El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Classification of Cell Line Halm Machine Data in Solar Energy Panel Production Factories Using Artificial Intelligence Models

Murat Simsek[1a], Ozel Sebetci[2b], Irfan Yilmaz[3c]
[1]Artificial Intelligence Engineering Department, Ostim Technical University, Ankara, Türkiye
[2]Department of Computer Technologies, Aydin Adnan Menderes University, Aydin, Türkiye
[3] Software Engineering Department, Ostim Technical University, Ankara, Türkiye

osebetci@adu.edu.tr

DOI : 10.31202/ecjse.1580430
Received: 06.11.2024 Accepted: 09.01.2025
How to cite this article:
Murat Simsek, Ozel Sebetci, Irfan Yilmaz, " Classification of Cell Line Halm Machine Data in Solar Energy Panel Production Factories Using Artificial Intelligence Models", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 1, (2025), pp.(44-53).
ORCID: [a]0000-0002-8648-3693; [b]0000-0002-2996-0270, [c]0009-0007-6168-4580.

Abstract : This study presents a quality estimation method for photovoltaic cells in solar panels using advanced machine learning techniques, including traditional methods and convolutional neural networks (CNNs). Photovoltaic cells, which are primarily composed of crystalline silicon, are of critical importance for the conversion of sunlight into electrical energy. The efficiency of these cells directly affects the performance and lifespan of solar panels. The study's objective is to assess the electroluminescence values of cells using the HALM device, which is capable of measuring the key parameters that determine cell quality. To enhance the performance of the CNN model, hyperparameter tuning and optimization techniques were employed to improve the accuracy of visual evaluation and classification. The proposed method offers significant advantages, including the optimization of the cell production process, a reduction in costs, and an improvement in operational efficiency through the minimization of discrepancies between human and machine decisions. Furthermore, this approach facilitates real-time monitoring and dynamic management of production processes by integrating machine learning models with production line databases. The findings indicate the potential of artificial intelligence to enhance the detection and classification of cell defects, thereby supporting more efficient and high-quality solar panel production. The study underscores the importance of AI-driven methods in advancing production technologies and improving the sustainability of solar energy systems.

Keywords : Artificial Intelligence, Machine Learning, Solar Energy, Pv-Photovoltaics, Energy Quality.

1 Introduction

The global and national focus on green energy has intensified significantly, with the goal of achieving sustainable energy solutions by 2050. In this context, various installations and production investments have been initiated, with solar energy emerging as a cornerstone of these efforts. As of the first quarter of 2024, solar energy capacity in Turkey reached an impressive $12,000MW$, showcasing the rapid development and prioritization of this sector. Concurrently, industrial investments have also increased, with a particular focus on advancing cell and panel technologies, which play a critical role in enhancing energy efficiency in the sector.

With the growing population and advancing technological infrastructure, meeting the increasing energy demands of households has become essential for ensuring a comfortable lifestyle. Additionally, cumulative energy consumption rates have surged due to the parallel growth of production facility investments. The installation of photovoltaic (PV) electricity generation plants has become crucial in addressing these energy demands. Solar energy production involves a multi-stage process that ensures energy reaches end-users efficiently. This process includes the following four key stages: Silicon production: Silicon, the fundamental material for solar panels, is derived from sand or quartz through metallurgical and chemical purification methods. High-purity silicon is essential for the subsequent production stages. Ingot production: Using the Czochralski method, high-purity silicon is transformed into large cylindrical crystals with a head and tail. Alternatively, polycrystalline ingots are produced by cooling molten silicon in a crucible. Wafer production: The ingots are processed by cutting their head and tail sections using specialized machinery. The remaining cylinder is shaped into a rectangular prism. The edges are trimmed, and the surface is smoothed through chamfering to eliminate cutting marks. Thin slices, approximately as thick as a hair strand, are then cut using diamond wire. From a single ingot measuring $700\breve{}800mm$in length, approximately 3,000 wafers can be produced. Cell production: Wafers undergo doping with phosphorus or boron to create a p-n junction, enabling them to generate electricity when exposed to light. An anti-reflective coating is applied to minimize light reflection and maximize light absorption. Metal contacts are added to the front and back surfaces to collect electrical current. The cells are then soldered, coated with various chemicals, and baked to complete the manufacturing process. Following this, the cells are tested using the Halm machine to evaluate their efficiency and parameters. Based on these evaluations, the cells are categorized, labeled, and prepared for module production.

Module production: n this final stage, the cells are arranged in series and parallel configurations on glass substrates. Soldering connects the cells, and a holding material is applied to bind the cells and glass together. The module is laminated to enhance durability. After undergoing quality testing, the modules are labeled and equipped with a frame and j-box cabling, transforming them into ready-to-package solar panels. This study focuses on the application of machine learning in the Halm machine, a critical component in evaluating solar cells during production. By leveraging machine learning algorithms, the system can classify and predict outcomes without requiring explicit programming or manual operator intervention. This approach addresses the challenges of employee turnover, which often leads to disruptions in quality control processes. Machine learning ensures consistency and minimizes errors in interpreting machine data, enabling proactive evaluation and reducing reliance on human intervention. The study employed Decision Trees, Random Forests, K-Nearest Neighbors (KNN), and Linear Regression as primary methods. Balanced data was utilized to ensure reliable outcomes, as each solar cell has defined quality parameters. For instance, unread barcodes are reprocessed without compromising existing data integrity or stability.

The Halm machine is an IV measurement system designed to inspect and classify bifacial solar panel cells on production lines. It features three distinct flash boxes equipped with advanced lighting systems. Initially, the back of the cell is illuminated with $1,000W/m^2$ of irradiation. In the subsequent stage, the cell is adapted to outdoor lighting conditions by simultaneously flashing the front and back of the cell, with the front receiving $1,000W/m^2$and the back $200W/m^2$. Integrating machine learning algorithms into solar energy production processes offers significant advantages, including improved accuracy, reduced dependency on human expertise, and enhanced efficiency in quality control. By implementing technologies such as the Halm machine, the solar energy sector can achieve greater productivity and contribute more effectively to sustainable energy goals. These advancements underscore the importance of combining technological innovation with renewable energy investments to meet the growing global energy demand.

Figure 1 illustrates the irradiance ranges and time intervals used in the Cell IV Measurement System for evaluating bifacial solar cells. The system employs three distinct measurement conditions: rear STC, outdoor illumination condition, and front STC. Initially, the rear of the cell is exposed to $200W/m^2$to $1,000W/m^2$of irradiance for 30 milliseconds, simulating standard rear-side test conditions. Following this, the bifacial condition is tested under outdoor illumination with simultaneous exposure of the front and rear sides for the next 30 milliseconds, where the rear side receives $200W/m^2$, and the front side is exposed to $1,000W/m^2$. Finally, the front STC is measured for the remaining 30 milliseconds, focusing solely on the front side with $1,000W/m^2$ of irradiance. This time-segmented approach ensures comprehensive evaluation of the cell's performance under varying irradiance and directional conditions, providing critical data for classification and efficiency analysis.

Measurement's work integrated with PV Control software and hardware in a control cabinet. The cell measurement system

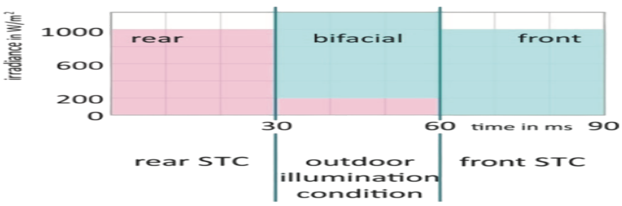

Figure 1: Cell IV Measurement System Ranges

was specially developed for industrial and light industrial use. Advanced hysteresis feature is used here. This is especially designed to be able to classify the next era of innovative cell technologies.Figure 2 depicts the HALM device, a state-of-the-art IV measurement system used for the evaluation and classification of bifacial solar cells in production lines. This device is equipped with advanced hardware and software components, including high-precision measurement units and integrated flash boxes, enabling detailed analysis of electrical parameters and efficiency. The HALM system features a user-friendly interface, as shown on the accompanying monitor, which facilitates real-time monitoring and data visualization. Its modular design ensures scalability and adaptability for various production environments, making it a critical tool in quality assurance and process optimization within the solar energy sector.

In Figure 3, measurements were made with two flash boxes, control units of two monitor cells and a HALM evaluation software PVC as an IV meter. Figure 3 shows a configuration of how test integration is done for a typical cell.

Figure 2: HALM device image

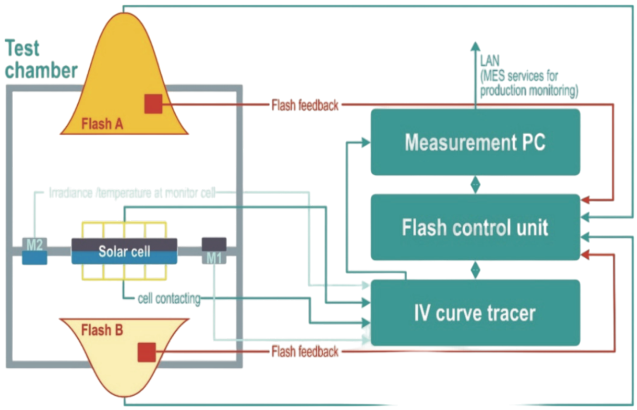

Figure 3: HALM Device Integration

2 Materials

Within the scope of this paper, minimizing operator errors in panel production lines and improving quality processes to increase production cycle efficiency are examined. The focus of the research is on understanding the challenges of faulty quality decisions and delays in production lines. Addressing these issues, the evaluation of machine parameters and quality class decisions presents an opportunity to optimize the workload of employees and improve organizational efficiency. These improvements have the potential to enhance production planning, reallocate human resources to other critical departments, and increase overall productivity. The research utilizes weekly production data from the Cell Production Department at the Kalyon PV factory, obtained using the HALM device. This dataset includes both visual evaluation criteria and efficiency-related parameters. The HALM device, equipped with an industrial automation camera, records images of the cells using a flash and conducts image analysis for visual evaluation. This automated process eliminates reliance on human error, ensuring consistent quality evaluation and accuracy, particularly in managing daily production volumes. By leveraging this technology, the study aims to streamline production processes and enhance overall operational efficiency.

Figure 4 above contains summary information of the processes and flows performed at all stages. The process, which starts with obtaining the data set, consists of pre-processing, analysis, model selection, training, model evaluation, training of the final model, monitoring the model and keeping it updated.

2.1 Provision and Promotion of the Data Set

In the study, the machine measurement parameters values in the data set of the Halm machine of the cell factory located in the Kalyon PV production facility were transferred to our table to be evaluated for inclusion in the production management system.

In the MES (Manufacturing Execution System), a data set comprising 1,288,007 cell parameters, representing one week's production data, was employed for evaluation purposes. This data set is used for cell quality classification and efficiency estimation based on the provided parameters.

The dataset consists of 30 variables identified for analysis, 20 of which are of type (int64) and 10 of type (varchar), all included

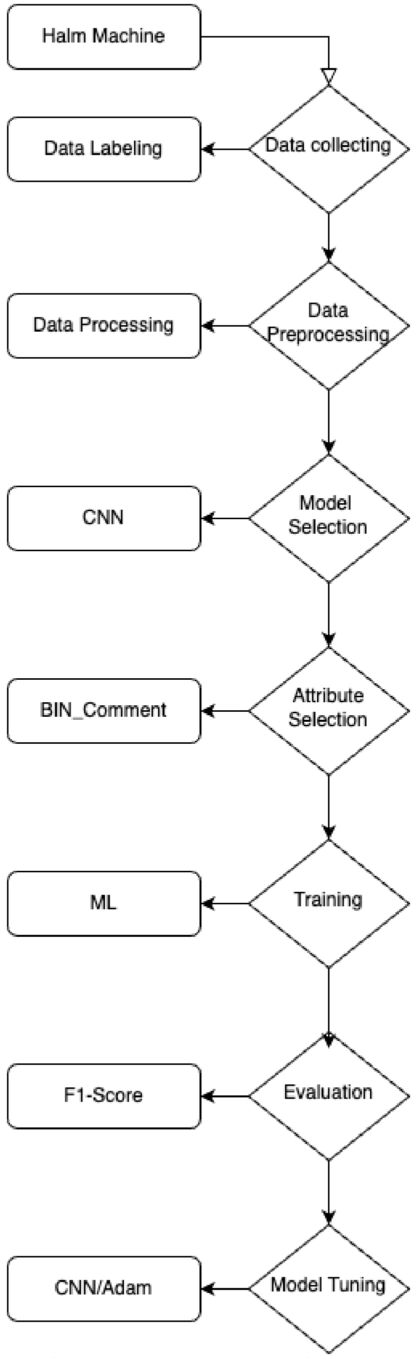

Figure 4: Process Flow Diagram

Table 1: Halm Machine Variables and Description

Variable Name (type)	Description
Measurement Time	Flash Burning Time
Multiple Flash Method	Multiple Flash Reflections
MPP Scanning Method	Full Directional IV Scan
Rheinland Dynamic Current Voltage Value	Cell Type and IV Chart Curve

as columns. Detailed descriptions and names of these columns are provided in the study. Additionally, images backed up within a shared folder structure are integrated into the system for visual evaluation. These images, coupled with the analysis of parameters, enable the identification of quality impacts on final panels and cells. Visual evaluations, performed through advanced image analysis, ensure reliable quality assessments free from human error, further enhancing production efficiency and product consistency.

Table 2: Halm Machine Variables and Description

Variable Name (type)	Description
UniqueID	Halm Makine ID
Title	Title
TestTimeDate	Test Time
BatchID	Batch ID
CellIDStr	Cell Soldering Machine ID
Operator	Operator
Comment	Line Information
BIN	Efficiency Value
Isc	Short Circuit Current
Jsc	Short Circuit Current Density
Uoc	Open Circuit Voltage
FF	Fill Factor
Eta	Efficiency
Impp	Current at Maximum Power Point
Umpp	Voltage at Maximum Power Point
Pmpp	Power at Maximum Power Point
Ivld1	Forward Light Curve
Jvld1	Measurement of Current Density
Ivld2	Measurement of Voltage Density
Jvld2	Measurement of Current Density 2
RshuntLf	Shunt/parallel resistor, forward light
RserLf	Series resistance forward light curve
$BIN_Comment$	Quality Class Value

2.2 Transferring Data to the System

After transferring the data set from the machine to the database, after determining the parameters affecting the quality criteria, the necessary database table for the parameters selected for machine learning was created, and with the help of a third-party application named FC, the data was transferred into the table named `h.a.l.m_machine_AI` on Microsoft SQL Server. PyCharm CE development environment was deemed appropriate, and developments were made on this IDE. Code management was carried out systematically via GitHub, as code versions needed to be considered and updated during the process. Additionally, in Microsoft SQL Server, `pandas` and `sqlalchemy` libraries were used to transfer data from the table, and `requests` libraries were used for the API. After these processes, it is brought to the appropriate `.csv` format ready for processing and made ready in the environment for processing. This process provides appropriate conditions for data processing and analysis.

2.3 Digitization of data

In this study, data is also divided into categorical and numerical data. For example, the quality value of a cell appears as categorical data such as (A, B, C). These values are evaluated numerically.

Table 3: Quality Data Set Digitization

Variable Name (type)	Value
A	1
B	2
C	3

2.4 Data visualizations

In order to transfer the log values of the detected errors in Figure 5 to the database and analyze the density of the data, error types can be grouped with the histograms in Figure 6 in order to understand the interpretations related to the numerical properties of the data set.

This section describes the algorithm to detect defects in the EL images of the cell and evaluate the image quality of the solar panels' cells. Quality standards are detailed in accordance with Kalyon PV production standards.

2.5 Pixel based Defects

As shown in Figure 7, pixel-based defects include cracks, finger cuts, and darkness errors. For each of these defects, the algorithm calculates a defect probability for each pixel of the cell. Different threshold evaluations are applied to distinguish various defect types. To evaluate the effect of the pixel being classified as a defect, defect areas are highlighted in white in the corresponding images.

2.6 Cell based defects

Figure 8 illustrates examples of cell-based defects, such as scratches on the cell, ghosting, and o-ring errors. These defects are identified and analyzed in the dataset using the hand image. The visualized examples demonstrate the specific characteristics

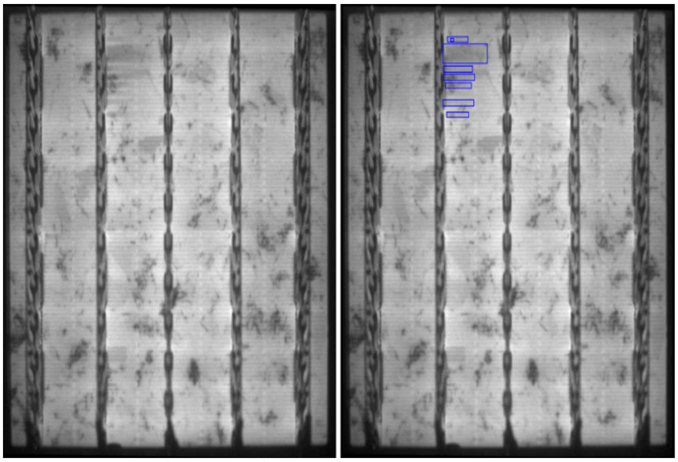
Figure 5: Cell Error Type Example

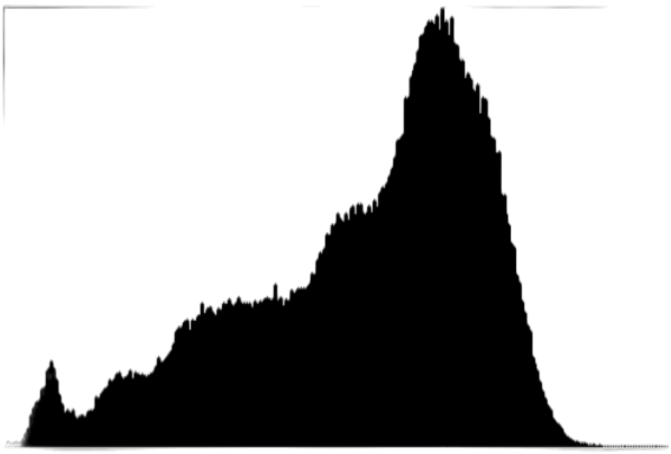
Figure 6: Distribution of Error Types

of each defect type, aiding in accurate defect categorization and evaluation.

3 Methods

3.1 Model Selection

The selection of models suitable for this study involves interpreting the data set and applying traditional machine learning classification and the CNN method to group the data and determine its quality. CNN is one of the most widely used deep learning models for processing visual data and is highly effective in detecting and classifying defects in microscopic images of solar panel cells. Transfer learning, which reuses a model previously trained on large data sets and adapts it to a new problem, is particularly useful for addressing challenges caused by data scarcity in specific areas of solar cell analysis. This approach reduces training time and generally achieves high accuracy rates.

3.2 Feature Extraction (Variable Selection)

In this study, the "BIN_Comment" variable was determined as the target variable, representing the output of the analysis from the evaluated data set. This involves identifying the factors affecting this variable and ensuring that the data set is appropriately formatted for modeling. Correlation is a statistical concept that measures the direction and strength of the relationship between two or more variables. It helps us understand whether a relationship exists between variables, and if so, its direction and strength. The correlation coefficient numerically represents this relationship [1]. **Positive Correlation:** A relationship between two variables where an increase in one variable corresponds to an increase in the other. The correlation coefficient takes a positive value (between 0 and +1).

Negative Correlation: A relationship between two variables where an increase in one variable corresponds to a decrease in the other. The correlation coefficient takes a negative value (between 0 and -1).

Zero Correlation: A lack of relationship between two variables, where the correlation coefficient is 0.

The red line correlation above helps us interpret its influence on the `BIN_Comment` variable in the evaluation of the cells. It can be used to make decisions regarding the factors affecting this parameter during the analysis stage. These values indicate

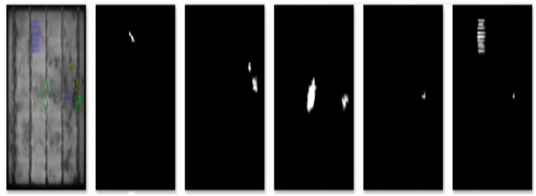

Figure 7: Pixel Based Error Examples

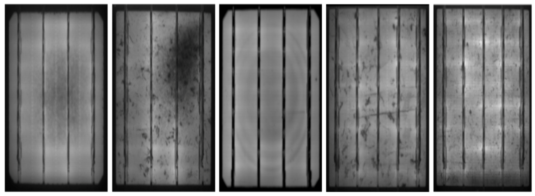

Figure 8: Cell-Based Error Examples

that they will play a significant role in providing insights into which factors have a greater impact on quality classification [2].

3.3 CNN Hyperparameter and Optimization Techniques

This method is a learning architecture that is extremely effective, especially in areas such as image processing and classification. Cell artificial intelligence evaluation can make predictions necessary for learning and identifying cell error types [3].

It consists of four main layers:

- **Convolution Layer:** This layer creates feature maps by applying various filters (kernels) on the input images. Filters detect edges and corners in the image.
- **Activation Function:** After each convolution layer, a non-linear activation function is applied, typically ReLU (Rectified Linear Unit). This helps the model learn complex features.
- **Pooling Layer:** This layer reduces the computational and work costs by decreasing the size of feature maps while preserving important features. The most common pooling operation is max pooling, which selects the maximum value in each subregion.
- **Fully Connected Layer:** This layer receives and evaluates the features obtained from the convolution and pooling layers, increasing the generalization and classification ability of the model.
- **Dropout Layer:** This layer disables random neurons during training to prevent overfitting.

The model visualized in this study was implemented with the help of the Python programming language and its powerful libraries `pandas`, `numpy`, and `sklearn`. These support libraries served as a cornerstone in the creation of the model, enabling the application of various statistical methods and traditional machine learning techniques. During the process, a laptop with a MacBook M2 Operating System and 16GB RAM was used as hardware, and the test environments were carried out on setups installed on this computer.

In this study, five different methods were tried. A total of$1,288,007$ data parameters and EL images were evaluated. Following the pre-processing stages specified within the scope of the study, classification analysis was carried out using machine learning methods. In this context, two different analysis methods were chosen to evaluate the results obtained and analyze their accuracy. The compatibility ratio of cell machine values was also evaluated as a criterion. Accuracy, recall, F1-score, and error rate matrices were used in the evaluation process.

Accuracy: It is the ratio of the model's correct predictions to the total predictions. It is used to evaluate overall performance. Its expression is as follows [4]:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Recall: It shows how many of the positive examples in real data the model correctly identifies. This is an important criterion, especially in unbalanced data sets [5].

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

F1-Score: It is the harmonic mean of precision and recall for a model. It is especially useful when working with unbalanced data sets where there are significant differences between classes, as it provides a balanced summary of both metrics [6].

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

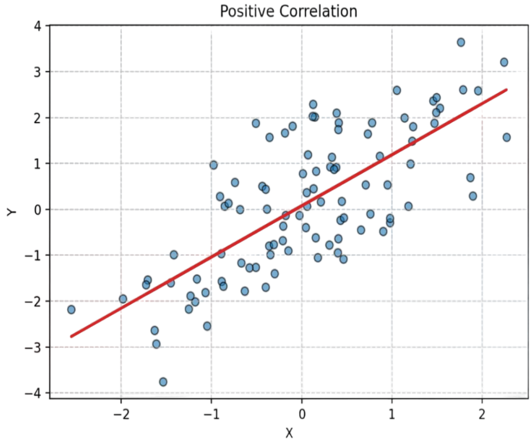

Figure 9: Correlation Heat Plot

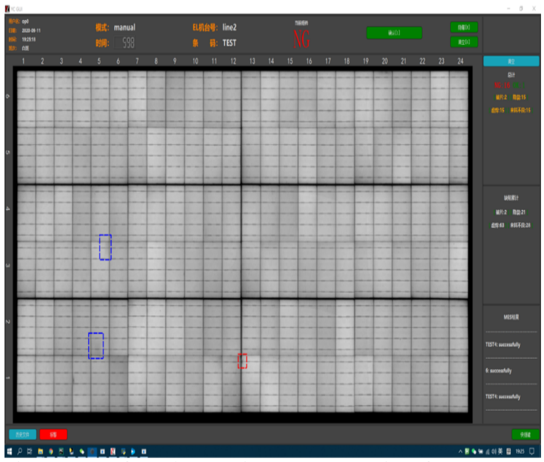

Figure 10: Cell Analysis on Panel with CNN model

4 Results and Discussion

We conducted within the scope of the metrics and methods we evaluated, we first prepared the quality groups to evaluate them numerically in the normalization phase, then EL was evaluated for the visual evaluation method, and after a printout was made and made ready for machine learning, machine learning methods were applied [7]. Two different evaluation tables were created to evaluate the methods used. Relevant analysis results for performance evaluation are given in Table 3.

Table 4: Comparison of Machine Learning Methods.

Method	Validation	Precision	Recall	F1 Score
Logistic Regression	0.76	0.75	0.76	0.76
SVM	0.61	0.59	0.61	0.61
Decision Tree	0.59	0.60	0.60	0.60
Random Forest	0.77	0.78	0.77	0.77
CNN	0.82	0.78	0.82	0.82

Evaluation of quality classes taught with real data and comparison of methods are given in Table 5.

As a first step in this study, the relevant machinery, equipment, operating principles, and integration design were detailed. The data for this quality assessment was obtained from the database of the HALM machine, which provides measurement and value assignment in EL images.

In the second stage, a normalization method was applied. Quality values were used to prepare the data for the numerical evaluation of letter outputs. The `BIN_Comment` field was selected as the key data column, representing the major decision point of the dataset output. This field was considered as an attribute of the data. By evaluating the visual quality output value from the HALM device alongside 20 additional parameter fields, this selection revealed both the quality and efficiency values of the cell. Various machine learning models were employed to evaluate these results [8].

Logistic Regression, Support Vector Machine (SVM), Random Forest, Decision Tree, and CNN methods were applied as traditional and deep learning approaches. The results of these analyses, including the compatibility rates of the data with the actual analysis values, are shown in Table 5. Among these, the CNN method demonstrated the highest compatibility,

Table 5: Comparison of Machine Learning and Deep Learning Methods

BIN_Comment	Logistic Regression	SVM	Random Forest	Decision Tree	CNN
22 A	70.00	69	60	59	71
22.1 A	70.00	64	61	57	74
22.2 A	75.00	63	66	65	84
22.3 A	73.00	65	62	58	69
22.4 A	70.00	67	64	63	69
22.5 A	75.00	69	61	64	78
22.6 A	73.00	66	62	57	83
22.7 A	75.00	68	61	57	74
22.8 A	75.00	62	61	57	70
22.9 A	74.00	68	63	62	81
23 A	74.00	69	61	62	79
23.1 A	74.00	69	65	64	74
23.2 A	71.00	63	62	59	72
23.3 A	74.00	68	64	56	79
23.4 A	72.00	68	61	63	77
23.5 A	74.00	69	65	64	72
23.6 A	74.00	68	61	59	82
23.7 A	70.00	68	61	59	72
23.8 A	70.00	69	62	56	77
23.9 A	71.00	62	61	62	72
24 A	70.00	62	61	62	72

achieving a suitability rate of 78%. This method provided significant contributions to the study by enabling visual evaluation and interpretation of the estimated values [9].

Convolutional Neural Networks (CNNs) are particularly powerful for visual data analysis and classification. Performance is enhanced by selecting appropriate hyperparameters and applying suitable optimization techniques [10]. The CNN model used in this study has the following configuration:

- The first convolutional layer contains 32 filters of size 3×3.
- The second convolutional layer contains 64 filters of size 3×3.
- The third convolutional layer contains 128 filters of size 3×3.
- A 2×2 MaxPooling layer was applied for downsampling.
- A dropout ratio of 0.5 was used to randomly disable 50% of neurons in the fully connected layer to reduce overfitting.
- The `ReLU` activation function was used in the convolution and flatten layers.
- A `sigmoid` activation function was used in the output layer for binary classification.
- The input size (`input_shape`) of the model is 64×64 pixels with 3-channel RGB images.
- The Adam optimizer was employed with a learning rate of 0.001 to provide stabilization and increase performance [11] [12].

To diversify the training data, the following data augmentation techniques were applied:

- **Rescaling:** Images were normalized to a range of 0 to 1 by multiplying pixel values by $1/255$.
- **Shear Range:** Images were distorted using shear transformations.
- **Zoom Range:** Images were randomly zoomed.
- **Horizontal Flip:** Images were randomly flipped horizontally.

The training of the model was carried out using the `model.fit` function. The training data was supplied via the `train_generator` object. The following parameters were used:

- `steps_per_epoch`: Calculated by dividing the total number of training data by the batch size.
- `epochs`: The number of times the model trains on the entire dataset, set to 25.
- `validation_steps`: The number of steps per validation step, calculated by dividing the number of validation data by the batch size [13], [14].

Table 4 presents the comparative analysis of various machine learning and deep learning methods for quality assessment in EL image evaluation. Among traditional machine learning methods, Random Forest and Logistic Regression performed well with validation scores of 0.77 and 0.76, respectively. However, CNN outperformed all other methods, achieving the highest validation score of 0.82.

Table 5 further illustrates a detailed comparison across different quality classes. CNN consistently yielded higher accuracy rates, surpassing 82% for several quality bins, while other methods generally achieved scores between 60% and 75%.

The study's methodology, incorporating normalization, hyperparameter optimization, and data augmentation techniques, significantly contributed to the robustness of the CNN model. The findings indicate that integrating advanced machine learning techniques into production processes can enhance quality control mechanisms, improve resource utilization, and increase production efficiency in the photovoltaic industry [15]–[18].

Authors' Contributions

Murat Simsek: Conceptualization, Methodology, Software. Ozel Sebetci: Data curation, Software, Writing - original draft. Irfan Yilmaz: Software, Writing - original draft.

Competing Interests

The authors declare that they have no conflict of interest.

References

[1] W.-C. Hong, P.-F. Pai, Y.-Y. Huang, and S.-L. Yang, "Application of support vector machines in predicting employee turnover based on job performance," in *Advances in Natural Computation*, 2005, pp. 668–674.
[2] B. L. Aylak, O. Oral, and K. Yazıcı, "Yapay zeka ve makine Öğrenmesi tekniklerinin lojistik sektöründe kullanımı," *El-Cezeri Journal of Science and Engineering*, vol. 8, no. 1, pp. 74–93, 2021.
[3] M. A. Green, *Solar Cells: Operating Principles, Technology, and System Applications*. University of New South Wales Press, 2018.
[4] E. E. Antunez, J. Gonzalez-Hernandez, and A. Dominguez, "Artificial neural network modeling of a photovoltaic panel considering temperature effects," *International Journal of Energy Research*, vol. 43, no. 12, pp. 5939–5950, 2019.
[5] H. H. Al-Kayiem and Z. S. Al-Khafaji, "Modeling and optimization of photovoltaic cells using artificial neural networks: A review," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 1811–1820, 2018.
[6] Y. Ma, Y. Yang, X. Yu, Y. Chen, and F. Blaabjerg, "Artificial intelligence for energy management in future smart grids: A review," *Renewable and Sustainable Energy Reviews*, vol. 104, pp. 62–72, 2019.
[7] X. e. a. Liu, "Automated defect detection for photovoltaic modules using convolutional neural networks," *IEEE Transactions on Industrial Informatics*, 2020.
[8] M. Garcia *et al.*, "Principal component analysis for solar cell defect detection," in *Procedia Manufacturing*, 2018, pp. 34–49.
[9] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," 2018.
[10] S. Sharma *et al.*, "Random forest-based approach for solar cell quality prediction," *Renewable Energy*, pp. 3–7, 2017.
[11] J. Zhang *et al.*, "Transfer learning for defect detection in solar cells using deep convolutional neural networks," *Solar Energy*, 2019.
[12] D. P. Kingma, "Adam: A method for stochastic optimization," 2014, arXiv preprint.
[13] F. Itano, M. A. de Abreu de Sousa, and E. Del-Moral-Hernandez, "Extending mlp ann hyperparameters optimization by using genetic algorithm," in *Proceedings of the 2018 International Joint Conference*, 2018, pp. 66–67.
[14] K. Ramspeck, S. Schenk, L. Komp, A. Metz, and M. Meixner, "Accurate efficiency measurements on very high efficiency silicon solar cells using pulsed light sources," in *Proceedings of the 29th European Photovoltaic Solar Energy Conference and Exhibition*, 2014, pp. 1253–1256.
[15] Z. Lu and J. Wu, "A review of artificial intelligence applications in marine renewable energy systems," *Renewable and Sustainable Energy Reviews*, vol. 117, p. 109469, 2020.
[16] C. Monokroussos, R. Gottschalg, A. N. Tiwari, G. Friesen, D. Chianese, and S. Mau, "The effects of solar cell capacitance on calibration accuracy when using a flash simulator," in *Proceedings of the 4th WCPSEC*, vol. 2, 2006, pp. 2231–2234.
[17] C. Y. J. Peng, K. L. Lee, and G. M. Ingersoll, "An introduction to logistic regression analysis and reporting," *The Journal of Educational Research*, vol. 96, no. 1, pp. 3–14, 2002.
[18] G. Rong, G. Mendez, E. B. Assi, B. Bajic, and N. V. Chawla, "Understanding sbml's flat line: An analysis of loss functions for oversampling-based classifiers in imbalanced datasets," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 5372–5379.