropout rates. Bayesian optimization, which iterates over a range of these hyperparameters, is utilized to find the optimal set of values that minimize classification error [25]. The algorithm focuses on minimizing the test accuracy loss while ensuring efficient convergence through regularization and fine-tuning of activation functions. Once the optimal hyperparameters are determined, the LSTM model is trained, and various performance metrics are computed, including precision, recall, F1 score, and confusion matrices for both training and testing datasets (see equations (6–9)). This approach ensures that the predictive model is both accurate and robust, helping to drive advancements in solar cell technology through a data-driven analysis of material performance. Lastly, it should be noted that the dataset was split into 60 % for training and 40 % for testing, ensuring a balanced representation of the data and enabling a robust evaluation of the model's performance.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{9}$$

The results of this work will be presented in a structured format, beginning with an evaluation of the training loss. The initial step will include monitoring the loss and accuracy metrics during training. The optimal hyperparameters for the LSTM model will be identified through Bayesian optimization, with the objective function focused on minimizing test accuracy loss. Following this, the model will be trained using the optimal parameters, and the predicted outcomes will be compared against the ground truth labels, providing insights into the model's performance. The performance will be assessed based on a variety of metrics, including training and testing accuracy, precision, recall, and F1 score, calculated from confusion matrices. The metrics will be presented for both the training and testing sets, ensuring a comprehensive evaluation of the model's generalization capabilities. Additionally, the results will include training time and the evaluation of the convergence behavior of the model during optimization. Finally, confusion matrices and other key performance indicators will be displayed to visualize how well the model distinguishes between different classes in both the training.

Fig. 9 presents the learning and testing behaviors of the deep learning model for both top and bottom sub-cell datasets. The model's training loss and performance metrics, including Receiver Operator Curve (ROC) curves and Area Under the Curve (AUC) values, are visualized to show how the model converges and how well it generalizes to the data. In Fig. 9(a), the training loss for the bottom sub-cell data is observed to take approximately 450 epochs to converge to a solution with a lower final loss value than the top sub-cell data. The model for the top sub-cell dataset, however, converges much more quickly, within 50–60 epochs, but with a higher final loss compared to the bottom sub-cell data. This discrepancy in convergence rates and final loss values can
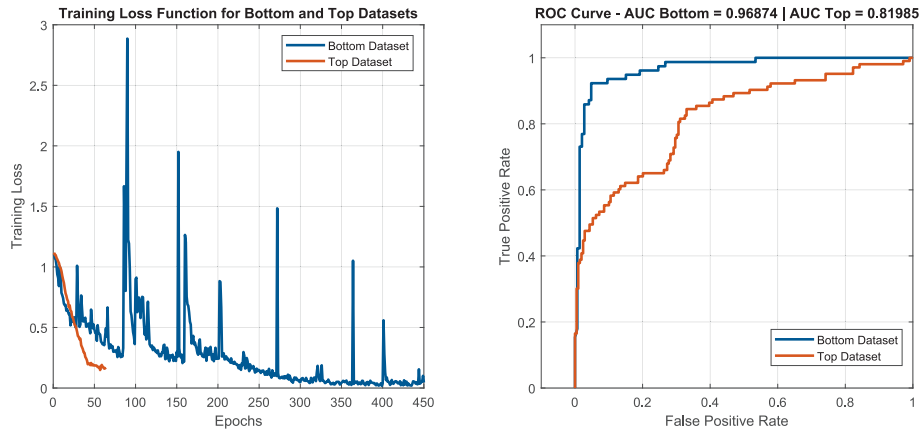
Fig. 9. Training and Testing Performance of the Deep Learning Model for Top and Bottom Sub-cell Data: (a) Loss Function During Training; (b) ROC Curves and AUC Values.

be attributed to the inherent complexity of the data and environmental conditions encountered by each sub-cell type. The bottom sub-cell data likely contains more distinct or easily separable features that allow the model to optimize more slowly and achieve a lower loss. In contrast, the top sub-cell data may be more variable or noisy, requiring faster adaptation by the model to arrive at a solution, but resulting in a higher loss due to greater data complexity or environmental factors. In Fig. 9(b), the ROC curves further highlight the differences between the two sub-cell datasets. The ROC curve for the top sub-cell data covers a smaller area, with an AUC of 0.8198, indicating a relatively weaker performance. This could be due to the more challenging nature of the top sub-cell data, which may include overlapping or less discernible features that make classification more difficult. On the other hand, the bottom sub-cell data exhibits an AUC of 0.9687, suggesting a better model performance, likely due to clearer, more separable features or better environmental conditions for the model's task. The AUC values reflect the complexity of the data and how well the model is able to distinguish between classes under varying conditions.

The increased model complexity of the top sub-cell can be inferred from several performance indicators. As shown in Fig. 9(a), the top sub-cell model converged more rapidly but reached a higher final loss compared to the bottom sub-cell, suggesting a more complex or noisy feature space. Furthermore, the lower AUC in Fig. 9(b) and the overlapping class boundaries in the confusion matrix (Fig. 10b) provide additional evidence of data entanglement. These behaviors point to greater variability in the top sub-cell dataset, likely due to the wider range of material combinations and design parameters, which increases the difficulty of class separation and contributes to model overfitting.

While correlation metrics were not explicitly calculated, these observed trends serve as indirect indicators of higher complexity.

Fig. 10 presents the confusion matrices for the classification performance on two datasets. For the bottom sub-cell dataset (Fig. 10 (a)), the confusion matrix shows that the model performs relatively well but still has some misclassifications. The first row indicates that 75 instances of class 1 were correctly predicted, with 1 misclassified as class 2 and 5 misclassified as class 3. The second row shows that 66 instances of class 2 were correctly predicted, while 4 instances were misclassified as class 1 and 8 as class 3. The third row reveals that 60 instances of class 3 were correctly predicted, with only 2 misclassified as class 1 and 3 as class 2. Overall, the model shows good performance with relatively low misclassification, especially for class 2 and class 3. In the top sub-cell dataset (Fig. 10 (b)), the confusion matrix highlights some different trends. The first row indicates that 34 instances of class 1 were correctly classified, but 3 were misclassified as class 2 and 67 as class 3, showing a significant misclassification into class 3. The second row shows that 47 instances of class 2 were correctly predicted, with 5 misclassified as class 1 and 51 as class 3. The third row shows that 102 instances of class 3 were correctly predicted, with only 3 misclassified as class 2 and no instances misclassified as class 1. This matrix suggests that class 3 is more frequently misidentified as class 1 and class 2, indicating a challenge for the model in distinguishing between these classes. Overall, the confusion matrices provide insights into how the model generalizes to different sub-cell datasets and highlight the areas where it struggles, particularly with misclassifications between certain classes in the top sub-cell dataset.

Table 4 presents the performance metrics and time statistics for both
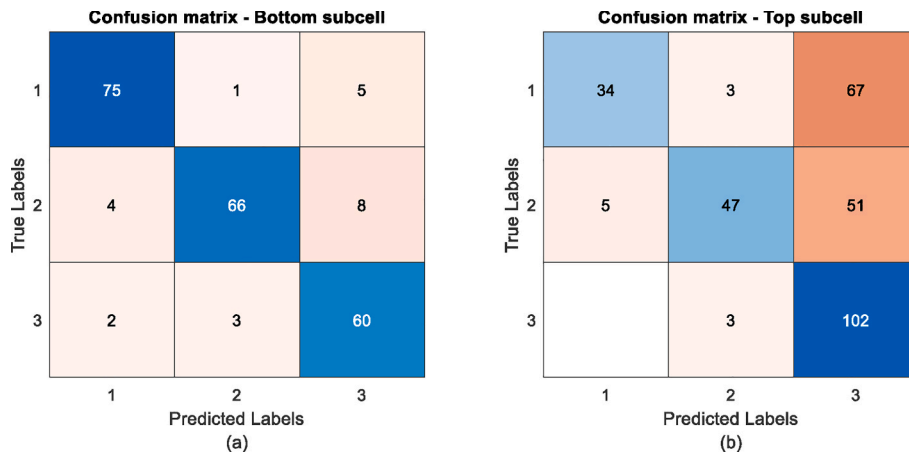
Fig. 10. Confusion matrices for both datasets (generalization): (a) Bottom sub-cell data; (b) Top sub-cell data.

Table 4
Performance metrics and time statistics for bottom and top sub-cell datasets.

Metric	Bottom Sub-cell	Top Sub-cell
Training accuracy	0.95	0.953
Training precision	0.94	0.9545
Training recall	0.93	0.953
Training F1 score	0.94	0.9533
Training time	0.5	0.504
Testing accuracy	0.58	0.5865
Testing precision	0.74	0.7407
Testing recall	0.57	0.5849
Testing F1 score	0.56	0.5686

Table 5
Hyperparameter settings for bottom and top sub-cell datasets.

Metric	Bottom Sub-cell	Top Sub-cell
Sub-cell type	Bottom	Top
Num. hidden units	27	29
Max-epochs	821	64
Mini-batch size	45	106
Initial learning rate	0.074533	0.027822
Gradient threshold	0.68549	0.87681
L2 regularization	0.031804	0.069145
Dropout rate	0.22061	0.46444
Recurrent dropout rate	0.10759	0.074699
State activation function	tanh	softsign
Gate activation function	hard-sigmoid	hard-sigmoid
Sequence length	38	12

the bottom and top sub-cell datasets, providing insights into the model's training and testing performance. In terms of training performance, the model shows high accuracy for both datasets, with 95 % for the bottom sub-cell and 95.3 % for the top sub-cell. Precision and recall are also comparable, with the top sub-cell dataset slightly outperforming the bottom sub-cell dataset in both metrics. Specifically, the precision is 94 % for the bottom sub-cell and 95.45 % for the top, while recall is 93 % for the bottom sub-cell and 95.3 % for the top. This indicates that the model is well-trained on both datasets, with a slightly better balance of precision and recall for the top sub-cell data. The F1 score, which balances precision and recall, is also similar for both datasets, showing values of 94 % for the bottom sub-cell and 95.33 % for the top sub-cell, reflecting strong overall performance. Regarding training time, the model requires approximately 0.5 s for both datasets, with a slight increase to 0.504 s for the top sub-cell dataset, indicating negligible difference in training duration between the two datasets. However, there is a noticeable drop in performance during testing. The testing accuracy is lower than the training accuracy for both datasets, with values of 58 % for the bottom sub-cell and 58.65 % for the top sub-cell, suggesting that the model struggles to generalize to unseen data. Despite high testing precision (74 % for both datasets), the testing recall is lower, at 57 % for the bottom sub-cell and 58.49 % for the top sub-cell, indicating that the model has difficulty identifying all positive instances during testing. The testing F1 score mirrors this trend, with values of 56 % for the bottom sub-cell and 56.86 % for the top sub-cell, showing that while the model maintains some accuracy in precision, recall and overall performance can be further improved. In summary, Table 4 highlights that while the model performs well during training with high accuracy, precision, and recall, it experiences a significant performance drop during testing, particularly in accuracy and F1 score. This suggests that further improvements are needed to enhance the model's ability to generalize effectively to new, unseen data. It is acknowledged that the high training accuracy contrasted with relatively low-test accuracy suggests overfitting. However, the scope of this study is exploratory and diagnostic rather than prescriptive. The objective was to apply and observe ML/DL methods for parameter sensitivity analysis in 4 T tandem cells, not to develop optimized predictive models. This benchmarking lays a foundation for future studies, where model refinement and overfitting mitigation can be strategically addressed as part of advanced design optimization efforts.

Table 5 compares the hyperparameter settings for the bottom and top sub-cell datasets used in the model. Both datasets have similar configurations, with slight differences in key parameters. The bottom sub-cell dataset uses 27 hidden units, 821 epochs, and a mini-batch size of 45, while the top sub-cell has 29 hidden units, 64 epochs, and a larger mini-batch size of 106. The bottom sub-cell also has a higher initial learning rate (0.074533) and a more relaxed gradient threshold (0.68549) compared to the top sub-cell's values of 0.027822 and 0.87681, respectively. For regularization, the bottom sub-cell uses lower L2 regularization (0.031804) and dropout rate (0.22061), while the top sub-cell has higher values (0.069145 and 0.46444, respectively). Both datasets use the "hard-sigmoid" gate activation function, but the state activation function differs, with "tanh" used for the bottom sub-cell and

"softsign" for the top sub-cell. Lastly, the bottom sub-cell dataset has a longer sequence length (38) compared to the top sub-cell's 12, indicating longer input sequences in the bottom dataset. These variations highlight the different training setups and strategies for each dataset.

Machine learning and deep learning models, particularly LSTM networks, play a crucial role in handling data complexities in photovoltaic research. LSTM networks are particularly effective in addressing sequential dependencies and complexities found in temporal or sequence-based data, which makes them ideal for modeling the intricate relationships between material properties and efficiency metrics in solar cells. However, the complexity of the data from different sub-cells, such as the top Perovskite sub-cell and the bottom CIGS sub-cell, can vary significantly, and this has important implications for model performance. In this work, the LSTM model was employed to analyze the relationship between various material properties and efficiency metrics across two sub-cell types, but the complexities inherent in the data differed between the two. The bottom sub-cell dataset demonstrated a more distinct feature space, with clearer separability between efficiency categories. This allowed the LSTM model to converge more slowly but ultimately achieve a lower final loss, as it was able to more easily learn from the data with less noise. Conversely, the top sub-cell dataset exhibited greater data complexity, characterized by more noise and feature overlap, making the task of distinguishing between efficiency categories more difficult. This increased complexity in the top sub-cell dataset meant that the LSTM model required faster adaptation to the data, leading to quicker convergence but a higher final loss compared to the bottom sub-cell. These differences in data complexity between the top and bottom sub-cells highlight the challenges that machine learning models, especially LSTM networks, face when dealing with heterogeneous datasets. For the top sub-cell, the model had to account for more variability and noise, which made the learning process more challenging and less efficient. This is reflected in the slower optimization and the higher final loss value for the top sub-cell, indicating that the model struggled to capture the underlying patterns as effectively as for the bottom sub-cell. The key takeaway is that while LSTM networks can effectively model complex relationships within data, the inherent complexity and noise in datasets, such as those from different sub-cells, require careful consideration when applying deep learning models. Data preprocessing, such as normalization and random oversampling, along with the fine-tuning of model hyperparameters, are essential steps in reducing data complexity and ensuring that the LSTM model can successfully learn and generalize from the data, regardless of its inherent challenges.

4. Conclusion

This research emphasizes the potential of 4 T Perovskite/CIGS tandem solar cells as a sustainable and efficient solution for next-generation photovoltaics. To address the challenges involved in studying these devices, SCAPS-1D numerical simulations were integrated with machine and deep learning to analyze the impact of design parameters on solar

cell performance. The study validates SCAPS-1D simulations through close alignment with experimental data and demonstrates how data-driven methods can enhance key metrics such as PCE, Voc, and Jsc. The optimized results indicate that the bottom sub-cell achieves an efficiency of 15.47 %, while the top sub-cell reaches 18.80 %, reflecting their complementary roles in total energy conversion. Machine learning analysis identified the buffer layer and its donor density as the most influential features for improving Jsc and FF in the bottom sub-cell, whereas ETL and its donor density were most critical in the top sub-cell. Deep learning classification models achieved training accuracy above 95 % for both sub-cells, although test accuracy was moderate (∼58 %), indicating model limitations due to data complexity and potential overfitting. These findings confirm the viability of using ML/DL tools for early-stage analysis and benchmarking. The proposed approach provides a foundational framework for guiding future fabrication and performance optimization of lead-free 4 T tandem thin-film solar cells. Future work will focus on refining ML/DL models for improved generalization, integrating larger and more diverse datasets, and extending the methodology to real-world fabrication and experimental validation. Besides, in future studies, incorporating SCAPS-based physical diagnostics such as capacitance–voltage, quantum efficiency, and recombination analysis will provide deeper insight into the underlying mechanisms and further validate the machine learning outcomes. Such forward-looking efforts will bridge the gap between numerical modeling and experimental realization, ultimately accelerating the commercialization of high.

CRediT authorship contribution statement

A. Maoucha: Writing – original draft, Software, Investigation, Formal analysis. **T. Berghout:** Writing – original draft, Software, Investigation, Conceptualization. **F. Djeffal:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

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

No data was used for the research described in the article.

References

[1] P.K. Nayak, S. Mahesh, H.J. Snaith, D. Cahen, Photovoltaic solar cell technologies: analysing the state of the art, Nature Review Materials 4 (2019) 269–285.

[2] V. Tyag, et al., Progress in Solar PV technology: Research and achievement, Renewable and Sustainable Reviews 20 (2013) 443–461.

[3] M.A. Green, et al., Solar cell efficiency tables (Version 60), Prog. Photovolt. 30 (2022) 687–701.

[4] K. Yoshikawa, et al., Silicon heterojunction solar cell with interdigitated back contacts for a photoconversion efficiency over 26%, Nat. Energy 2 (2017) (2017) 17032.

[5] W. Shockley, H.J. Queisser, Detailed balance limit of efficiency of p-n junction solar cells, J. Appl. Phys. 32 (1961) 510–519.

[6] A. Maoucha, F. Djeffal, H. Ferhati, H. AbdelMalek, Eco-friendly perovskite/CZTSSe tandem cell exceeding 28% efficiency through current matching and bandgap optimization: a numerical investigation, The European Physical Journal plus 138 (2023) 620.

[7] A. Maoucha, F. Djeffal, H. Ferhati, SCAPS-FDTD simulation of 20.1 % efficient Perovskite-SnS tandem solar cell based on alternative charge transport layers and Au-nanoparticles, Phys. Scr. 99 (2023) 015919.

[8] M.F.M. Noh, et al., Perovskite/CIGS tandem solar cells: progressive advances from technical perspectives, Mater. Today Energy 39 (2023) 101473.

[9] M. Yu, A. Los, G. Xiong, Thin film absorbers for tandem solar cells: an industrial perspective, JPhys Energy 5 (4) (2023) 042002.

[10] S. Ašmontas, M. Mujahid, Recent Progress in Perovskite Tandem Solar Cells, Nanomaterials 13 (2023) 1886.

[11] N. Shrivastav, et al., Perovskite-CIGS Monolithic Tandem Solar Cells with 29.7% Efficiency: a Numerical Study, Energy Fuel 37 (2023) 3083–3090.

[12] M.H. Azar, et al., SCAPS Empowered Machine Learning Modelling of Perovskite Solar Cells: Predictive Design of Active Layer and Hole Transport Materials, Photonics 10 (2023) 271.

[13] A. Maoucha, T. Berghout, F. Djeffal, H. Ferhati, Machine learning-assisted investigation of CIGS thin-film solar cell degradation using deep learning analysis, J. Phys. Chem. Solid 199 (Apr. 2025) 112526.

[14] M. Burgelman, K. Decock, A. Niemegeers, J. Verschraegen, and S. Degrave, "SCAPS manual." 2024.

[15] Atlas User's manual, SILVACO TCAD, 2012.

[16] L. Greenstein, G.R. Hutchison, Screening Efficient Tandem Organic Solar Cells with Machine Learning and Genetic Algorithms, J. Phys. Chem. C 127 (2023) 6179–6191.

[17] H. Q. Tan, X. Zhao, Akhil Ambardekar, E. Birgersson, and H. Xue, "Exploring the optimal design space of transparent perovskite solar cells for four-terminal tandem applications through Pareto front optimization," APL Machine Learning, vol. 2, pp 026111, 2024.

[18] A. Maoucha, F. Djeffal, and H. Ferhati, "Numerical Investigation of a New Double-Absorber Lead-free Perovskite Solar Cell via SCAPS-1D," 2023 20th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE), Oct. 2023. DOI: 10.1109/CCE60043.2023.10332813.

[19] A. Maoucha, F. Djeffal, T. Berghout, H. Ferhati, Photoresponsivity Enhancement of SnS-Based Devices using Machine Learning and SCAPS Simulations, Engineering Proceedings 58 (2023) 128.

[20] M. Burgelman, P. Nollet, S. Degrave, Modelling polycrystalline semiconductor solar cells, Thin Solid Films 361–362 (2000) 527–532.

[21] T. Wu, et al., Heterogeneous FASnI$_3$ Absorber with Enhanced Electric Field for High-Performance Lead-Free Perovskite Solar Cells, Nano-Micro Letters 14 (2022) 99.

[22] J. Hwang, et al., Improved carrier transport in CIGS solar cells induced by Ag treatment, J. Alloy. Compd. 886 (2021) 161193.

[23] K. Kim, et al., Simulations of chalcopyrite/c-Si tandem cells using SCAPS-1D, Sol. Energy 145 (Jan. 2017) 52–58.

[24] D. Kumari, S.K. Pandey, Effect of an ultra-thin 2D transport layer on eco-friendly Perovskite/CIGS tandem solar cell: a numerical study, Micro and Nanostructures 170 (2022) 207398.

[25] T. Berghout, M. Benbouzid, UBO-EREX: uncertainty Bayesian-Optimized Extreme Recurrent EXpansion for Degradation Assessment of Wind Turbine Bearings, Electron 13 (2024) 2419, https://doi.org/10.3390/electronics13122419.

[26] F. Kadri, et al., Neural classification method in fault detection and diagnosis for voltage source inverter in variable speed drive with induction motor, in: IEEE Eighth International Conference and Exhibition on Ecological Vehicles and Renewable Energies (EVER), 2013, pp. 1–5, https://doi.org/10.1109/EVER.2013.6521549.

[27] T. Berghout, Joint image Processing with Learning-Driven Data Representation and Model Behavior for Non-Intrusive Anemia Diagnosis in Pediatric patients, J. Imaging 10 (2024) 245, https://doi.org/10.3390/jimaging10100245.

[28] T. Berghout, E. Bechhoefer, F. Djeffal, W.H. Lim, Integrating Learning-Driven Model Behavior and Data Representation for Enhanced remaining Useful Life Prediction in Rotating Machinery, Machines 12 (2024) 729, https://doi.org/10.3390/machines12100729.